Effect of an Internet-Based Multimedia Teacher Development Program in Enhancing Teachers' Assessment Literacy

Nan Huai
Department of Educational Psychology/
Wisconsin Center for Education Research
University of Wisconsin–Madison
nhuai@wisc.edu

Jeffery P. Braden
Department of Educational Psychology/
Wisconsin Center for Education Research
University of Wisconsin–Madison
jeff_braden@ncsu.edu

Jennifer White
Department of Educational Psychology/
Wisconsin Center for Education Research
University of Wisconsin–Madison
Jlwhite2@wisc.edu

Stephen N. Elliott
Department of Educational Psychology/
Wisconsin Center for Education Research
University of Wisconsin–Madison
snelliot@wisc.edu

Effect of an Internet-Based Multimedia Professional Development Program in Enhancing Teachers' Assessment Literacy

Nan Huai, Jeffery P. Braden, Jennifer White, and Stephen N. Elliott

Introduction and Literature Review

The goal of this study was to evaluate the efficacy of an Internet-based multimedia professional development course in improving teachers' knowledge and skills in educational assessment and inclusive practice in assessment. A secondary goal was to investigate teachers' use and acceptance of the format of professional development (e.g., traditional workshop vs. online course).

The course, Assessing One and All (AOA; Elliot & Braden, 2001), aims to enhance parents' and teachers' general assessment literacy, knowledge about and tactics for administering testing accommodations, and understanding of alternate assessment. The online AOA course consists of five components:

- Cases;

- Unit 1: Educational Assessment Principles and Practices;

- Unit 2: Large-Scale Assessment;

- Unit 3: Accommodations and Alternate Assessments; and

- An activity book.

The case component of AOA is a multimedia presentation of three cases. Using video of teacher and parent interviews and of student performance in testing, the cases illustrate the inclusion of students with and without disabilities in statewide accountability assessment. AOA provides hyperlinks to related informational Web sites and texts. A book, *Assessing One and All* (Elliott, Braden, & White, 2001), accompanies the Web-based course.

Presuppositions for Evaluation Study of AOA

AOA is a professional development program incorporating new technological features that were suggested by prior researchers (e.g., Putnam & Borko, 2000; Mott, 2000; Killion, 2000) to bring about effective professional learning. Accordingly, in this section we review relevant literature on professional development and online adult learning. We also discuss the features of AOA in relation to the models and guidelines derived from prior research.

Professional Development and Inclusive Assessments

In a review of situated learning and its application to effective teacher education, Putnam and Borko (2000) drew the following conclusions: First, practicing teachers learn best when the

learning task is an integral part of the physical and social context. Both real classroom practice and case-based learning serve as appropriate contexts for teacher professional development. Second, learning is effective when teachers have the opportunity to join a professional community with shared goals. Professional communities shape the way teachers view and practice their craft. Third, appropriate tools assist teachers' learning. For example, pedagogical tools such as multimedia systems, hypertexts, and hyperlinks provide new and flexible ways to present and connect information. Putnam and Borko specifically pointed out that case-based activities presented by hypermedia tools "allow teachers to explore the richness and complexity of genuine pedagogical problems" (p. 8).

Mott (2000), in a review of several professional development models, stated that successful continuing education programs (a) reflect the changing environments present in the professional practice, (b) emphasize relevant formative and summative self-assessment, (c) are situated in and based on practice, (d) promote collaboration and stress the community of practice, and (e) address present as well as future needs.

The models proposed by Putnam and Borko (2000) and Mott (2000) address similar issues, such as collaboration and situated learning, while also complementing each other. For example, Putnam and Borko captured the marriage between technology and cognitive theory, whereas Mott pointed out the importance of change and needs in social context. Together, the two models offer comprehensive criteria for evaluating AOA. Table 1 presents elements of the two models and illustrates application of the models to AOA.

Table 1
Professional Development Models and AOA

Key elements of professional development	AOA	Putnam & Borko model	Mott model
1. Situated learning	[a]√	√	√
2. Pedagogical tools as facilitators of learning	√	√	[b]×
3. Collaborative learning in a learners' community	√	×	√
4. Emphasis on formative and summative self-assessment	√	×	√
5. Changing practice environments reflected	√	×	√
6. Both present and future needs addressed	[c]?	√	√

[a]√ = the element is addressed. [b]× = the element is not addressed. [c]? = the element is not addressed exactly as described in the theoretical model(s).

Situated learning. It has been widely accepted that adult learning should be situated in genuine contexts (e.g., Daley, 2000; DuFour, 2001; Mott, 2000; Putnam & Borko, 2000). Case examples, as well as daily classroom practice, provide valuable contexts for effective teacher learning (Daley, 2000; Putnam & Borko, 2000).

AOA incorporates three cases representing students with and without disabilities situated in different states. Case materials include students' products, assessments, and individualized

education program (IEPs) and video interviews with parents and teachers. The AOA Activity Book provides keys to the activities based on the three cases as models to assist participants in developing their own activity books. In the present study, participants were expected to explore the intended learning subjects (educational assessment principles and practices, large-scale assessment, accommodations and alternate assessments) through the AOA cases and eventually to transfer the acquired knowledge to their own contexts (i.e., developing their own activity books based on real cases). Working through the AOA Activity Book is considered an authentic task because it fosters the thinking and skills that are important in real-life situations and situates the activities in teachers' work settings.

Pedagogical tools as facilitators in learning. One of the features that distinguishes AOA from traditional professional development programs is its technological element. Putnam and Borko (2000) stated that multimedia systems, such as hyperlinks and hypertexts, serve as pedagogical tools by providing new and flexible ways to connect and present information. They specifically pointed out the advantages of using multimedia materials to present cases. Case materials presented this way are no longer isolated and unidimensional. Instead, they better approximate the complexity of real-life situations, and they allow learners to explore the richness of the intended subjects.

As an Internet-based course, AOA contains more than 50 links to internal and external Web pages related to any given topic. AOA's three cases offer rich information, ranging from videos of students' test performance to written permanent products. Hyperlinks to the cases are provided throughout the AOA site. Participants can easily refer to cases for context-related information as they explore the knowledge content (i.e., educational assessment principles and practices, large-scale assessment, accommodations, and alternate assessments).

Collaborative learning in a learners' community. Social cognition theory (Cobb, 1994) considers learning as a process, in which individuals come to know how to participate in the discourse and practice of a particular community. Teachers conceptualize and construct their knowledge and skills by participating in the school community. Collaboration among personnel with diverse expertise—that is, the IEP team—characterizes the decision-making process in inclusive assessment. The AOA Activity Book encourages learners to discuss assessment and inclusion issues with colleagues. However, it should be noted that AOA does not allocate time and space for structured collaborative activities.

Emphasis on formative and summative self-assessment. Activities in the AOA Activity Book are designed to facilitate participants' application of the knowledge acquired through the AOA course. In addition, based on the Activity Books provided for the three cases, participants might evaluate their answers to their own AOA Activity Books to assess their learning and understanding in the course. Activities 1.1–1.5, 2.1–2.4, and 3.1–3.3 address AOA's three course units: (a) educational assessment principles and practices, (b) large-scale assessment, and (c) accommodations and alternate assessments. Activities 4.1–4.3 then require comprehensive application of knowledge and skills acquired through the three course units to a real case.

Changing practice environments reflected. AOA was created in response to the mandatory inclusion of students with disabilities in educational accountability systems. Inclusive practice in instruction has long existed. However, students with disabilities often have been

excluded from large-scale assessment (Thurlow, 2000). As states started to adopt educational standards in early 1990s to respond to the call for higher and more rigorous standards for students (Thurlow, 2000), score-reporting systems and accountability mechanisms were also implemented to determine students' progress toward the standards via large-scale assessments. As more states tied high stakes (e.g., graduation) to testing, educators and parents became cognizant of the problems caused by the exclusion of students with disabilities (Thurlow, House, Scott, & Ysseldyke, 2000). The 1997 amendments to the Individuals With Disabilities Education Act (IDEA '97) require that IEPs mandate the inclusion of students with disabilities in large-scale assessments and specify appropriate accommodations or alternate assessments. However, state (e.g., Hollenbeck, Tindal, & Almond, 1998; Johnson, Kimball, Brown, & Anderson, 2001) and national (Plake, 1993) research revealed limitations in teachers' assessment competency. Within the context of current legal requirements and deficits in teachers' assessment literacy, AOA is appealing in that it addresses both the needs of the profession and the needs of individuals in the profession.

Both present and future needs addressed. AOA was created to address the need to improve teachers' assessment competency. As standards-based reform continues in the United States, large-scale assessments and the inclusion of students with disabilities will continue to be imperative components of educational accountability systems. As an online course, AOA can be easily updated to reflect policy changes and trends, thus dynamically addressing the changing needs in professional practice.

Online Education

Computer and computer-related technology have developed dramatically over recent years. Researchers generally agree that Internet-based instruction brings both promises and perils to designers and learners (Killion, 2000).

Successes. The development of information technology offers appealing new ways to meet professional education needs. For example, the Collaborative Teacher Education Program (CTEP) at Indiana University has adapted a Web-based course to enable educators in remote areas to access professional development programs without coming to campus (Rodes, Knapczyk, Chapman, & Chung, 2000). Investigators reported that after careful planning and gradual introduction of Web-based learning, distance education could be a successful vehicle for professional development programs.

Michigan trained more than 1,000 in-service teachers using a program that combined an Internet-based multimedia course with peer support (Hoffman & Thompson, 2000). Participants watched a video of an intended practice and then shared their thoughts with other teachers by going to a virtual discussion area built into the course. This program, named Teaching for Tomorrow (TFT), offered flexibility (in both time and content) for teachers, schools, and administrators (Hoffman & Thompson, 2000). Participants indicated their preference for the flexible learning arrangement and peer support.

Researchers have achieved consensus on the beneficial features of Web-based education, such as convenient access and flexible time and pace (Burgstahler, 1995; Killion, 2000; Matthews, 1999). Specifically, online professional development programs can benefit educators

in remote areas, who may find it difficult to attend traditional workshops due to various constraints (e.g., time, location, financial costs) (Killion, 2000). In addition, an online course is available 24 hours a day. Thus, learners can flexibly schedule their time and pace of learning according to their needs (Killion, 2000; Matthews, 1999).

Researchers (e.g., Burgstahler, 1995; Killion, 2000; Sujo de Montes & Gonzales, 2000) have also pointed out that the Internet offers a new and convenient vehicle for communication and collaboration (e.g., e-mail and bulletin boards). Successful Internet-based professional development programs such as CTEP at Indiana University (Rodes et al., 2000) and TFT in Michigan (Hoffman & Thompson, 2000) have incorporated virtual collaborative activities as an integral part of the course.

Although AOA encourages learners to discuss issues with their colleagues, it does not allocate virtual space (e.g., a bulletin board or an electronic mailing list) for group activities. This omission may limit the effectiveness of the online AOA course.

Challenges. Although technology offers new opportunities for professional development programs, several issues may pose significant obstacles to the effectiveness of such programs. First, Killion (2000) pointed out that not all professional development programs can be adapted to online programs. Programs requiring constant practice, coaching, and modeling are likely to be less successful in an online format than in traditional live instruction.

Second, the effectiveness of distance education depends, to a great extent, on teachers' readiness to enter an online learning environment (Killion, 2000). Lacking a live instructor, online learning is largely a self-directed process. Hyperlinks enable teachers to assume full control of the pace and content of a course. Whereas this feature serves as an advantage with regard to flexibility and access to information, it can also create confusion, to the degree that it creates a learning experience lacking in logical and hierarchical order (Killion, 2000). Rodes et al. (2000), reflecting on the CTEP program, pointed out that most teachers are not well prepared to take an active and self-directed role in Web-based programs, and thus designers of such programs must design activities that ease teachers' transition from passive to active learners.

Failure to actively participate in online learning is often associated with lack of skill in technology (Rodes et al., 2000). Moreover, technology skills may become another learning task for teachers to master before they can learn the intended content of the professional development program. Oelrich (2001) acknowledged that teachers might be resistant to online delivery methods due to a lack of confidence in their technology skills. Reviews of several successful teacher professional development programs (e.g., Bohnenkamp & McMahon, 2001; Rodes et al., 2000) have shown that it is important to provide technology skill training and scaffolding at the beginning of online learning process.

Synthesis of Presuppositions

The AOA program was developed in response to changes in education policy—namely, the mandatory inclusion of students with disabilities in accountability systems and the consequent demand for expertise in assessment. The program incorporates three simulated cases to create meaningful settings, which facilitate the application and exploration of the intended

knowledge and skills. The AOA Activity Book provides both practice-based authentic tasks and formative and summative assessment.

Technology is a central feature of AOA. Hypertext, hyperlinks, and multimedia enrich the information presented and offer opportunities for participants to explore the subjects in a variety of ways. The delivery of AOA via the Internet provides participants with flexibility regarding their time, space, and progress of study. The content of AOA can be easily updated to meet new needs.

However, potential limits of AOA should also be noted. First, although the developers encourage participants to discuss with colleagues several topics in the program, there are no specifically allocated virtual or real resources for collaboration. Second, AOA does not include a phase addressing learners' readiness for online learning, which is considered essential by Web-based education program developers (e.g., Rodes et al., 2000).

 In the current study, we used pre- and post-learning assessment surveys, permanent product reviews (e.g., IEPs, Activity Book), focus group discussions, and individual interviews to assess whether the AOA program (a) improved participating teachers' knowledge of and self-efficacy in educational assessment principles and practices, large-scale assessment, and accommodations and alternate assessments, (b) enhanced participants' appropriate use of testing accommodations and/or alternate assessments with students with disabilities, and (c) is an acceptable tool for professional learning.

Suppositions

Supposition #1

We began with the assumption that AOA would improve participants' knowledge of assessment principles and practices, large-scale assessment, and accommodations and alternate assessments. We expected that (a) for participants who accessed the AOA course (i.e., the instructional group), their knowledge and attitude scores would show more improvement from pre- to post-test than scores obtained by those who did not access the course (the comparison group), (b) controlling for knowledge of and self-efficacy in assessment issues prior to participation in AOA learning, the instructional group would score higher than the comparison group on the core activities, and (c) in the instructional group, participants who were more experienced with the Internet would improve their attitude and knowledge more than those with less Internet experience.

Supposition #2

We assumed that AOA would increase participants' appropriate use of testing accommodations and/or alternate assessments as evidenced by their IEPs for students with disabilities. We expected that IEPs revised by participants after completing the AOA course would show better decision making with regard to students' participation in large-scale assessment than IEPs written before participants' exposure to AOA.

Supposition #3

We assumed that practicing teachers would favor the Web-based AOA course for its user-friendly features such as flexible time and learning pace. We also expected that instructional group participants would accept an Internet hypermedia approach to continuing professional development as a desirable alternative to traditional learning tools.

The methods used to test these suppositions are described next.

Methods

Participants

Education leaders at three sites in three states (Arizona, South Carolina, and Wisconsin) recruited educators to participate in the study. Initially, a total of 89 people were recruited. Thirty-six participants joined the study in Arizona, with 15 volunteering for the comparison group and 21 for the instructional group. A total of 42 participants were recruited from South Carolina, with 21 volunteering for the comparison group and 21 for the instructional group. The third sample included 11 participants from Wisconsin, where a comparison group was not recruited due to the limited availability of participants. Due to participant dropout, the overall final sample included 55 participants, 26 in the comparison group and 29 in the instructional group. The attrition rate was 27.8% for the comparison group and 45.3% for the instructional group. At an alpha level of .05, the attrition rates for the two groups did not differ significantly: $(\chi^2 (1, N = 89) = 2.783, p = .095)$.

Demographic information was obtained from the pre- and post-learning surveys returned by participants. A total of 55 participants from the three sites reported their demographic information, which is summarized in Table 2.

The participants were also asked about the frequency of their participation in the IEP process and their access to the Internet. In both the instructional and the comparison groups, the majority of participants (88.5% of the comparison group, 96.6% of the instructional group) reported that they had participated in the IEP process *frequently* or *sometimes*. Among the comparison group participants, 23.1% had access to the Internet either at home or at the office but not both, and 76.9% had access at both locations. In the instructional group, 17.2 % had access to the Internet at only one location, whereas 82.8% had access at both locations.

Materials and Instruments

The AOA course was offered through the Council for Exceptional Children (CEC) Web site. The CEC webmaster assigned each of the instructional group participants a password and a username, which they used to access the AOA Web site for 3 months. Instructional participants were also given the book that accompanies the site, *Assessing One and All* (Elliott et al., 2001). They were asked not to share access to the Web site or the book with participants in the comparison group.

Table 2
Demographic Information

Group	Gender		Ethnicity[a]			Profession		
	Female	Male	Caucasian	African American	Other	Teaching staff[b]	Support staff[c]	Administrator[d]
Comparison group								
N	24	2	23	1	1	17	8	1
% within group	92.3%	7.7%	88.4%	3.8%	3.8%	65.4%	30.8%	3.8%
Instructional group								
N	26	3	24	5	0	16	6	7
% within group	89.7%	10.3%	82.8%	17.2%	0%	55.2%	20.7%	24.1%

[a]One participant in the comparison group failed to provide information on race (3.8% of 26). [b]Teaching staff include any position that involves delivery of instruction to students (e.g., general and special education teachers). [c]Support staff include those who do not directly participate in academic instruction, such as school psychologists, social workers, and counselors. [d]Administrators include persons in leadership positions, such as principals, assistant principals, and special education coordinators.

Pre- and Post-Learning Assessment Survey

The pre- and post-learning assessment surveys measured participants' knowledge and self-rated competency level in general assessment and inclusive assessment practices. The survey also required the participants to report demographic information such as gender, ethnicity, and profession. The survey items were classified into three groups:

1. *Items probing participants' knowledge on assessment issues*. The first group included Items 1 to 15, requiring *true*, *false*, or *not sure* responses, and Items 26 and 27, requiring short written responses. Scores derived from this group had a possible score range of 0 to 19. The internal consistency for this group of items was .67.

2. *Items measuring self-efficacy with regard to assessment knowledge and skills.* The second group included Items 16 to 23, using four-level Likert scale ratings. The Likert ratings were labeled with terms drawn from states' proficiency standards so that the terminology would be consistent with both the AOA content and state standards. The score derived from this group ranged from 0 to 24. The internal consistency for the items in this group was .84.

3. *Items addressing participants' use and experience with the Internet.* The third group consisted of Items 24 and 25. The possible score range for this group was 0 to 6. The internal consistency for the two items was .68.

The post-learning survey for the instructional group contained six additional items that asked the participants to rate the degree to which the features of the AOA course met their needs.

Focus Group Discussion Protocols and Individual Interview Protocols

Interview questions were generated to probe participants' general perceptions and experiences with educational assessment and related professional development programs. The protocol provided a structure for both the group discussions and the individual interviews. Discussions and interviews were conducted prior to and after the 3-month instruction time. The instructional group discussion protocol administered after participants completed the AOA course included additional questions on the participants' use of and satisfaction with the online course.

Core Activities

The AOA Activity Book is an integral part of the AOA Web site. It consists of activities related to educational assessment principles and practices, large-scale assessment, and accommodations and alternate assessments. Participants in both the instructional group and the comparison group completed three core activities drawn from the AOA Activity Book (Activities 1.4, 2.2, and 3.1). The core activities comprise the essentials of educational assessment principles and practices, large-scale assessment, and accommodations and alternate assessments.

Rubric to Score Core Activities

Scoring rubrics were developed for the core activities. The answer key to Core Activity 1.4, which addresses general concepts in educational assessment and accountability system, was constructed based on the book *Assessing One and All* (Elliot et al., 2001) and the online AOA course. Core Activity 2.2 prompts participants to seek information about their state and district policies on educational assessments and inclusive practices. Due to the differences among state assessment and inclusion polices, the answer key to Core Activity 2.2 was based on information unique to the particular state in which the site was located. Core Activity 3.1 assesses both generic knowledge about inclusive assessment practices and relevant state policies. Its answer key was based on the book and state-specific information. For items concerning state policies, we also verified the scoring keys with the administrators in each site. The rubric for each site was constructed while the instructional group was engaged in the AOA course. Core Activities 1.4, 2.2, and 3.1 had possible score ranges of 0–21, 0–16, and 0–14, respectively. The total score for the core activities ranged from 0 to 51. The internal reliability of the core activity items was .90.

IEP Evaluation Framework

DiPerna, McKevitt, Elliott, and Braden (1999) constructed a framework for evaluating IEPs. In the current study, we used the framework's series of yes/no questions to evaluate whether IEPs included statements of students' participation in large-scale assessment and applicable accommodations and alternate assessments. Three ratings (*explicit*, *vague*, and *none*) were used to describe whether the statement supplied information specific enough to enable reviewers to understand what accommodations and alternate assessments were used.

Only participants in the instructional group were required to provide a copy of an IEP constructed before the AOA course. Instructional group participants were encouraged to note any changes they wanted to make on the IEP during and after taking the course as a means of capturing how they might change their practice after taking the course.

We did not ask participants in the comparison group to supply IEPs for two reasons:

1. IEP making is a group process in which participants from the comparison group might be teamed with participants from the instructional group, thus contaminating the products and obscuring the quality of participants' decision making.

2. We expected a higher level of resistance from comparison group participants due to the time and efforts involved in obtaining an IEP.

Diary of Ongoing Professional Development

We provided participants with a structured diary to help them record their participation in any assessment-related professional development programs. This diary was intended to collect information about contextual variables that might influence the participants' knowledge and skills on educational assessment issues.

All participants kept a diary about the ongoing professional development programs addressing assessment and inclusive practices in assessment. Every 2 weeks during the AOA instruction, we prompted participants via e-mail messages to record ongoing assessment-related professional development programs in the diary. This procedure increased our understanding of AOA's effectiveness by identifying factors external to AOA that may have affected learning in both groups.

Procedures

Local educational administrators within the three school districts in different states helped us recruit the participants. We abandoned our initial plan to randomly assign participants to the instructional or the comparison group after hearing from site administrators that educators would be more likely to participate if they had a choice over their group membership. Participants in the three sites volunteered to join either the comparison or the instructional group. The three sites started the course at different times: South Carolina began in January 2002, Arizona in April 2003, and Wisconsin in November 2002. To compensate for the lack of random assignment, we planned to carry out a time-series design, in which the data in the three sites would be analyzed separately to rule out possible confounding variables and to demonstrate learning gains across the settings. However, due to previously mentioned attrition rates, the sample size at each site was not sufficient to ensure adequate power for these analyses. Therefore, we aggregated the data across the three sites in the final analysis. Table 3 presents the summarized return rate of products.

Table 3
Number and Rate of Research Products Returned

Group	N^c	Number[a] and proportion[b] of returned products			
		Pre- and post-learning assessment survey	Core activities[d]	IEP sample	Professional development diary
Comparison	36	26 (72.2%)	24 (66.7%)	N/A	27 (75%)
Instructional	53	29 (54.7%)	20 (37.7%)	20 (37.7%)	21 (39.6%)
Total	89	55 (61.8%)	44 (49.4%)	N/A	48 (53.9%)

[a]The number of returned products is listed outside of the parentheses. [b]The return rate is listed within the parentheses. [c]The original number of volunteering participants. [d]The Wisconsin participants' surveys, diaries, and post-learning core activities were included in the instructional group.

All participants were asked to (a) complete the pre- and post-learning assessment survey before and after the instructional group completed the AOA course materials, (b) finish the core activities, and (c) keep a record of any professional development programs attended while the instructional group was engaged in the course. Due to the limited availability of participants, we were unable to recruit a comparison group in Wisconsin. Therefore, the Wisconsin participants finished the core activities before receiving the course materials (hereafter referred to as *pre-*

learning core activities). They then completed the core activities again during and after the course (hereafter referred to as *post-learning core activities*). The instructional group participants at all sites were also required to hand in an IEP sample.

Our research grant paid for the instructional groups' access to the online course and the book *Assessing One and All* (Elliott et al., 2001), which together have a retail value of $498. After instructional group participants completed the course, participants in the comparison group who had finished all the required tasks were given the opportunity to take the course free of charge. In South Carolina and Arizona, participants who completed the required tasks were given $50. In Wisconsin, the participants were offered $ 100 because they had to complete more tasks (i.e., the pre-learning core activities).

Administration and Scoring of Pre- and Post-Learning Assessment Survey

One or more researchers visited the sites before the start of the instructional condition and again after the instruction was completed. During the site visits, the researchers administered pre- and post-learning assessment surveys to all participants before or after group discussions. Two researchers scored the surveys. The agreement percentage was calculated to ensure accuracy. On items requiring multiple-choice or four-level rating, inter-rater agreement was 100%. Two researchers practiced scoring the constructed-response items using responses collected from 10 participants. The inter-rater agreement was 90%. Without knowing the group membership, two researchers then independently scored the responses from all the participants in the present study. After the independent scoring, the two researchers reviewed any disputed items together. If the discrepancy could not be resolved through discussion, a third researcher with more expertise in testing accommodations and alternate assessments scored the section independently, and the expert's score was considered final. In the present study, the inter-rater agreement of independent scoring was 91% prior to dispute resolution.

Conduct of Focus Group Discussions and Individual Interviews

One or more members of the research team conducted focus group discussions with the participants before and after the instructional group's completion of AOA. The discussions were conducted either in person or by conference call. To give participants enough time and opportunity to participate in the discussion, the comparison group and the instructional group were divided into two groups of fewer than 10 people each. The group discussion leaders also called on individuals who were silent to ensure all members' participation. Participants were encouraged to respond to open-ended questions. Each session lasted about 45 minutes. Audiocassette recorders were used to record the discussion. A trained staff member completed transcriptions that provided qualitative data for describing context. In addition to the group discussion, individual interviews were held with 6 participants (3 from the instructional group and 3 from the comparison group). The interviews were conducted either by phone or in person. Each individual interview lasted about 20 minutes. All the discussions and individual interviews were structured following the protocols.

Scoring of Core Activities

Two researchers scored all core activities using the scoring rubric independently and without knowing the group status of the person completing the activities. Prior to scoring, the two researchers practiced scoring using 10 core activities completed by the participants. The percentage of agreement for item-by-item scoring was 85%. In the present study, the inter-rater agreement for independent scoring of all products was 92%. After the independent scoring, the researchers met to review the disputed items. If the discrepancy between scores on an item could not be resolved through discussion, the item was directed to the expert researcher, whose score was considered final.

Scoring of IEP Samples

Participants handed in 20 IEP samples. None of the participants indicated any changes after completing the AOA course. It should be noted that the research sites provided guidelines for the IEPs, which reduced participants' ability to influence IEPs. Thus, there was no variability in the completeness and quality of the statements on accommodations. The analysis of IEPs was subsequently omitted from the study because of the lack of meaningful data.

Design and Analysis

This study intended to evaluate (a) the effectiveness of a Web-based professional development program in enhancing teachers' knowledge and skills in general and inclusive assessment and (b) the teachers' use of and attitudes toward the program. Due to the multidimensional nature of the research questions, we decided to include descriptive, qualitative, and quantitative data in the study.

The independent variable was participation in the AOA program. The statistical analyses were performed at a significance level of .05 ($\alpha = .05$), using one-tailed tests. The dependent variables included (a) scores on pre- and post-learning assessment surveys and (b) scores on core activities. Descriptive data from the instructional group's post-learning survey helped clarify how the participants perceived the features of AOA course. The information derived from the diaries of concurrent professional development programs provided contextual data that described influences other than AOA. The focus group discussions and individual interviews reflected the participants' interaction with the online professional development program and were used to complement the quantitative results.

Wisconsin participants' surveys and diaries were included in the instructional group analysis involving repeated measures. However, due to the difference in recruiting methods between the Wisconsin site and the other two sites, a matched-pair t test was performed separately to analyze the pre- and post-learning core activities completed by Wisconsin participants.

Repeated Measures MANOVA

Items on the pre- and post-learning assessment survey were classified to yield knowledge, self-efficacy, and Internet experience scores. The scores were analyzed using the pre- and post-learning assessment surveys as within-subject measures and the group membership as the blocking variable. The changes between the pre- and post-learning knowledge scores and the efficacy scores were treated as dependent variables. The correlation between pre-knowledge and efficacy scores was .63, and the correlation between post-knowledge and efficacy scores was .65. These correlations reached the criterion of .30 for applying MANOVA. The repeated measures MANOVA was chosen over the simple t test because it offered higher power and reduced Type I error. The Group Membership × Knowledge Score and Group Membership × Self-Efficacy interaction effects were examined to determine whether there were significant differences between the changes demonstrated by individuals in the instructional group and the changes demonstrated by individuals in the comparison group over time.

Repeated MANCOVA

Researchers (e.g., Killion, 2000; Rodes et al., 2000) have pointed out that learners' experience with the Internet greatly influences the effectiveness of online learning. In the current study, the instructional group's changes in knowledge and attitude scores on the pre- and post-learning assessment surveys were treated as dependent variables, and Internet experience (derived from Items 24 and 25) was treated as a covariate (continuous variable). The interactions between pre- to post-learning changes and the covariate were examined to determine whether individuals with more Internet experience improved more than those with less experience.

ANCOVA

Participants' prior knowledge and self-efficacy might influence their scores on the core activities. Therefore, they could be considered covariates for the analysis of core activities scores. The correlation of the core activities score and the prior knowledge score was minimal: $r = .18$. The correlation between the core activities score and the self-efficacy score was $r = .31$. Through the comparison between the full and reduced linear model, we found that the self-efficacy score was not an effective covariate; the assumption of ANCOVA was not fulfilled. Therefore, an ANOVA was used in place of the ANCOVA to analyze participants' scores on core activities.

Matched-Pair T-Test

Due to the lack of an independent comparison sample in Wisconsin, we did not include the pre– and post–core activities from Wisconsin participants in the ANOVA. Instead, a matched-pair t-test was used.

Predictions

Prediction #1

Participants in the instructional group will show greater improvement in their knowledge and self-efficacy scores on the pre-learning to post-learning assessment survey than participants in the comparison group.

Prediction #2

Participants who are more experienced and comfortable with the Internet will demonstrate more improvement on the post-learning assessment survey than those who are less experienced and comfortable with the Internet.

Prediction #3

Instructional group participants will obtain higher scores on the core activities after taking the course than comparison group participants, and Wisconsin participants' scores on the core activities will improve after their exposure to the AOA course.

Prediction #4

In focus group discussions and individual interviews, participants in the instructional group will show more positive perceptions that are positive and increased self-efficacy with regard to general and inclusive educational assessment issues than comparison group participants. They will also report a preference for the Web-based AOA course as a desirable alternative to traditional professional development programs due to its flexibility, convenience, and other favorable features.

Results

Quantitative Results

Descriptive Data

Table 4 presents descriptive data on the knowledge, self-efficacy, and core activities scores on the pre- and post-learning assessment survey. Table 5 summarizes instructional group participants' responses to the six additional items on the post-learning assessment survey. Wisconsin participants' pre- and post-learning assessment surveys were included in the instructional group. Their core activities scores were separated from those of the other two sites due to differences in recruitment methods. Wisconsin participants' mean score on the pre-learning core activities was 15.94 ($SD = 5.67$); on the post-learning core activities, their mean score was 32.40 ($SD = 4.18$).

Table 4

Scores on Knowledge, Self-Efficacy, and Core Activities for Comparison and Instruction Groups

Group	$N_1{}^a$	Knowledge		Self-efficacy		$N_2{}^c$	Core activities
		Pre-learning	Post-learning	Pre-learning	Post-learning		
		M *(SD)*[b]	*M* *(SD)*	*M* *(SD)*	*M* *(SD)*		*M* *(SD)*
Comparison	26	11.33 (2.63)	12.69 (2.85)	9.96 (3.96)	10.04 (4.72)	24	11.75 (8.68)
Instruction	29	12.12 (2.81)	14.16 (2.44)	11.62 (3.79)	15.17 (3.21)	15	26.07 (6.77)

[a]N_1 denotes the number of participants who submitted surveys. [b]The standard deviation appears in parentheses. [c]N_2 denotes the number of participants who submitted core activities.

Table 5
Participants' Rating of the Features in the AOA Course

Items	Responses		
	Not at all	Somewhat	Very much
	$N^{a.}$	N	N
	$(\%)^b$	(%)	(%)
1. The case study facilitated my understanding of the content.	1	13	12
	(3.4%)	(44.8%)	(41.4%)
2. The hyperlinks and hypertexts in the course facilitated my learning.	1	14	11
	(3.4%)	(48.3%)	(37.9%)
3. The content of the course reflects the changes that are happening in the profession.	1	9	16
	(3.4%)	(31%)	(55.2%)
4. The course addresses my present as well as my future needs.	1	12	13
	(3.4%)	(41.4%)	(44.8%)
5. The core activities helped me check the knowledge and skills I learned in the course	1	14	12
	(3.4%)	(48.3%)	(41.4%)
6. Collaboration with colleagues can make the course more effective.	0	6	21
	0%	(20.7%)	(72.4%)

[a]Among instructional group participants, 2 (6.9% of 29) failed to provide responses to all six items; 1 participant (3.4% of 29) provided responses to the last two items, but not to the first four items. [b]Percentage calculated based on the total number (29) of instructional group participants.

Knowledge and Self-Efficacy

Prediction #1: The instructional group will improve more than the comparison group on the post-learning measures: Supported. A repeated MANOVA supported the hypothesis that the instructional group would demonstrate more improvement than the comparison group on the post-learning knowledge and self-efficacy scores. The knowledge and self-efficacy scores on the pre- and post-learning assessment surveys were used as multiple dependent variables. The group membership was used as the blocking variable. The repeated MANOVA was performed as a one-tailed test at an alpha level of .05. The Knowledge × Group Membership effect was significant, $F(1, 53) = 8.77$, $p = .005$, partial $\eta^2 = .14$, as was the Self-Efficacy × Group Membership effect, $F(1, 53) = 14.67$, $p = .000$, partial $\eta^2 = .22$. For the instructional group, the effect size expressed by the ratio of mean differences to the standard deviation of pre-learning score was 0.73 for the knowledge gains and 0.94 for the self-efficacy gains. The three-way interaction among knowledge, self-efficacy, and group membership was also significant, $F(1, 53) = 17.42$, $p = .000$, partial $\eta^2 = .25$. According to the commonly held standard for effect size (Cohen, 1977, as cited by Grimm & Yarnold, 1998), the Knowledge × Group Membership and Self-Efficacy × Group Membership interactions fell in the medium to large effect-size range. The effect size of the three-way interaction was large. As shown in Figure 1, with different starting points on both measures, the two groups changed at different rates on the measures, which accounted for the significant Group Membership × Knowledge × Self-Efficacy interaction.
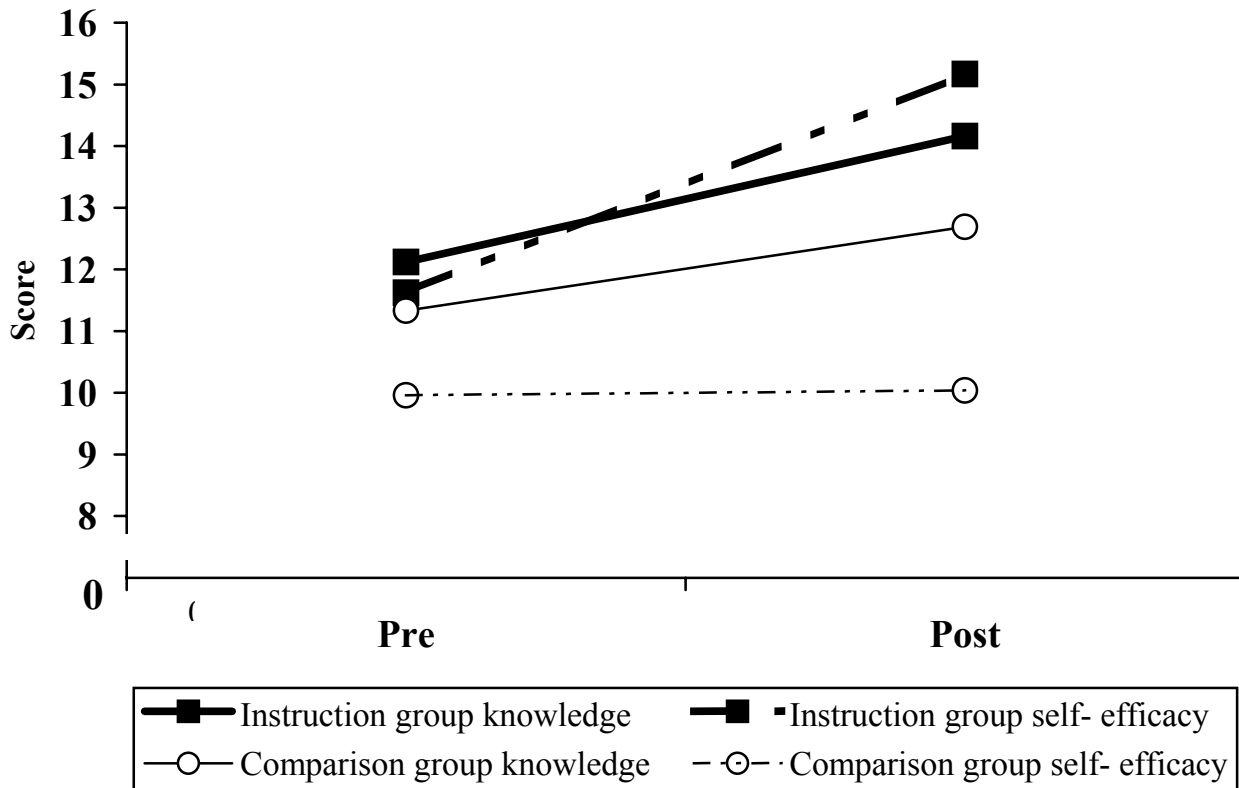
Prediction #2. Prior experience with the Internet will influence the instructional group participants' learning: Not supported. Participants in the instructional group had a mean of 4.21 ($SD = 1.32$) on the Internet experience scores. The interaction effect of Internet Experience × Knowledge at an alpha level of .05 was not significant, $F(1, 27) = .17$, $p = .69$. The interaction effect of Internet Experience × Self-Efficacy was also not significant, $F(1, 27) = 2.97$, $p = .10$. Inspection of the scatter plots of Internet experience and gains on the knowledge and self-efficacy scores revealed that no relationship existed among the variables. We further examined the relationship between the instructional group participants' scores on core activities (another measure of knowledge) and Internet experience. The correlation was not significant, $r = -0.15$. Therefore, the finding reported by previous researchers that individuals' prior Internet experience is related to online learning was not supported by these results.

Core Activities

Prediction #3: The instructional group will score higher on the core activities than the comparison group; and Wisconsin participants' scores on the core activities will improve after they access the AOA course: Supported. The between-group (instructional group vs. comparison group) one-tailed ANOVA performed at the .05 alpha level yielded a statistically significant result with a large effect size, $F(1, 37) = 29.47$, $p = .000$, $\eta^2 = .44$. The effect size (expressed as the ratio of the mean difference to the pre-test standard deviation of the comparison group) was 1.65. The matched-pair t test comparing the Wisconsin participants' scores on the pre- and post-learning core activities was significant, $t(.05, 4) = -6.83$, $p = .000$. The effect size (expressed by the ratio of the mean difference from the pre- to post-core activities to the pre-instruction standard deviation) was 2.90.

Figure 1. Means by group membership and pre- and post-learning measures.

Contextual Information

A total of 48 participants handed in diaries on professional development programs (27 in the instructional group, 21 in the comparison group). The comparison group participants spent 0 to 14.8 hours ($M = 2.48$, $SD = 3.98$) on assessment-related, non-AOA professional development activities. The instructional group participants spent 0 to 18 hours ($M = 3.81$, $SD = 4.98$) on assessment-related, non-AOA professional development programs. The time spent on the AOA course ranged from 6 to 62 hours ($M = 32.71$, $SD = 14.55$).

Qualitative Results

Prediction #4 was examined through the qualitative results, which indicated partial support for the hypothesis. In general, instructional participants showed improved perceptions and self-efficacy with regard to large-scale assessments and inclusive practices. While most of the instructional group participants reported their interaction with the AOA course as a positive experience, a few participants pointed out difficulties involved in online learning.

Inspection of the transcribed focus group discussions and individual interviews showed no difference in participants' responses in these two contexts. Therefore, data from the group discussions and individual interviews were combined in this analysis.

State Policies

In pre-learning focus group discussions, participants across the three sites mentioned the adoption of high-stakes testing in their respective states. They also reported the increased inclusion of students with disabilities in the large-scale assessments. State policies for accommodations and alternate assessments were discussed by the participants. Participants from Arizona emphasized how the guidelines provided by the state had helped them implement testing accommodations.

In the post-learning discussion, participants did not report changes in major state policy with regard to large-scale assessments. Participants from Arizona and Wisconsin mentioned ongoing policy making about the inclusion of English language learners.

Perception of Large-Scale Assessments and Inclusive Practice

We counted the number of times that participants expressed their opinions on large-assessment and inclusive assessment practice using (a) explicitly positive statements (e.g., "it is useful," "it is meaningful"), (b) explicitly negative statements (e.g., "it is meaningless," "it is not appropriate"), and (c) positive statements with reservations (e.g., "it could be good if it is valid," "it could be useful for some students, but not all"). Table 6 presents the percentage of each type of statement made by the two groups.

Table 6

Percentages of Positive Statements, Negative Statements, and Positive Statements With Reservations in Pre-and Post-Learning Discussions

	Comparison group		Instructional group	
	Pre-learning discussions	Post-learning discussions	Pre-learning discussions	Post-learning discussions
Positive statements	32%	24%	20%	52%
Negative statements	67%	68%	64%	45%
Positive statements with reservations	1%	8%	15%	3%

In the pre-learning group discussions, the opinions expressed across groups and sites were overwhelmingly negative about large-scale assessments and inclusive assessment practices. Participants complained about the large amount of time devoted to test preparation and the anxiety aroused by high-stakes testing and the overemphasis on test scores. Other major concerns were that (a) the large-scale assessments were "unfair" to students who were from disadvantaged backgrounds or did not function well under testing situations; (b) evaluations based on the results of large-scale assessments penalized some schools and teachers, regardless of student demographic characteristics such as disabilities or socioeconomic status; (c) the testing instruments failed to align with educational standards; (d) the assessments did not address the learning process or instructional needs of regular or special education students; and (e) the states' policies on alternate assessments and testing accommodations were questionable.

Participants did not identify the positive aspects of large-scale assessments and inclusive educational assessments until prompted by the discussion leaders. Participants generally acknowledged that test results could be used to compare students, schools, and districts. They also recognized that accountability systems were necessary and large-scale assessments were "worth considering" in that they could be used to set up comparable standards across schools and districts. In addition, they conceded that tests could help set high expectations for all students and that inclusive assessment practices might help measure students' academic abilities. It should be noted, however, that when discussing the positive aspects of large-scale assessments and inclusive practices, the instructional group participants repeatedly emphasized qualifications such as "if they are valid" and "if they are used correctly."

As evidenced in Table 6, instructional group participants expressed increased positive attitudes after going through the AOA course. One instructional group participant in Wisconsin

said: "[The course] changed my perception about the large scale assessments and the accountability system. . . . They [the assessments] are necessary." Another participant said: "We are more open to large-scale assessments; we see the positive side of it." They further discussed how they or their co-workers use the assessment results in (a) identifying students' difficulties, (b) designing instruction, (c) monitoring students' growth, and (d) communicating students' needs to colleagues and parents. The instructional group participants also discussed the usefulness of the accommodations and alternate assessments. They pointed out that inclusive assessments could be useful for gauging students' progress and informing instructional interventions.

Notwithstanding this increase in positive comments on large-scale assessments and inclusive educational assessments, the instructional group participants still maintained that (a) the alignment of curriculum and large-scale tests was not thorough, (b) the policy of comparing schools based on test results was highly questionable, and (c) teachers' ability to design and use testing accommodations and alternate assessments was highly variable—or, as one participant put it, the benefits of testing students with disabilities "all depend on the team" that designed the accommodations or alternate assessments.

Notably, comparison group participants' opinions of large-scale assessments and inclusive assessment practices expressed in post-instruction discussions did not differ from the opinions they expressed in pre-instruction discussions. No meaningful shifts in attitude, awareness, or knowledge were detected in post-instruction discussions.

Professional Development Programs

Participants' experience with assessment-related professional development programs varied. At least 20 participants reported no prior training in assessment-related issues at all. The rest of the participants reported that they had attended workshops conducted before statewide tests. Such workshops focused on administration issues particular to the specific tests. In general, participants felt that they were not well prepared with regard to educational assessment knowledge and skills. Participants from Arizona were very concerned about the great variability in teachers' knowledge of accountability and assessment. One participant expressed the consensus of the group by saying, "everyone [needs] to be on board at the same level" to ensure the veracity of the accountability system.

Participants reported that other professional development programs they had experienced were summer workshops lasting a few hours. The program characteristics they considered most desirable were relevance to their daily practice, provision of ongoing support, flexible scheduling, and promotion of collaboration.

Use of the AOA Course

With the exception of one participant (discussed below) who found the AOA course frustrating, instructional group participants across the sites considered the course useful in enhancing their knowledge and skills in educational assessments. The participants explicitly mentioned that the course was highly relevant to their practice. At least 5 participants mentioned that they used both the book and the online course materials to help themselves and colleagues

interpret and communicate scores. It appeared that participants chose the media of learning (i.e., book or Web) according to their preference. Some used the book routinely as "bed time reading," whereas others preferred Internet surfing. The participants further pointed out that the course was flexible in that it allowed access at their convenience. Furthermore, the hyperlinks offered them the opportunity to choose information according to their needs and interests. The participants appreciated the cases as meaningful examples of applying knowledge to practice. They especially liked the discussion about the three cases throughout the course. In addition, the core activities helped the learners to review and organize the information.

Although acknowledging the richness of the online resources, at least 4 participants also reported that the hyperlinks could be confusing and distracting. Some participants, particularly those in South Carolina, mentioned that access to the course materials and the hyperlinks was sometimes difficult due to technical problems (i.e., Internet access). Some participants reported that Internet speed limited the accessibility of the video clips.

The participant who reported great frustration considered the format of the Internet course her major obstacle. She acknowledged that the course materials were well organized and well presented, but characterizing herself as a traditional learner, she said she needed "to be at [a] certain place at a certain time, and have an instructor and have assignments due on a certain date."

Participants considered our frequent e-mail messages helpful. They also reported that informal discussions with colleagues in the instructional group supported the learning process. Finally, they suggested that more collaboration and discussion facilitated by the course designer might be beneficial.

Additional Information

After completing the data analysis, we contacted the CEC webmaster about the participants' use of the AOA Web course. We were interested in whether those who dropped out of the study had tried the course. Of the 24 instructional group participants who did not complete the study, 37.5% ($n = 9$) had logged on and viewed some parts of the course, whereas 62.5% ($n = 15$) had never logged on at all. Thus, it is unlikely that the course content influenced the majority (15) of the dropouts from the instructional group. The CEC webmaster was unable to provide additional information about the 9 participants who had logged on but had not completed the course (e.g., the duration or frequency with which they had interacted with AOA), making it difficult to interpret whether course content might have influenced these participants' decision to withdraw from the study.

Discussion

We used both quantitative and qualitative methods to examine the efficacy of the online AOA course in enhancing educators' knowledge about and self-efficacy in administering large-scale assessments, testing accommodations, and alternate assessments. We also investigated the educators' use of the online professional development course. Finally, we analyzed participants' pre- and post-learning assessment surveys, core activities, focus group discussions, and individual interviews.

Major Findings

The present study suggests that the multimedia, Web-based AOA course and book were effective in improving participants' knowledge and self-efficacy with regard to general and inclusive educational assessments. Participants showed significant improvement in their understanding of general assessment concepts and their self-rated level of efficacy in dealing with assessment-related practices. Knowledge gains were further evidenced by the increase in the participants' scores on the core activities, which directly address the application of AOA course content to professional practices. The knowledge gains of Wisconsin participants as measured by pre- and post-learning core activities (within-individual comparison) were larger than the knowledge gains measured by the between-group comparisons at the other two sites. It is possible that the pre-learning core activities cued Wisconsin participants to attend to relevant information in the course, and thus, they demonstrated more progress on the post-learning measure. Another possible explanation is that the within-individual comparison (matched-pair *t* test) may have more accurately reflected the learning gains due to its ability to rule out errors external to the participants. However, given the small sample size, the generalization of this finding is limited.

The high correlation between knowledge and self-efficacy indicates that improvement on either of the variables could influence the other. In other words, the AOA course was related both to teachers' knowledge gains on educational assessments and to teachers' perceived ability to conduct general and inclusive assessments. Interestingly, the participants changed more with regard to self-efficacy than with regard to knowledge as measured by the assessment survey. It is possible that the participants' learning through AOA enhanced their self-efficacy more than their knowledge because the course offered them a sense of empowerment by directly addressing areas in which they lacked familiarity and proficiency. However, given the internal reliability of .67 for the items that yielded the knowledge score, the instrument might have measured a construct or constructs other than knowledge. Thus, the instrument may not have accurately reflected participants' knowledge gains.

Consistent with the research on effective professional learning (Putnam & Borko, 2000; Mott, 2000), the present study suggests that several features may have contributed to successful learning in the present study. First, given the nationwide movement toward accountability systems, the participants were aware of the discrepancy between their assessment competency and the demands imposed by various policies. The AOA course may have met participants' needs by directly addressing the basic concepts and practice of educational assessments. Second, AOA's modeled application of relevant information in the simulated cases may have helped the participants put the instruction into a meaningful context and thus facilitated the connection between knowledge and practice. Third, the participants were able to acquire, explore, and connect to information through the distributed learning tools such as hypertexts, multimedia tools, and the book. In addition, the core activities may have provided meaningful formative assessments that assisted participants' review and organization of acquired knowledge. Finally, the participants responded favorably to the informal collaboration with colleagues recommended and encouraged by the course.

Consistent with prior research findings on Web-based distance learning (e.g., Killion, 2000; Hoffman & Thompson, 2000), the features of the online course—such as flexible

schedule, unlimited access, and self-directed learning—were beneficial for some participants but created obstacles for others. On the one hand, most participants found the freedom to choose the time, place, content, and pace of their learning desirable. They also reported that the vast amount of information on the Internet provided rich resources for the learning process. On the other hand, a few participants reported difficulties with the online course due to technical issues (i.e., access), the self-directed format of online learning, and the use of hypertext to browse and obtain information.

Interestingly, our study failed to confirm the hypothesis that people with more Internet experience achieve greater learning gains through online courses than those with less Internet experience. No relationship was found between the participants' Internet experience and measures of knowledge and self-efficacy. However, the two survey items that assessed Internet experience had an internal consistency of .68. It is possible that the items did not adequately measure the construct of Internet experience and therefore failed to reliably measure its influence. It is also possible that the accompanying textbook, *Assessing One and All* (Elliott et al., 2000), designed as one of the distributed pedagogy tools, helped participants with less Internet experience achieve gains comparable to those of participants with more Internet experience. Thus, some participants may have relied more on the textbook to acquire information and consequently showed no relationship between Internet experience and learning. Another explanation could be that the AOA online course is structured in such a user-friendly manner that learners' prior experience with the Internet was not an important factor.

Limitations of the Study

Participants

In light of the strong effects demonstrated by the present study, its limitations should also be recognized and interpreted carefully. The generalizability of the present findings is certainly limited due to the participant recruitment method. Because availability was the key determinant for recruitment, the participants may not adequately represent professionals who were not interested in participating or who were unavailable to participate.

In addition, the recruiting methods and the attrition rate might have caused self-selection that influenced the outcomes. Specifically, participants in the present study were volunteers. And they volunteered for specific incentives, such as professional development credits and monetary rewards. They were neither randomly selected nor randomly assigned to groups. It is possible that the individuals who volunteered for the instruction were more prone to improvement because they were more motivated and open to learning from the very beginning. Those who volunteered for the comparison group might have failed to show gains because they were less interested in changing their professional behavior and orientation. Furthermore, a relatively high attrition rate occurred in the study. However, the majority of the instructional group participants who dropped out never logged on to the Web course; in other words, their attrition was not related to the Web course itself. In addition, the attrition rates for the instructional group and the comparison group were similar. Still, attrition warrants careful consideration. It could be that those who could not or who did not benefit from the AOA course eventually dropped out and the large learning gains were found among the remaining participants. Thus, the present results could overestimate the actual effect of AOA learning.

Measurements

The measurements such as the pre- and post-learning assessment surveys and the AOA core activities were constructed by the developers of the AOA course and were directly linked to the course materials. However, the items relating to Internet experience and knowledge gains only have moderate internal consistency, indicating the items may mot adequately assess the intended constructs. This could cause problems in significance testing. For example, the lack of relationship between prior Internet experience and learning gains could result from unreliable measurement of Internet experience.

Application of Knowledge and Skills

The initial design called for an analysis of changes made to participants' IEP samples during the AOA course as a means of probing the embedded application of the knowledge and skills addressed by the course. However, this analysis was not performed. The checklists of allowable accommodations provided by the states eliminated the participants' latitude in making decisions, and therefore it was not possible to tell from the IEP samples how much participants' learning through the AOA course had influenced their actual practice. Thus, generalization of the skills and knowledge to actual IEP decision making was not captured by this study.

Strengths of the Study

As a multidimensional program evaluation, the present study employed a combination of quantitative and qualitative analysis that made possible a comprehensive understanding of the efficacy of the AOA program. Specifically, we used the MANOVA, ANOVA, and matched-pair *t*-test to assess whether participants improved their assessment knowledge and practice to a statistically significant extent because of the AOA course. The focus group discussions and individual interviews provided qualitative data that explained how the improvement, if any, happened and what might be the underlying mechanism for the change.

The matched-pair *t*-test conducted with Wisconsin participants was an unintended design resulting from recruiting difficulty. However, through the analysis of both between-group differences (ANOVA) and within-individual differences (matched-pair *t*-test), the study demonstrated strong evidence of learners' knowledge gains.

This study is also socially valid. We conducted research with participants in their typical work settings, as they engaged in AOA instruction and other continuing professional development while continuing their usual professional routines and responsibilities. This facet of the study lends credibility to the hypothesis that learning gains would be similar in "real-world" settings.

The amazing advances of technology in recent years are changing the landscape of education. More and more educational programs such as professional development courses and alternative certificate systems are being offered in a Web-based format. This study contributed to our understanding of the efficacy of online professional development programs. It also shed light on the features of Web-based programs that are potentially most effective.

Implications of the Study

This study provides information on the effectiveness of the Web-based AOA professional development program in promoting educators' knowledge and self-efficacy in general educational assessments, testing accommodations, and alternate assessments. It suggests that integrating effective professional learning principles with Internet-based features effectively improved the participants' understanding and self-efficacy in general and inclusive assessments.

As the No Child Left Behind Act and IDEA '97 mandate educational accountability systems for all students, large-scale assessments and the inclusion of students with disabilities in assessments have become more salient issues for educators in all states. It is essential for educators to master the basics of educational assessment and to acquire the knowledge and strategies necessary to achieve a meaningfully inclusive assessment practice. The Web-based AOA course could be an efficient alternative to traditional preservice and in-service professional training programs across the nation.

With respect to the further development of the AOA course, we conclude that it would be valuable to establish a mechanism to facilitate learners' dialogue and collaboration. A virtual forum or an electronic mailing list could provide such a mechanism. This kind of mechanism would also enable the course designers to keep regular contact with learners, responding to questions and providing assistance needed to better structure the learning process, particularly for those who are more used to a traditional instructor-directed learning format.

Given the limitations of the present study, future research on the efficacy of the AOA course (or any similar Internet-based multimedia teacher development program) should address the following issues: First, random selection and assignment of participants, perhaps with a time-series design across sites, should be implemented to increase the power and generalizability of the study. Second, participants' application of knowledge (i.e., in IEP decision making and assessment practices) should be examined to ascertain whether the online course improved the learners' skills in practice. Third, future research should use more accurate measures of participants' prior Internet experience and their actual use of the Internet over the course of the study to investigate (a) the influence of prior experience on AOA Web-based learning and (b) the characteristics of effective online learners. Finally, data are needed on the relative cost-effectiveness of Web-based and traditional professional development programs to inform the allocation of resources. For example, Web-based instruction, to the degree that it eliminates the need for release time and substitute teachers, may appear to be less expensive than other forms of professional development. However, these savings may be offset by increased costs for technology and Internet access. Analyses are also needed of the relative cost-effectiveness of Web-based and more traditional professional development in achieving learning gains for similar content.

Conclusions

We conclude that the Web-based multimedia AOA course effectively increased teachers' knowledge and self-efficacy in dealing with large-scale assessments, testing accommodations, and alternate assessments. Combining the principles of successful professional development programs with the course's online delivery features appears to support learners' progress.

However, certain technical difficulties and the self-directed learning process associated with Web-based learning could be obstacles for some learners.

In the context of today's heightened concerns about the educational outcomes of students with disabilities and the adequacy of educators' assessment knowledge and skills, the AOA course could be an effective alternative to the traditional forms of professional development for educators. This study provides promising results and suggests avenues for continued research in this area.

References

Bohnenkamp, J., & McMahon, J. (2001). Click: Community of learning, information, communication and knowledge. *T.H.E. Journal*, *28*(1), 20–24.

Burgstahler, S. E. (1995). Distance learning and the information highway. *Journal of Rehabilitation Administration, 19*(4), 271–276. Retrieved June 19, 2003, from http://www.washington.edu/doit/Brochures/Technology/jra.95.html

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, *23*(7), 13–20.

Daley, B. J. (2000). Learning in professional practice. *New Directions for Adult and Continuing Education*, *86*, 24–33.

DiPerna, J. C., McKevitt, B. C., Elliott, S. N., & Braden, J. P. (1999). *CESA #1 student assessment project midproject evaluation report*. Unpublished manuscript.

DuFour, R. (2001). In the right context. *Journal of Staff Development*, *22*(1), 14–17.

Elliott, S. N., Braden, J. P., & White, J. L. (2001). *Assessing one and all: Educational accountability for students with disabilities*. Reston, VA: Council for Exceptional Children.

Grimm, G. L., & Yarnold, R. P. (1998). *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.

Hoffman, E., & Thompson, G. (2000). Putting the research to work: Professional development models from Michigan. *Tech Trends*, *44*(2), 20–23.

Hollenbeck, K., Tindal, G., & Almond, P. (1998). Teachers' knowledge of accommodations as a validity issue in high-stakes testing. *Journal of Special Education*, *32*(3), 175–183.

Johnson, E., Kimball, K., Brown, S. O., & Anderson, D. (2001). A statewide review of the use of accommodations in large-scale, high-stakes assessments. *Exceptional Children*, *67*(2), 251–264.

Killion, J. (2000). Log on to learn. *Journal of Staff Development*, *21*(3), 48–53.

Matthews, D. (1999). The origins of distance education and its use in the United States. *T.H.E. Journal*, *27*(2), 54, 56, 58, 60, 62, 64, 66–67.

Mott, V. W. (2000). The development of professional expertise in the workplace. *New Directions for Adult and Continuing Education*, *86*, 23–31.

Oelrich, K. (2001). Virtual school: A 21st century strategy for teacher professional development. *T.H.E. Journal*, *28*(11), 48–50.

Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Midwestern Educational Researcher*, *6*(1), 21–27.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, *29*(1), 4–15.

Rodes, P., Knapczyk, D., Chapman, C., & Chung, H. (2000). Involving teachers in web-based professional development. *T.H.E. Journal*, *27*(10), 95, 96, 98, 100, 102.

Sujo de Montes, L. E., & Gonzales, C. L. (2000). Been there, done that: Reaching teachers through distance education. *Journal of Technology and Teacher Education*, *8*(4), 351–371.

Thurlow, M. L. (2000). Standards-based reform and students with disabilities: Reflections on a decade of change. *Focus on Exceptional Children*, *33*(3), 1–16.

Thurlow, M. L., House, A. L., Scott, D. L., & Ysseldyke, J. E. (2000). Students with disabilities in large-scale assessments: State participation and accommodation policies. *Journal of Special Education, 34*(3), 154-163.