

# Impact of Self-Assessment Questions and Learning Styles in Web-Based Learning: A Randomized, Controlled, Crossover Trial

David A. Cook, MD, MHPE, Warren G. Thompson, MD, Kris G. Thomas, MD, Matthew R. Thomas, MD, and V. Shane Pankratz, PhD

## Abstract

### Purpose

To determine the effect of self-assessment questions on learners' knowledge and format preference in a Web-based course, and investigate associations between learning styles and outcomes.

### Method

The authors conducted a randomized, controlled, crossover trial in the continuity clinics of the Mayo-Rochester internal medicine residency program during the 2003–04 academic year. Case-based self-assessment questions were added to Web-based modules covering topics in ambulatory internal medicine. Participants completed two modules with questions and two modules without questions, with sequence randomly assigned. Outcomes included knowledge assessed after each

module, format preference, and learning style assessed using the Index of Learning Styles.

### Results

A total of 121 of 146 residents (83%) consented. Residents had higher test scores when using the question format (mean  $\pm$  standard error, 78.9%  $\pm$  1.0) than when using the standard format (76.2%  $\pm$  1.0,  $p = .006$ ). Residents preferring the question format scored higher (79.7%  $\pm$  1.1) than those preferring standard (69.5%  $\pm$  2.3,  $p < .001$ ). Learning styles did not affect scores except that visual-verbal "intermediate" learners (80.6%  $\pm$  1.4) and visual learners (77.5%  $\pm$  1.3) did better than verbal learners (70.9%  $\pm$  3.0,  $p = .003$  and  $p = .033$ , respectively). Sixty-five of 78 residents (83.3%, 95% CI 73.2–90.8%) preferred

the question format. Learning styles were not associated with preference ( $p > .384$ ). Although the question format took longer than the standard format (60.4  $\pm$  3.6 versus 44.3  $\pm$  3.3 minutes,  $p < .001$ ), 55 of 77 residents (71.4%, 60.0–81.2%) reported that it was more efficient.

### Conclusions

Instructional methods that actively engage learners improve learning outcomes. These findings hold implications for both Web-based learning and "traditional" educational activities. Future research, in both Web-based learning and other teaching modalities, should focus on further defining the effectiveness of selected instructional methods in specific learning contexts.

Acad Med. 2006; 81:231–238.

Logistic difficulties and demands for increasing faculty productivity challenge teaching in ambulatory settings.<sup>1</sup> Simultaneously, academic health centers are charged to reform residents' education<sup>2</sup> and improve teaching and assessment for them.<sup>3</sup> These challenges,

**Dr. Cook** is assistant professor of medicine, Mayo Clinic College of Medicine, Rochester, Minnesota.

**Dr. Thompson** is associate professor of medicine, Mayo Clinic College of Medicine, Rochester, Minnesota.

**Dr. K. G. Thomas** is assistant professor of medicine, Mayo Clinic College of Medicine, Rochester, Minnesota.

**Dr. M. R. Thomas** is assistant professor of medicine, Mayo Clinic College of Medicine, Rochester, Minnesota.

**Dr. Pankratz** is assistant professor of biostatistics, Mayo Clinic College of Medicine, Rochester, Minnesota.

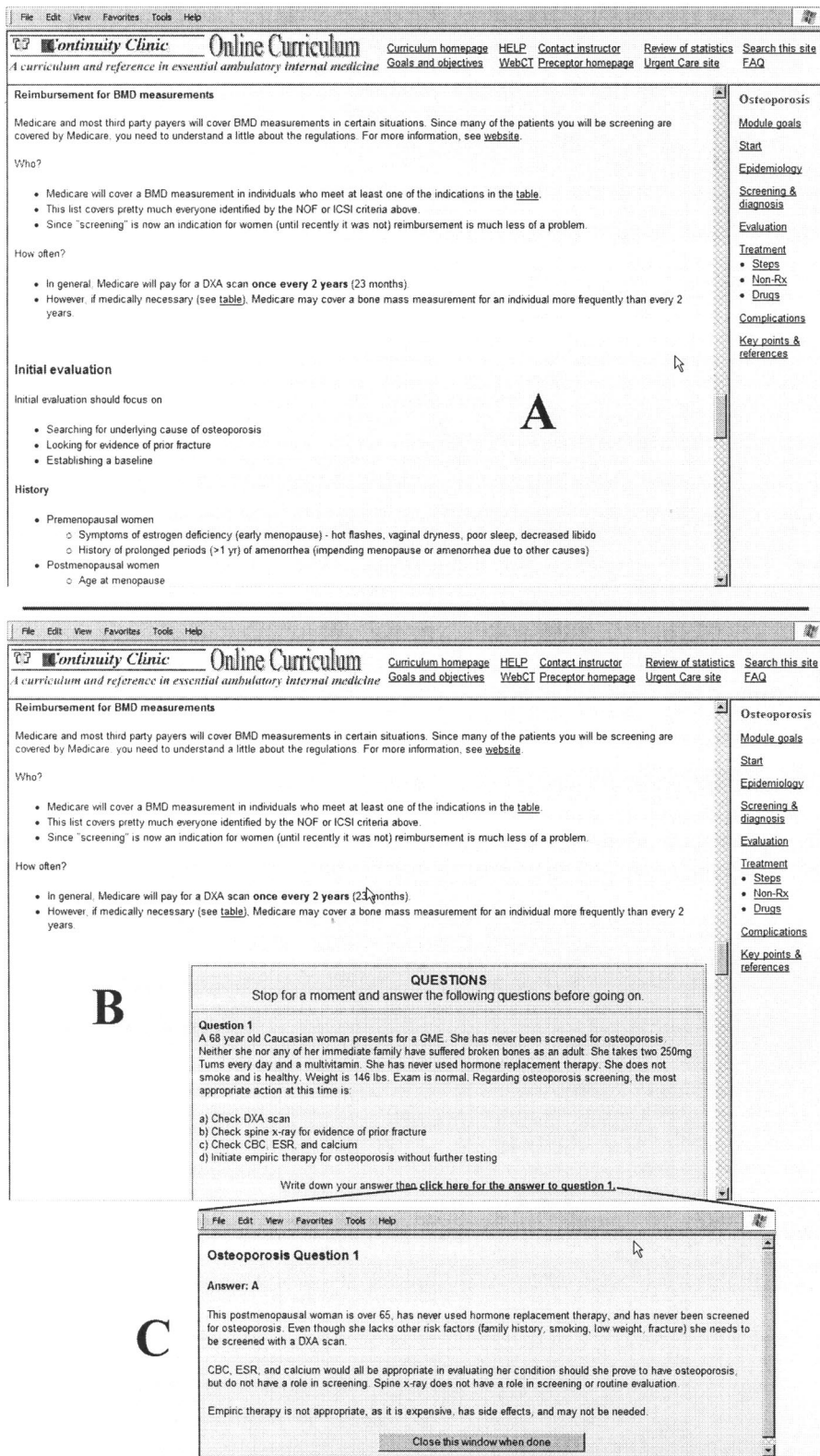
Correspondence should be addressed to Dr. Cook, Division of General Internal Medicine, Mayo Clinic College of Medicine, Baldwin 4-A, 200 First Street SW, Rochester, MN 55905; e-mail: (cook.david33@mayo.edu).

coupled with resident duty hour restrictions, heighten the need to maximize learning within available opportunities. Web-based learning (WBL) offers a potential solution.<sup>4,5</sup> Studies show that WBL improves learning as much or more than traditional teaching methods<sup>6–8</sup> or no intervention,<sup>9–13</sup> facilitates more efficient learning than traditional methods,<sup>14–16</sup> and overcomes barriers of distance and scheduling constraints.<sup>17–21</sup> Yet little research has investigated ways to optimize WBL methods.

Multiple authors have suggested that not all educational Web sites are equally effective,<sup>22–24</sup> and that variations in instructional design can influence learning outcomes.<sup>25–27</sup> However, there is little evidence to support this argument, and even less evidence to guide educators regarding what *does* work. WBL course enhancements are not free; they often require increased faculty and programmer effort, computer

infrastructure, and learner time. Yet if evidence supports their use, WBL educators will need to consider adapting their Web-based instructional approaches accordingly.

One technique to stimulate learning in live lectures is the use of self-assessment questions and feedback.<sup>28</sup> Similar practices would presumably be effective in Web-based teaching,<sup>24,29</sup> but objective evaluations are few. One uncontrolled study found anecdotal evidence that medical students preferred WBL modules that had self-assessment questions.<sup>18</sup> Other studies in medical education comparing self-assessment questions to a less interactive computer-assisted learning (CAL) format showed no difference<sup>30,31</sup> or favored the less interactive format,<sup>32</sup> but a study of college students found significant benefit from the question format.<sup>33</sup> Thus, while self-assessment questions and feedback hold theoretical promise, evidence is inconclusive.



**Figure 1** Shots of computer screens depicting a representative Web-based module used in the curriculum of the Mayo-Rochester internal medicine residency program during the 2003–04 academic year. The top screen (A) is from a standard format module, while the bottom screen (B) illustrates the question format and accompanying feedback (C). Note that the interposition of a case-based question does not affect the module content or format.

Another technique that may enhance Web-based learning is adaptation of instruction to the characteristics, such as learning style, of individual learners.<sup>34–36</sup> A recent review of learning styles in CAL<sup>36</sup> found only four studies evaluating medical students or physicians, and the only study<sup>37</sup> comparing alternate CAL instructional methods found no interaction with learning style. However, studies outside of medicine<sup>36</sup> have found significant effects, and further investigation appears to be warranted.

We hypothesized that internal medicine residents using WBL modules with self-assessment questions and feedback would have higher test scores than those using WBL modules without questions. We also sought to determine which format residents prefer, whether there is an association between format preference and test score, and whether associations exist between residents' learning styles and test score or format preference. To do this, we compared WBL modules with and without self-assessment questions in a randomized, controlled crossover trial in which each participant used both formats to study topics in ambulatory medicine.

**Method**

**Setting and sample**

All 146 categorical residents in the Mayo-Rochester internal medicine residency program were invited to participate. This study took place in the residents' weekly continuity clinic during the 2003–04 academic year. Our institutional review board approved this study. Consent was obtained from all participants.

**Interventions and randomization**

In a previous study, we found that WBL modules on ambulatory medicine topics were superior to a paper format.<sup>16</sup> The WBL format became a standard part of the internal medicine residency ambulatory clinic curriculum for the 2003–04 academic year. As topics for this curriculum, we selected four common clinical problems: cervical cancer screening, dementia, osteoporosis, and dyspepsia. For each topic, a WBL module was developed in both standard and intervention ("question") format (see Figure 1). Content and layout were identical between the two formats except for case-based multiple-choice questions

embedded periodically throughout question format modules. Residents were encouraged, but not required, to answer each question. After answering a question, residents were able to see the correct answer along with a rationale (feedback). The feedback did not introduce new content information.

Development and structure of the standard module format have been previously described.<sup>16</sup> To summarize, each evidence-based module consisted of several Web pages of text and tables, with hyperlinks to online resources, including full-text journal articles. In the present study, instructional methods common to both formats included activating prior knowledge by asking residents to reflect on a patient they had seen with this problem, facilitating self-directed learning through hyperlinks to additional resources, focusing on clinically relevant information, and repeating key concepts. Residents accessed the modules through WebCT (version 3.8), which automatically controlled format presentation. Accessing WebCT required a password (the same password used to access the local network). The standard version of each module was available without a password for use as a reference.

Participants completed two modules using the standard format and two modules using the format with questions. Sequence was determined randomly using MINIM (version 1.5, London Hospital Medical College, London, UK), with stratification by postgraduate year and continuity clinic site. Modules were released every six to eight weeks, but residents completed them on their own schedules. Blinding was not possible in this study.

\*Since some of these learning style dimensions may be unfamiliar, a brief review is in order. *Active* learners prefer practical application or exercise of information they have received. *Reflective* learners, on the other hand, internalize information—observing before passing judgment, examining from different perspectives, and looking for meaning as they create new knowledge. *Sensing* learners prefer what is real (facts, data, and experimentation), while *intuitive* learners look for patterns and meaning (principles and theories). *Visual* learners prefer the written or spoken word, while *verbal* learners learn best from pictures, demonstrations, and displays. *Sequential* learners follow a linear process of logical steps when they learn, while *global* learners seem to make large leaps, occasionally struggling until suddenly they “get it.” For a more detailed discussion see Felder and Silverman’s original description,<sup>63</sup> and also Cook’s review<sup>36</sup> and accompanying references.

## Instruments and outcomes

One primary outcome, knowledge, was determined by score on a post-test at the end of each module. Case-based test questions were developed to address each module’s objectives using the method we have described previously,<sup>16</sup> which involves expert review and piloting on internal medicine faculty. Questions were designed to assess application of knowledge.<sup>38</sup>

At the end of the academic year, residents completed a cumulative test (“delayed test”) composed of the questions from each post-test. Residents received test scores, answers, and feedback only following the delayed test. To obtain an estimate of knowledge retention over time, we included for analysis only scores from delayed tests completed at least three weeks after the post-test.

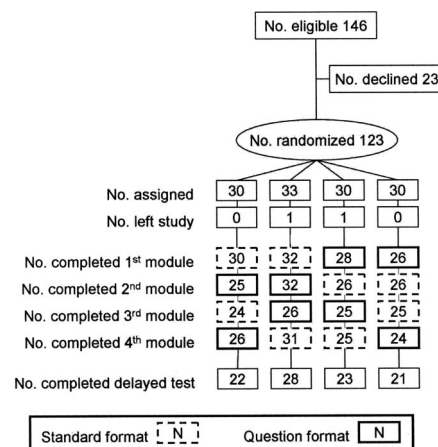
The other primary outcome, format preference, was assessed on an end-of-course questionnaire using a scale ranging from 1 = “Strongly prefer questions and feedback” to 6 = “Strongly prefer no questions and feedback.” The questionnaire included additional comparisons of the two formats (efficiency, effectiveness, and time spent), and questions regarding the number of self-assessment questions actually answered, continued use of the modules, and technical difficulties. We recorded Web site hits at two-week intervals.

Learning styles were assessed using Felder and Solomon’s Index of Learning Styles (ILS).<sup>39</sup> The ILS provides a separate score for each of four dimensions (active-reflective, visual-verbal, sensing-intuitive, and sequential-global),\* with scores ranging from -11 to +11 in increments of 2 (-11, -9, ... 7, 9, 11). Our preliminary data suggested acceptable reliability for these scores.<sup>40</sup> Because the effects of learning styles are more pronounced for learners with scores at the extremes of the scale,<sup>41</sup> scores between -3 and +3 were considered “intermediate,” while scores above or below this were classified according to the corresponding style.

All tests and questionnaires were administered using WebCT.

## Statistical analysis

Test scores were compared over time and between the two formats using a mixed-effects analysis of variance (ANOVA)



**Figure 2** Study enrollment and number of participants completing module tests at each study time point. The setting was a Web-based course in ambulatory medicine in the Mayo-Rochester internal medicine residency program during 2003–04. Two modules had case-based, self-assessment questions; two modules (“standard format”) did not have these questions. All data available at each time point were included in the analysis.

accounting for repeated measurements on each subject and for differences among modules. For the comparison between formats, additional adjustments were planned for group assignment, postgraduate year, gender, clinic site, and the number of self-assessment questions answered. To determine the effect of format preference and learning style on test scores, the ANOVA was repeated adjusting separately for format preference (treated as a dichotomous variable using the center of the possible range as the cutpoint) and learning styles. The ANOVA was repeated for the delayed test, with additional adjustment for time from post-test to delayed test.

Format preference was analyzed using the Wilcoxon signed rank test, testing the null hypothesis that there was no preference. Either the Wilcoxon rank sum test or the Kruskal-Wallis test was used for comparisons among two or more groups. Generalized linear models were used to evaluate the possibility of simultaneous effects from multiple learning styles on format preference. The fit of this parametric model was assessed and deemed adequate for analysis of this ordinal variable. Other questionnaire responses were analyzed using the Student’s *t*, Wilcoxon signed-rank, or Wilcoxon rank sum test as appropriate.

All analyses were performed using intention-to-treat and a two-sided alpha

Table 1

**Characteristics of 121 Study Participants, Mayo-Rochester Internal Medicine Residency Program, 2003–04\***

Characteristic	No. of participants <sup>†</sup>			
	Group 1 (n = 30)	Group 2 (n = 32)	Group 3 (n = 29)	Group 4 (n = 30)
	No. (%)	No. (%)	No. (%)	No. (%)
<b>Male gender</b>	20 (67)	21 (66)	20 (69)	15 (50)
<b>Postgraduate year</b>				
1	12 (40)	14 (44)	11 (38)	12 (40)
2	10 (33)	11 (34)	10 (34)	10 (33)
3	8 (27)	7 (22)	8 (28)	8 (27)
<b>Ethnic group<sup>‡</sup></b>				
African American	0 (0)	0 (0)	1 (5)	1 (6)
Arabic	0 (0)	1 (5)	0 (0)	1 (6)
Asian or Asian American	3 (15)	3 (14)	1 (5)	1 (6)
Hispanic/Latino	2 (10)	0 (0)	1 (5)	2 (11)
East Indian	2 (10)	1 (5)	3 (14)	2 (11)
White	11 (55)	15 (68)	12 (57)	9 (50)
Other	0 (0)	0 (0)	1 (5)	1 (6)
Prefer not to answer	2 (10)	2 (9)	2 (10)	1 (6)
<b>Prior postgraduate medical training</b>				
None	16 (76)	27 (100)	16 (70)	17 (71)
1 year	1 (5)	0 (0)	4 (17)	2 (8)
2–4 years	4 (19)	0 (0)	2 (9)	3 (13)
More than 4 years	0 (0)	0 (0)	1 (4)	2 (8)
<b>Post-residency plans</b>				
General internal medicine	2 (10)	3 (14)	2 (10)	5 (28)
Internal medicine subspecialty	16 (80)	19 (86)	15 (71)	13 (72)
Non-internal medicine specialty	2 (10)	0 (0)	1 (5)	0 (0)
Undecided	0 (0)	0 (0)	3 (14)	0 (0)
<b>Prior experience with Web-based learning</b>				
None	3 (14)	2 (7)	2 (9)	3 (13)
1 or 2 courses	6 (29)	12 (44)	6 (27)	7 (29)
3–5 courses	10 (48)	8 (30)	11 (50)	9 (38)
6 or more courses	2 (10)	5 (19)	3 (14)	5 (21)
<b>Comfort using the Internet</b>				
Uncomfortable	2 (10)	4 (15)	3 (13)	6 (25)
Neutral	1 (5)	3 (11)	1 (4)	2 (8)
Comfortable	18 (85)	20 (73)	19 (82)	16 (50)

\* A total of 121 residents participated. Each completed two modules with questions and two modules without questions in a crossover fashion, with sequence randomized.

<sup>†</sup> Number of responses varies because information was obtained from different questionnaires and not all respondents answered all questions. All percentages are calculated for group totals i.e. for the column.

<sup>‡</sup> Self-reported ethnic group.

level of .05. The expected sample size of 86 participants was to provide 90% power to detect a difference of 0.5 points in preference and a 7% change in test score. All analyses were performed using SAS 8.02.

## Results

One hundred twenty-three residents consented to participate and were randomized (see Figure 2), for a response rate of 84%. One resident subsequently left the program, and one withdrew from

the study citing difficulty completing tests in WebCT, leaving 121 residents for final analysis. One hundred nineteen residents (98.3%) completed at least one module, 90 (74.4%) completed all modules, and 78 (64.4%) completed the final survey. Demographics and other characteristics of the participants are summarized in Table 1.

## Knowledge

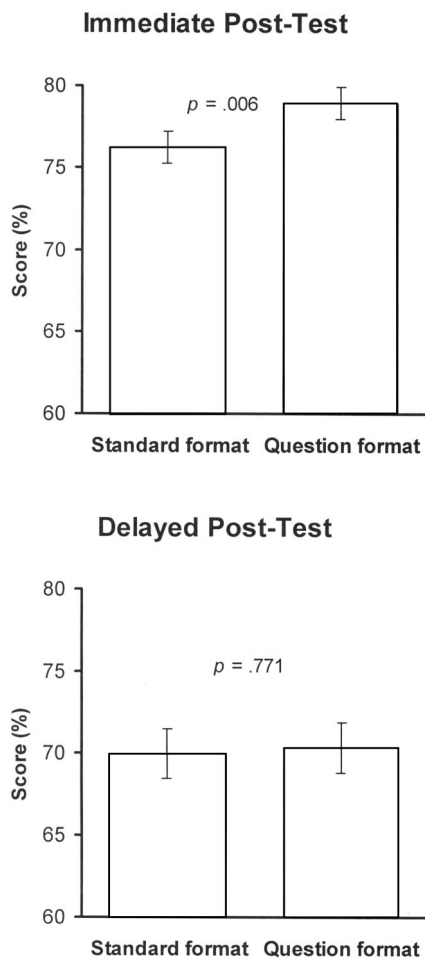
After adjusting for differences among modules, post-test scores were higher for the question format (mean  $\pm$  standard error, 78.9%  $\pm$  1.0%) than for the standard format (76.2%  $\pm$  1.0%,  $p = .006$ ; see Figure 3). This difference persisted after adjusting for study group, gender, postgraduate year, and clinic site ( $p = .005$ ). Adjustment for additional covariates including ethnic group, comfort using the Internet, experiences with WBL, or perceptions of technical problems yielded similar results. Adjustment for the number of self-assessment questions answered revealed a trend ( $p = .066$ ) suggesting that those who had answered more questions scored higher. After adjusting for postgraduate year and clinic site, women had higher scores (81.0%  $\pm$  1.7%) than men (76.7%  $\pm$  1.5%,  $p = .02$ ) regardless of format, although without multivariate adjustment this difference did not reach statistical significance.

Format preference influenced test scores, with those preferring the question format scoring higher (79.7%  $\pm$  1.1%) than those preferring standard (69.5%  $\pm$  2.3%,  $p < .001$ ). Preference also interacted with format ( $p = .031$ ) with the lowest scores occurring among those using the standard format who reported preference for that format.

Scores on the delayed test were not significantly different between the question (70.3%  $\pm$  1.6%) and standard (69.9%  $\pm$  1.5%) formats both before ( $p = .771$ ) and after ( $p = .873$ ) multivariate adjustment, including time from post-test to delayed test. Once again, residents preferring the format with questions performed better than those who preferred the standard format ( $p < .001$ ).

## Format preference and use of modules

Learners strongly preferred the question format (2.18  $\pm$  0.18,  $p < .001$ ), with 65 of 78 (83.3%, 95% CI 73.2–90.8%)



**Figure 3** Test scores on immediate and delayed (at least three weeks) post-tests, after adjustment for differences among Web-based modules and study groups. Participants were 121 residents in the Mayo-Rochester internal medicine residency program during 2003–04.

preferring this format. Subgroup comparisons revealed no statistically significant difference in preference between men and women, ethnic groups, postgraduate years, degree of comfort using the Internet, or perceptions of technical problems ( $p > .15$ ). Residents with intermediate experience with WBL reported lower preference for the question format ( $2.60 \pm 0.47$  for one or two WBL courses,  $2.38 \pm 0.28$  for three to five courses) than did those with significant experience (six or more courses) or no experience ( $1.31 \pm 0.17$  and  $1.2 \pm 0.2$ , respectively), although this did not reach statistical significance ( $p = .054$ ).

Self-reported time to complete the modules was greater for the question format ( $60.4 \pm 3.6$  minutes) than for the standard format ( $44.3 \pm 3.3$ ,  $p < .001$ ). Despite the greater time required by the

question format, 63 of 76 residents (82.9%, 72.5–90.6%) felt that the question format was more effective and 55 of 77 (71.4%, 60.0–81.2%) reported that it was more efficient.

Fifty-seven of 75 residents (76.0%, 64.8–85.1%) reported returning to use the Web-based modules after completing the module, with 25 (33.3%, 22.9–45.2%) returning more than three times. Forty-four residents (58.7%, 46.7–70.0%) used hyperlinks to access full-text journal articles.

Twenty-nine of 74 residents (39.2%, 28.0–51.2%) experienced significant technical problems at the beginning of the course, and 19 of 73 (26.0%, 16.5–37.6%) felt these were still significant at the end of the course. Eleven residents (14.9%, 7.7–25.0%) reported difficulty with passwords.

Hits to module web pages are reported in Figure 4.

#### Associations with learning styles

Learning style scores (summarized in Figure 5) had no significant association, individually or combined in multivariate analysis, with format preference ( $p > .384$ ). Post-test scores were not significantly associated with scores in the active-reflective ( $p = .198$ ), sensing-intuitive ( $p = .522$ ), or sequential-global ( $p = .305$ ) dimensions. However, visual-verbal styles were associated with immediate post-test scores ( $p = .009$  after adjusting for topic and format). Learners with intermediate visual-verbal style scored highest ( $80.6\% \pm 1.4\%$ ), followed by visual ( $77.5\% \pm 1.3\%$ ) and verbal ( $70.9\% \pm 3.0\%$ ) learners. This difference was significant between intermediate and verbal ( $p = .003$ ) and visual and verbal ( $p = .033$ ) styles. There was no significant interaction ( $p = .922$ ) between visual-verbal styles and format when looking at post-test scores. Similar results were found after further adjusting for study group, gender, postgraduate year, and clinic site. In simultaneous adjustment for all learning styles, only the visual-verbal dimension was significantly associated with scores ( $p = .003$ ).

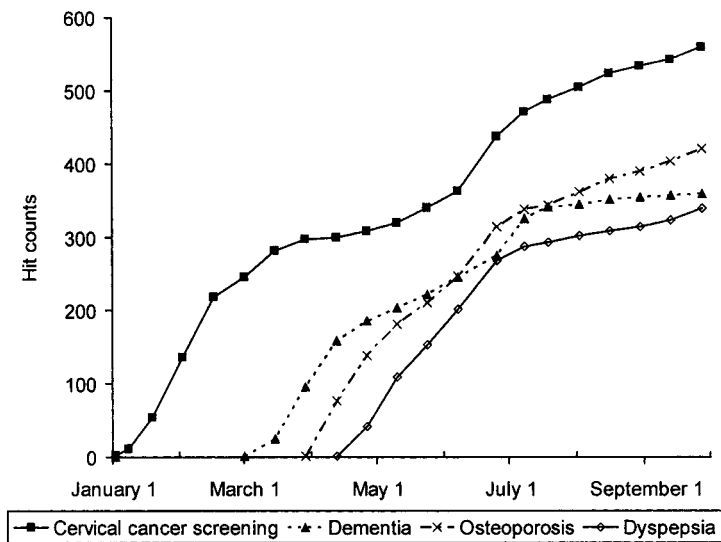
#### Discussion

In a crossover trial comparing WBL modules with self-assessment questions and feedback to modules without

questions, we found that internal medicine residents had higher test scores when using the format with questions. Furthermore, residents strongly preferred the question format and felt it facilitated more effective and efficient learning even though it required more time to complete. These effects remained stable after adjusting for learner characteristics including gender, ethnic group, postgraduate year, prior experience with WBL, and learning styles.

While these data support the theory that instructional methods promoting learner interaction are more effective than less active methods, they contradict prior studies of CAL in medical education. One study found no difference between two case-based CAL formats with varying levels of interactivity, but the variation between formats was poorly defined.<sup>30</sup> Another study found that basic science medical students using a CAL format with case-based questions and answers had lower scores than did those using a less interactive CAL format.<sup>32</sup> The study's authors attributed this unexpected finding to the outcome measure, which tested recall rather than application of knowledge, and lack of familiarity of the learners with case-based learning. A follow-up report studying the same CAL formats and the same group of learners in a clinical setting (neuroradiology) found no difference between formats,<sup>31</sup> but in this case potential differences between study groups might have been diluted by learning from other sources during the eight-week lapse between pre-test and post-test. The present study avoids the limitations of this prior research, and corroborates a study of college students<sup>33</sup> that demonstrated that self-assessment questions with feedback significantly improved post-test scores, and a study of medical students<sup>18</sup> who preferred case-based modules over less interactive formats. Future research could confirm our findings and investigate alternate means of engaging learners in Web-based environments.<sup>27</sup>

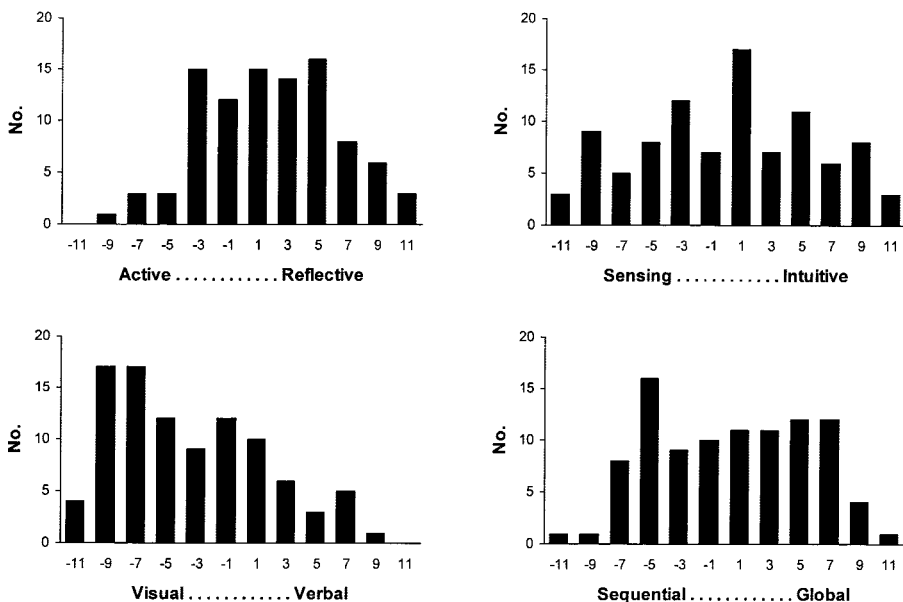
These results have relevance beyond WBL, as they support the use of active learning methods in general. Although the literature abounds with courses and curricula using “active learning methods,” few studies have rigorously compared active methods with alternative instructional techniques. For



**Figure 4** Hits to Web-based modules used in the ambulatory medicine curriculum of the Mayo-Rochester internal medicine residency program during 2003–04, recorded at two-week intervals. The academic year ended on June 25, after which continued use of the modules was voluntary.

example, many comparative studies in medical education are limited by the use of multifactorial educational interventions<sup>42–45</sup> that confound the attribution of effect or lack thereof to any specific method or process.<sup>46,47</sup> Other studies make comparisons with no-intervention controls,<sup>42–45,48,49</sup> thus failing to inform selection of effective methods from multiple available options. We suggest that instead of

studying *whether* medical students and physicians can learn using a designated method or combination of methods, we should study *how best to facilitate* learning. By carefully controlling for confounding variables, the present study has demonstrated that variations in instructional method—namely, methods that actively engage the learner—can positively influence learning.



**Figure 5** Scores of study participants on four dimensions of the Index of Learning Styles. As the figure indicates, scores ranged from –11 to +11 in increments of two points. For the purposes of data analysis, scores between –3 and +3 were considered intermediate, while scores above or below this range were classified according to the corresponding learning style. Participants were 96 residents in the Mayo-Rochester internal medicine residency program during 2003–04.

Although residents felt that the format with questions required more time, they preferred this format and felt that it was more efficient. We ascribe this to their perception that this format was more effective. Additionally, learners who preferred the active learning format had higher test scores than did those preferring the standard format. While unmeasured factors such as motivation may be playing a role here, these findings should reassure educators concerned about employing instructional methods that demand more from the learners.

It is important to consider the clinical (educational) significance of these findings. Although the treatment effect of test scores is modest, the true difference is likely attenuated by other sources of variance among the study groups.<sup>46</sup> It is also possible that additional learning gains were realized yet unmeasured by our assessment. The observed effect compares favorably with findings of other education studies using active-intervention comparison groups, where differences are typically nonsignificant<sup>14–16,50–53</sup> or small.<sup>6</sup> Since learners preferred the more effective instructional method, the concordance among outcomes suggests that self-assessment questions with feedback do have a clinically significant benefit on learning.

With the exception of lower test scores among residents with verbal style, learning styles did not affect test scores or format preference. There were no aptitude-treatment interactions<sup>54</sup> between learning style and format. This lack of effect is not wholly unexpected given the central importance of instructional method in facilitating learning, and supports the argument that use of effective instructional methods should be ensured before considering the influence of learning styles.<sup>36,55</sup> Future studies might investigate theory- and evidence-based predictions<sup>36</sup> regarding the adaptation of sound instructional methods to individual learning styles. While the association between visual-verbal styles and test scores is interesting, we caution that this finding should be considered preliminary: it was not predicted by theory, prior research using this style dimension has yielded inconsistent results,<sup>36</sup> and a recent study showed poor test-retest reliability for ILS visual-verbal scores.<sup>56</sup>

The significant difference between formats was no longer present when residents were tested after a delay of several months. Although we are disappointed that the effect did not persist, knowledge retention is challenging in all educational settings, including CAL and WBL,<sup>14,57,58</sup> and should be a focus of further research.

Our previous study suggested that passwords impede use of WBL modules.<sup>16</sup> Although residents must still use a password to access WebCT, it is the same password used to access the institution network. Furthermore, we made the standard version of the course available without a password. With this change, fewer residents reported problems with passwords (15% now, compared with 61% previously), hit counts to the sites increased nearly ten-fold, and self-reported ongoing use increased by 40%. These findings are congruent with our experience with a WBL site that was not password protected,<sup>20</sup> and support our previous proposition that passwords impede use of WBL.<sup>16</sup> However, use is still lower than might be expected, suggesting again that learners likely use other reference resources. This implies that when WBL developers face trade-offs between effective instructional methods and features to enhance usability as a reference, the former should prevail.

Our study has limitations. We did not assess patient-related outcomes.<sup>59</sup> Also, it is possible that the case-based questions in the intervention gave learners an advantage on the post-test, which also used case-based questions. However, we made intentional effort to avoid similarities among the cases and questions in the intervention and those in the assessment. Finally, the study involved a single training program, and the composition of our program may not be representative of that of other programs.

This report occupies a sparsely populated niche in WBL literature. The vast majority of publications to date are descriptive, akin to the clinical case report. While these demonstrate the feasibility of an intervention, they do little to inform practice. Almost all evaluative studies are limited either by the use of no-intervention controls or by comparison across different media (e.g., comparison of WBL to lecture or

textbook). Inasmuch as authors have consistently denounced media-comparative research for at least 20 years,<sup>25–27,60,61</sup> it is time for research in WBL to move forward through a line of research that produces generalizable knowledge and builds upon the past. This study provides a model of such research. Specific directions for research suggested by this study include comparisons of alternate instructional methods to engage learners, theory-based investigations of cognitive and learning styles, and the role of learner motivation in Web-based learning. We further suggest that comparing carefully selected variations in instructional method will provide more meaningful and generalizable results in medical education studies, regardless of the medium, than will other widely prevalent study designs.<sup>46</sup> Such studies will answer the recent plea for rigorous controlled trials in medical education.<sup>62</sup>

In conclusion, we found that self-assessment questions and feedback enhanced learning for internal medicine residents using a Web-based course in ambulatory medicine. We suggest that these findings hold implications for “traditional” educational activities as well as WBL—namely, that teachers must incorporate methods to actively engage learners in the learning process. As educators struggle to assist learners in the face of a rapidly growing body of information and decreasing time in which to learn, it will be increasingly important to identify effective educational practices. Future research, both in WBL and other teaching modalities, should focus on further defining the effectiveness of selected instructional methods in specific learning contexts.

## References

- Bowen JL, Irby DM. Assessing quality and costs of education in the ambulatory setting: a review of the literature. *Acad Med.* 2002;77:621–80.
- Cohen JJ. Reassert the ‘E’ in GME ([www.aamc.org/newsroom/reporter/jan05/word.htm](http://www.aamc.org/newsroom/reporter/jan05/word.htm)). Accessed 2 December 2005. Washington, DC, Association of American Medical Colleges.
- The ACGME Outcome Project (<http://www.acgme.org/outcome>). Accessed 2 December 2005. Accreditation Council for Graduate Medical Education.
- MacKenzie JD, Greenes RA. The World Wide Web: redefining medical education. *JAMA.* 1997;278:1785–86.
- Zucker S, White JA, Fabri PJ, Khonsari LS. Instructional intranets in graduate medical education. *Acad Med.* 1998;73:1072–75.
- Grundman J, Wigton R, Nickol D. A controlled trial of an interactive, Web-based virtual reality program for teaching physical diagnosis skills to medical students. *Acad Med.* 2000;75(10 suppl):S47–S49.
- Kumta SM, Tsang PL, Hung LK, Cheng JCY. Fostering critical thinking skills through a Web-based tutorial programme for final year medical students: a randomized, controlled study. *J Educ Multimedia Hypermedia.* 2003;12:267–73.
- Leong SL, Baldwin CD, Adelman AM. Integrating Web-based computer cases into a required clerkship: development and evaluation. *Acad Med.* 2003;78:295–301.
- Balcezak TJ, Lynch P, Jackson S, Richter J, Jaffe CC, Cadman EC. A web-based risk management and medical-legal curriculum for graduate medical education. *J Biocommun.* 1998;25(4):2–5.
- Curran VR, Hoekman T, Gulliver W, Landells I, Hatcher L. Web-based continuing medical education. (II): Evaluation study of computer-mediated continuing medical education. *J Contin Educ Health Prof.* 2000;20:106–19.
- Lipman AJ, Sade RM, Glotzbach AL, Lancaster CJ, Marshall MF. The incremental value of Internet-based instruction as an adjunct to classroom instruction: a prospective randomized study. *Acad Med.* 2001;76:1060–64.
- Gerbert B, Bronstone A, Maurer T, Berger T, McPhee S, Caspers N. The effectiveness of an Internet-based tutorial in improving primary care physicians’ skin cancer triage skills. *J Cancer Educ.* 2002;17:7–11.
- Harris JM, Jr., Kutob RM, Surprenant ZJ, Maiuro RD, Delate TA. Can Internet-based education improve physician confidence in dealing with domestic violence? *Fam Med.* 2002;34:287–92.
- Bell DS, Fonarow GC, Hays RD, Mangione CM. Self-study from web-based and printed guideline materials: a randomized, controlled trial among resident physicians. *Ann Intern Med.* 2000;132:938–46.
- Spickard A, III, Alrajeh N, Cordray D, Gigante J. Learning about screening using an online or live lecture: does it matter? *J Gen Intern Med.* 2002;17:540–45.
- Cook DA, Dupras DM, Thompson WG, Pankratz VS. Web-based learning in resident continuity clinics: a randomized, controlled trial. *Acad Med.* 2005;80:90–97.
- Kronz JD, Silberman MA, Allsbrook WC, Epstein JI. A web-based tutorial improves practicing pathologists’ Gleason grading of images of prostate carcinoma specimens obtained by needle biopsy: validation of a new medical education paradigm. *Cancer.* 2000;89:1818–23.
- Swagerty Jr. D, Studenski S, Laird R, Rigler S. A case-oriented web-based curriculum in geriatrics for third-year medical students. *J Am Geriatr Soc.* 2000;48:1507–12.

- 19 Harris JM, Salasche SJ, Harris RB. Can Internet-based continuing medical education improve physicians' skin cancer knowledge and skills? *J Gen Intern Med.* 2001;16:50–56.
- 20 Cook DA, Dupras DM. Teaching on the Web: Automated online instruction and assessment of residents in an acute care clinic. *Med Teach.* 2004;26:599–603.
- 21 Sisson SD, Hughes MT, Levine D, Brancati FL. Effect of an Internet-based curriculum on postgraduate education. A multicenter intervention. *J Gen Intern Med.* 2004;19:505–9.
- 22 Merrill MD. First principles of instruction. *Educ Tech Res Dev.* 2002;50(3):43–59.
- 23 Alur P, Fatima K, Joseph R. Medical teaching websites: do they reflect the learning paradigm? *Med Teach.* 2002;24:422–24.
- 24 Cook DA, Dupras DM. A practical guide to developing effective Web-based learning. *J Gen Intern Med.* 2004;19:698–707.
- 25 Clark R. Confounding in educational computing research. *J Educ Comput Res.* 1985;1:28–42.
- 26 Keane D, Norman G, Vickers J. The inadequacy of recent research on computer-assisted instruction. *Acad Med.* 1991;66:444–48.
- 27 Cook DA. The research we still are not doing: an agenda for the study of computer-based learning. *Acad Med.* 2005;80:541–48.
- 28 Brown G, Manogue M. AMEE Medical Education Guide No. 22: Refreshing lecturing: a guide for lecturers. *Med Teach.* 2001;23:231–44.
- 29 Kennedy GE. Promoting cognition in multimedia interactivity research. *J Interact Learn Res.* 2004;15:43–61.
- 30 Maleck M, Fischer MR, Kammer B, et al. Do computers teach better? A media comparison study for case-based teaching in radiology. *Radiographics.* 2001;21:1025–32.
- 31 Hudson JN. Computer-aided learning in the real world of medical education: does the quality of interaction with the computer affect student learning? *Med Educ.* 2004;38:887–95.
- 32 Devitt P, Palmer E. Computer-aided learning: an overvalued educational resource? *Med Educ.* 1999;33:136–39.
- 33 Gao T, Lehman JD. The effects of different levels of interaction on the achievement and motivational perceptions of college students in a Web-based learning environment. *Journal of Interactive Learning Research.* 2003;14:367–86.
- 34 Dillon A, Gabbard RB. Hypermedia as an educational technology: a review of the quantitative research literature on learner comprehension, control, and style. *Rev Educ Res.* 1998;68:322–49.
- 35 Chen SY, Paul RJ. Individual differences in web-based instruction: an overview [editorial]. *Br J Educ Technol.* 2003;34:385–92.
- 36 Cook DA. Learning and cognitive styles in Web-based learning: theory, evidence, and application. *Acad Med.* 2005;80:266–78.
- 37 Lieberman G, Abramson R, Volkan K, McArdle PJ. Tutor versus computer: a prospective comparison of interactive tutorial and computer-assisted instruction in radiology education. *Acad Radiol.* 2002;9:40–49.
- 38 Case SM, Swanson DB. Constructing Written Test Questions for the Basic and Clinical Sciences, 3rd ed. Philadelphia: National Board of Medical Examiners, 2001.
- 39 Felder RM, Soloman BA. Index of Learning Styles (<http://www.ncsu.edu/felder-public/ILSpage.html>). Accessed 2 December 2005.
- 40 Cook DA, Smith AJ. Assessment of convergent and divergent validity of scores for Kolb's Learning Style Inventory, Felder's Index of Learning Styles, and Riding's Cognitive Styles Analysis using the Multitrait Multimethod Matrix. *Educ Méd.* 2004;7(3):191.
- 41 Curry L. Individual differences in cognitive style, learning style, and instructional preference in medical education. In: Norman G, Van der Vleuten C, Newble D (eds). *International Handbook of Research in Medical Education.* Dordrecht: Kluwer Academic Publishers, 2002:263–76.
- 42 Fender GR, Prentice A, Gorst T, et al. Randomised controlled trial of educational package on management of menorrhagia in primary care: the Anglia menorrhagia education study. *BMJ.* 1999;318:1246–50.
- 43 Cornuz J, Humair JP, Seematter L, et al. Efficacy of resident training in smoking cessation: a randomized, controlled trial of a program based on application of behavioral theory and practice with standardized patients. *Ann Intern Med.* 2002;136:429–37.
- 44 Brown R, Bratton SL, Cabana MD, Kaciroti N, Clark NM. Physician asthma education program improves outcomes for children of low-income families. *Chest.* 2004;126:369–74.
- 45 Holmboe ES, Hawkins RE, Huot SJ. Effects of training in direct observation of medical residents' clinical competence: a randomized trial. *Ann Intern Med.* 2004;140:874–81.
- 46 Norman G. RCT = results confounded and trivial: the perils of grand educational experiments. *Med Educ.* 2003;37:582–84.
- 47 Beckman TJ, Cook DA. Educational epidemiology. *JAMA.* 2004;292:2969.
- 48 Roy MJ, Herbers JE, Seidman A, Kroenke K. Improving patient satisfaction with the transfer of care: a randomized controlled trial. *J Gen Intern Med.* 2003;18:364–69.
- 49 Sloan DA, Plymale MA, Donnelly MB, Schwartz RW, Edwards MJ, Bland KI. Enhancing the clinical skills of surgical residents through structured cancer education. *Ann Surg.* 2004;239:561–66.
- 50 Graham HJ, Seabrook MA, Woodfield SJ. Structured packs for independent learning: a comparison of learning outcome and acceptability with conventional teaching. *Med Educ.* 1999;33:579–84.
- 51 Steele DJ, Medder JD, Turner P. A comparison of learning outcomes and attitudes in student- versus faculty-led problem-based learning: an experimental study. *Med Educ.* 2000;34:23–29.
- 52 Haidet P, Morgan RO, O'Malley K, et al. A controlled trial of active versus passive learning strategies in a large group setting. *Adv Health Sci Educ.* 2004;9:15–27.
- 53 Ochsendorf FR, Boehncke WH, Boer A, Kaufmann R. Prospective randomised comparison of traditional, personal bedside and problem-oriented practical dermatology courses. *Med Educ.* 2004;38:652–58.
- 54 Jonassen DH, Grabowski BL. *Handbook of Individual Differences, Learning, and Instruction.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1993.
- 55 Merrill MD. Instructional strategies and learning styles: which takes precedence? In: Reiser R, Dempsey JV (eds). *Trends and Issues in Instructional Design and Technology.* Upper Saddle River, NJ: Merrill/Prentice Hall, 2002.
- 56 Cook DA. Reliability and validity of scores from the Index of Learning Styles. *Acad Med.* 2005;80(10 suppl):S97–S101.
- 57 Lynch TG, Steele DJ, Johnson Palensky JE, Lacy NL, Duffy SW. Learning preferences, computer attitudes, and test performance with computer-aided instruction. *Am J Surg.* 2001;181:368–71.
- 58 Naidr JP, Adla T, Janda A, Feberova J, Kasal P, Hladikova M. Long-term retention of knowledge after a distance course in medical informatics at Charles University Prague. *Teach Learn Med.* 2004;16:255–59.
- 59 Prystowsky JB, Bordage G. An outcomes research perspective on medical education: the predominance of trainee assessment and satisfaction. *Med Educ.* 2001;35:331–36.
- 60 Friedman C. The research we should be doing. *Acad Med.* 1994;69:455–57.
- 61 Adler MD, Johnson KB. Quantifying the literature of computer-aided instruction in medical education. *Acad Med.* 2000;75:1025–28.
- 62 Carney PA, Nierenberg DW, Pipas CF, Brooks WB, Stukel TA, Keller AM. Educational epidemiology: applying population-based design and analytic approaches to study medical education. *JAMA.* 2004;292:1044–50.
- 63 Felder RM, Silverman LK. Learning and teaching styles in engineering education. *Eng Educ.* 1988;78:674–81.