

## *Review question formats and web design usability in computer-assisted instruction*

**Rebecca S. Green, Marion A. Eppler, Marsha Ironsmith,  
and Karl L. Wuensch**

*Rebecca Green is a doctoral student in Industrial/Organizational Psychology at North Carolina State University. This research was carried out at East Carolina University. Marion Eppler, Marsha Ironsmith and Karl Wuensch are based in the Psychology Department at East Carolina University. Address for correspondence: Marion Eppler, Department of Psychology, East Carolina University, Greenville, NC, USA 27858–2025. Email: epplerm@ecu.edu*

### **Abstract**

We tested the effects of two embedded review question formats and the application of web design guidelines in a computer-assisted mastery learning course in developmental psychology. Students used either a branching review question format that redirected them to relevant portions of the study module after incorrect answers or a linear format that only provided the correct answer and then continued to the next review question. Students who used the branching format scored higher on their first attempt to pass the mastery quiz, and they required fewer tries to achieve the 90% mastery criterion. The effectiveness of web design guidelines was evaluated based on students' opinions. Students with positive opinions about the readability and navigational usability of the study module scored higher on their first quiz.

### **Introduction**

Effective instruction in a distance-learning environment depends on two factors: (1) application of empirically based learning principles and (2) attention to the usability of the technology, particularly web design principles. Instructors have access to research on the benefits of computer-assisted instruction (CAI) as a supplement to traditional classroom instruction (Cotton, 1991). Regarding the technology component, there is now ample advice on effective web design (Nielsen, 2000), but research addressing how web design usability influences student learning is scarce (Chalmers, 2003). One purpose of our study was to assess the effects of different review question formats on student learning. The second purpose of our study was to assess the relationship between students' opinions of web design usability and their retention of study materials.

Repeated testing enhances long-term retention (Roediger & Karpicke, 2006), but what kinds of review formats produce the highest level of mastery? One advantage of CAI

over classroom instruction is that CAI affords instructors the opportunity to provide students with multiple review questions that can be individualised to suit each student's level of mastery. Graham (1994, 1997) developed a computerised branching programme that forced students to review material until they have answered all questions with 100% accuracy. Incorrect answers directed students to a variety of branches designed to clarify misconceptions. Student opinions of this review format were highly favourable. Graham (2006) also compared the test scores of students who completed the tutorials electronically to students who had access to the same materials in printed form without the programmed accuracy requirement and branching features. Students who used the electronic materials scored 13% higher than students who used the printed materials.

Electronic pedagogical materials can be effective tools for student learning. However, instructors must take into account usability issues. CAI conveys information for learning, but the user interface or the look and feel of the learning environment may affect how well students learn the material. For example, disorientation is a common problem, when computer users do not have a sense of the overall organisation of information or they lose track of where they have been and where they can go (Chalmers, 2003). Even the aesthetics of screen design can have an impact on whether or not students complete a study module and how quickly they complete it (Szabo & Kanuka, 1998). Most research on web design guidelines has focused on users' perceptions, and studies examining the impact on learning are rare (Chalmers, 2003). Computer technology offers unique opportunities for instructors to enhance education, but we do not know yet which of the many factors that facilitate learning in more traditional educational settings will transfer to this medium.

#### *Purpose and design*

The first purpose of our study was to assess the effects of two versions of an electronic study module on students' quiz performance. Students in a mastery learning developmental psychology course were randomly assigned to complete either a linear or a branching electronic study module while preparing for a chapter quiz. The study modules were identical except for the linear versus branching format for embedded review questions. Both study modules required students to answer questions as they read a textbook chapter, and both had the same sets of review questions embedded at the same points throughout the module (four sets of five questions). The linear feedback module simply provided the correct response for each review question and allowed students to continue to the next question. The branching feedback module required students to re-examine relevant material after an incorrect response and to correct their answers before proceeding to the next review question. We hypothesised that students using the branching review format would score higher on their first attempt to pass a mastery quiz compared to students using the linear review format, and also that they would require fewer attempts to achieve a 90% mastery criterion.

The second purpose of our study was to assess the relationship between student opinions of usability and quiz performance. We used a compilation of web design guidelines

from industry and educational research to develop our electronic study module. We focused on three primary usability factors: (1) organisational guidelines, (2) navigational guidelines and (3) readability guidelines. Organisational guidelines focus on screen layout and the structure of the content. Two important organisational guidelines are the use of a table of contents and a site glossary, both directly linked to the content of the document (Chalmers, 2003; Nielsen, 2000). These features should help users to develop elaborate schemas that serve to guide acquisition and retrieval of information (Chalmers, 2003). Navigational guidelines include features that reduce user disorientation and clarify where users need to go next as they move from screen to screen (Chalmers, 2003). For example, navigational buttons should have a location and purpose that is immediately clear to the user, and the functions should be consistent across screens (Nielsen, 2000). Readability guidelines focus on ease of use including legibility (size and style of font, contrasting colours, etc), amount of text on a screen, time to load a screen and the general difficulty of reading information on a computer screen compared to that of reading information on printed text (Nielsen, 2000). Both electronic study modules incorporated the same web design usability guidelines, and we used questionnaire data to examine the relationship between students' perceptions of the usability of our electronic study modules and their performance on the mastery quizzes.

## Method

### *Participants*

Eighty students enrolled in a developmental psychology course participated in the study. The students included 70% European Americans, 19% African Americans and 11% other racial/ethnic identities. Ten percent were freshmen, 52% sophomores, 15% juniors, 22% seniors and 1% graduate or nondegree students. Students' self-reported grade point averages (GPA) ranged from 1.68 to 4.00 on a 4-point scale ( $M = 2.94$ ,  $SD = 0.62$ ). The students were randomly assigned to either a linear or a branching format group, and there was no significant difference in the GPA between the two groups,  $t(73.9) = -0.16$ ,  $p = 0.87$ . Thirty students completed fewer than four out of five sets of review questions (20 in the branching group and 10 in the linear group) and their data were therefore omitted from analyses. Some of these students inadvertently returned to an older version of the document because their browser cache was not updated, so they failed to complete sufficient electronic review questions. More students in the branching quiz group failed to complete all of the review questions because of a programming error that was corrected early in the testing week.

### *Materials and procedure*

#### Original course format

The course was a modified version of the Personalised System of Instruction (PSI) described by Keller (1968). PSI is a self-paced mastery format that features frequent quizzes on small units of material, immediate feedback, review units and, most importantly, a high mastery criterion (Buskist, Cush & DeGrandpre, 1991; Kulik, Jaksa & Kulik, 1978). Students completed one study module per week covering one or two textbook chapters. All study modules were available on the Internet with spaces for

students to type answers and notes, but the modules were not interactive and students generally printed them out and used a paper version to study. The study modules all consisted of learning objectives for the chapter, questions on the textbook material for students to answer and supplemental information. After completing each study module, students checked their answers against a key and then took a self-graded, 20-item paper-and-pencil practice quiz. When they felt ready, students proceeded to take computerised mastery quizzes in our testing lab. They received immediate feedback on their quiz performance and had up to four opportunities to achieve the required 90% mastery criterion by the weekly deadline. Students were limited to take no more than two quiz attempts per day, which means that if they waited until the deadline day then they had only two opportunities to take the quiz.

#### Modified course format

The learning module being evaluated in this study covered a chapter on the development of self-concept and personality. We converted our typical study module into an interactive module that was divided into five sections, each covering two–six learning objectives. The sections consisted of fill-in and short-answer questions for students to answer while reading the textbook chapter. After typing an answer, students could activate a button that provided the correct answer for them to check their work. Instead of taking a separate practice quiz, the practice quiz questions were embedded in the electronic module as review questions (four questions for each of the five sections for a total of 20 questions).

#### Review question formats

The review questions were incorporated directly into the electronic study module using Java scripting. Students were randomly assigned to use either a linear or a branching format for the review questions. Both formats provided immediate feedback as soon as students answered a review question. If the students' answers were correct, then they proceeded to the next review question. The two formats differed in the case of incorrect answers.

Students in the linear format group received a message stating the error, the correct answer and a list of the relevant sections of the study module that they needed to review. They were then allowed to proceed to the next review question. Students in the branching format group went through iterative rounds that required them to return to the relevant section of the study module before being allowed to return to the review questions. They were not given the correct answer to any question but were redirected to the same section of the module that they had studied earlier. These students had to answer each review question correctly before moving on to the next question.

#### Usability guidelines

The new study modules were designed to follow recommended web design usability guidelines to maximise organisation, navigation and readability. The major features for facilitating organisational usability included a table of contents that listed specific learning objectives and a glossary of definitions, both with direct links to relevant sections of

the study modules. We also used heading formats to organise information and to signify important concepts. Features for facilitating navigational usability included a consistent placement and function of action buttons across screens and attention to links and transitions. We used readability guidelines to direct our choices of font style and size, use of colour and amount of information presented on a screen or in a window. We also paid careful attention to file size, which influences the amount of time required to access information. We strove to keep the presentation of information simple and straightforward, without crowding of information or excessive use of animation.

At the end of the electronic study module, all students completed an online survey assessing their opinions about various features of the study module. Students were also asked to identify the best and worst features of the module from a list of choices.

## Results

We tested the effects of the review question format on students' performance on the first mastery quiz and on the number of attempts required to achieve the 90% mastery criterion with a pair of Welch-Satterthwaite separate-variances *t*-tests. The branching group's performance was significantly better than the linear group's performance for both dependent measures. The branching group scored significantly higher ( $M = 85.23$ ,  $SD = 11.18$ ,  $n = 22$ ) than the linear group ( $M = 80.0$ ,  $SD = 10.24$ ,  $n = 32$ ) on the first mastery quiz,  $t(42.6) = -1.75$ ,  $p = 0.045$ ,  $d = 0.5$ . The branching group also required fewer attempts to pass the quiz ( $M = 1.73$ ,  $SD = 0.90$ ,  $n = 11$ ) than the linear group ( $M = 2.37$ ,  $SD = 1.12$ ,  $n = 19$ ),  $t(24.7) = 1.71$ ,  $p = 0.0495$ ,  $d = 0.7$ . The analysis for the number of attempts excluded students (13 linear, 11 branching) who waited until the deadline day to take the mastery quizzes as these students had only two opportunities to pass the quiz.

An item analysis of the student opinion survey yielded four factors: (1) organisational usability, which included items on helpfulness of instructions, use of prompts and help features, clarity and usefulness of the table of contents and site glossary ( $\alpha = 0.71$ ), (2) navigational usability, which included items on clarity and effectiveness of the navigational buttons and links ( $\alpha = 0.87$ ), (3) readability, which included items on the amount of information on each page and the difficulty of using the electronic version compared to that of using the print versions of the study module ( $\alpha = 0.75$ ) and (4) a fourth factor that seems to reflect more general usability issues such as confidence, how much they felt the module helped their performance, preference for this format over the paper versions and the overall level of difficulty ( $\alpha = 0.83$ ).

We created four composite scores that were the average of the items for each factor, and these composite scores were used in subsequent analyses. Table 1 includes the mean ratings for each of the four usability factors presented overall and separately for the two review question format groups. Students' perceptions of web design features were related to quiz performance, but only for some design features. The performance on the first mastery quiz was significantly correlated with student ratings of navigational

Table 1: Mean student ratings of web design features

Module feature	Overall		Linear		Branching	
	M*	SD	M*	SD	M*	SD
General usability	3.57	0.85	3.81	0.63	3.28	1.00
Navigational usability	4.17	0.60	4.16	0.54	4.18	0.68
Organisational usability	3.82	0.71	3.96	0.60	3.64	0.80
Readability	3.43	0.58	3.59	0.39	3.26	0.71

\*Based on a 5-point scale: 5 = Strongest positive rating

usability ( $r = 0.24$ ,  $p < 0.034$ ) and readability ( $r = 0.31$ ,  $p < 0.008$ ), but not with organisational usability ( $r = -0.16$ ,  $p < 0.11$ ) or general usability ( $r = 0.17$ ,  $p < 0.10$ ).

When students identified the best and worst features of the electronic modules, the review questions received the highest number of nominations for both the best and the worst features. A majority of the linear group (56%) identified the review questions as the best feature of the module, while 36% perceived the quizzes to be the worst feature. In contrast, 54% of students in the branching group identified them as the worst feature while 23% identified them as the best feature. None of the other features were identified as best or worst by more than 25% of the students.

## Discussion

As predicted, students in the branching condition scored higher on their first mastery quiz and required fewer attempts to reach mastery than students in the linear condition. Students' opinions of readability and navigational usability predicted quiz performance, but opinions of organisational and general usability were not related to quiz scores.

Ample evidence attests that the structure and organisation of materials in computer-assisted and online instruction affects the usability of the materials as well as students' perceptions of these materials (Chalmers, 2003; Nielsen, 2000). However, few studies have assessed the impact of usability features on students' actual learning and performance. One contribution of our study was to offer empirical evidence of a link between specific web design features and students' quiz performance. Students who gave higher ratings for navigational usability and readability earned higher quiz scores. Our study is exploratory and the results do not allow us to give specific advice about how design features impact student outcomes. However, we provide preliminary data to suggest that perceptions of usability relate to student performance.

The embedded review questions were the most salient feature of the modules for students as students were most likely to nominate them as the best or worst feature. The linear group tended to perceive the review questions to be the best feature of the online module, but the branching group perceived them to be the worst feature. One possible reason why the branching group disliked the review questions may be related to pro-

gramming glitches, which made it difficult for some students in this group to access all of the review questions. Alternatively, the branching group may have been frustrated by having to search for the correct answer rather than being given immediate feedback. Ironically, whether they liked them or not, the review questions allowed the branching group to perform better on the quiz and to pass in fewer tries. Higher attrition in the branching group raises the possibility that students who completed the branching module were more persistent, started the module earlier or were better equipped to cope with the technological challenges.

Assessment of pedagogical materials needs to include both students' perception of the materials and measures of performance. When assessing student opinions of design features, it would be useful to ask specific questions about individual features and to include open-ended questions to determine why students like or dislike a feature. When relating student opinions to performance, it is important to bear in mind that a particular feature could be unpopular but still result in better mastery of the material. Future research might use within-subjects designs to help students see firsthand the changes in their performance under different conditions. For example, students could be exposed to two or more review question formats and then see firsthand which helps them learn the material better. One question is whether or not they would come to appreciate the benefits of searching for answers.

Technological issues such as those that we encountered pose a challenge for online instruction and empirical assessment. Despite these challenges, an increasing reliance on computer technology and a greater emphasis on accountability in education (National Commission on Accountability in Higher Education, 2005; Moran & Malott, 2004; Whitehurst, 2003) require researchers to demonstrate the efficacy of pedagogical and technological design on students' learning. This study demonstrates that both the review question format and the usability features of web design influence student performance.

### Acknowledgements

Thanks are due to Ashley Hicks, Joe Dunkelman, Steve Timmons, Shannon Lee, Angi Bosso and the students of our Developmental Psychology class.

### References

- Buskist, W., Cush, D. & DeGrandpre, R. J. (1991). The life and times of PSI. *Journal of Behavioral Education, 1*, 215–234.
- Chalmers, P. A. (2003). User interface improvements in computer-assisted instruction, the challenge. *Computers in Human Behavior, 16*, 507–517.
- Cotton, K. (1991). *Computer-assisted instruction*. Retrieved June 5, 2006, from <http://www.nwrel.org/scpd/sirs/5/cu10.html>
- Graham, R. B. (1994). A computer tutorial for psychology of learning courses. *Teaching of Psychology, 21*, 115–116.
- Graham, R. B. (1997). A computer tutorial on consequences in operant learning. *Teaching of Psychology, 24*, 216–217.
- Graham, R. B. (2006). *Controlling student study behavior through computerization of course content*. Unpublished manuscript, East Carolina University.

- Keller, F. S. (1968). "Good-bye, teacher...". *Journal of Applied Behavior Analysis*, 1, 79–89.
- Kulik, J. A., Jaks, P. & Kulik, C.-L. C. (1978). Research on component features of Keller's personalized system of instruction. *Journal of Personalized Instruction*, 3, 2–14.
- Moran, D. J. & Malott, R. W. (Eds) (2004). *Evidence-based educational methods*. San Diego, CA: Elsevier.
- National Commission on Accountability in Higher Education (2005). *Accountability for better results: a national imperative for higher education*. Retrieved June 5, 2006, from <http://www.shceo.org/account/accountability.pdf>
- Nielsen, J. (2000). *Designing web usability*. Indianapolis, IN: New Riders Publishing.
- Roediger, H. L. & Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255.
- Szabo, M. & Kanuka, H. (1998). Effects of violating screen design principles of balance, unity and focus on recall, learning, study time and completion rates. *Journal of Multimedia and Hypermedia*, 8, 23–42.
- Whitehurst, G. J. (2003). Rigor, relevance, and utilization. *APS Observer*, 16, 1–2.