The Psychological Record (2018) 68:489–499
https://doi.org/10.1007/s40732-018-0295-y

ORIGINAL ARTICLE

Using Feedback and Postfeedback Delays to Improve Performance with Online Lessons

Anna L. McCalpin[1] · Douglas A. Johnson[2] · Thomas N. Ferragut[2]

Published online: 29 May 2018

Abstract

Computer-based instruction has become an increasingly popular tool in both business and education throughout the last decade. Despite the various benefits of using computer-based instruction, there are several challenges that accompany this mode of instruction, including computer-based racing. Computer-based racing occurs when learners respond so quickly that frequent mistakes are made. The purpose of the current study was to investigate the effect of postfeedback delays and different feedback options on performance with online lessons conducted in uncontrolled settings. Six different computer-based instructional formats were assessed in terms of learner performance using a between-group pretest-posttest design. Statistically significant differences were observed for both the delay and feedback variables. The results of this study extend the current literature on postfeedback delays by suggesting that an overt form of self-evaluation during a delay may not be necessary for postfeedback delays, and that postfeedback delays may be effective in uncontrolled environments.

Keywords Computer-based instruction · Computer-based racing · Postfeedback delays · Feedback · Self-evaluation · Online education · Training

Computer-based instruction (CBI) has become an increasingly common tool for instruction in both education and industry (Marroletti & Johnson, 2014). In recent years, there has been a significant shift from traditional in-class interactions to computer-based and computer-assisted instruction (i.e., both standalone and supplemental models) in higher education. Despite overall declines in enrollment for higher education, distance education continued its growth, with 5.8 million distance education students enrolled during fall 2014 (Allen & Seaman, 2016). This trend is likely to continue due to the potential for reaching students not easily served by face-to-face education, thanks to a greater degree of flexibility in location, time and resources (Robinson, 2017; Wei, Peng, & Chou, 2015).

Early behaviorists such as B. F. Skinner were proponents of automated forms of instruction (Johnson, 2014). Although he pioneered mechanized instruction that may appear unsophisticated by today's standards (e.g., the need to clumsily rotate knobs and turn levers), many of Skinner's recommended best practices remain relevant for modern forms of computerized instruction. It is also notable that Skinner was one of the earliest advocates for using automated forms of education to give learners the ability to progress through the material at their own pace (Skinner, 1958). Much like the teaching machines of the 1950's, computer-based instruction (CBI) has the ability to provide immediate and individualized feedback for responding, regardless of class size or the time at which an assignment is completed. A review by Johnson and Rubin (2011) summarized 12 years of comparative research on CBI between 1995 and 2007. Specifically, the reviewers looked at research on interactive CBI relevant to employee training techniques. Interactive CBI means the learner's response is demonstrative, requiring the student to emit an overt and accurate response relevant to the subject matter before proceeding to new material. The authors reported that interactive CBI was found to be comparable, or even superior to,

✉ Anna L. McCalpin
annaleahconard@gmail.com

Douglas A. Johnson
behavioranalyst@gmail.com

Thomas N. Ferragut
thomas.n.ferragut@wmich.edu

1 Ardent Learning, Inc., 300 State St. #401, Rochester, NY 14614, USA

2 Department of Psychology, Western Michigan University, 1903 West Michigan Ave, Kalamazoo, MI 49008-5439, USA

instructional alternatives (e.g., classroom instruction, textbook/manuals, etc.) 95.2% of the time. Specifically, the authors found that 64.3% of instructional comparisons demonstrated improvements through the use of interactive CBI, while 31% reported no difference or mixed results. There are a number of CBI applications common in both business and education, with one popular application being eLearning. eLearning is an online tool for delivering instruction using numerous types of media (text, audio, images, etc.). This type of CBI application may be used as standalone instruction or may supplement other forms of instruction.

Despite the tremendous potential for eLearning, there is also considerable reason for concern. Such courses often face high levels of attrition (Angelino, Williams, & Natvig, 2007; Willging & Johnson, 2004; Xu & Jaggars, 2011) and significant upfront investments for development (Chapman, 2010), social isolation (Woods, 2002), and lack of training for implementation (Allen & Seaman, 2011; Bodner, 1997). Much of the current use of CBI is often a replication of traditional teaching techniques and the required responding is largely passive in nature (Johnson, 2014; Markle, 1990). According to Johnson and Rubin (2011), while many current computer-based programs claim to be interactive, this level of interactivity may be no more advanced than simply progressing through a textbook. An interactive CBI program should allow the student to not only progress at his or her own pace, but also allow the student to demonstrate his or her knowledge and understanding of the material. The nature of these interactions is likely a key consideration in developing successful CBI.

Although much has been written about best practices when it comes to CBI (Johnson & Rubin, 2011; Mayer, 2014; Vargas & Vargas, 1991), there are still many considerations to investigate, especially when one considers the increasing trend in usage and potential drawbacks when done poorly. One consideration regarding computer-based interactions is how the learner paces his or her responding. While self-pacing is a frequently cited advantage of CBI, some types of self-pacing may have detrimental effects on learning. There are two types of self-pacing in CBI: overall course pacing and within unit pacing (Johnson & Dickinson, 2012). Overall course pacing refers to the deadlines in which learners are expected to complete the assigned material. Within unit self-pacing refers to the time spent studying within a specific instructional unit. From a behavioral perspective, previous research suggests that self-pacing should not be used for overall course pacing, as learners are typically poor managers of their own time and more likely to procrastinate. Conversely, self-pacing within instructional units remains a primary benefit of CBI.

When CBI is poorly designed, there is a high likelihood that the learner will try to avoid or escape the instructional environment. As such, it is important to try to make CBI more palatable to learners to promote approaching and orienting behaviors (Marroletti & Johnson, 2014). However, a problem remains even if the aversive elements of instruction have been minimized: CBI is typically still competing against an array of activities with more reinforcing value. When learners are allowed to self-pace in a CBI environment (one of the promoted benefits of CBI), these learners are also likely to respond quickly in order to move on to a more reinforcing set of conditions (i.e., a non-instructional environment). Unfortunately, the learner's responding may become too rapid for learning to take place, a phenomenon termed computer-based racing (Johnson & Dickinson, 2012). When such racing occurs, students are not attending to the instructional material and move through the lesson at a detrimentally fast pace. As previously mentioned, one of the most significant contributions of CBI is the ability to engage the learner in meaningful responding. By hurrying through the unit without attending to the instructional material, learners are no longer able to engage in such responses.

One investigated method for reducing computer-based racing is the use of postfeedback delays (Crosbie & Kelly, 1994; Dubuque, 2012; Johnson & Dickinson, 2012). When using CBI with postfeedback delays, the learner is presented with some form of feedback following a correct or incorrect response. After presenting this feedback, the computer will enforce a delay for a predetermined amount of time before the learner can continue through the material. Postfeedback delays appear to be effective in that they foster additional exposure time to the material. Feedback is a frequently utilized and beneficial variable in many applications involving training and performance (Alvero, Bucklin, & Austin, 2001; VanStelle et al., 2012), however, the ubiquity of feedback does not elevate it to the status of a behavioral principle (Peterson, 1982). Feedback is broad stimulus class that can serve many different functions depending on the particular components, pre-existing history, or method of delivery employed in each application. This has led some to call for research to more carefully examine the features of feedback that make it more or less potent (Johnson, 2013). Postfeedback delays represent potentially important delivery consideration for the use of feedback.

Crosbie and Kelly (1994) was one of the first studies to investigate postfeedback delays with lengthy and well-tested instructional material. Specifically, they investigated the possibility of postfeedback delays functioning as punishers for incorrect responding. Instructional content was drawn from the Holland and Skinner (1961) textbook, which required readers to supply missing words in reaction to frequent overt response requests. The experimental task replicated this by requiring participants to type missing words in the same manner. Furthermore, participants had to self-score their answers as correct or incorrect after receiving feedback. The authors compared performance under three conditions: 10s delay following all answers (correct and incorrect), 10s delay following incorrect answers only, and no delay. The authors found no substantial difference in responding between the no delay condition and the 10s delay following only incorrect answers. They did, however, report that the 10s delay following all answers improved performance over the no delay condition. These findings suggest that punishment was not the mechanism of action for the effectiveness of postfeedback delays. In a second experiment utilizing the same task, Crosbie and Kelly compared a blank-screen delay, in which no material was presented throughout the delay, to a postfeedback delay and no delay condition. The authors found no significant difference between the no delay condition and the 10s blank-screen delay condition. They did, however, find that the 10s postfeedback delay condition resulted in higher performance than the other two conditions. These findings suggest that extra time spent looking at the content during the 10s delay may have resulted in the increased performance.

A later study by Kelly and Crosbie (1997) confirmed their previous finding that an opportunity to review instructional content during a forced postfeedback delay results in better performance. The authors utilized the same experimental task as the previous studies but modified their previous study by shortening the length of experimental sessions and adding pretest, posttest, and follow-up tests to assess the impact of their interventions over time. For this experiment, subjects were exposed to either a 10s postfeedback delay for each question or no postfeedback delay. Similar to their previous findings, the authors found that the postfeedback delay improved performance, and these improvements were maintained and even increased throughout the remainder of the experiment.

A later study by Johnson and Dickinson (2012) evaluated performance using three different formats: postfeedback delay, incentives/disincentives, and control. Similar to previous studies, the instructional material was drawn from the Holland and Skinner textbook (1961) and utilized frequent typed response requirements. Unlike the previous studies, Johnson and Dickinson applied a shorter delay of just 5s. For the postfeedback delay condition, participants were paid 5 cents for each question they completed, regardless of accuracy, and encountered a 5s delay in which the question, feedback, and the participants' responses were displayed on the screen. For the incentives/disincentives condition, participants were paid 5 cents for each question they answered correctly and lost 5 cents for each incorrect response. Unlike the postfeedback delay condition, there was no delay in this condition. For the control condition, participants were paid 5 cents for each response regardless of accuracy, and again, did not encounter a delay. The authors found that posttest scores increased with the use of postfeedback delays as compared with the other experimental conditions.

Dubuque (2012) also investigated the use of postfeedback delays in three separate experiments which used order of operation math problems as the instructional content. For the first experiment, participants were exposed to three conditions: control, contingent delay, and contingent interactive delay. Problems answered correctly for all three conditions produced immediate feedback, no enforced delay, and access to the next problem. Incorrect answers in the control condition produced the same consequence as correct answers. Problems answered incorrectly for the contingent delay condition resulted in immediate feedback and an enforced 60-s delay following the response. Incorrect answers under the contingent interactive delay condition required subjects to click a button every 5s throughout the 60s delay in order to continue through the material. Results from the first study suggested a ceiling effect, as the authors reported high levels of responding under all three conditions. This finding suggested the importance of using challenging instructional material in order to detect differences among experimental conditions. In order to address this, Dubuque completed a second experiment with more difficult questions and found a significant increase in performance for the contingent delay condition when compared to the other two conditions. The third and final experiment decreased the delay from 60s to 30s for the contingent delay condition and requiring the participant to click a button (within 3s) at a random time throughout the delay for the contingent interactive delay condition. Unlike Crosbie and Kelly (1994), results from these three experiments found that delays could be effective even in the absence of further exposure to instructional material. However, it is possible that this finding may be limited to only long postfeedback durations and shorter delays may still require instructional content to remain visible in order to be effective.

While previous research on postfeedback delays has provided valuable contributions in reducing racing during CBI, there are several limitations that warrant attention. Firstly, the majority of previous studies have been conducted in highly controlled laboratory settings (Crosbie & Kelly, 1994; Johnson & Dickinson, 2012; Kelly & Crosbie, 1997), thus removing most real-life competing activities that likely have an impact on performance. It is possible that the participants in previous research had nothing better to do than to study the material presented on the screen during the delay. It is further possible that, if completed in a location of their choosing, learners may engage in competing behaviors rather than study the material. Given this, CBI research conducted in a controlled lab setting may prevent findings from generalizing to an actual online lesson.

Another limitation in previous research is the length of the postfeedback delay. Longer postfeedback delays (e.g., 10, 30, and 60s) greatly increase instructional time and make a potentially aversive situation even more aversive (going against the intent of Skinner's automated instruction). Kelly and Crosbie (1997) found that while a longer postfeedback delay (10s) significantly improved performance, training took 20%

longer and some subjects complained that, regardless of the correctness of their answer, their progress was unfairly delayed. Similarly, Dubuque (2012) found that the use of much longer delays (30s and 60s) significantly improved performance, such lengthy delays may create a considerable barrier for adoption in training and education.

Finally, many of the previous studies (Crosbie & Kelly, 1994; Johnson & Dickinson, 2012; Kelly & Crosbie, 1997; Munson & Crosbie, 1998) required participants to self-score their responses as correct or incorrect. However, none of these studies tested whether this was an important component for the effectiveness of the experimental conditions.

In order to more closely simulate the typical use of CBI, the lessons used in this study were completed online in a location of the participant's choosing. The current study also utilized shorter durations of postfeedback delays to improve the potential acceptability of postfeedback delay procedures for most educational and business settings. Finally, this study would test the necessity of self-scoring on the effectiveness of feedback. In sum, the purpose of the current study was to investigate the use of short postfeedback delays on performance when lessons are completed in an uncontrolled environment.

Method

Participants and Setting

193 undergraduate students were recruited using recruitment flyers and in-class recruitment from a large university. 159 participants completed the study in its entirety and attrition was roughly equally distributed across conditions. Two sessions (pretest and posttest) were conducted in a university laboratory containing four workstations. Cubicle walls separated each workstation and participants were never placed adjacent to each other to prevent participants from viewing the work of other participants. Instructional modules were completed online in a location of the participant's choosing (home, coffee shop, computer lab, etc.).

Instructional Material

A computer program, using instructional material from *The Analysis of Behavior* (Holland & Skinner, 1961), was used for the instructional modules. This research program was privately available to participants only and involved a modification of the instructional modules used in Johnson and Dickinson (2012). The pretest and posttest was paper-based and also developed using Holland and Skinner's text. Sets 1-16 and 18-22 (21 total) were used for the instructional modules. Questions from sets 17 and 29 were combined to construct the pretest and posttest, as these sets are cumulative reviews of all previous units. As such, the pretest and posttest closely corresponded to the content and format of the instructional content. The instructional content covered topics such as respondent conditioning, reinforcement, cumulative records, shaping, schedules of reinforcement, and stimulus control. For 16 of the instructional modules, additional "exhibit" printouts (based on the exhibits used in the Holland and Skinner textbook) supplemented each of the instructional sets. Instructional material from *The Analysis of Behavior* was used in order to avoid ceiling effects, as 50% correct responding was a typical outcome for posttest measures in a previous study (Johnson & Dickinson). Additionally, previous research has successfully applied this text for investigating computer-based racing (Crosbie & Kelly, 1994; Johnson & Dickinson; Kelly & Crosbie, 1997).

All instructional sets and a tutorial were emailed to participants within 24 h of the introductory/pretest session. The program tutorial allowed the participant to become familiar with the CBI format before beginning the instructional sets. The tutorial did not include any of the material from Holland and Skinner's (1961) text, but was simply used as a tool to familiarize participants with the navigation and format of the modules.

Upon opening the program and selecting a specific unit, participants clicked on the "Begin Unit" button to begin the lesson. The program was presented on the computer screen, which displayed the unit number and total number of questions for the unit. The slides used for each instructional module included short, incomplete statements that required participants to type a response. After typing a response into the appropriate field, participants clicked the "Submit Your Answer" button immediately below or next to the response field. For participants in feedback conditions, the screen displayed the correct answer and participants viewed feedback for an unlimited and unmeasured duration. At the end of each instructional set, participants were required to click on the "End Unit" button displayed on the screen to complete the lesson. Sample images from the program can be seen in the appendix Fig. 3.

Pretest and Posttest Measures

The pretest measure consisted of a 51-question paper-based test and the subject's score on the pretest was used as a covariate measure of performance. Additionally, the pretest was used to filter out any subjects who were fluent with the material. The questions used for the pretest included short, incomplete statements that required participants to write a response. Answers were scored as correct irrespective of minor spelling errors and if responses were considered reasonably synonymous. Prior to completing the pretest, subjects were told that they would earn $5 cash by scoring higher than 65% on the test. Any subject who met this criterion was immediately excluded from further participation in the study; however, participants were not informed of this prior to completing the pretest. No participants met this criterion and therefore no participants were removed on the basis of the pretest. The same set of 51 questions was used as the posttest measure, and again, subjects were told that they would earn $5 cash by scoring higher than 65% on the test. Further, regardless of test scores, participants earned $15 for simply completing all 21 modules and the posttest within a three-week period. The incentives were used to ensure that participants made genuine attempts to do well on the tests, and do so within the time interval.

Dependent and Independent Variables

Posttest scores were used to assess differences between six CBI groups. Following completion of the posttest, the experimenter immediately evaluated and recorded the participant's score on his or her personal record sheet. Participant record sheets were filed in a secure location and only accessible to the experimenter and research assistants. In addition, post-participation surveys were administered to all subjects following their participation in this study. In order to confirm the participant's completion of all 21 modules within a three-week period, the lead experimenter had exclusive access to a webpage with the participant number and the time stamp associated with each unit. The exact amount of time spent within each unit was not tracked.

The independent variables investigated in this study were the delay length (5s or no delay) and feedback type (self-evaluative feedback, feedback only, or no feedback). Participants were randomly assigned to one of the six experimental groups. See Fig. 1 for a step-by-step comparison of the different procedures for each group.

Feedback Only with 5s Delay At the start of each instructional unit, participants clicked the "Begin Unit" button to begin the instructional set. After clicking the "Begin Unit" button, the screen displayed a question, a response box, and a "Submit Your Answer" button. After reading the question, participants responded by typing their answer into the appropriate response field and clicking the "Submit Your Answer" button. Participants in this condition were then presented with a screen in which the correct answer was displayed directly next to the participant's typed response. In addition to the question, participant response, and correct answer, a horizontal countdown bar that gradually decreased in size remained visible on the screen throughout the 5s delay. After the countdown bar disappeared, participants were able to click the "Proceed to Next Question" button and continue through the lesson. Participants were able to view this screen for as long as they liked until they clicked the "Proceed to Next Question" button.

Feedback Only with No Delay This condition was identical to the previous condition, except participants were not exposed to the enforced delay of 5s. Rather, participants assigned to this condition had immediate access to the "Proceed to Next Question" button following the presentation of the question, correct answer, and participant response. Again, participants were able to view this screen for as long as they preferred until they clicked the "Proceed to Next Question" button.

Self-Evaluative Feedback with 5s Delay This condition was identical to the previous 5s delay condition, except participants were exposed to a self-evaluative feedback component. Similar to the previous conditions, the question, participant response, and correct answer were displayed immediately after submitting a response. However, following the presentation of this material, subjects were required to score the correctness of their response by typing either "C" (correct) or "I" (incorrect) into the appropriate field. After they reviewed the correct answer and scored their response, participants clicked on the "Submit Scoring" button. After they submitted a score, a 5s countdown immediately began in which the participant was unable to continue until after the countdown bar disappeared. Participants were able to review the content for as long as they preferred, until they clicked the "Proceed to Next Question" button and moved on to the next question.

Self-Evaluative Feedback with No Delay This condition was identical to the previous condition, except participants were not exposed to the 5s enforced delay. Rather, the "Proceed to Next Question" button became immediately available after the participant scored his or her response. Again, participants were able to review the content for as long as they preferred, until they clicked the "Proceed to Next Question" button and moved on to the next question.

No Feedback with 5s Delay After the participant typed a response and clicked the "Submit Your Answer" button, a screen appeared with only the question and countdown bar in view. Unlike the previous conditions, participants did not have access to the correct answer, nor were they able to view their response after they submitted an answer. When the countdown bar disappeared, a "Proceed to Next Question" button immediately appeared and clicking the button allowed the participant to advance to the next question.

No Feedback with No Delay This condition was identical to the previous condition, except the "Proceed to Next Question" button became immediately available after the participant submitted a response. Again, the participant did not have access to his or her response after they clicked the "Submit Your Response" button.

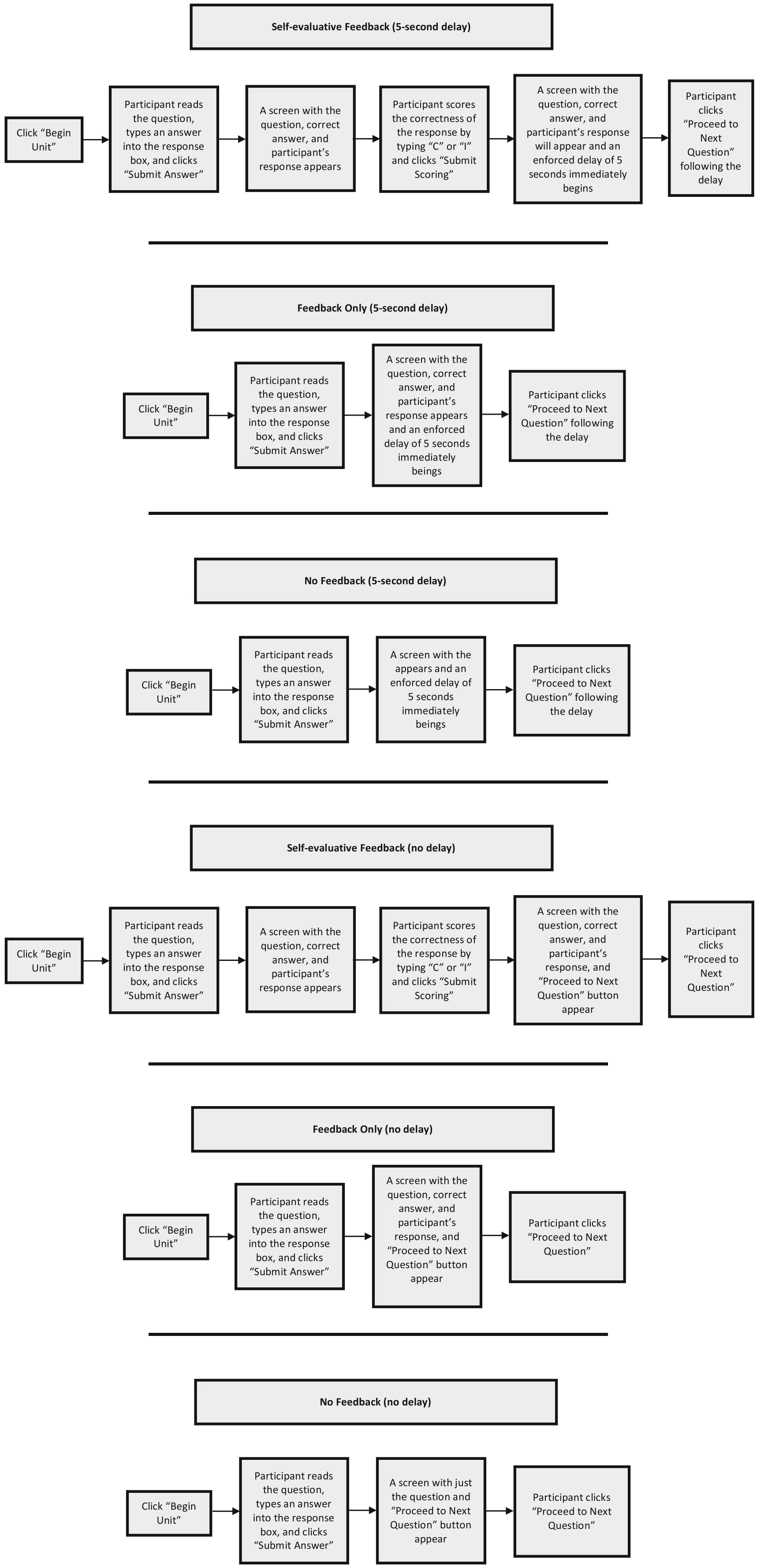


Fig. 1 Comparison of experimental groups

Experimental Design

A between-group pretest-posttest design was used and participants were randomly assigned to one of six experimental groups: a) Self-evaluative feedback with a 5s delay (24 participants), b) Feedback only with a 5s delay (30 participants), c) No feedback with a 5s delay (23 participants), d) Self-evaluative feedback with no delay (31 participants), e) Feedback only with no delay (25 participants), or f) No feedback with no delay (26 participants).

Experimental Procedures

Pretest Session During the pretest/introductory session, the experimenter briefly explained the study, including details such as where the study will take place, the required time commitments, and the need to complete instructional materials and tests. The experimenter then randomly assigned the participant to one of six experimental conditions, and had the participant complete the pretest.

After the participant completed the pretest, the posttest/debriefing session was scheduled approximately three weeks following the introductory session. The experimenter informed the participant that he or she will receive an e-mail with the instructional sets and program tutorial within 24 h. The participant was also told that the modules could be completed at a location of their choosing, and all instructional sets had to be completed before taking the posttest. The experimenter provided the participant with a folder composed of 16 instructional "exhibits" (which supplied graphs and other support materials needed for the instructional sets).

Posttest Session The content and administration of the posttest was also identical to the pretest. Participants earned an additional $5 for scoring higher than 65% on the test. Participants were also debriefed following the posttest.

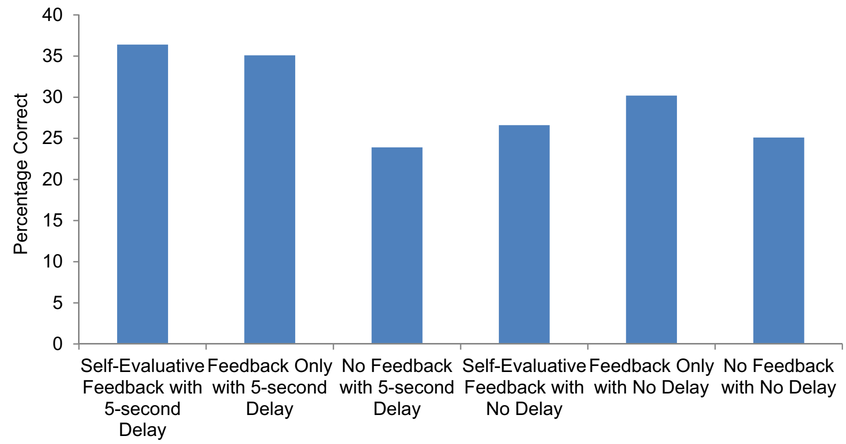
Results

A two-factor ANCOVA was used to analyze posttest scores for all participants in the six experimental groups, with the pretest score from the introductory session serving as the covariate measure. An ANCOVA was selected due to its higher power relative to an ANOVA and its ability to reduce bias from differences based on pre-intervention baseline measures (Huitema, 2011). As is standard with an ANCOVA, adjusted means were compared in order to allow the dependent variable (i.e., posttest scores) to be adjusted to control for any existing group difference related to the baseline measure (i.e., pretest scores). Figure 2 displays the adjusted means for percentage correct on the posttest for all 6 conditions. The assumptions for the use of an ANCOVA were met. There was a statistically significant main effect for both the feedback variable ($F(2, 152) = 4.99$, $p = 0.008$) and the delay variable ($F(1, 152) = 4.09$, $p = 0.045$). Tukey pairwise comparisons revealed statistically significant differences between the self-evaluative feedback with 5s delay condition and the no feedback with 5s delay condition ($p = 0.031$), the self-evaluative feedback with 5s delay conditions and the no feedback with no delay condition ($p = 0.053$), and the feedback only with 5s delay condition and the no feedback with 5s delay condition ($p = 0.052$). No other individual group differences were discovered. Effect size calculations indicated a Cohen's d of 0.81 for the self-evaluative feedback with a 5s delay condition, 0.71 for the feedback only with a 5s delay condition, 0.36 for the feedback only with no delay condition. 0.11 for the self-evaluative feedback with no delay condition, and -0.08 for the no feedback with a 5s delay condition. Inter-scorer agreement for both pretest and posttest scoring averaged 98.1% (ranging between 94.2% and 100%) and was collected for all pretests and posttests. There was an average agreement of 97.3% for the pretest scoring and 98.9% agreement for the posttest scoring. To determine inter-scorer agreement, two experimenters scored each test independently, marking each response as correct or incorrect. A second experimenter then independently scored the accuracy of responses at a later time. To calculate inter-scorer agreement, the number of agreements was divided by the number of agreements plus disagreements, and then multiplied by 100. In the interests of expediting grading for the purposes of paying incentives, graders were not blind to experimental conditions. According to survey data collected after the posttest, the three most frequently reported alternative activities while completing the modules were talking/texting on the phone (81%), checking social media (60%), and socializing with friends (53%).

Discussion

While previous research on the use of postfeedback delays has demonstrated their effectiveness for increasing the retention of instructional material, a majority of studies have been conducted in highly controlled laboratory settings. Further, some of the previous research investigated longer postfeedback delays (i.e., 60s), possibly limiting adoptability for most educational and business settings. Given these limitations, the purpose of the current study was to investigate the use of brief postfeedback delays to improve the retention of CBI within realistic environments. As indicated by the results, the two conditions that incorporated feedback and an enforced delay had the most significant effects in comparison to conditions that lacked this combination. Of particular interest is the fact that both of these conditions had significant differences compared to the condition that had an enforced delay but did not provide any kind of feedback to participants.

Fig. 2 Adjusted means for percentage correct on posttest



The implication of these analyses is that postfeedback delays are effective at improving performance, even when used in uncontrolled environments, but delays in general are not effective. That is, delays only have functional value in the presence of feedback. However, requiring learners to overtly self-evaluate the accuracy of their responding appears to be irrelevant in the context of postfeedback delays. In other words, feedback matters, but self-assessment of that feedback does not. These findings are largely in alignment with previous studies examining the efficacy of postfeedback delays in reducing computer-based racing and retaining instructional material in CBI (Crosbie & Kelly, 1994; Dubuque, 2012; Johnson & Dickinson, 2012).

The investigation of postfeedback delays in uncontrolled environments was a key component of the current study and one that distinguishes it from other research in this area. Investigating the effectiveness of postfeedback delays in real-world settings offers insight to the practical worth of such delays in CBI. As such, the results are noteworthy given all of the unknown and potential competing variables introduced by the natural environment. Johnson and Dickinson (2012) conducted a study that also used 5s feedback with identical materials and similar methodology, but took place in a controlled lab setting over seven sessions with limited availability of distractions. The current study was able to reproduce the effect of the 5s postfeedback delay in an uncontrolled environment. In this study, participants were able to access many potential alternative activities, and content was self-paced rather than dispersed over the course of seven sessions. Overall, these results are suggestive that at-home CBI can be made more effective by the inclusion of both feedback and enforced delays in spite of access to competing activities and the freedom to self-pace the course. Although a 5s delay was effective, it still remains unknown if this delay can be shortened further in controlled or uncontrolled environments. It is plausible that too short of a duration will cause delays to lose their effectiveness by not encouraging remediation or rehearsal of the material.

The nature of potential competing variables is important for understanding the practical value of postfeedback delays and designing future research. For example, the most commonly reported alternative activity was talking and texting on the phone. It is possible that some participants exposed to feedback throughout the delay simply attended to other stimuli throughout the duration of the delay. While a delay of 5s does not provide a large interval for engaging in many types of alternative activities, simple tasks such as texting can easily be completed in less than 5s (thus enabling the learner to return to their progress as soon as possible) and this activity is readily available at all times for someone who carries a cell phone. In other words, the home environment may allow for fast and relatively simple competing contingencies to occur, unlike many controlled lab settings. Participants engaging with such alternatives may discount feedback entirely during the delay intervals. Studies such as Dubuque (2012) attempted to control for such competing activities by introducing the requirement for participants to continually click a moving button throughout the delay interval, but reported there was no further reduction in racing from such an innovation.

The findings also suggest ways to increase the validity of future lab research. For example, laboratory studies may consider allowing participants the opportunity to use cell phones for texting purposes during sessions. This would introduce a realistic and probable competing contingency, but would do so in a relatively controlled fashion, unlike the home environment that introduces a large amount of variability in the data. Particularly given that the next two most common alternative activities were checking social media and socializing with friends, reducing the impact on the data of the availability and degree of distraction introduced by these social interactions may be quite valuable.

When examining postfeedback delays, it appears that the different types of feedback conditions are irrelevant. Given this, an overt evaluation of response accuracy might not be a key characteristic of the success of a postfeedback delay. One important practical implication of this finding is that it might make the programming of instructional materials easier. If it is true that there are no performance differences between feedback only and self-evaluative feedback, programmers can avoid the extra work involved with requiring overt self-evaluation following the provision of feedback. However, it appears that it would still be important to both provide feedback and enforce a delay when that feedback appears. Covert self-evaluation seems plausible and likely when an opportunity to remediate is enforced, which may account for why no differences were obtained during self-evaluative feedback and feedback only conditions. That is, participants may always self-evaluate, even within feedback only conditions, when presented with an enforced postfeedback delay.

Further, the current findings suggest a way of improving learning outcomes without the need for additional human evaluators. Again, the findings suggest that performance may improve in the absence of any evaluation, human or machine. That is, there were no statements such as "that's correct" or "that's incorrect" in this study. Instead, the program simply stated the correct answer and in the self-evaluate feedback conditions, asked the participant to self-evaluate. If it is the case that simply presenting feedback during a delay increases performance without external evaluation, it may be argued that there is not a need for human observers to evaluate the accuracy of a response. Rather, the instructional material may simply provide the learner with information on how they should have responded. With simple responses, such as a single word or sentence, machines may provide the evaluation by simply displaying the correct response. However, it is likely that human evaluators will be needed for evaluating more complex responses, such as lengthy essays or papers.

Although the current study provides some evidence in support of previous findings in that postfeedback delays increase performance, this does not mean that postfeedback delays will always be equally effective. It is important that researchers continue to pursue such investigations to determine the formats under which postfeedback delays can perform most effectively. In addition to measures of test performance, future researchers should collect data on the differences in total training time across groups. When some form of online training is used, either in a business or academic setting, it is likely going to be important that the program is completed within a specified length of time. That is, a measure of training duration is another key variable for determining the effectiveness and adoptability of CBI.

One possible limitation of this study concerns the low posttest scores (ranging on average between 23.5% and 36.9%). That is, participants were not particularly knowledgeable about the material even after completing the instructional material. However, several considerations should be kept in mind when examining these scores. These scores still represent an improvement over pretest measures (ranging on average between 8.5% and 11.8%) and are in line with previous research (Johnson & Dickinson, 2012). As mentioned earlier, it was important to select instructional material that would avoid a ceiling effect and in that regard lower scores are useful for discovering differences among experimental conditions. It is also reasonable to speculate that participants probably spent little time rehearsing material outside of the program itself since no academic grades or employment conditions were tied to the mastery of the material. Finally, the purpose of the current study was not to validate the effectiveness of the instructional content, but to see if differences in instructional formats could lead to differences in retention. In that regard, the differences in posttest data are informative.

Future researchers should also collect self-report measures of covert responding for those in the feedback only groups. If it is the case that the delay is effective because of covert responding, future research should consider collecting information on the type of responding that is occurring during the delay with the use of participant self-reports. Ideally, this information will be collected in real-time, immediately following the delay. Inquiries such as, "what were you doing throughout the delay?" or "were you asking yourself clarifying questions throughout the delay?" may provide some valuable information on the use of postfeedback delays without an overt self-evaluative component.

In summary, the present study shows the potential to improve learning outcomes with computer-based instruction, even in uncontrolled settings representative of typical learning conditions. Research, including this study, has consistently shown the potential for greater performance improvements when instructional lessons are explicitly designed with user behavior under consideration, such as using feedback and enforced delays to regulate proper pacing under conditions of remediation. Rather than simply looking at what technological innovations are possible, it may be more beneficial to look at how technological innovations can align with known principles of human behavior.

Acknowledgements The authors would like to thank Continuous Learning Group for their financial assistance with this study.

Funding This study was funded by Continuous Learning Group (CLG).

Compliance with Ethical Standards

Conflict of Interest Author A declares that he/she has no conflict of interest. Author B declares that he/she has no conflict of interest. Author C declares that he/she has no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Appendix

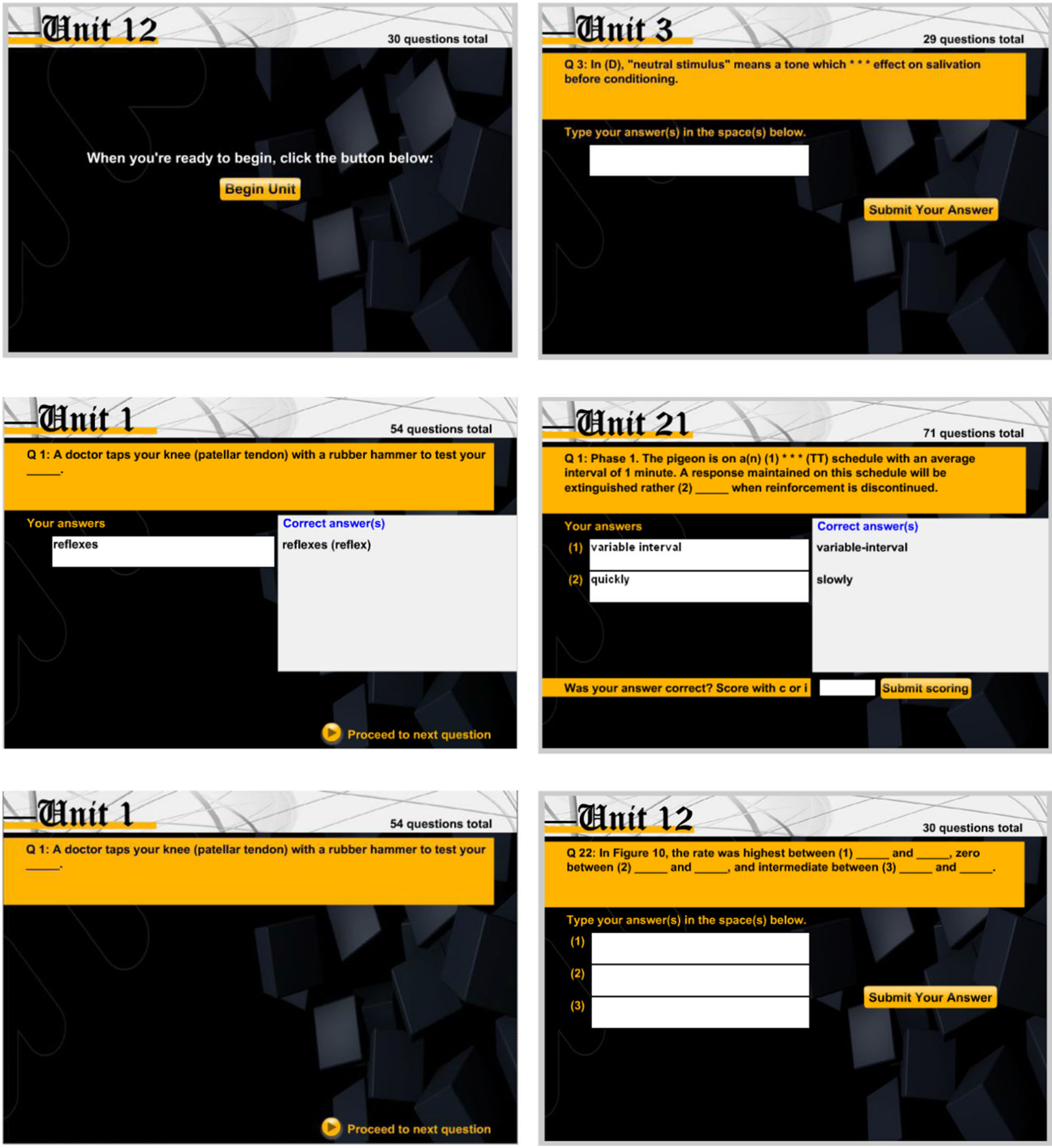


Fig. 3 Sample screenshots of instructional material

References

Allen, I. E., & Seaman, J. (2011). *Going the distance: Online education in the United States, 2011*. Babson Park, MA: Babson Survey Research Group.

Allen, I. E., & Seaman, J. (2016). *Online report card: Tracking online education in the United States*. Babson Park, MA: Babson Survey Research Group.

Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985-1998). *Journal of*

Organizational Behavior Management, 21, 3–29. https://doi.org/10.1300/J075v21n01_02.

Angelino, L. M., Williams, F. K., & Natvig, D. (2007). Strategies to engage online students and reduce attrition rates. *The Journal of Educators Online, 4*(2), 1–14.

Bodner, G. (1997). Confessions of a modern luddite: A critique of computer-based instruction. *UniServe Science News, 6*(3), 10–12.

Chapman, B. (2010). How long does it take to create learning? Chapman Alliance.

Crosbie, J., & Kelly, G. (1994). Effects of imposed postfeedback delays in programmed instruction. *Journal of Applied Behavior Analysis, 27*, 483–491. https://doi.org/10.1901/jaba.1994.27-483.

Dubuque, E. M. (2012). *Reducing racing during online instruction with contingent postfeedback delays* (Unpublished doctoral dissertation). University of Nevada, Reno, Reno, NV.

Holland, J. G., & Skinner, B. F. (1961). *The analysis of behavior*. New York, NY: McGraw-Hill.

Huitema, B. E. (2011). *The analysis of covariance and alternatives: Statistical methods for experiments, quasi-experiments, and single-case studies* (2nd ed.). Hoboken, NJ: Wiley.

Johnson, D. A. (2013). A component analysis of the impact of evaluative and objective feedback on performance. *Journal of Organizational Behavior Management, 33*, 89–103. https://doi.org/10.1080/01608061.2013.785879.

Johnson, D. A. (2014). The need for an integration of technology, behavior-based instructional design, and contingency management: An opportunity for behavior analysis. *Mexican Journal of Behavior Analysis, 40*(2), 58–72.

Johnson, D. A., & Dickinson, A. M. (2012). Using postfeedback delays to improve retention of computer-based instruction. *Psychological Record, 62*, 485–496.

Johnson, D. A., & Rubin, S. (2011). Effectiveness of interactive computer-based instruction: A review of studies published between 1995 and 2007. *Journal of Organizational Behavior Management, 31*, 55–94. https://doi.org/10.1080/01608061.2010.541821.

Kelly, G., & Crosbie, J. (1997). Immediate and delayed effects of imposed postfeedback delays in computerized programmed instruction. *Psychological Record, 47*, 687–698.

Markle, S. M. (1990). *Designs for instructional designers*. Champaign, IL: Stipes.

Marroletti, K., & Johnson, D. A. (2014). Current best practices for creating effective and palatable eLearning. *Mexican Journal of Behavior Analysis, 40*, 73–84.

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Munson, K. J., & Crosbie, J. (1998). Effects of response cost in computerized programmed instruction. *Psychological Record, 48*, 233–250.

Peterson, N. (1982). Feedback is not a new principle of behavior. *Behavior Analyst, 5*, 101–102.

Robinson, L. (2017). Embracing online education: Exploring options for success. *Journal of Marketing for Higher Education, 27*, 99–111. https://doi.org/10.1080/08841241.2016.1261978.

Skinner, B. F. (1958). Teaching machines. *Science, 128*, 969–977. https://doi.org/10.1126/science.128.3330.969.

VanStelle, S. E., Vicars, S. M., Harr, V., Miguel, C. F., Koerber, J. L., Kazbour, R., & Austin, J. (2012). The publication history of the *Journal of Organizational Behavior Management*: An objective review and analysis: 1998–2009. *Journal of Organizational Behavior Management, 32*, 93–123. https://doi.org/10.1080/01608061.2012.675864.

Vargas, E. A., & Vargas, J. S. (1991). Programmed instruction: What it is and how to do it. *Journal of Behavioral Education, 1*, 235–251. https://doi.org/10.1007/BF00957006.

Wei, H. C., Peng, H., & Chou, C. (2015). Can more interactivity improve learning achievement in an online course? Effects of college students' perception and actual use of a course-management system on their learning achievement. *Computers & Education, 83*, 10–21. https://doi.org/10.1016/j.compedu.2014.12.013.

Willging, P. A., & Johnson, S. D. (2004). Factors that influence students' decision to dropout of online courses. *Journal of Asynchronous Learning Network, 8*(4), 105–118.

Woods, R. H. (2002). How much communication is enough in online courses? Exploring the relationship between frequency of instructor-initiated personal email and learners' perceptions of and participation in online learning. *International Journal of Instructional Media, 29*(4), 377–394.

Xu, D., & Jaggars, S. S. (2011). The effectiveness of distance education across Virginia's community colleges: Evidence from introductory college-level math and English courses. *Educational Evaluation & Policy Analysis, 33*(3), 360–377. https://doi.org/10.3102/01623737114138.