The Psychological Record, 2005, 55, 401-418

EFFICACY OF ERROR FOR THE CORRECTION OF INITIALLY INCORRECT ASSUMPTIONS AND OF FEEDBACK FOR THE AFFIRMATION OF CORRECT RESPONDING: LEARNING IN THE CLASSROOM

GARY M. BROSVIC, MICHAEL L. EPSTEIN, and MICHAEL J. COOK
Rider University

ROBERTA E. DIHOFF
Rowan University

Participants completed 5 classroom examinations during which the timing of knowledge of results (no feedback: Scantron form; delayed feedback: end-of-test, 24 hour delay; immediate feedback: educator, response form) and iterative responding (1 response, up to 4 responses) were manipulated. At the end of the semester, each participant completed a 100-item final examination which included 10 items randomly selected from each classroom examination, plus 50 entirely new items. Neither the source of feedback nor the number of responses permitted influenced performance on classroom examinations but both factors interacted significantly to enhance the final examination performance of participants provided with immediate feedback and iterative responding. The correction of initially inaccurate strategies by combining immediate feedback with iterative responding was not differentially effective as a function of information source: educator or the Immediate Feedback Assessment Technique (IF AT) form. For these participants, response identification accuracy, confidence ratings, and retention were higher and inaccurate perseverative responding was lower. Performance on the final examination permits the preliminary quantification of how immediate feedback coupled with iterative responding, when used during classroom examinations that contain items that will be repeated on a cumulative final examination, not only assesses student knowledge but also teaches in a manner that promotes the retention of course materials.

Many of the earliest studies conducted in the psychological sciences were dedicated to examining changes in the performance of learners provided with information (i.e., feedback) that either affirmed a correct response or corrected an error (e.g., Thorndike, 1913, 1927). This corrective information, viewed initially within an associationistic framework as a contingent event, strengthened correct responses

Address correspondence to Gary M. Brosvic, Department of Psychology, Rider University, Lawrenceville, NJ 08343. (E-mail: Brosvic@Rider.edu).

through positive reinforcement and weakened incorrect responses through nonreinforcement. This mechanistic perspective emphasized minimizing errors, but provided neither the means for, nor an insight into, their correction. Indeed, the efficacy of errors and their inclusion within the learning processes did not enter mainstream psychology until the ascendance of information processing within which errors were seen as valuable resources that learners could use to evaluate their understanding of test materials and to correct their initially inaccurate assumptions (e.g., Kluger & DeNisi, 1996; Kulhavy & Stock, 1989).

One of the most divisive points in the literature on feedback concerns the optimal timing of feedback following the learner's response: immediately, or after the elapse of a prescribed interval. The definitions of immediate feedback range from seconds after a response (Epstein, Epstein, & Brosvic, 2001) to the next weekly meeting of a class (Robin, 1978). Definitions of delayed feedback range from the end of a test (Dihoff, Brosvic, & Epstein, 2003, 2004) to delays of 7 days (Bruning, Schraw, & Ronning, 1999; Robin, 1978). Even though the considerable overlap of these definitions precludes the direct numerical comparison of experimental outcomes, the rationales advanced for preferring either immediate or delayed feedback are more exclusive than the intervals that define the timing of feedback. Proponents of immediate feedback recommend the correction of errant responses and the acquisition of the correct response before exiting a test problem or test session (Epstein et al, 2001). In comparison, proponents of delayed feedback recommend the imposition of a delay of 24 to 48 hours to facilitate the forgetting of errant responses and the acquisition of correct responses in the absence of the interference that immediate feedback on an item-by-item basis generates (Kulhavy & Stock, 1989). Reports that immediate feedback is more effective than delayed feedback during applied, but not during laboratory studies (Gick & Holyoak, 1987) perhaps accounts for why the literature on the effects of the timing of feedback, shows approximately an even number of reports for the superiority of immediate over delayed feedback, and for delayed over immediate feedback.

The classroom is the one testing environment that has been included in both applied and laboratory studies, and in which immediate and delayed feedback have been directly compared. Typically, the provision of immediate feedback on examinations completed in the classroom has required the use of either computer-assisted instruction (CAI) or the assignment of one evaluator per learner. In the absence of computers, course-specific software, or programming skills, the delivery of feedback is delayed—if it is provided at all. In the laboratory, the use of CAI enables the presentation of supporting materials, branching for additional instruction and assessment, and the measurement of response times to index the processing of test materials. Until recently, the simple and practical means by which immediate feedback might be provided in the classroom in the absence of computers has not been available. The tool that has been refined and validated in our classrooms and laboratories is

the Immediate Feedback Assessment Technique, or IF AT[1] (Epstein et al., 2001, 2002).

The IF AT manifests the theoretical and practical foundations of the teaching-testing machines described by Pressey (1950) and Skinner (1968), transforming the passive receiver of information into the active demonstrator of skills and knowledge. The IF AT form (see Figure 1) is a multiple-choice answer sheet with rows of rectangular answer spaces (e.g., A, B, C, D) that is nearly identical in layout to the ubiquitous machine-scored answer sheet

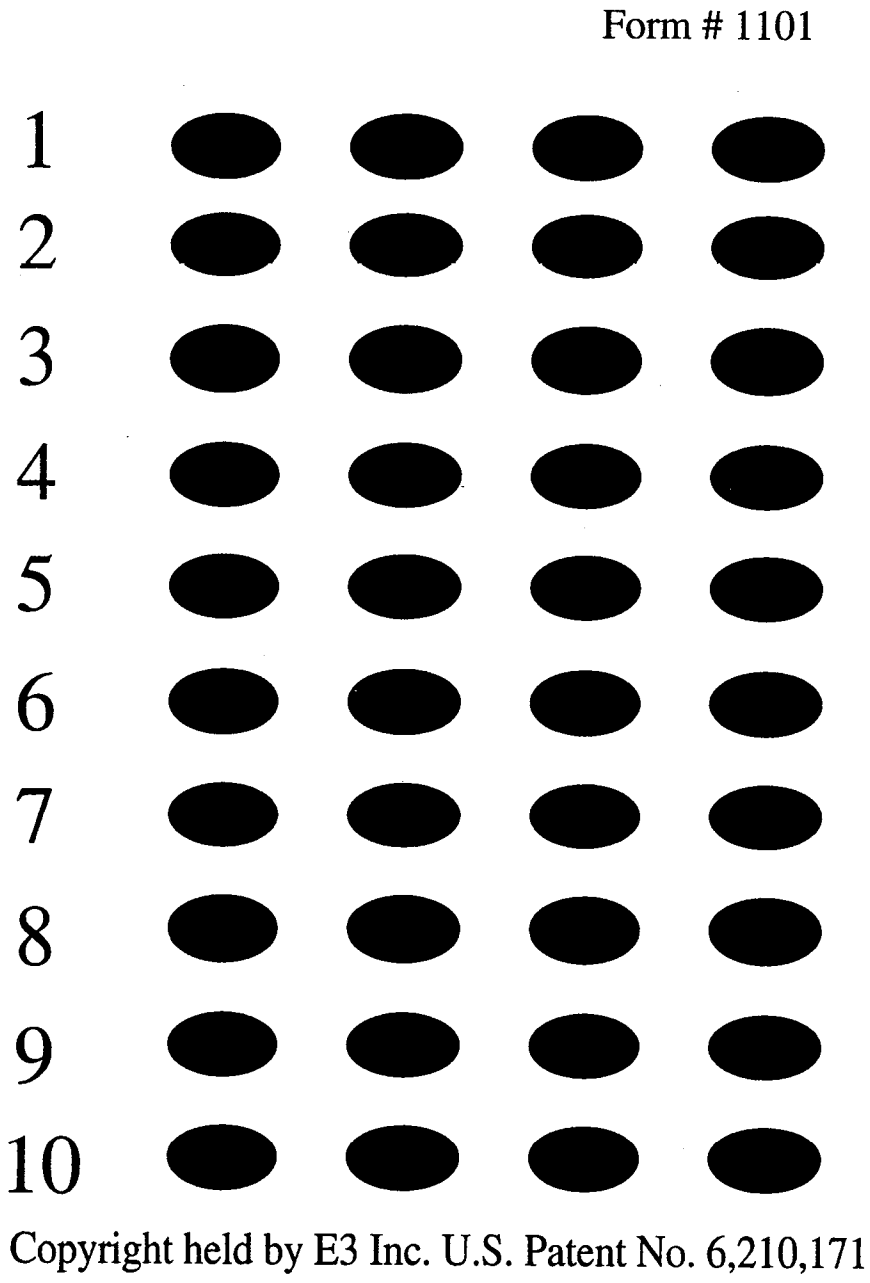

Figure 1. Sample portion of the Immediate Feedback Assessment Technique (IF AT) form. Patent is held by E3 Corporation.

[1]IF AT is a registered trademark of E3 Corporation.

available from Scantron Corporation. Participants scrape off an opaque, waxy coating covering an answer space on the IF AT form to record their answer. If a symbol (e.g., a star) is printed beneath the covering, the student receives instant feedback that a correct choice was made; the absence of a symbol provides instant feedback that an incorrect choice was made. However, rather than simply exiting the question, the student reviews the remaining response options, continues to respond until the correct answer is discovered (a self-correction procedure), and thus exits each question knowing the correct answer.

Indeed, without corrective feedback, the learner likely exits an examination assuming that incorrect responses were actually correct; thus, an examination that does not employ feedback may promote misconceptions (e.g., Brown, Schilling, & Hockensmith, 1999). In comparison, the IF AT is an effective tool for promoting the correction of initially inaccurate response strategies and the retention of learning, and its use has been shown to facilitate the learning and retention of participants with developmental delays (Epstein, Brosvic, Dihoff, Lazarus, & Costner, 2003). In recent studies we found that the IF AT form could be easily used by preschool children with developmental delays studying academic readiness materials (provided sufficient motor skill was present), by junior high school students classified with mild mental retardation studying a life skills curriculum (Epstein et al., 2003), and by elementary school students classified with mild mental retardation acquiring mathematical fact series (Dihoff Brosvic, Epstein, & Cook, 2005), with more favorable outcomes observed when feedback was provided by an educator than by the IF AT form. The concurrent presentation of both sources of immediate feedback (educator and the IF AT) was more effective than the presentation of either source separately, and thus the inclusion of the IF AT facilitated the teaching-learning process to promote higher levels of independent learning and retention. These outcomes raise several interesting practical and theoretical issues for learning and retention that require further study and form the bases of the present study: the timing of feedback, the source of feedback, and the opportunity to engage in iterative responding (i.e., to answer until correct).

Method

Participants

Thirty-one male and 79 female students enrolled in a liberal arts and sciences undergraduate course served as voluntary participants. The modal participant was a Caucasian female, approximately 20 years of age and majoring in the liberal arts and sciences.

Materials

Five classroom examinations were prepared from a publisher-supplied test bank, each with 50 questions, and each question with four response options (i.e., A, B, C, D). The final examination consisted of 100

items which included 10 randomly selected items from each classroom examination, plus 50 entirely new items.

Design and Procedures

At least 11 participants (3 male and 8 female) were randomly assigned to each experimental group generated from the factorial combination of five feedback groups (Scantron, Delayed, End-of-Test, Immediate/Educator, and Immediate/Form) and two levels of iterative responding (a maximum of one response permitted, a maximum of four responses permitted). Prior to recording an answer for each test item, participants reviewed the four answers that were provided, and from those answers, recorded the number of answers that they determined could be correct; participants then selected the "best" answer and rated confidence in their selection on a scale ranging from 1 (no confidence) to 100 (complete confidence). This process was repeated, in the event of either an initial or subsequent incorrect response when iterative responding was permitted, for each test item.

In the no-feedback control group, answers were recorded with a pencil on a Scantron form. In the end-of-test-feedback group, answers were recorded with a pencil on a Scantron form, and upon the completion of the test, all writing implements were removed and participants were permitted to review the examination, the correct solutions, and their answer sheets for 30 minutes. Participants in the other groups were requested to remain seated and to read noncourse materials until the end of the test period under the supervision of test monitors. As with the aforementioned condition during the review process, all participants were required to remain silent and were permitted neither to share their materials nor to ask any questions of the test monitors. In the delayed-feedback group, participants recorded their answers in pencil on a Scantron form and, on the following day, these participants reviewed the examination, the correct solutions, and their corrected answer sheets for 30 minutes. Participants in the other groups were requested to remain seated and to read noncourse materials until the end of the test period. During this review process all participants were required to remain silent and were permitted neither to share their materials nor to ask any questions of the test monitors. Participants in the immediate/educator group were seated in pairs, and an experimental assistant sat between the members of each pair. After the participants recorded their answers on Scantron forms, the experimental assistant indicated by holding up 3-x 5-inch index cards if the response was correct; if an incorrect response was made and iterative responding was available, additional cards identifying already selected responses (e.g., A, B) were made visible to each participant who continued to select responses until the correct answer was discovered. In the immediate/form group, answers were recorded using the IF AT form (E3 Corporation) which enabled participants to receive immediate affirming or corrective feedback; if appropriate to the experimental group the participant was permitted to continue selecting responses until the correct answer was discovered.

After completing the fifth classroom examination, each participant then completed a 15-item questionnaire assessing ease of understanding and ease of completing response requirements, perceived fairness of and preference for an answer-until-correct procedure, and involvement in the test-taking process, as described previously by Epstein and Brosvic (2002). The final examination was administered 1 week after completion of the fifth classroom examination, and at the time, all participants used Scantron forms to record their answers. Once the final examination was completed, participants reviewed each examination item and identified those items they believed were repeated from one of the classroom examinations, and their initial responses to those items, both correct and incorrect, and then rated confidence in the accuracy of their identifications on a scale ranging from 1 (no confidence) to 100 (complete confidence). Performance on the items carried over from classroom examinations to the final examination served as the primary measure of retention. Although the IF AT method enables the assignment of partial credit (i.e., correct responding on the first attempt is assigned 100% of item credit whereas correct responding on the second, third, or fourth attempt may be assigned reduced percentages according to instructor discretion), this procedure was not used, and the results described below were based upon the accuracy of initial responses.

Results

There were no differences in SAT scores, current semester classroom performance, overall GPA, or any other dependent measure described below as a function of sex of participant, feedback group, the opportunity to engage in iterative responding (IR), or their interaction, all $F < 1$, all $p > .5$. There were no differences in any dependent measure between the delayed-feedback and the end-of-test-feedback groups, all $F < 1$, all $p > .5$; responses were aggregated and hereafter referred to as delayed feedback. A similar lack of differences between the immediate/educator and the immediate/form groups was observed, all $F < 1$, all $p > .5$; responses were aggregated and hereafter referred to as immediate feedback.

Scores on classroom examinations. Potential differences in mean scores (see Figure 2) were examined using a 3 (feedback group) x 2 (IR) x 5 (classroom examination) ANOVA, with significance observed for neither the main effects nor their interactions, all $F < 1$, all $p > .5$.

Final examination scores. Potential differences in mean scores (see Figure 3) were examined using a 3 (feedback group) x 2 (IR) ANOVA, with significance observed for each main and their interaction, all $F > 17.92$, all $p < .001$. Scheffe comparisons indicated that scores were significantly (a) higher for the immediate-feedback than for the delayed-feedback and control groups, (b) higher for the delayed-feedback than for the control group, and (c) highest when immediate feedback was combined with IR, all $p < .005$.

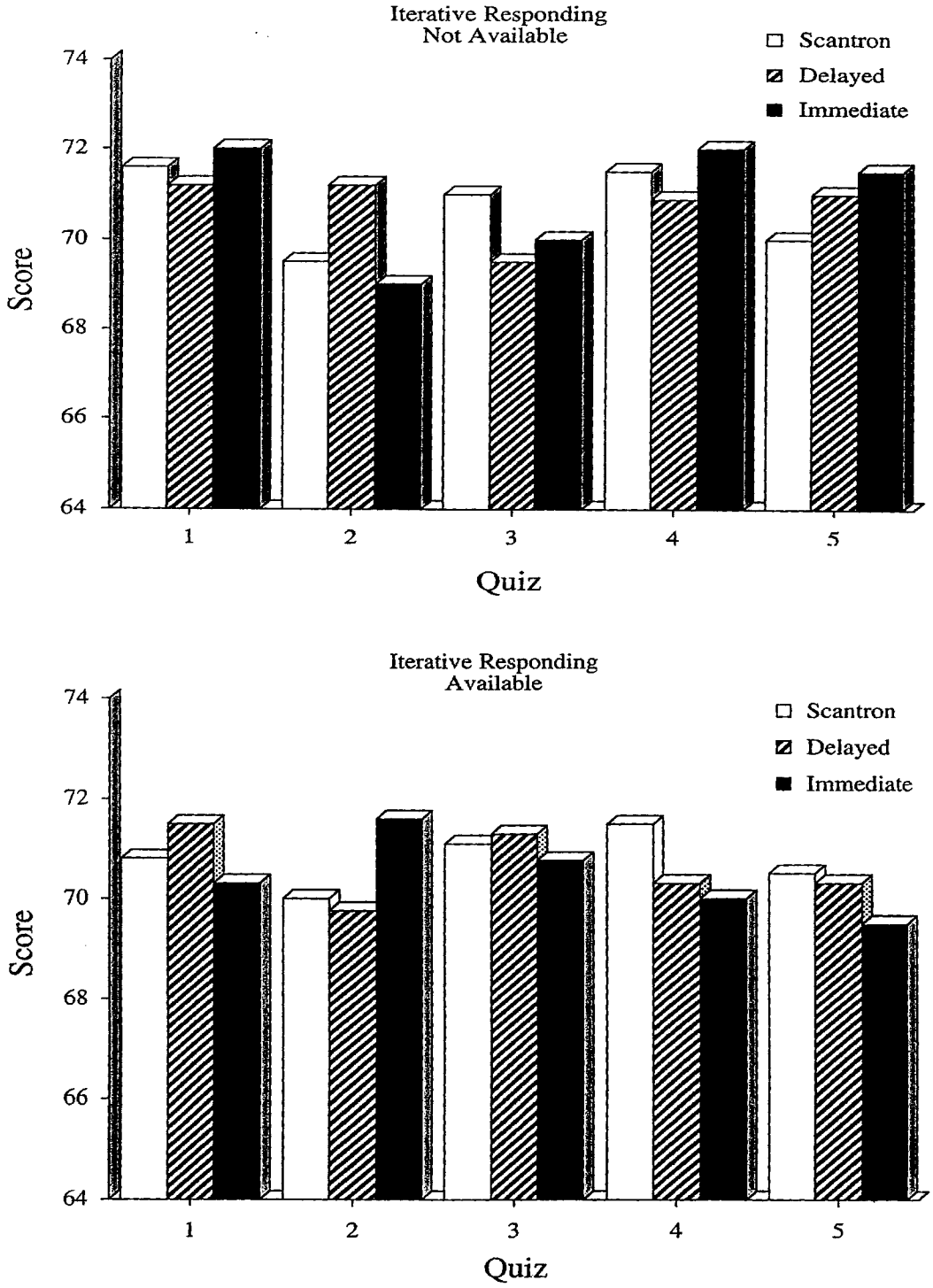

Figure 2. Scores on the five classroom examinations for the control and feedback groups when IR was (lower panel) or was not (upper panel) available.

Role of prior experience with test items. The percentage of correct responding on the 50 test items previously included on classroom examinations is presented in Figure 4. Potential differences in percentages of correct responding were examined using a 3 (feedback group) x 2 (IR) ANOVA, with significance observed for each main and their interaction, all *F*

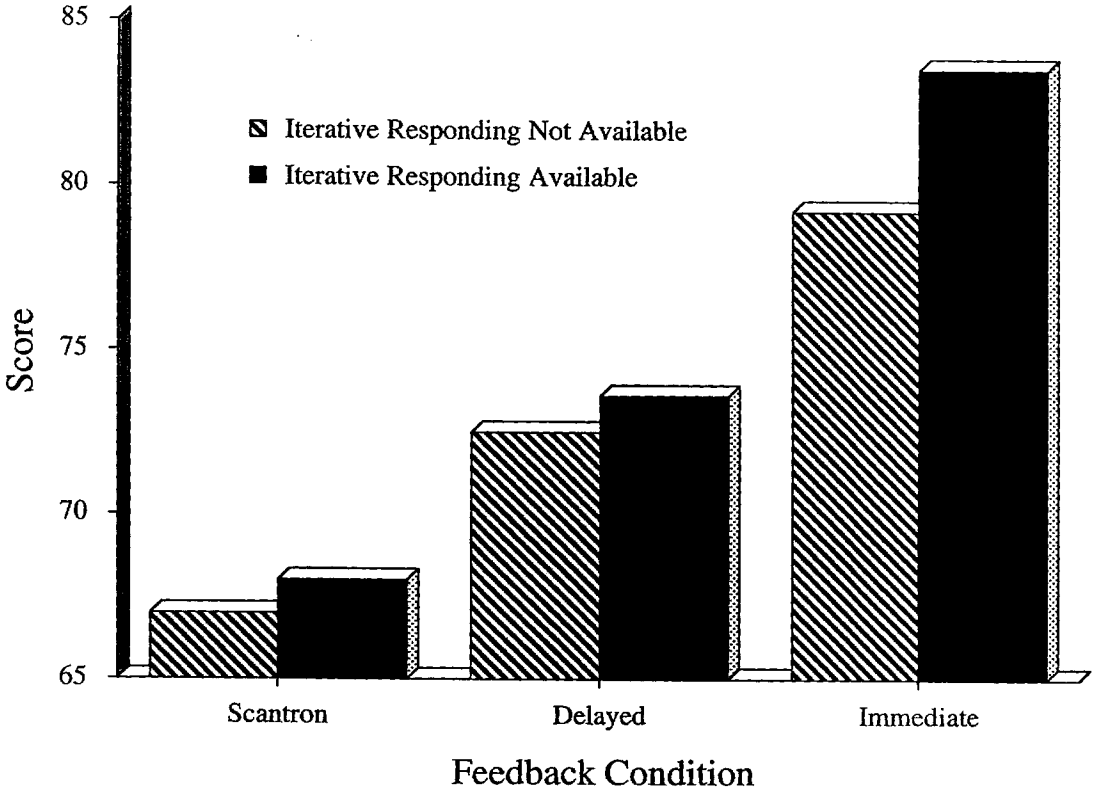

Figure 3. Final examination scores for the control and feedback groups when IR was available.

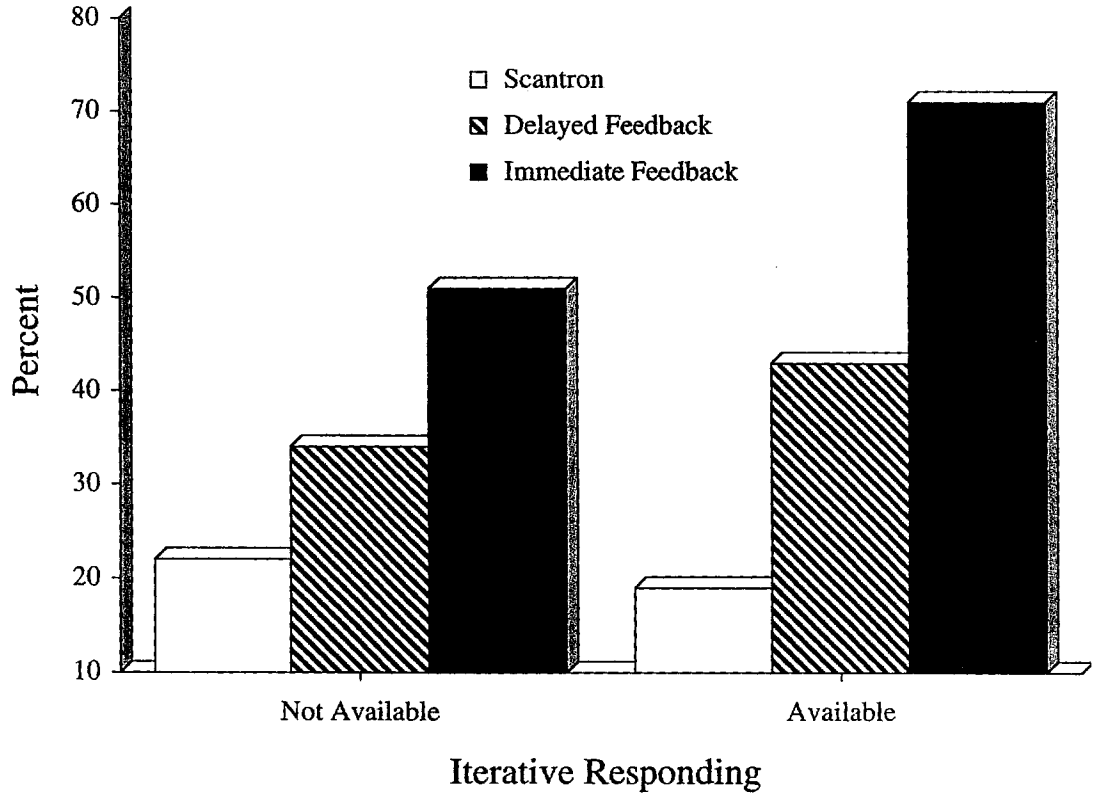

Figure 4. Percentages of correct responding on final examination items that were previously included on one of the classroom examinations.