The Psychological Record, 1998, 48, 601-616

AUGMENTING COMPUTER-INTERACTIVE SELF-ASSESSMENT WITH AND WITHOUT FEEDBACK

H. A. CHRIS NINNESS
Stephen F. Austin State University

SHARON K. NINNESS and SANDRA SHERMAN
Nacogdoches Independent School District

CHUCK SCHOTTA
Northwest Independent School District

Following a preexperimental assessment of computer-interactive math performance during VR 6 reinforcement and extinction, 4 regular education students and 2 students identified as behaviorally disordered participated in an A-BC-D-BC withdrawal of treatment design. Subsequent to baseline observations of math performance during self-assessment with and without accuracy feedback, students were trained in self-assessment procedures by way of a series of computer-interactive tutorials. During treatment, students were provided computer-displayed accuracy feedback plus reinforcement for correct self-assessments of their math performance. Reinforcement and feedback were gradually leaned, and in the final treatment condition, accuracy feedback was terminated; however, monetary reinforcement for correct self-assessment was sustained. Following treatment, students were given opportunities to perform math problems in the absence of reinforcement while self-assessing their performances *with and without accuracy feedback*. This was succeeded by a withdrawal condition and a final session in which students, again, were given an opportunity to self-assess with and without feedback from the computer. Outcomes suggest that subsequent to training computer-interactive self-assessment *with feedback* may facilitate high rates and long durations of math performance even in the absence of compensation. Implications regarding the *augmental* as a type of rule-governed behavior and the necessary and sufficient conditions for sustaining self-assessment as a learned reinforcer are discussed.

There is a growing body of evidence indicating that accuracy *feedback* enhances the acquisition of some academic and social skills (e.g., Gettinger, 1985), and it has been suggested that teaching students to self-assess their own behavior may be one way to facilitate this

This investigation was supported in part by a Stephen F. Austin State University Faculty Research Grant 1-14065. The authors thank Alan West for his technical support in the development of this project. Correspondence concerning this article should be directed to H. A. Chris Ninness, Director of Graduate Training in School and Behavioral Psychology, PO Box 13019 SFA Station, Stephen F. Austin State University, Nacogdoches, TX 75962 (E-mail: cninness@titan.sfasu.edu).

feedback process (Fatuzzo & Rohrbeck, 1992). Self-assessment has been characterized as one component of self-management in which students evaluate their own ongoing performance and furnish a form of self-feedback in accordance with the standards set forth by some external mechanism or socializing agent (Glynn & Thomas, 1974). For at least the last 25 years, self-assessment has been a popular intervention strategy and a source of highly diversified applied research in school settings (Ninness & Glenn, 1988).

To cite only a few of the many examples, self-assessment strategies have demonstrated diversified positive outcomes such as enhancing attentive behaviors and reducing disruptive behaviors of children identified as hyperactive (Christie, Hiss, & Lozanoff, 1984); establishing generalization of students' on-task behavior across academic environments (Rhode, Morgan, & Young, 1983); enhancing students' skills at acquiring teacher praise for appropriate behavior (Connell, Carta, & Baer, 1993); increasing the rate and accuracy of assignments (McLaughlin, Burgess, & Sackville-West, 1982); and improving the test performance of elementary students (Wall & Bryant, 1979). Self-assessment of tape recordings of new words has increased student opportunities to produce accurate pronunciation (Lalli & Shapiro, 1990). Recent outcomes have also confirmed that junior high school students with behavior disorders can effectively self-assess social skills and on-task behavior (Houghton, 1991; Kern-Dunlap, Dunlap, Clarke, Childs, White, & Stewart, 1992; Ninness, Fuerst, & Rutherford, 1995; Ninness, Fuerst, Rutherford, & Glenn, 1991; Smith, Nelson, Young, & West, 1992) and aggression control strategies (Ninness, Ellis, Miller, Baker, & Rutherford, 1995) in the absence of supervision. Self-assessment interventions have improved preschool children's participation in group instruction (Miller, Strain, Boyd, & Jarzynka, 1993) and facilitated academic efficiency of students with learning disabilities in regular education (Maag, Reid, & DiGangi, 1993). Moreover, because self-assessment is relatively unobtrusive, students endorse this strategy over peer and other forms of classroom evaluations (Turco & Elliot, 1986).

Despite this diversified array of inspiring self-assessment investigations, the internal activities that support self-assessment have proven difficult to characterize and quantify. Although there is evidence that external instructions interact with self-instructions and self-evaluations to influence the rate and duration of task-related activities (Flora, Pavlik, & Pittenger, 1990; Ninness & Ninness, 1998; Rosenfarb, Newland, Brannon, & Howey, 1992), attempts to identify the necessary and sufficient conditions for maintaining effective self-assessment have engendered some argument (Ninness, Glenn, & Ellis, 1993).

Most particularly, the role of feedback during self-assessment remains ambiguous. For example, Broden, Hall, and Mitts (1971) demonstrated that simply having a student spontaneously self-record her own on-task behavior at various unscheduled times throughout the school day appeared to be all that was necessary to substantially improve her academic performance. Independent of accuracy feedback, Glynn, Thomas, and Shee (1973) had eight second-grade students simply place a check on a piece of prepared data sheet paper at irregular intervals if they were on-task at the moment a beeper sounded. On-task behaviors increased dramatically, and independent observers indicated that 76% of the students' self-assessed on-task recordings agreed with those of trained observers. Glynn and Thomas (1974) replicated this strategy using an intermittent audio signal to cue self-assessment intervals. Likewise, Maag et al. (1993) found that improved academic accuracy or academic productivity was associated with self-monitoring even when no feedback for correct self-recordings was provided. These findings seem to support the assumption that external accuracy feedback systems (social or mechanical) may be irrelevant to the development of effective and durable self-assessment protocols.

In contrast, it has been argued that some kind of external feedback regarding the accuracy of self-assessment may be an integral component to sustaining the reinforcing value of this procedure (Van Houten, 1984). For example, Connell et al. (1993) found that self-assessment and accuracy feedback improved the academic engagement of some students. Ninness, Ellis, and Ninness (in press) contrasted computer-interactive academic performance with and without self-assessment and found improved rates and durations were associated with self-assessment with feedback; however, no attempt to isolate the effects of feedback was made. At this time, the role of feedback in supporting effective self-assessment remains unresolved (Malott, 1986; Newman, Buffington, Hemmes, & Rosen, 1996; Skinner & Smith, 1992), and the question remains as to whether *accuracy feedback* serves a critical function in sustaining self-assessment as a motivational variable. Thus, the experimental preparations in the following study were designed to investigate the effects of computer-interactive self-assessments *with* and *without accuracy feedback* among regular education and behaviorally disordered students.

Method

Participants, Settings, and Apparatus

Four students from a regular education fifth-grade class and two sixth-grade special education students from a self-contained social adjustment class were invited to participate in the experiment. The 2 special education students had been identified as behaviorally disordered. Selection of participants was random but limited to students whose classrooms had immediate access to the necessary computer facilities throughout the entire school day. The 4 regular education students were of average or above average intelligence, and their math skills were at or above grade level. The special education students were of average intelligence; however, they performed more than one year below grade level in basic math skills. Moreover, these students had a well-documented history of drifting off-task during academic instruction.

The four regular education students performed all experimental sessions during the regular school day in their homeroom. The 2 students identified as behaviorally disordered performed within their self-contained social adjustment class. All aspects of the experiment were performed on IBM compatible computers connected to laser printers. The software was written by Chris Ninness in QBASIC for IBM PC compatible machines.

Data Collection and Reliability Checks

Following computer interactive self-assessment strategies developed by Ninness et al. (in press), students were given an opportunity to perform a series of multiplication problems. Subsequent to the opening message "DO YOU WANT TO PLAY?" displayed at the top center of the screen, students could respond by typing "Y" or "N." If "N" were typed the program terminated; if "Y" were typed, a multiplication problem appeared, and it could be solved by typing the number key/s followed by the enter key. Immediately, a new problem was presented to the student according to the same format. In this context, a unit of measured behavior always entailed a combination of three or four key presses. Software automatically calculated performance measurements as correct answers/min. However, if a student exited the program early by typing "N," the interval that elapsed after exiting was scored as zero. Prior to and subsequent to the experiment, the computer output was cross-validated by having an observer directly count correct problems per minute as one of the researchers interacted with the program.

Preexperimental Assessment

Preexperimental assessments were conducted in order to ascertain the students' optimal rates of correct problems/min as measured by variable ratio 6 (VR 6) schedules of reinforcement. The second part of this assessment gauged the students' rates and durations of math performance in the complete absence of reinforcement (extinction).

VR 6. At the beginning of this session, the computer screen displayed the following accurate rule for maximizing reinforcement during the VR 6 schedule: "THE FASTER YOU WORK, THE MORE POINTS YOU EARN. POINTS MAY BE EXCHANGED FOR MONEY AFTER THE SESSION. TYPE ENTER IF YOU UNDERSTAND." In this context, the above rule *accurately* described the relationship between solving problems quickly and gaining access to more points (money) at the end of the session. Reinforcement consisted of a brief (.5-s), flashing message on the computer screen indicating the earning of five cents. In the event that no key strokes occurred over a period of 7min, the computer program automatically terminated.

Extinction. After 25 min, the VR 6 reinforcement schedule ended; however, the session was not interrupted, and multiplication problems continued to become available on the computer screen. Starting on the 26th min, solving problems simply resulted in more math problems being displayed on the computer screen; however, no form of reinforcement was provided for solving problems, and students were not notified that this extinction procedure had gone into effect.

Experimental Design

After the preexperimental assessment, students participated in an A-BC-D-BC withdrawal of treatment design (Hersen & Barlow, 1976). During baseline observations, in which no form of extrinsic reinforcement was accessible during problem-performing sessions at the computer, students were given an opportunity to self-assess *with* accuracy feedback. Subsequently, students were trained in self-assessment procedures by way of a series of computer-interactive tutorial programs. Following training, students were given an opportunity to perform math problems while self-assessing their performances *with and without feedback in the absence of extrinsic reinforcement*. During a withdrawal of treatment condition, students were provided another occasion to perform problems at the computer, but <u>no</u> opportunities to self-assess or obtain external reinforcement were provided. This was succeeded by a final experimental condition in which students, once again, were provided opportunities to self-assess with and without accuracy feedback.

Fidelity of self-assessments. Concurrent with calculating the number of correct answers/min, the computer program kept a running record of the percentage of accurate self-assessments performed throughout each session. Correct self-assessments and correct answers/min were automatically saved on disk, and student records were printed following each experimental session.

Baseline. During baseline observations, students were given an opportunity to self-assess at 2-min intervals. In this condition, each self-assessment was accompanied by accuracy feedback from the computer. At the beginning of the session, the computer screen posted the following message: "YOU MAY SCORE YOURSELF (4 to 1) WHEN THE COMPUTER ASKS. 4 = EXCELLENT SPEED & ACCURACY, 3 = GOOD SPEED & ACCURACY, 2 = FAIR SPEED & ACCURACY, 1 = POOR SPEED & ACCURACY. THE COMPUTER <u>WILL TELL</u> YOU IF YOU HAVE SCORED YOURSELF CORRECTLY, BUT YOU WILL NOT EARN ANY MONEY [OR OTHER REWARDS] FOR SCORING YOURSELF CORRECTLY. PRESS ENTER IF YOU UNDERSTAND."

Self-assessment tutorials: Three phases. Extending the format employed during baseline, the tutorial programs afforded students accuracy feedback after each self-assessment. Self-assessment training was introduced to the students by incorporating computer-displayed reinforcement in the form of points for accurate self-assessments while students were performing multiplication problems.

Preexperimental assessments of student proficiency under VR 6 contingencies provided information regarding the students' pretraining rates of correct problems/min. This data furnished a basis for indexing each student's highest rate of correct problems/min during the tutorial sessions. For example, if a student demonstrated an asymptote of approximately 17 problems/min during the preexperimental assessment, self-assessment tutorials were predicated on that standard during the initial stages of training; however, as students' math rates increased during training, their criteria for correct problems/min was gradually advanced.

During self-assessment tutorials, students were instructed to score themselves in accordance with a Lykert scale such that a score of 4 indicated excellent speed and accuracy, 3 indicated good speed and accuracy, 2 indicated a fair speed and accuracy, and 1 indicated poor speed and accuracy (Young, West, Smith, & Morgan, 1991). Software specifications indexed rates at or above 90% of the students' preexperimental asymptote as a score of 4, 80% to 89% as 3, 70% to 79% as 2, and 60% to 69% as 1. Rates below 60% of the students' preexperimental asymptote were below all criteria for correct matches with the computer program. Although students were told *how fast* they had performed math problems during the preexperimental session, they were not specifically informed of the particular problems/min requirement that corresponded with each point on the Lykert scale. At the beginning of each tutorial session, the computer screen simply displayed the following instructions: "YOU MAY SCORE YOURSELF (4 to 1) WHEN THE COMPUTER ASKS. 4 = EXCELLENT SPEED & ACCURACY, 3 = GOOD SPEED & ACCURACY, 2 = FAIR SPEED & ACCURACY, 1 = POOR SPEED & ACCURACY. IF YOU SCORE YOURSELF CORRECTLY, YOU MAY EXCHANGE YOUR POINTS FOR AN EQUAL NUMBER OF PENNIES [OR OTHER REWARDS] AFTER THE SESSION. PRESS ENTER IF YOU UNDERSTAND."

Student self-assessments that matched the computer's calculations resulted in the screen strobing the words, "RIGHT" for .5-s; incorrect self-assessments resulted in the screen strobing the words, "THE CORRECT ANSWER IS [correct answer]." for 1-s. Following the session, the total number of correct self-assessments points were exchanged for an equal number of pennies and the computer printed the student's average correct answers/min. This outcome was used in setting the students' individual response criteria for the following tutorial session. As students performed problems more quickly, their self-assessment standards for scoring speed and accuracy were gradually raised. When students achieved 90% correct self-assessment and 90% of their required correct problems/min for at least 3 consecutive training sessions, they moved to the next phase of the tutorial program. Tutorials were conducted during 20-min daily (or semidaily) sessions over a period of approximately 3 to 4 weeks.

Students in the self-contained classrooms performed under slightly different contingencies. Rather than pennies, these students exchanged their self-assessment points for access to tangible (e.g., cokes or comic books) or activity-based (e.g., free-time or games) reinforcers. Because these 2 students functioned at a lower level of math proficiency, they required additional training sessions before advancing to the next tutorial phase. Tutorial sessions were conducted (3 to 5 days per week) over a period of approximately 2 months. Although these students advanced through the self-assessment tutorials more slowly than the regular education students, their individualized criteria for progressing through the tutorials were the same.

Phase 1: Reinforcement for matching computer assessments. Initially, the program enabled students to evaluate their own performances at the end of each 2-min interval. Correctly matching the computer assessment allowed students to earn points throughout the session. Points were exchanged for an equivalent number of pennies (or tangible/activity-based reinforcers). Following a minimum of 3 sessions of performing at or above 90% of their preexperimental asymptote and achieving at least 90% correct self-assessment, students progressed to the next training phase.

Phase 2: Leaning self-assessments. In Phase 2, the frequency of opportunities to self-assess was gradually leaned (Mace, Brown, & West, 1987). Leaning was based on students demonstrating performance levels at or above 90% of their previous session's rate of correct problems/min and achieving 90% correct self-assessments. Leaning was accomplished by expanding self-assessment intervals from 2 min, to 4 min, and, finally, 8 min in duration (Fisher, Ninness, Piazza, & Owen-DeSchryver, 1996).

Phase 3: Terminating feedback. In the final three 20-min training sessions, students were afforded opportunities to self-assess their math performances at 2-min, 4-min, and 8-min intervals, respectively. As during the previous phases, advancing to a wider self-assessment interval required that the student perform at or above 90% of their previous session's rate and that they continue to achieve at least 90% correct self-assessments. During these tutorial sessions, the computer did not furnish accuracy feedback of self-assessments; however, correct self-assessment points were tallied by the computer and were exchanged for an equivalent number of pennies (or tangible/activity-based reinforcers). Thus, in this condition, students had no way of knowing how many points they had earned until the session ended. At the beginning of each session the screen displayed the following message: "YOU MAY SCORE YOURSELF (4 to 1) WHEN THE COMPUTER ASKS. IF YOU SCORE YOURSELF CORRECTLY, YOU MAY EXCHANGE YOUR POINTS AFTER THE SESSION. THIS TIME THE COMPUTER WILL NOT TELL YOU IF YOU SCORE YOURSELF CORRECTLY. PRESS ENTER IF YOU UNDERSTAND."

Experimental Condition 1: Posttutorial self-assessment. In the first part of this condition, students were afforded opportunities to self-assess at 2-min intervals; however, no accuracy feedback from the computer was furnished. At the beginning of the session, the computer screen presented the following message: "YOU MAY SCORE YOURSELF (4 to 1) WHEN THE COMPUTER ASKS. THE COMPUTER WILL NOT TELL YOU IF YOU SCORED YOURSELF CORRECTLY. YOU WILL NOT

EARN ANY MONEY [OR OTHER REWARDS] FOR SCORING YOURSELF CORRECTLY. PRESS ENTER IF YOU UNDERSTAND."

In the second part of this condition, students were advised that self-assessments would not result in any form of financial exchange or other compensation; however, students were informed that the computer program would display feedback for correct self-assessments throughout the session. At the beginning of the session, the computer screen posted the following message: "YOU MAY SCORE YOURSELF (4 to 1) WHEN THE COMPUTER ASKS. THIS TIME IF YOU SCORE YOURSELF CORRECTLY, THE COMPUTER WILL TELL YOU. HOWEVER, NO MONEY [OR OTHER REWARDS] CAN BE EARNED DURING THIS SESSION. PRESS ENTER IF YOU UNDERSTAND." As during the previous session, the program displayed a continuing series of multiplication problems. Although this experimental phase did not yield compensation, opportunities to *self-assess with accuracy feedback* were provided at 2-min intervals.

Withdrawal: Removal of all self-assessment contingencies. At the outset of this phase, the students were advised that they could not score themselves and that they could not earn any money or other rewards for correctly performing math problems. The monitor displayed the following message: "YOU MAY NOT SCORE YOURSELF OR EARN ANY MONEY [OR OTHER REWARDS] DURING THIS SESSION. PRESS ENTER IF YOU UNDERSTAND." As during the prior phase, the computer generated a continuing series of multiplication problems.

Experimental Condition 2: Posttutorial self-assessment. In the first part of this condition, students were afforded opportunities to self-assess at 2-min intervals; however, no accuracy feedback from the computer was furnished. At the beginning of the session, the computer screen presented the following message: "YOU MAY SCORE YOURSELF (4 to 1) WHEN THE COMPUTER ASKS. HOWEVER, THE COMPUTER WILL NOT TELL YOU IF YOU SCORED YOURSELF CORRECTLY. AND YOU WILL NOT EARN ANY MONEY [OR OTHER REWARDS] FOR SCORING YOURSELF CORRECTLY. PRESS ENTER IF YOU UNDERSTAND."

On the following day, students repeated the above condition. Students were advised that correct self-assessments would not result in any form of monetary compensation, but that they *would* obtain instantaneous feedback regarding the accuracy of each self-assessment as they performed math problems. At the beginning of the session, the computer screen posted the following: "YOU MAY SCORE YOURSELF (4 to 1) WHEN THE COMPUTER ASKS. THIS TIME, IF YOU SCORE YOURSELF CORRECTLY, THE COMPUTER WILL TELL YOU. HOWEVER, NO MONEY [OR OTHER REWARDS] CAN BE EARNED DURING THIS SESSION. PRESS ENTER IF YOU UNDERSTAND."

Results

Preexperimental assessment outcomes. Figure 1 suggests that following the first few minutes of responding during the preexperimental assessment, all 6 students exhibited a gradual elevation in the number of correct answers/min. This was followed by a moderate rate of relatively

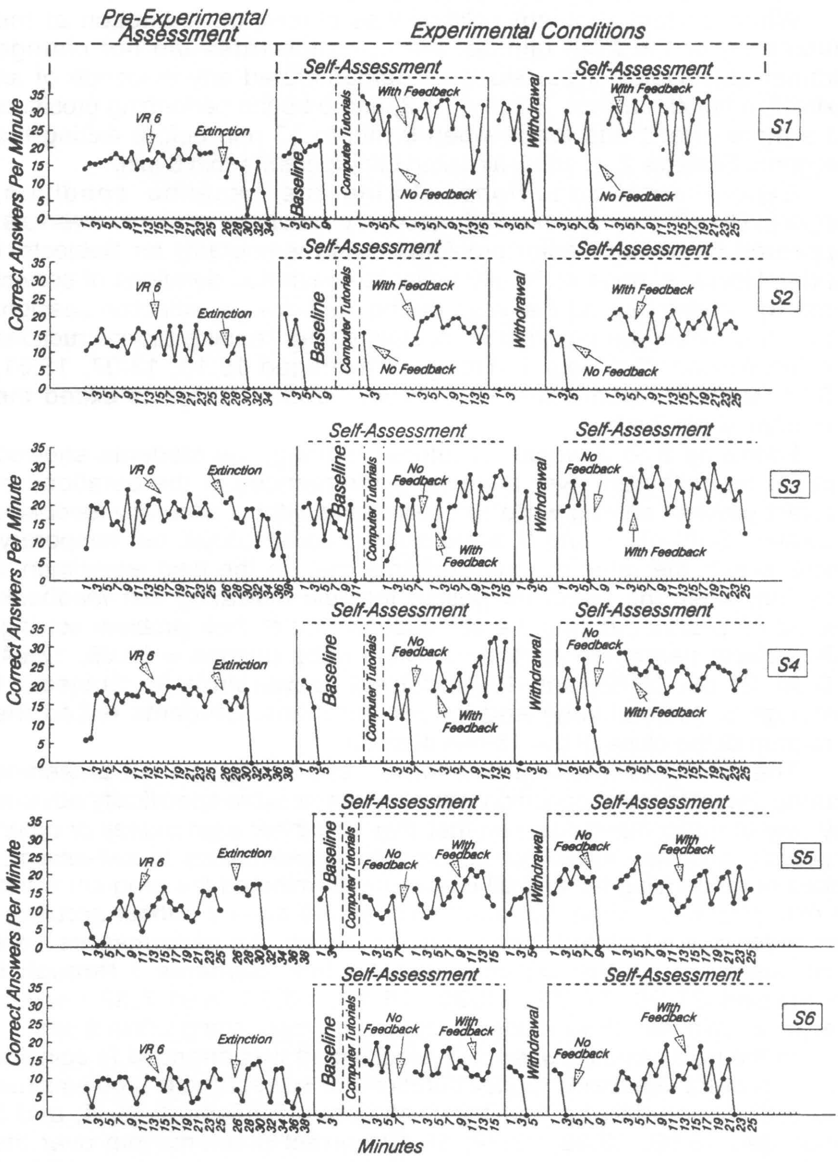

Figure 1. Preexperimental and experimental conditions for Subjects 1 through 6.

stable responding for 5 of the 6 students while the VR 6 contingencies were in effect. Subject 2 demonstrated somewhat more min-by-min variation in his rate of correct problem solving during the final 10 min of this session. Students 1-6 averaged 16.98, 13.17, 17.92, 16.78, 9.66, and 8.27 correct problems/min respectively.

When correct problem solving was placed on extinction at the beginning of the 26th minute, performance rates did not change dramatically. None of the students demonstrated any evidence of an extinction burst; however, Subjects 1, 3, and 6 began performing problems at a more erratic rate for between 9 min to 13 min before exiting the program. Subjects 2, 4, and 5 all exited the program within 6 min.

Experimental conditions. During the baseline condition, opportunities to self-assess *with accuracy feedback at 2-min intervals* appeared to promote faster problem solving temporarily for Subjects 1 and 3. However, none of the students demonstrated durations of correct problem solving beyond that seen during the previous extinction session. Subject 6 exited the program immediately after reading the instructions on the screen. Subjects 1 through 5 averaged 18.13, 13.07, 13.81, 13.76, 9.81, problems/min, respectively, and all subjects exited the program within 7 min.

Following 3 to 4 weeks of tutorial training, the students showed limited improvement over baseline performances in the durations of correct problem solving while *self-assessing without accuracy feedback*; however, Subjects 1 and 6 demonstrated conspicuous, but temporary, increases in the rates of their performances. In the next experimental session, when the computer posted the rule indicating that *feedback would be provided* during the self-assessment of their problem solving, all subjects performed at relatively high rates (means = 28.06, 17.64, 20.36, 21.69, 15.08, and 13.42 correct answers/min for Subjects 1 through 6, respectively) and longer durations. Students exited the program at the close of the 15-min session.

This improved rate and duration of responding was not preserved during the withdrawal condition. When students were specifically advised by way of the computer screen that *they could not earn money or other rewards, and they would not be provided opportunities to self-assess,* rates of responding fell and all 6 students terminated the program within 8 min. Moreover, when the opportunity to *self-assess without accuracy feedback* was reinstated, long durations of high rate performances were not sustained. In the absence of feedback, Students 1 through 6 averaged 21.96, 10.72, 18.30, 16.12, 16.73, and 7.89 correct response/min, but all six students terminated responding within 9 min.

In the next session, when self-assessment was arranged to coincide *with accuracy feedback*, longer durations of correct math problems/min were demonstrated by 4 of the 6 students. Subjects 2, 3, 4, and 5 averaged 18.09, 23.38, 23.32, 16.48 correct problems/min over the entire 25-min session, respectively. Subjects 1 and 6 averaged 23.10 and 10.57 correct problems/min, but terminated responding by the end of 21 min and 22 min, respectively. Nevertheless, the rates and durations of correct problems/min for these 2 students were beyond what they had accomplished under the *baseline self-assessment with feedback* condition.

Accuracy of self-assessments. Tutorial training appears to have had a positive influence on the *precision of student self-assessments with and without accuracy feedback.* While Subjects 2 and 4 produced 100% accurate self-assessments throughout baseline, Figure 2 demonstrates that Subject 1 obtained 50% accurate self-assessments, and that Subjects 3, 5, and 6 were inaccurate throughout all baseline self-assessments. After 3 to

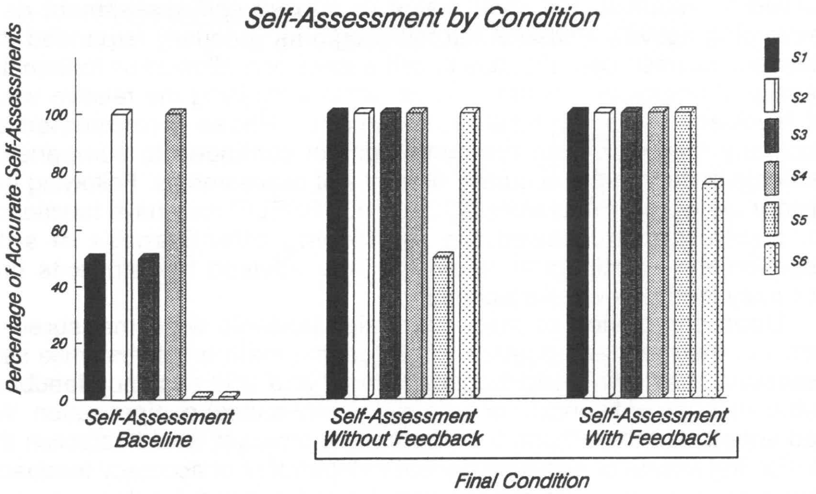

Figure 2. Percentages of accurate self-assessments during baseline and the second experimental condition for Subjects 1 through 6.

4 weeks of tutorial training, students demonstrated conspicuous progress in the percentage of accurate self-assessments. The final experimental condition is most revealing. Although Subject 5 obtained only 50% accuracy during the *self-assessment without feedback condition,* he was 100% accurate during the *self-assessment with feedback condition.* Moreover, 5 of the 6 students performed at 100% correct self-assessment throughout the duration of this final session.

Discussion

Our data appear to advance the theory that computer-interactive tutorials may provide a vehicle to establish self-assessment as a learned reinforcer while enhancing academic performance in the absence of exchange for financial or tangible/activity-based reinforcement. Additionally, these outcomes suggest that the *rules* regarding correct self-assessment gradually become somewhat more effective following a particular history. According to Hayes, Zettle, and Rosenfarb (1989), a rule that functions to increase the reinforcing effectiveness of a given event is termed an *augmental.* However, Hayes et al. point out that "it is not clear how various psychological processes could combine to produce augmenting, even after the rule is understood" (p. 207). Perhaps, elements of the training history provided by our tutorials recommend one format by which this process might occur.

Phase 1 tutorial programs accelerated student opportunities to self-assess and provided frequent compensation based on correct matches between student and computer assessments. This procedure may have served to establish the initial value of *correct self-assessment* as a reinforcing activity. Phase 2 tutorial programs gradually expanded the intervals between opportunities to self-assess and allowed us to lean the density of reinforcement per problem while sustaining the relative value of feedback for correct self-assessments. Phase 3 terminated all accuracy feedback from the computer but continued to compensate students based on the accuracy of their self assessments. Following this history the words, "YOU MAY SCORE YOURSELF" may have functioned as a rule that increased the reinforcing effectiveness of self-assessment—particularly, when the rule advised the students that accuracy feedback was forthcoming.

Upon completion of these tutorials, students were measured in terms of their rate and duration of performing math problems while self-assessing their own performances *without and with* accuracy feedback in the absence of financial (or tangible/activity-based) compensation. We had anticipated that Phase 3 of our tutorial program would establish the reinforcing effects of self-assessment independent of accuracy feedback. However, our data only provide very limited support for this outcome. While Subjects 1, 3, 4, and 5 showed elevations in their rates of responding following tutorials, their *enthusiasm* for performing multiplication problems *while self-assessing without feedback* was short-lived. During the first experimental condition that began by advising students that they would not be given accuracy feedback or compensation while performing math problems, all students terminated responding within 7 min. However, in the next experimental session when students were given opportunities to engage in self-assessment and receive accuracy feedback from the computer, this procedure appeared sufficiently *enticing* to support relatively higher rates and longer durations of math behavior.

During the withdrawal condition, we removed all opportunities for self-assessment, and math performance quickly diminished across students. Opportunities to engage in self-assessment did not seem to "reinvigorate" their enthusiasm until the computer posted the rule stating that it would provide *accuracy feedback for correct self-assessment* of math performance. Thus, it appears that although our tutorial programs seem to have *augmented* the effectiveness of rules regarding the opportunity to *self-assess with feedback*, this was not true of rules that described opportunities to *self-assessment without feedback*. These findings are consistent with outcomes in many applied settings in which termination of feedback for correct self-assessment resulted in regression of self-managed behaviors (e.g., Smith et al., 1992).

Perhaps, this should not be entirely surprising. Dickinson (1989) notes that all forms of *self-motivated behavior* may be grounded in some form of social exchange. She points out that behavior that is intrinsically reinforcing is often a form of learned reinforcement that has been established by approval from others. The reinforcing value of such rules may become enhanced by the individual's coming to understand and manipulate the environment, but the environment usually continues as, at least, an occasional source of acknowledgement.

During debriefing, students were asked why they had stopped solving problems during the *self-assessment without feedback* condition and why they had continued performing during the *self-assessment with feedback condition*. Subject 3 indicated, "doing problems and not knowing how well I was doing got very boring." When asked why he had performed when the computer gave him feedback, he remarked, "the whole thing got to be more like a game." All five of the other students provided answers that reflected essentially the same theme. It is interesting to point out, however, that these students did not appear to "enjoy this game" until after they had completed the tutorial process.

The behaviorally disordered students. Importantly, these procedures appear to have been as effective with students, identified as behaviorally disordered, as they were with regular education students. Although both behaviorally disordered students started at lower rates of correct problems/min and lower levels of self-assessment accuracy, they both demonstrated substantial progress following approximately 2 months of tutorial training. This was particularly conspicuous in the case of Subject 5 who almost doubled his rate of correct problems/min and tripled the number of minutes he continued to perform math problems in the absence of tangible reinforcement. Although Subject 6 did not demonstrate such overwhelming gains, he clearly benefited from learning to self-assess his own performance.

One caveat regarding this research presents itself. Characterizing our research as an A-BC-D-BC withdrawal of treatments design requires some amendment. Treatment (BC conditions) entailed *self-assessment without feedback* followed by *self-assessment with feedback*. Did our students' history of thinning financial reinforcement during computer-tutorial training act to curtail performance, because of a contrast effect, when the following no-feedback phase of the self-assessment condition was introduced after tutorials? Staggering presentation of treatment, or using some form of counterbalancing tactic, might have been helpful in precluding the influence of contrast effects on our results.

On balance, having our students perform during *self-assessment without feedback* first might have introduced some element of fatigue in the following *self-assessment with feedback* phase. Thus, our subjects may have operated at some physical disadvantage during their most productive *self-assessment with feedback* phase. Nevertheless, we are not able to rule out, unequivocally, the possibility of contrast effects within our methodology. And, although this study seems to suggest that establishing self-assessment as a completely independent reinforcing activity may be unlikely, it does not decisively preclude the possibility.

Notwithstanding, these results may have implications for self-assessment strategies in applied settings. Rule-following behavior regarding self-assessment procedures appeared to generalize across time most effectively when supported by some level of feedback. But, providing students with occasional accuracy feedback should not represent a serious obstacle to maintaining most self-assessment programs. Future research might explore a variety of self-assessment procedures that augment the value of rule-following in basic and applied research settings.

References

BRODEN, M., HALL, R. V., & MITTS, B. (1971). The effects of self-recording on the classroom behavior of two eighth-grade students. *Journal of Applied Behavior Analysis*, 3, 191-199.

CHRISTIE, D. J., HISS, M., & LOZANOFF, B. (1984). Modification of inattentive classroom behavior: Hyperactive children's use of self-recording with teacher guidance. *Behavior Modification*, 8, 391-406.

CONNELL, M. C., CARTA, J. J., & BAER, D. M. (1993). Programming generalization of in-class transition skills: Teaching preschoolers with developmental delays to self-assess and recruit contingent teacher praise. *Journal of Applied Behavior Analysis*, 26, 345-357.

DICKINSON, A. M. (1989). The detrimental effects of extrinsic reinforcement on intrinsic motivation. *The Behavior Analyst*, 12, 1-15.

FATUZZO, J. W., & ROHRBECK, C. A. (1992). Self-managed groups: Fitting self-management approaches into classroom systems. *School Psychology Review*, 21, 255-263.

FISHER, W. W., NINNESS, H. A. C., PIAZZA, C. C., & OWEN-DESCHRYVER, J. S. (1996). On the reinforcing effects of the content of verbal attention. *Journal of Applied Behavior Analysis*, 29, 235-238.

FLORA, S. R., PAVLIK, W. B., & PITTENGER, D. J. (1990). Effects of a masking task on schedule discrimination and extinction in humans. *The Psychological Record*, 40, 83-104.

GETTINGER, M. (1985). Effects of teacher directed versus student directed instruction and cues versus no cues for improving spelling performance. *Journal of Applied Behavior Analysis*, 18, 167-171.

GLYNN, E. L., & THOMAS, J. D. (1974). Effect of cueing on self-control of classroom behavior. *Journal of Applied Analysis*, 7, 299-306.

GLYNN, E. L., THOMAS, J. D., & SHEE, S. M. (1973). Behavioral self-control of on-task behavior in an elementary classroom. *Journal of Applied Behavior Analysis*, 6, 105-113.

HAYES, S. C., ZETTLE, R. D., & ROSENFARB, I. (1989). Rule-following. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 191-220). New York: Plenum.

HERSEN, M., & BARLOW, D. (1976). *Single-case experimental designs: Strategies for studying behavior change.* New York: Pergamon Press.

HOUGHTON, S. J. (1991). Promoting generalization of appropriate behaviour across special and mainstream settings: A case study. *Educational Psychology in Practice*, 7, 49-54.

KERN-DUNLAP, L., DUNLAP, G., CLARKE, S., CHILDS, K. E., WHITE, R. L., & STEWART, M. P. (1992). The effects of a videotape feedback package on peer interactions of children with serious behavioral and emotional challenges. *Journal of Applied Behavior Analysis*, 25, 355-364.

LALLI, E. P., & SHAPIRO, E. S. (1990). The effects of self-monitoring and contingent reward on sight word application. *Education and Treatment of Children*, 13, 129-141.

MAAG, J. W., REID, R., & DIGANGI, S. A. (1993). Differential effects of self-monitoring attention, accuracy, and productivity. *Journal of Applied Behavior Analysis*, 26, 329-344.

MACE, F. C., BROWN, D. K,. & WEST, B. J. (1987). Behavioral self-management in education. In C. A. Maher & J. E. Zins (Eds.), *Psychoeducational interventions in the schools. Methods and procedures for enhancing student competence* (pp. 160-176). New York: Pergamon.

MALOTT, R. W. (1986). Self-management, rule-governed behavior, and everyday life. In H. W. Reese & L. J. Parrott (Eds.), *Behavioral science: Philosophical, methodological, and empirical advances* (pp. 207-228). Hillsdale, NJ: Lawrence Erlbaum.

MCLAUGHLIN, T. F., BURGESS, N., & SACKVILLE-WEST, L. (1982). Effects of self-recording + matching on academic performance. *Child Behavior Therapy*, 3(2/3), 17-27.

MILLER, L. J., STRAIN, P. S., BOYD, K., & JARZYNKA, J. (1993). The effects of classwide self-assessment on preschool children's engagement in transition, free play, and small group instruction. *Early Education and Development*, 4, 162-181.

NEWMAN, B., BUFFINGTON, D. M., HEMMES, N. S., & ROSEN, D. (1996). Answering objections to self-management and related concepts. *Behavior and Social Issues,* 6, 85-95.

NINNESS, H. A. C., ELLIS, J., MILLER, B., BAKER, D., & RUTHERFORD, R. (1995). The effects of self-management on the transfer of aggression replacement skills in the absence of supervision. *Behavior Modification*, 19, 464-490.

NINNESS, H. A. C., ELLIS, J., & NINNESS, S. K. (in press). Self-assessment as a conditioned reinforcer during computer interactive math performance: An experimental analysis. *Behavior Modification.*

NINNESS, H. A. C., FUERST, J., & RUTHERFORD, R. (1995). A descriptive analysis of disruptive behavior during pre- and post-unsupervised self-management by students with SED: A within study replication. *Journal of Emotional and Behavioral Disorders*, 3, 230-240.

NINNESS, H. A. C., FUERST, J., RUTHERFORD, R. D., & GLENN, S. S. (1991). The effects of self-management training and reinforcement on the transfer of improved conduct in the absence of supervision. *Journal of Applied Behavior Analysis*, 24, 499-508.

NINNESS, H. A. C., & GLENN, S. S. (1988). *Applied behavior analysis and school psychology: A research guide to principles and procedures*. Westport, CT: Greenwood Press.

NINNESS, H. A. C., GLENN, S., & ELLIS, J. (1993). *Assessment and treatment of emotional or behavioral disorders*. Westport, CT: Praeger Publishers.

NINNESS, H. A. C., & NINNESS, S. K. (1998). Superstitious math performance: Interactions between rules and scheduled contingencies. *The Psychological Record*, 48, 45-62.

RHODE, G., MORGAN, D. P., & YOUNG, K. R. (1983). Generalization and maintenance of treatment gains of behaviorally handicapped student from resource room to regular classroom using self-evaluation procedures. *Journal of Applied Behavior Analysis*, 16, 171-188.

ROSENFARB, I. S., NEWLAND, M. C., BRANNON, S. E., & HOWEY, D. S. (1992). Effects of self-generated rules on the development of schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior*, 58, 107-121.

SKINNER, C. H., & SMITH, E. S. (1992). Issues surrounding the use of self-management interventions for increasing academic performance. *School Psychology Review*, 21, 202-210.

SMITH, D. J., NELSON, R. J., YOUNG, K. R., & WEST, R. P. (1992). The effect of a self-management procedure on the classroom and academic behavior of students with mild handicaps. *School Psychology Review*, 21, 59-72.

TURCO, T. L., & ELLIOT, S. N. (1986). Assessment of students' acceptability ratings of teacher-initiated interventions for classroom misbehavior. *Journal of School Psychology*, 24, 277-283.

WALL, S. M., & BRYANT, N. D. (1979). Behavioral self-management of academic test performances in elementary classrooms. *Psychology in the Schools*, 16, 558-567.

YOUNG, K. R., WEST, R. P., SMITH, D. J., & MORGAN, D. P. (1991). *Teaching self-management strategies to adolescents*. Longmont, CO: Sopris West.

VAN HOUTEN, R. (1984). Setting up performance feedback systems in the classroom. In W. L. Heward, T. E. Heron, J. Trap-Porter, & D. S. Hill (Eds.), *Focus upon applied behavior analysis in education* (pp. 112-125). Columbus, OH; Merrill.