TECHNOLOGICAL ADVANCES IN PRECISION TEACHING: A COMPARISON BETWEEN COMPUTER-TESTING AND SAFMEDS

Claudia E. McDade
Dorothy M. Austin
Charles P. Olander
Center for Individualized Instruction
Jacksonville State University

A look at the programs for the annual meeting of the Association for Behavior Analysis and the National Precision Teaching Conference for the past several years indicates that precision teachers are relying more frequently on computer-assisted instruction. The Center for Individualized Instruction at Jacksonville State University offers a program of "precision courses", some of which use a computer-generated, frequency-based testing program (Merbitz & Olander, 1980b; Olander, Yaracs, & Merbitz, 1980). Modifications to the program have been empirically based since its inception in 1979 (Merbitz & Olander, 1980a; McDade & Olander, 1984; Olander, McDade, Caine, & Merbitz, 1981; Olander, McDade, Grimsley, Yaracs, & Merbitz, 1981). One major question which was addressed in an earlier study (McDade, Olander, & Lea, 1983) but not fully answered is whether computer-based precision teaching is preferable over SAFMEDS.

SAFMEDS refers to a frequency-oriented learning system where a deck of cards is made with a question on one side and the answer on the other side of each card. The student's task is to proceed through the deck giving verbal responses to the questions as rapidly as possible, while a tester checks answers and "keeps track of the recording interval". Lindsley (1984) coined the term SAFMEDS to stand for "Say All Fast A Minute Each Day Shuffled."

The Center for Individualized Instruction offered two precision undergraduate psychology courses during the 1984-85 academic year. A study was undertaken in these courses to compare two frequency-based test formats. The first was the precision teaching technique of SAFMEDS, a card deck of at least one hundred free recall questions per unit. SAFMEDS were given to the student with unlimited sort time before the student was required to answer the questions verbally. The second testing format was computer-generated in a frequency based testing program which selects items and their alternatives at random from a test item pool of at least one hundred items per unit. Questions were taken from identical material in both testing formats. Students were given free access to both testing formats whenever a microcomputer or the instructor/psychology advisors were available.

METHOD

Fifteen undergraduate students successfully completing PSY 410: History and Theories in Psychology in fall semester-1984 and eighteen undergraduate students successfully completing PSY 335: Theories of Personality in spring semester-1985 served as subjects. All volunteered to participate in this study.

Course policies were exactly the same for both classes. Students were assigned units of material(i.e., 15 units in PSY 410; 14 in PSY 335) to master at their own pace, with the semester's ending date as the only limit on their progress. Optional discussion groups led by the instructor were held twice weekly with no external contingency on student attendance. Course requirements were described to students in the following section of the syllabus:

> Unit quizzes: Short answer identification
>
> All unit quizzes are mastery-based, allowing the student to retake any quiz without penalty until mastery is reached, or until the student is satisfied with his/her beyond mastery performance. Mastery is defined as at least 15 correct responses per minute. The student may take a unit quiz in the Center for Individualized Instruction with the instructor or a psychology advisor.
>
> Although all questions are taken from a large test item pool, they are presented in two formats-- concealed multiple choice format on a microcomputer or recall format on file cards. Using the Findley forced-choice procedure (Lockhart, Sexton, & Lea, 1975), tests will be given to all students in both formats. The Daily Progress Chart indicates to the student the testing format for the first unit. The student is then required to test in the other format on the second unit. Thereafter, the student may choose to test in either format. If, however, the student chooses the same format for three successive units, s/he **must change** to the other format on the following unit. While this procedure allows the student to maintain some control over his/her performance conditions, it also requires the student to use both test formats to

determine if one is preferable for the individual.

Unit quizzes are ten points each. Performance on each quiz is based on an accuracy score which counts only after the minimum of 15 correct responses per minute is reached. For example, the student may choose a score of 70% at 15 correct per minute to count toward his/her grade or the student may elect to continue working toward a higher grade, for instance, 92% correct at 35 correct per minute. In the first case, the student would receive 7 out of 10 points on the quiz; in the second, 9.2 out of 10 points.

Midterm and final exams: Essay

The student may take the midterm exam when s/he has mastered the first seven units and the final exam when s/he has mastered the last seven units.

All computer-generated tests were provided on Apple II microcomputers using a one-minute testing program designed within the Center for Individualized Instruction (McDade, 1985; Olander & Merbitz, 1980). All questions were designed by the instructor in the concealed multiple-choice format (Bowles, 1978). SAFMEDS were designed by the instructor with questions on one side and answers on the other in a free recall format. Students were given ten cards drawn at random from the card deck and tested verbally for one minute, with the answers checked by the instructor or a psychology advisor. A student evaluation of the course, the IDEA System (1981) was used. Additional instructor-made items were used to assess specific aspects of the courses.

RESULTS

The data were analyzed in two ways. From the statistical analysis perspective, each class was treated as a separate study using non-parametric comparisons for dependent samples, since sample sizes were small. Then the classes were combined into one group, using parametric comparisons for dependent samples. From the experimental (behavioral) analysis perspective, separate standard celeration charts for each testing format were plotted for each student.

Statistical Analysis of Results: The hypothesis that the highest best performances were no different in either testing format was rejected both with the Wilcoxon Matched-Pairs Test for individual classes and the t test for dependent samples for the combined classes (i.e., T[PSY 335]= 34 with n= 18 and T[PSY 410]= 9.5 with n= 15, $p<.01$; t= -4.88 with 32 df, $p<.01$). Higher best performances, evaluated by frequency of correct responses, tended to occur on SAFMEDS.

The hypothesis that the number of trials on the computer was no different from the number of trials on SAFMEDS was rejected at the .001 level (i.e.,T[PSY 335]= 1 with n= 18; T[PSY 410]= 2.5 with n= 15; t= 9.48 with 32 df). The mean total number of trials was more than three times greater using computer-generated tests (i.e., Mean = 34.0) than SAFMEDS (i.e., Mean = 10.1).

The hypothesis that the mean number of attempts to mastery was no different in either testing format was accepted (i.e., T[PSY 335]= 40 with n= 18; T[PSY 410]= 64 with n= 15; t= 0.67 with 32 df). The hypothesis that the mean number of attempts after mastery was no different in either testing format was rejected at the .02 level in the individual classes and at the .001 level with combined classes (i.e., T[PSY 335]= 18.5 with n= 18; T[PSY 410]= 0 with n= 15; t= 6.52 with 32 df). Students tended to test past mastery more frequently on the computer.

Experimental Analysis of Results: Charts 1 and 2 display data for one student in PSY 410. In PSY 335 fourteen of eighteen students and in PSY 410 thirteen of fifteen students showed higher best performances on SAFMEDS.

Only one student in each class used more trials on SAFMEDS than on computers. Fourteen of fifteen students in PSY 410 used the computer past mastery while only ten used SAFMEDS past mastery. In PSY 335 all eighteen students used the computer past mastery, while only eight used SAFMEDS past mastery.

DISCUSSION

Both SAFMEDS and computer-generated tests resulted in relatively high fluencies for both classes. Best performances were typically double the minimum fluency criterion for mastery. Since the number of attempts to mastery does not vary in either testing format, teachers can expect high levels of student performance using either SAFMEDS or computer-generated tests.

When higher fluencies, especially verbal ones, are desired, SAFMEDS should be preferable to computer-generated tests.

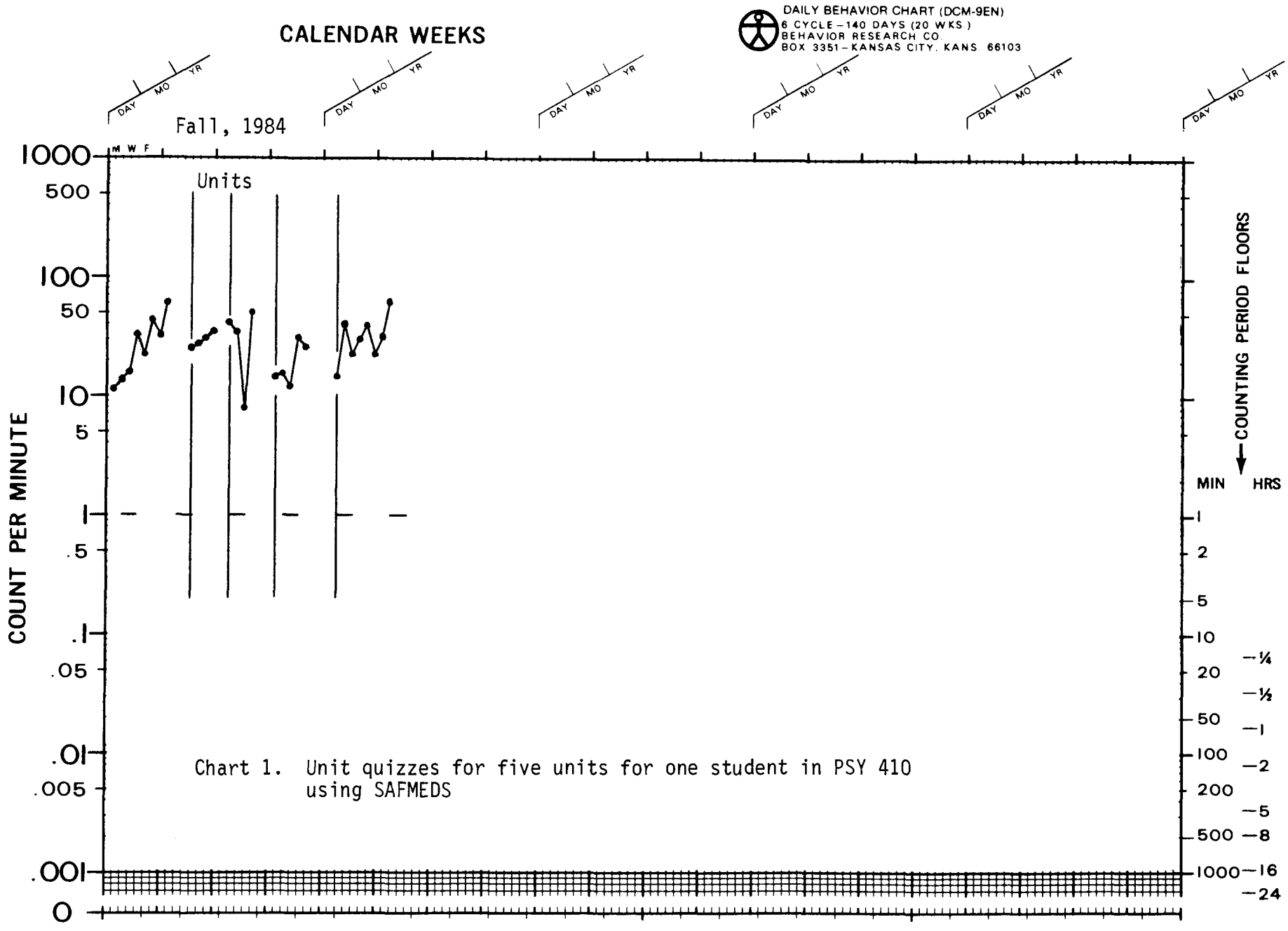

Chart 1. Unit quizzes for five units for one student in PSY 410 using SAFMEDS

CALENDAR DAYS

McDade	Wigley			S.S.			see-say answers
SUPERVISOR	ADVISER	MANAGER		BEHAVER	AGE	LABEL	COUNTED

Center for Individualized Instruction Jacksonville State University Jacksonville, AL

AGENCY

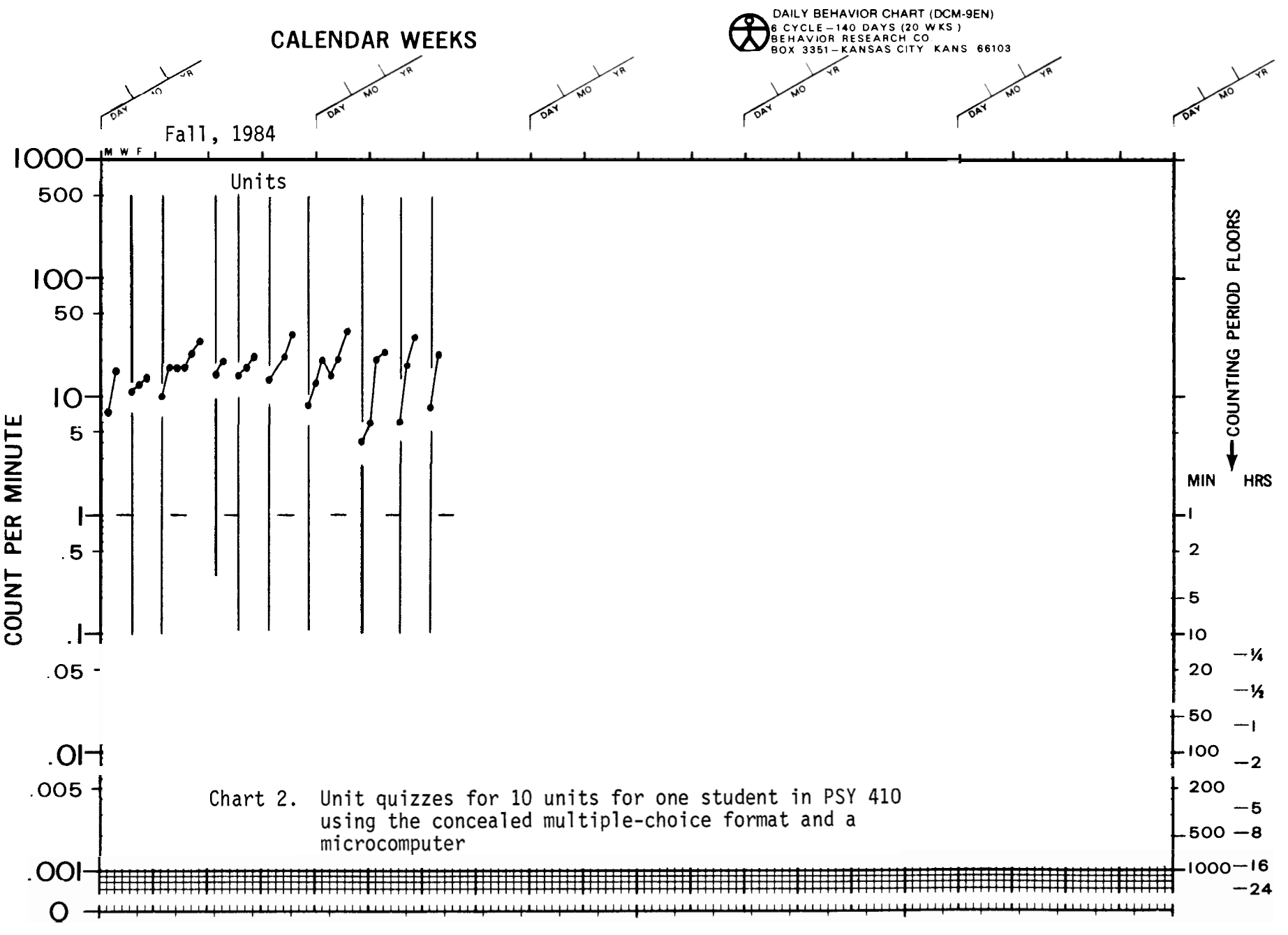

Chart 2. Unit quizzes for 10 units for one student in PSY 410 using the concealed multiple-choice format and a microcomputer

Limits to student performance on computer-generated tests may exist as an artifact of the particular program used. Further research with the testing program used in the Center for Individualized Instruction will attempt to determine whether such limits exist.

The findings that students are more likely to use the computer-generated tests than SAFMEDS and are more likely to test past mastery on the computer are consistent with the feedback students give to Center staff about the differences between the two testing formats. Computers are available from 8 a.m. to 7:30 p.m., while the instructor/psychology advisors are available for SAFMEDS testing fewer hours daily. Student evaluations of the courses indicated that they studied more for SAFMEDS than for computer-generated tests, because they viewed SAFMEDS as more "anxiety-producing". Students reported that the computer did not evaluate them "personally" for poor performances, the computer was never moody, and the computer was more accurate than a person. They said that they could learn more from the computer, they liked the immediate feedback given by the computer, and they found the computer tests equivalent to an enjoyable game. Teachers can encourage students to use the computer as a teaching device for just these reasons.

The best news is that effective precision teaching does not require expensive equipment. These data indicate that computer-based precision teaching is not preferable over SAFMEDS. Student learning can be managed quite well with SAFMEDS. However, precision teaching with computers can also be effective, especially for students to learn, study, and practice until high proficiency levels are reached.

REFERENCES

Bowles, K. L. (1978). Microcomputer based mass education. **Journal of Personalized Instruction, 3**, 151-156.

IDEA Survey - Student reactions to instruction and courses. (1981). Manhattan, KS: Center for Faculty Evaluation & Development.

Lindsley, O. R. (1984, April). **Antecedents and consequences of precision teaching and standard celeration charting.** Keynote address, National Precision Teaching Conference, Park City, UT.

Lockhart, K. A., Sexton, J., & Lea, C. (1975). The Findley procedure: A method for examining choice-making behavior in academic settings. In J. M. Johnston (Ed.), **Behavior Research and Technology in Higher Education.** Springfield, IL: Charles C. Thomas.

McDade, C. E. (1985). Computer-assisted instruction within the Center for Individualized Instruction. In **Issues in College Learning Centers.** Brooklyn: Long Island University.

McDade, C. E., & Olander, C. P., (1984, May). **What's wrong with higher education and what we are doing about it.** Invited address at the meeting of the Association for Behavior Analysis, Nashville, TN.

McDade, C.E., Olander, C. P., & Lea, C. R. (1983, May). **Flip cards vs. computer-assisted testing: A comparison of performance.** Invited paper in poster format at the meeting of the Association for Behavior Analysis, Milwaukee, WI.

Merbitz, C. T., & Olander, C. P. (1980). Precision teaching in a university center. **Journal of Precision Teaching, 1**(1), 18-25.

Merbitz, C. T., & Olander, C. P. (1980). Precision management of instructional technology: A model. **Technological Horizons in Education Journal, 7**(4), 32-35.

Olander, C. P., McDade, C. E., Grimsley, J. D., Yaracs, R. K., & Merbitz, C. T. (1981, May). **The effects of feedback timing in a computer-assisted PSI course.** Invited paper in poster format at the meeting of the Association for Behavior Analysis, Milwaukee, WI.

Olander, C. P., McDade, C. E., Caine, C. A., & Merbitz, C. T. (1981, May). **The effects of student preference of test type on performance.** Invited paper in poster format at the meeting of the Association for Behavior Analysis, Milwaukee, WI.

Olander, C. P., & Merbitz, C. T. (1980). Using technologies to teach: A CAI, A/V, PSI course. **Educational Technology, 10**(5), 50-52.

Olander, C. P., Yaracs, R. K., & Merbitz, C. T. (1980). An apple for the teacher. **The Apple Peel, 2**(1), 3-4.

Claudia McDade, Charles Olander, and Dorothy Austin are on the faculty of the Center for

Individualized Instruction, Jacksonville State University, Jacksonville, AL 36265.

IMPROVEMENT PICTURES WITH LOW CELERATIONS: AN EARLY FORAY INTO THE USE OF SAFMEDS

John W. Eshleman
West Virginia University

Ogden Lindsley coined the neologism "SAFMEDS" to stand for "Say All Fast a Minute Each Day Shuffled" (e.g. Lindsley, 1983). SAFMEDS is a functional substitute for the word "flashcards", for SAFMEDS specifies a procedure. Above and beyond this specification, SAFMEDS implies a particular instructional system. This system is relatively simple. First, one creates a set of SAFMED cards. These are usually 3 x 5 inch cards, with a problem, question, or premise written on the front and an answer written on the back of each card. Next, one "does" daily timings with the cards every day for several weeks. Prior to the timings, the cards are shuffled, so that one doesn't learn responses by the order of occurrence. During these timings, which are usually one minute in duration, one goes through as much of the deck as possible, looking at the front of a card, and attempting to say aloud what's on the back. After an answer is attempted, the card may be turned over, the answer checked, and the card put into either a "corrects" or "misses" pile. Provided a countdown timer is available, a timing can be done alone or with a friend. After the timing the number of cards in each pile is counted and the respective frequencies are plotted on standard celeration charts.

One of the distinct advantages of a SAFMEDS instructional system is that costs remain very low; a pack of 3 x 5 cards costs much less than a microcomputer. At the same time the system of SAFMEDS packs an educational "wallop". Moreover, it serves as a convenient way to study human operant behavior, specifically that category of verbal behavior that Skinner (1957) identified as "intraverbal". In addition, SAFMEDS can be used with any subject and educational level.

College students are perhaps the most difficult group to have as subjects in a SAFMEDS study. They already have an education history of at least 15 years. They may be quite "set in their ways", insofar as their learning and studying repertoires. After all, in 15 years, one is going to find the "right" way to succeed in school, and the "best" way to study. So, instead of going through the entire deck at once, one might decide to peel off the top 10 and learn them well, and then the next 10 and so on (Lindsley, 1980, 1983). Some might decide to study the cards for a half-hour once a week, going through them slowly, while "flashing" back and forth between front and back (Lindsley, 1980, 1983). Plus, given a choice, college students will probably create SAFMEDS on a topic that they already know something about, rather than on an alien subject. These are only some of the reasons why they seem to be the most difficult subjects.

Procedure

Keeping the preceding precautions in mind, I conducted a SAFMEDS study where seven undergraduate education majors served as subjects. All were enrolled in one section of a large introductory behavior analysis course. Students had a choice of participating in several projects for credit, and these seven chose the "Precision Teaching Project". They received course credit for both participating and for improving their performance.

I met weekly with them. We did an in-class timing at each meeting, students pairing up with each other, with one the behaver and the other the recorder, and then switching roles. Frequencies were charted immediately after the timings. During the remainder of the meeting we covered some aspect of charting, decided upon interventions, and set aims. For the intervening days they were instructed to do the timings in the SAFMEDS style and chart their responses. To encourage honesty, they were allowed to choose the topic or subject matter of their SAFMEDS. Not too surprisingly, all chose topics relevant to courses they were taking. Also, they were instructed not to put a data point on their chart on days they missed--the same convention as "no-chance days." (q.v., Pennypacker, Koenig, & Lindsley, 1972). The weekly meetings and periodic timings with me as recorder calibrated honest charting, at least to the point where "faking it" would make little difference.

The range of topics for the SAFMEDS included: (1) atomic element symbols, (2) French vocabulary, (3) physics formulae, (4) physical education training terms, (5) voice concepts, (6) agricultural mechanics terms, and (7) herbicide names. The latter two topics were selected by two foreign students from central Africa who planned to go back to