

**PARALLEL BETWEEN FREQUENCY TESTING
AND PERFORMANCE ON ESSAY QUESTIONS
IN A THEORIES OF PERSONALITY COURSE**

**C. E. McDade
S. B. Rubenstein
C. P. Olander**

Center for Individualized Instruction

Abstract

Reinforcing high frequency of correct responses in a personality theories course resulted in student fluency with those theorists studied. This study examined the relationship between frequency testing of basic concepts and application of those concepts on essay questions. A consistent concomitant increase in frequency of correct basic concepts described or identified on flash cards and frequency of correct concepts written on essay questions was observed. When concepts were identified through an identical method on a microcomputer, similar results were observed.

In recent years there has been an emergence of alternative measurement and instructional strategies such as Precision Teaching (PT), Personalized Systems of Instruction (PSI) and Computer-assisted Instruction (CAI). Each system has produced valid principles which can be used to improve instruction. The Center for Individualized Instruction (CII), Jacksonville State University, has experimented with frequency testing, a component of PT, in conjunction with CAI and PSI. Frequency testing refers to assessing student accuracy and fluency with academic material during short counting periods.

Frequency testing key basic concepts in a discipline can result in a proficient understanding of these concepts by the student, as well as, positive reinforcement for the instructor (Merbitz & Olander, 1980). Both student and professor develop a common ground--a working fluency with the terminology--which allows them to discuss freely and confidently the more complex ideas in the discipline.

A combination of three innovative teaching methods--PT, PSI, and CAI--allows a student to learn through self-paced instruction, with no penalty for repetition of material or for being "slower" than other students. Large lessons are divided into modules which the student learns one at a time, achieving a fluency criterion determined by the instructor.

Microcomputers can be used for frequency testing students, since they can minimize bias and completely eliminate scoring errors. Scanlon

(1981) suggested that the computer is not meant to replace the teacher and should rarely be used without advisor or instructor supervision. In the CII, microcomputers are used to provide testing, exercises, or simulations, most requiring advisor-student interaction to interpret computer-generated feedback.

The unique nature of frequency testing makes convincing students of the validity of the method difficult at first, but this is soon overcome. When a student experiences feelings of accomplishment from reaching a goal, s/he is apt to seek further opportunities to realize additional achievement-oriented behaviors (Porter, Lawler, & Hackman, 1975).

The present study compares frequency testing under two conditions: flash cards versus computer-generated terms. It also examines the validity of frequency testing by observing parallels between frequency test performance and essay question performance in university undergraduates enrolled in a theories of personality course.

Method

Subjects and Setting

Six undergraduate students enrolled in PSY 335: Theories of Personality at Jacksonville State University (JSU) served as subjects in this study. All subjects enrolled in the class were unaware of any experimental design incorporated into their course policy, although they were informed that the policies of the course would be modified in response to student performance.

The course was structured under the Personalized Systems of Instruction (PSI) method used at the Center for Individualized Instruction. Student performance sessions were monitored by the instructor or by graduate and undergraduate advisors with knowledge of the course material. Students were allowed to come to the Center as often as they liked to seek help on material.

Procedure

Students taking PSY 335 were required to become fluent with the ideas of one theorist before proceeding to subsequent theorists. After fluency with three or four theorists was demonstrated, the student took a review test, contributing one-fourth of the course grade.

Student evaluation exercises were of two types: (1) description or identification of basic concepts, terms or definitions, and (2) composition of an essay. Frequency testing of the descriptions or identifications required two fluency criteria: (1) an accuracy criterion of

80%, and (2) a minimum correct frequency of 10 concepts per minute. A randomly chosen essay question was used to assess student knowledge of more complex concepts (i.e., application, integration, synthesis, or evaluation). If the number of correct concepts written exceeded the number of incorrect concepts (i.e., those omitted or definitely incorrect), the essay was scored, "Pass!" If not, the essay was scored, "Try again!" Once a student chose to frequency test, s/he was required to attempt an essay question immediately after the frequency session, regardless of whether fluency was reached.

Course material was organized into four units, each with three or four individual theorists. The frequency testing component of evaluation was presented on flash cards for the first and last quarters of the course, while an Apple II microcomputer presented it for the second and third quarters. The flash cards allowed the student to see and sort all ten terms to be defined and required the student to give a verbal response to an advisor. It was the advisor's responsibility to time the student and assess his/her accuracy. When the microcomputer controlled the frequency testing, it presented questions in the concealed multiple choice format, randomizing both questions and individual foils. The computer determined the student's accuracy and fluency. In both forms of frequency testing, the student received immediate performance feedback. The difference between the forms of frequency testing could be described as recall vs. recognition. After each testing period, the students were required to chart their frequencies correct and incorrect for both essay performance and the flash card or computer frequency testing. All charting was done on Standard Celeration Charts, usually with the assistance of an advisor.

Results

Number of attempts to read fluency with both types of frequency testing--the flash card technique and the computer-assisted testing--were compared. No significant difference was observed. The mean number of attempts to reach fluency on the material by recall on flash cards was 9.0 and by recognition on the computer was 10.2 (Mann Whitney U = 15.5, $p = .38$). Additionally, the individual quarters of the course were compared with a Kruskal-Wallis One Way ANOVA. No significant difference in the number of attempts to reach fluency was found ($H = 1.68$, $\gamma = 3$).

A consistent, concomitant increase was seen in frequency of correct basic concepts described or identified on both frequency testing methods (i.e., flash cards and microcomputer) and frequency of correct concepts written on essay

questions. Charts 1 and 2 display one student's performance on frequency testing of basic concepts and essay questions. Generally (for all students), as the frequency of correct concepts on the frequency testing component increased, the frequency of correct concepts on the essay questions also increased. The opposite also holds true for the frequency of incorrect concepts.

Typically, also, the frequency of words and concepts written per minute on the essay questions increased over the course. A comparison between the last essay question on the second unit and the thirteenth unit showed a mean increase in frequency of words of 5.8 ($t = 2$, $p < .05$), as well as a mean increase in frequency of correct concepts of 0.69 ($t = 0$, $p < .05$). The authors also observed that students were answering the essay questions concisely, without excessive verbage.

Discussion

The present study has two implications for improving college instruction. First, frequency testing key basic concepts, resulting in the student becoming fluent as well as accurate, facilitates student use of those concepts on essays. Not only do the students apply the concepts better as they identify them fluently, but they also do so more concisely.

Secondly, frequency testing requiring recall and frequency testing requiring recognition appear to have the same effect on essay performance. With no difference in the mean number of attempts to reach fluency, both formats of frequency testing are efficacious. Instructors with access to microcomputers can enjoy the benefits of computerizing frequency testing without fear of the machine "de-humanizing" the classroom. Those who do not have the technological advances in their settings can create flash card decks which will assess student performance just as well.

The validity of frequency for shaping and evaluating student performance is being supported by a growing body of empirical data. Given the charge of teaching unique individuals with varying levels of intelligence, motivation, and past academic performance, the university instructor may find frequency indispensable for accomplishing the task.

REFERENCES

- Bloom, B. S., & Broder, L. **Problem solving processes of college students.** Chicago: University of Chicago Press, 1950.
- Lochhead, J., & Clement, J. (Eds.). **Cognitive process instruction: Research on teaching**



DAILY BEHAVIOR CHART (DCM-9EN)
 6 CYCLE-140 DAYS (20 WKS)
 BEHAVIOR RESEARCH CO
 BOX 3351-KANSAS CITY, KANS 66103

CALENDAR WEEKS

10 Jan 82
 DAY MO YR

7 Feb 82
 DAY MO YR

7 Mar 82
 DAY MO YR

4 Apr 82
 DAY MO YR

DAY MO YR

DAY MO YR

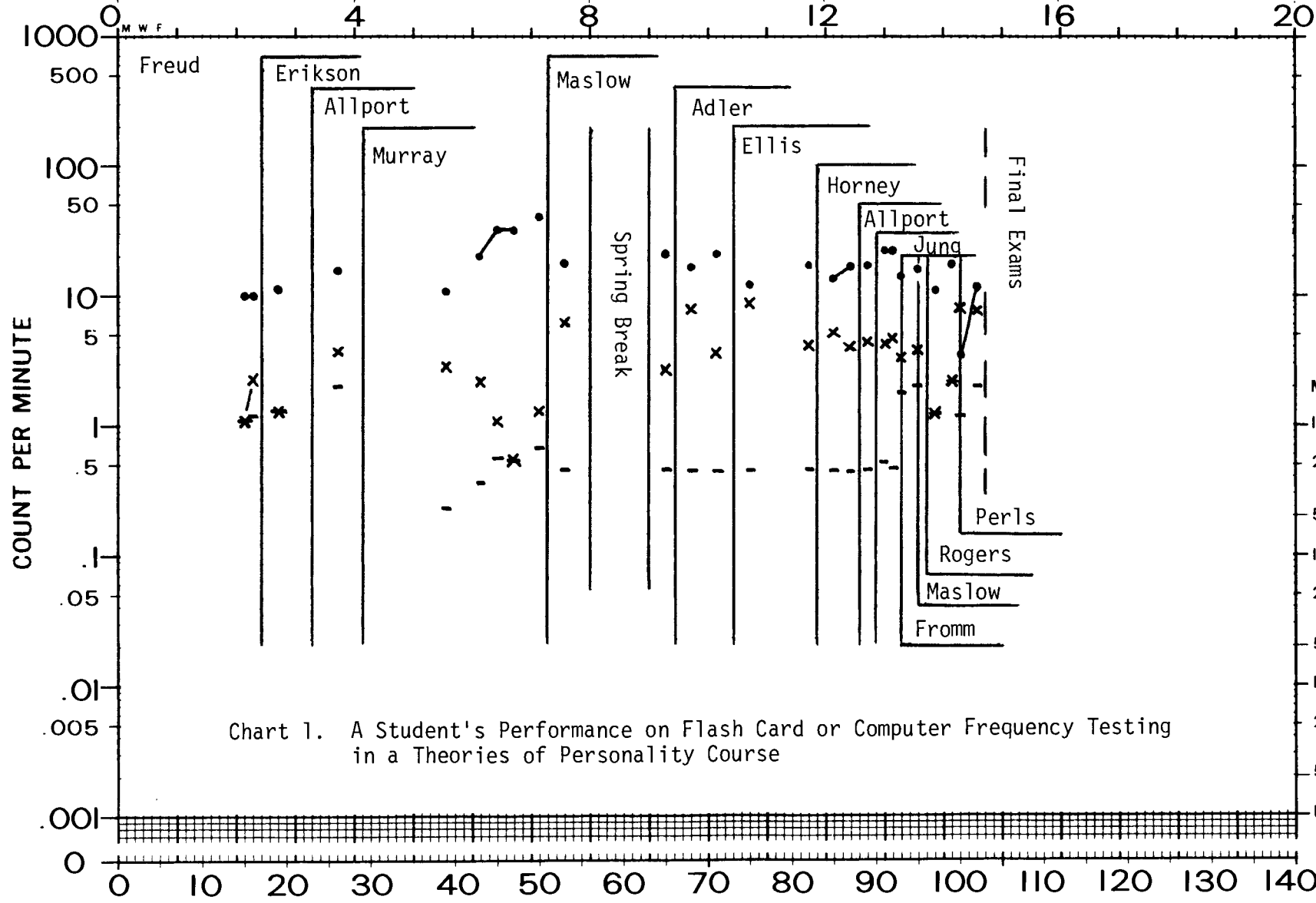


Chart 1. A Student's Performance on Flash Card or Computer Frequency Testing in a Theories of Personality Course

McDade, C.E., Rubenstein, S.B., and Olander, C.P. Parallel between frequency testing and performance on essay questions in a theories of personality course. *Journal of Precision Teaching*, Volume IV, Number 1, Spring, 1983.

McDade Olander Rubenstein **SUCCESSIVE CALENDAR DAYS** C.P. sees and says
 SUPERVISOR ADVISER MANAGER BEHAVIOR AGE LABEL ANSWERS
 Center for Individualized Instruction Jacksonville State University Jacksonville, Alabama
 AGENCY

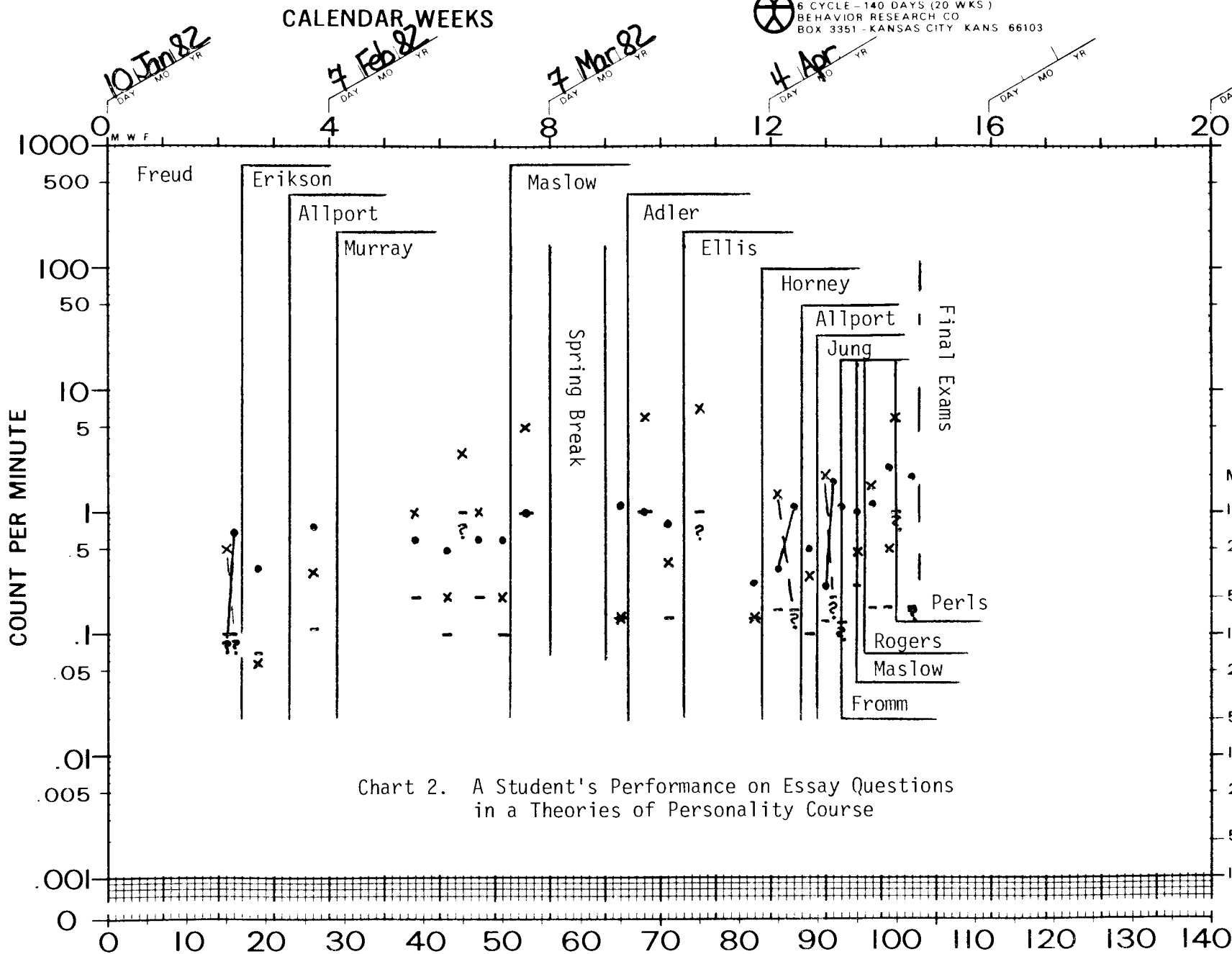


Chart 2. A Student's Performance on Essay Questions in a Theories of Personality Course

McDade, C.E., Rubenstein, S.B., and Olander, C.P. Parallel between frequency testing and performance on essay questions in a theories of personality course. *Journal of Precision Teaching*, Volume IV, Number 1, Spring, 1983.

thinking skills. Philadelphia, PA: Franklin Institute Press, 1979.

Merbitz, C. T., & Olander, C. P. Precision teaching in a university center. *Journal of Precision Teaching*, 1980, 1(1), 18-26.

Porter, L., Lawler, E., & Hackman, J. *Behavior in organizations.* New York: McGraw-Hill, 1975.

Scanlon, R. T. Computer-assisted instruction in Latin and in English vocabulary development. *The Classical Journal*, 1981.

C. E. McDade is the Director, S. B. Rubenstein is a graduate assistant and C. P. Olander is the Associate Director of the Center for Individualized Instruction, Jacksonville State University, Jacksonville, Alabama 36265.

VARIABILITY: AN AID IN THE ASSESSMENT OF THE EFFECTIVENESS OF TRAINING PROCEDURES¹

Fred Spooner

Western Carolina Center
and

University of North Carolina at Charlotte

Doreen Spooner

Western Carolina Center

Abstract

Pennypacker, Koenig, and Lindsley's variability procedure was used to illustrate the bounce in performance of eight severely/profoundly handicapped persons when they were trained to assemble two complex vocational tasks via two training procedures (total task presentation and backward chaining). From an analysis of the results, it can be concluded that subjects' correct performance under the backward chaining condition was significantly more variable than when correct responding was controlled by the total task condition. Practitioners are encouraged to quantify and analyze bounce to

¹The authors would like to acknowledge Owen R. White, University of Washington, for assistance in the analysis of the data; Jim Favell, Western Carolina Center, for editorial critique; Meda Smith, Western Carolina Center, for typing the final manuscript, as well as previous drafts; and Cynthia Stirewalt for assistance in re-drafting the design.

assist in making decisions about the effectiveness of training procedures.

The quantification of behavior change (celeration) and variability (bounce) are two of the many features of the Standard Celeration Chart that have a significant impact on the daily decision-making behavior of practitioners. The relationship between celeration and bounce and to what extent the bounce is due to celeration or uncontrolled variability is an important practical training issue. Frequently, this relationship is not used to its greatest practical utility--an aid in the assessment of the effectiveness of training procedures.

Pennypacker, Koenig, and Lindsley (1972) and White (Note 1) suggest that the more variability that can be explained by the effects of the celeration, the more effective the procedure, and the greater the predictive power of the procedure. In essence, measurement of variability during the treatment phase of a training program can be used to assist in assessing the effects of procedures on learning. The purpose of this paper is to explain Pennypacker et al.'s (1972) procedure for measuring and quantifying variability, and to apply this method to the assessment of the effectiveness of backward chaining (BC) and total task (TT) training procedures with severely handicapped persons.

Method

Subjects and Setting

The subjects were eight severely and profoundly handicapped individuals living in a state residential training facility. The five women and three men ranged in age from 14 to 58 years. Their I.Q.'s as measured by the Stanford Binet ranged from 14 to 27. Six of the subjects were enrolled in a vocational training program where they sorted plastic spoons, while two adolescent subjects were enrolled in an on-grounds school program.

The setting was a 5 by 4 meter room divided by a wall to provide two training rooms. The settings were tailored to be similar.

Apparatus

Two different items, a drain and a gate valve, were assembled by each subject. Each item consisted of seven different pieces. No pieces of the two items were identical or interchangeable. The drain was composed of a 7 cm by 3.5 cm drain head, a .65 cm rubber washer, a 6.5 cm hexagonal lock washer, a 4.5 cm by 10 cm pipe, a 4 cm slip nut, a 4.5 cm