USING REWARDS AND PENALTIES TO OBTAIN DESIRED SUBJECT PERFORMANCE

Marcia Cook, Henry R. Jex,
Anthony C. Stein and R. Wade Allen
Systems Technology, Inc.
Hawthorne, California 90250

SUMMARY

The Critical Tracking Test (CTT) is a psychomotor test that has proven in the past to allow reliable measurement of human operator performance, and has been shown to give sensitive decrements to a variety of stresses. Currently the CTT is being tested as a method of detecting human operator impairment. A pass level is set for each subject, based on that subject's asymptotic skill level while sober. It is critical that complete training take place before the individualized pass level is set in order that the impairment can be detected.

Some subjects have shown erratic learning trends typical of unmotivated performance. These subjects seemed more motivated to finish their training sessions quickly, rather than to receive monetary rewards for good performance.

This paper describes operant conditioning procedures, specifically the use of negative reinforcement, in achieving stable learning behavior. These results now provide a more general basis for the application of reward/penalty structures in manual control research.

INTRODUCTION

The Critical Tracking Test (CTT) is a psychomotor task based on manual control principles (Reference 1) that has been highly developed and applied to a variety of research problems (Reference 2). CTT performance is affected by individual differences and abilities, and critical instability scores (λ_c) vary across subjects. For a given trained subject, however, λ_c is quite consistent, and is sensitive to blood alcohol concentration. Thus, it is possible to set a pass/fail level required for a particular decision strategy (Reference 2). It is important that the pass criterion accurately reflect the subject's skill level. If λ_{pass} is too low, the operator will be able to pass while intoxicated. If λ_{pass} is too high, the sober operator will experience unnecessary and annoying delays while the equipment resets. One objective of this experiment was the development and refinement of a rational a priori procedure for setting pass levels based on training scores only. Because of motivational problems encountered in the original Experiment, a subsequent Training Experiment was carried out.

EXPERIMENT

Background

For the experiment, 24 subjects were obtained through the Los Angeles Municipal Courts, who had plead guilty to charges of Driving Under the Influence of Alcohol (DUI). The judge offered them the option of participating in this research project instead of the usual $350 traffic fine and traffic school. Even though the monetary rewards for participating were substantial, it was still a choice between the lesser of two evils. Subject participation was often reluctant, and in some cases subjects were effectively participating under duress.

The MMPI was administered and subjects with clinically abnormal profiles were eliminated from the population. The subjects were trained to "drive" in a simulator and to perform the critical tracking task. Three sessions of approximately 2 hours each on separate days were required for training. Following training, subjects participated in 3 experimental alcohol sessions, scheduled about a week apart, and lasting 10 to 12 hours. Each subject also participated in a follow-up session of approximately 1/2 hour. They received $3.10/hour for training and experimental sessions, and a bonus schedule was worked out to provide incentives for good performance and completion of the program (e.g., Reference 2). The subjects averaged about $220 total from wages and bonus money and the court cancelled a $350 traffic fine if they satisfactorily completed all requirements.

Learning curves were generated from training data to determine if asymptotic levels of performance had been attained. The 24 training plots were evaluated for indications of poor motivation or inadequate training. Half of the subjects showed lack of motivation and/or not enough training. These curves were characterized by unusually low scores in the last training session and a slope that was increasing even at the end of the last session (further discussion of this problem is found in section entitled Results). The remaining subjects showed asymptotic learning and consistent behavior, i.e., Figure 1. The simple pass criteria chosen for these 12 subjects correlated 0.862 with the "ideal" or hindsight criteria taken from the later experimental data as shown in Figure 2.

Methods

CTT training was done in 3 sessions on separate days. Each subject read the instructions (Appendix) and was informed of the bonus structure (at least 1 pass in a block of 4 trials = 75¢). Each training session consisted of 25 blocks of 4 trials, or 100 trials.

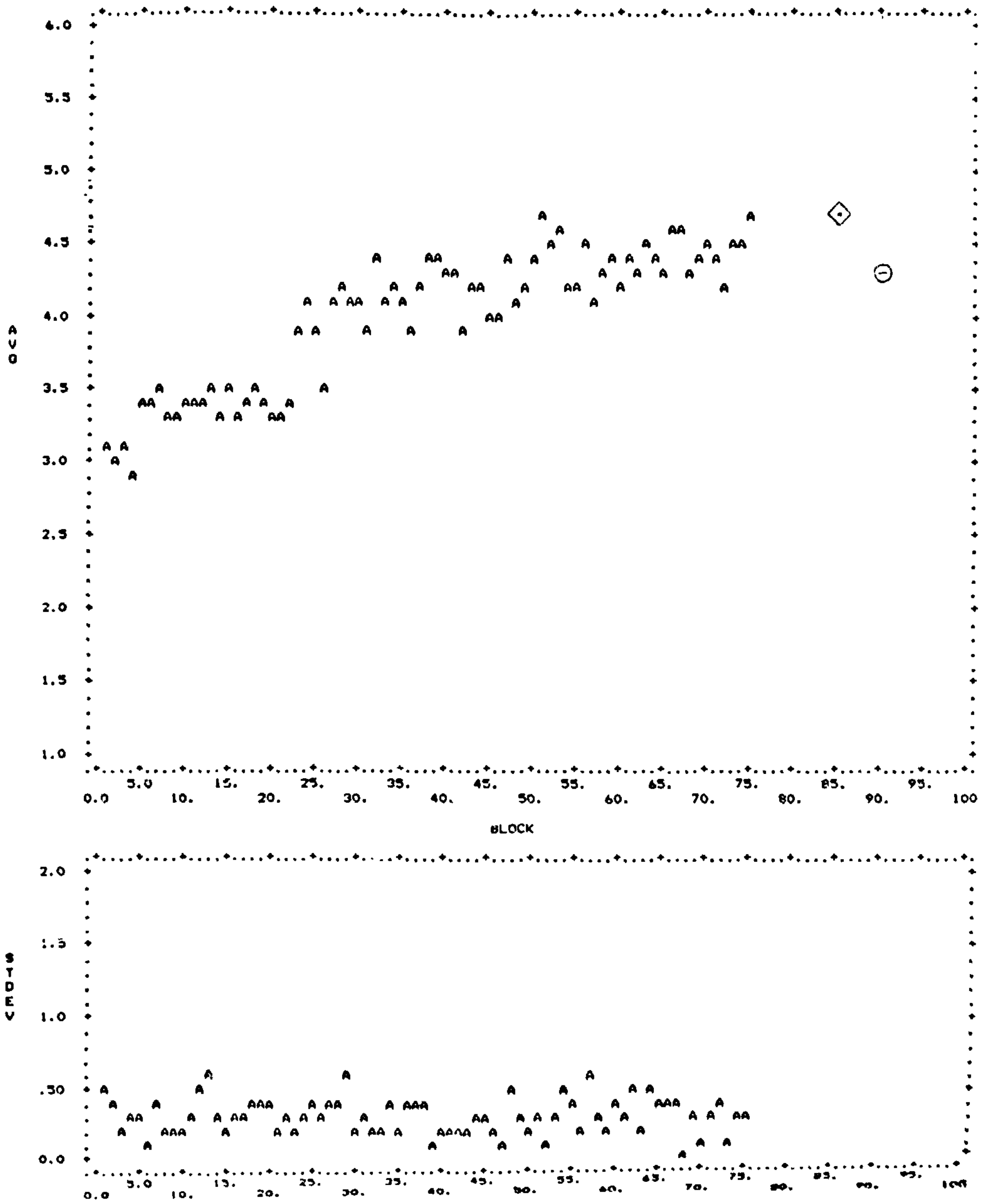

Figure 1. Good Asymptotic Training Exhibited by One Test Subject

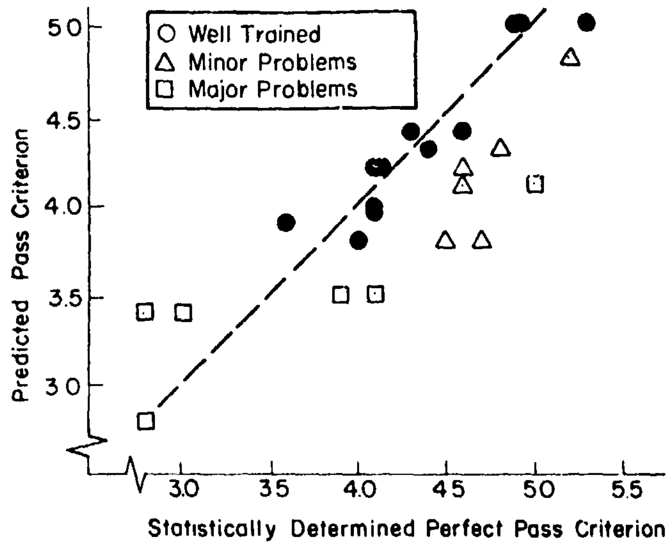

Figure 2. Pass Criterion Analysis

The initial pass criterion was set at an arbitrary low level that all subjects could pass within a few trials (i.e., $\lambda = 2.9$). If the subject passed 3 out of 4 trials in a block, the pass level was raised 0.1. If the subject passed all 4 trials the pass level was raised 0.2. If the subject failed all 4 trials the pass level was lowered 0.1. The pass level could go back to, but not below, a level where 3 out of 4 trials had previously been passed. This strategy tends to prevent deliberate "backsliding" by the subjects.

Breaks were given every 10 or 15 minutes. The mean λ_c for each block in the last training session was computed. The experimental λ_p was chosen by taking the median of the highest 3 block means and reducing this value by 0.3 (roughly 1/2 of σ_λ). This procedure was found to give a pass level consistent with a one trial probability of failure of 40 percent which was required for a one-pass-in-four-attempts decision strategy as discussed in Reference 2.

Results

For this experiment we were interested in determining the extent to which we were able to predict failure rates of impaired operators based on sober training data, and thus needed to predict an accurate pass score (λ_{pass}). The experimental (alcohol) session data (all sober baselines and placebo trials) were analyzed to determine in hindsight what the "ideal" (a posteriori) pass/fail score for each subject should have been to insure a 40 percent probability of failing a single trial when sober (as discussed in Reference 2). Comparing the experimental criterion set during this ideal criterion we found that in 15 cases the

λ_{pass} was too low (Figure 2). The solid circles indicate those subjects who were well trained. For these individuals our predicted pass level was remarkably consistent with the statistically determined perfect pass level. Problems were found when we compared predicted and perfect pass levels for subjects with both minor and major training problems (i.e., erratic learning curves). Subjects with minor problems exhibited a remarkable and consistent increase in λ_c scores as soon as there was no time benefit to failing. Subjects with major training problems had displayed problems consistently during training, and these problems persisted in the experiments. Even though the individual pass criteria were not perfect, the discriminability results agree with the statistical model described in the companion paper (Reference 2).

Discussion

Because 15 of the 24 subjects were assigned λ_{pass} criteria that were too low, the incentive structure was re-evaluated. The 75¢ bonus had been given for passing at least one trial in a block of 4. (This was chosen because it simulates the one pass out of four attempts strategy used in the formal validation experiment trials in the reanalysis of prior data (Reference 2). We found that most of the subjects did not even bother to learn the reward structure in order to maximize the bonuses earned (the bonus money was added across sessions and paid out at the completion of the experiment). They preferred to put forth a random effort and accept whatever total monies they happened to earn.

Research in operant conditioning (i.e., Reference 4) has shown that the more delayed such a "reinforcer" is, the less potent it is. Therefore, while the bonus money may have reinforced the subject for completing the experiment, it had little effect on each individual trial of the CTT. It was also observed during training that the subjects' primary motivation was to complete training sessions as soon as possible, even to the extent of foregoing rest breaks and failing the test quickly to speed up the trial repetition rate.

It was decided that in order to elicit a consistent and stable performance from the subject a more immediate reinforcer should be applied after each passed trial on the CTT. Reinforcement occurs when a rewarding stimulus follows a response or when a negative or punishing stimulus is avoided. It was decided that, since the training procedure is so tedious, a negative, or punishing, stimulus would be a 30 second "time out" condition added after each failed trial. By using the avoidance paradigm, the absence of the aversive stimulus (the "time out") becomes a reinforcer (Reference 4).

A subsequent study was conducted to verify if addition of this "time out" procedure would motivate the subjects in order that stable and complete training is obtained in three sessions.

TRAINING EXPERIMENT

Objectives

As described earlier, each subject is required to complete 300 trials on the CTT to establish an individualized pass/fail score. This is a tedious process done over 3 sessions, each lasting 2 to 3 hours.

Because of the poor performance and comments made during the experiment, we conducted a brief training experiment using the reinforcement strategy described above. In this situation a green "pass" light in the display of the CTT apparatus takes on new reinforcing properties, as it now signals the absence of a "time out." In the previous experiment the green light, in general, only provided information as to the outcome of the trial. The red "fail" light, on the other hand, now takes on aversive properties because it is present during the 30 second "time out." The red light was also merely informational in the previous experiment.

Procedure

Six additional subjects were contacted and enrolled through the Los Angeles Courts, as before. One hundred dollars of their $350 fine was dropped when they completed the project. The MMPI was administered as in the earlier Experiment. The first two subjects were given the printed instructions (Appendix) that outlined the objectives of the study and the bonus structure. They repeatedly expressed their surprise at being paid, because traffic schools do not pay for participation. Also, as in the earlier experiment, they chose not to learn the incentive structure and said, in effect, "You just keep track and pay me later." It was decided at that point that paying the subjects was superfluous and unnecessary, so the next four subjects received verbal instructions with no mention of wages or incentives.

The subjects received the same training procedure as in the earlier experiment (25 blocks of 4 trials each), except that each time a trial was failed there was a 30 second delay. [Note: These subjects all came after work and, because of the hour and the time of year, the testing, which took place in a parked car, was done in the dark. When the trial was completed, the display light went off with only the red "FAIL" or green "PASS" light remaining on, depending on the outcome of the trial. This meant that for a failed trial the subject waited in a dark car for 30 seconds, looking at a small red light that said "FAIL."]

Results

Some of the learning curves are shown in Figures 3-5 for the new condition using the "time out" procedure. There is no qualitative difference between those subjects that were paid and those that were not paid. The learning curves show the same asymptotic learning curves that

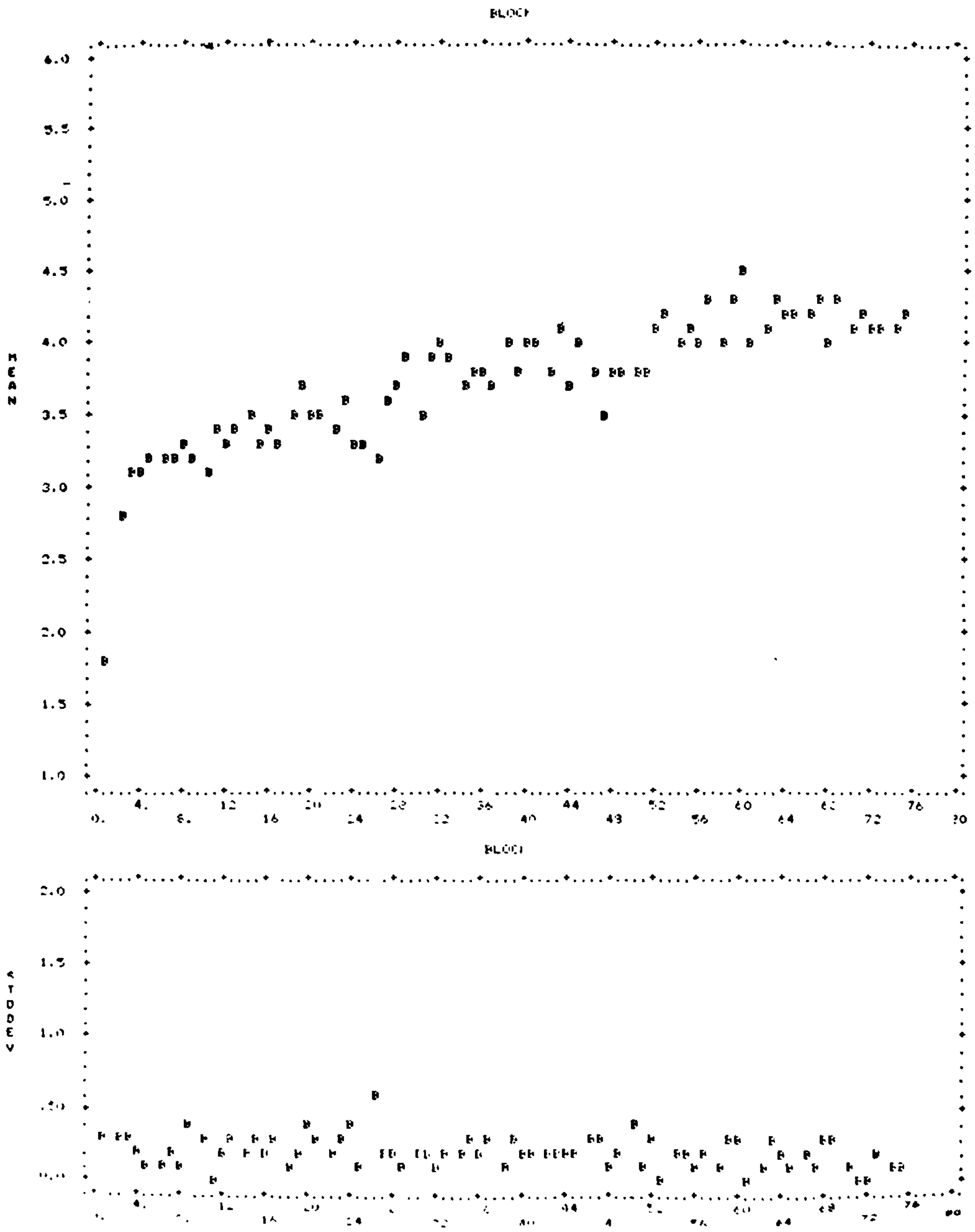

Figure 3. Learning Curve for a Subject Who Received Delayed Monetary Rewards for Task Performance in Addition to Immediate Wait Time Penalties

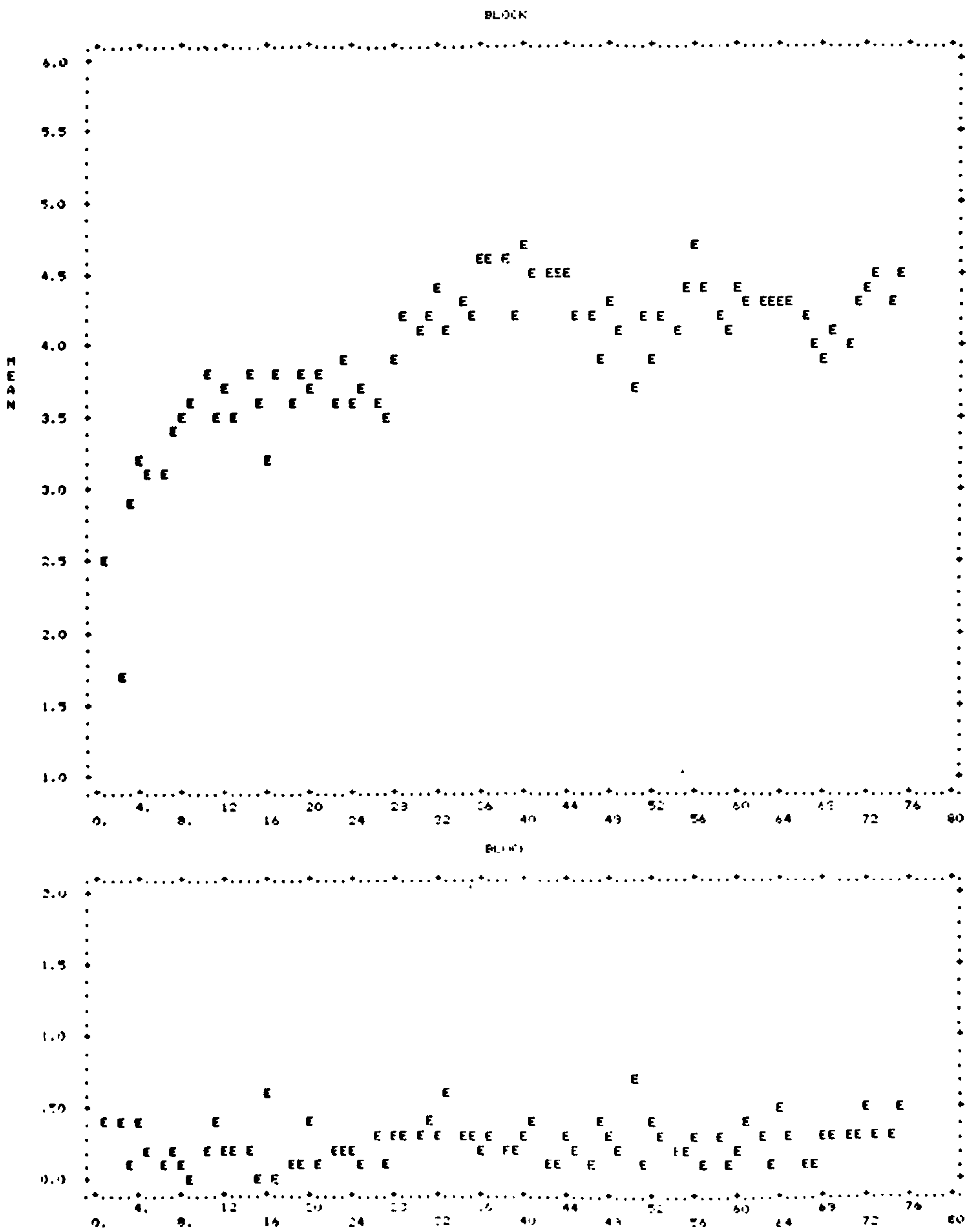

Figure 4. Learning Curve for a Subject Receiving No Monetary Rewards for Task Performance

ORIGINAL PAGE IS
OF POOR QUALITY

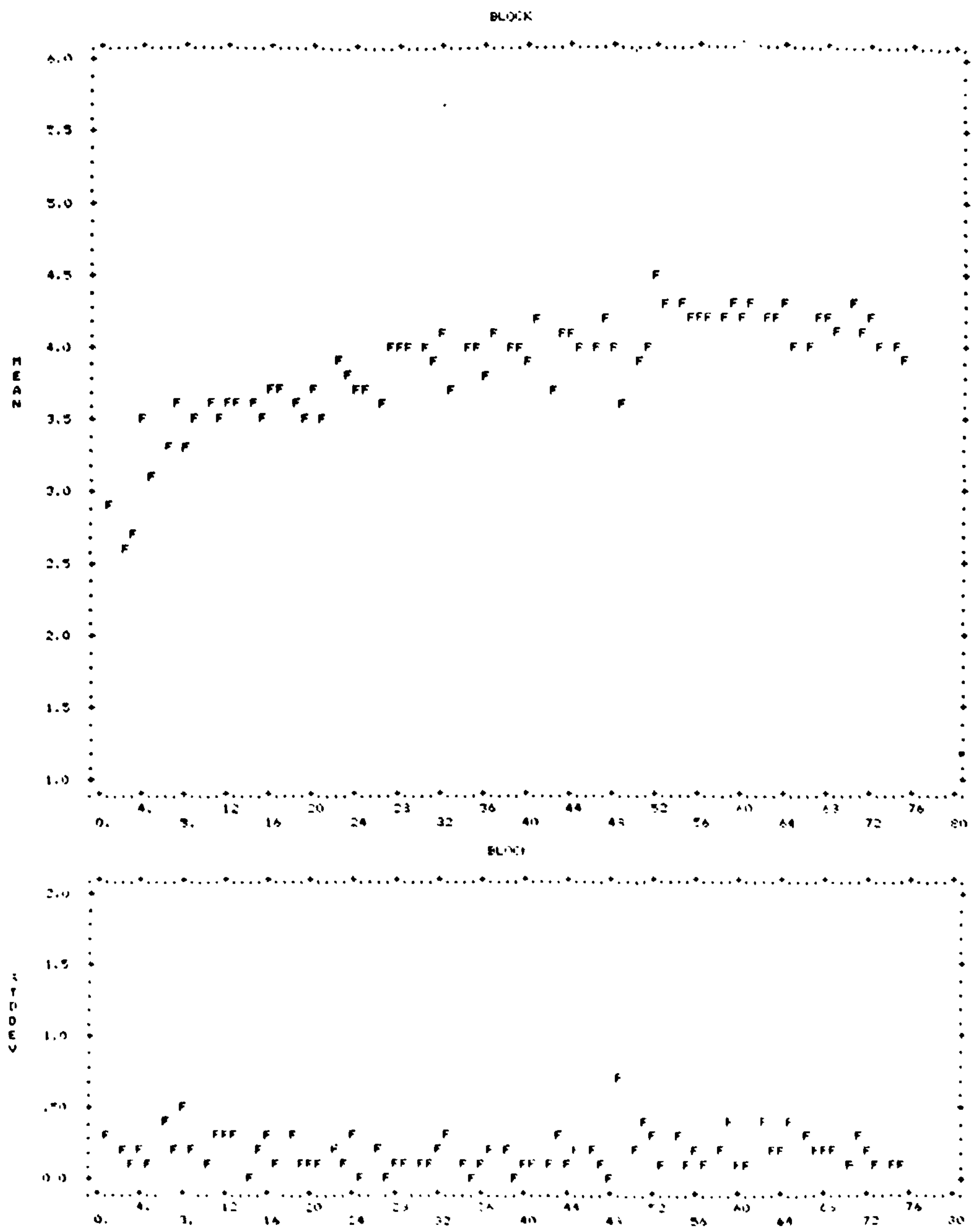

Figure 5. Learning Curve for a Second Subject Receiving
No Monetary Rewards for Task Performance

yielded accurate criteria predictions in the earlier experiment. Furthermore, the trial-to-trial consistency is better, as shown by the low standard deviations at the bottom of Figures 3-5.

CONCLUDING REMARKS

This study has shown that, to make an accurate prediction of a pass criterion, stable and complete learning must take place. It has also been shown that learning is facilitated when the subject is reinforced after each trial. In the past, subjects were given numerical feedback on each CTT trial as to what their actual scores were (i.e., References 5-7). This acted as an immediate and powerful reinforcer, because each trial provided an opportunity to better their last score. In addition, we were fortunate in working only with motivated subjects. For these reasons our earlier Reward/Penalty structures (Reference 3) were based on a positive reinforcement model only. In this experiment the subject was trained to "pass the test" without being aware that the eventual pass level was adjusted to their individual ability to perform. Since the normal reinforcement by display of CTT scores was not used, the 30 second time out after each fail was substituted.

In assessing what reinforcers are available for use with unmotivated subjects, money seems to be a neutral stimulus. The main motivator should be based on passing the test to avoid the aversive stimulus. In this way subjects will learn to pass the test without being aware that the eventual pass level is adjusted according to their individual ability to perform. The "30 second time out for fail" procedure seems to provide the desired immediate reinforcement after each trial and maintains consistent test scores. In addition, these results allow for a more general application of reward/penalty structures in manual control research depending on the basic motivation for subject participation.

REFERENCES

1. Jex, H. R., J. D. McDonnell, and A. V. Phatak, A "Critical" Tracking for Man-Machine Research Related to the Operator's Effective Delay Time. Part 1: Theory and Experiments with a First-Order Divergent Controlled Element, NASA CR-616, Nov. 1966.

2. Allen, R. Wade, Anthony C. Stein, and Henry R. Jex, "Detecting Human Operator Impairment with a Psychomotor Task," 1981 Annual Conference on Manual Control Proceedings, Jun. 1981, in press.

3. Stein, Anthony C., R. Wade Allen, and Stephen H. Schwartz, "Use of Reward-Penalty Structures in Human Experimentation," 14th Annual Conference on Manual Control, NASA CP-2060, Nov. 1978, pp. 267-278.

4. Reynolds, George S., A Primer of Operant Conditioning, Glenview, IL, Scott, Foresman, 1968.

5. Oates, John F., Jr., Experimental Evaluation of Second-Generation Alcohol Safety-Interlock Systems, DOT-TSC-NHTSA-73-9, 1973.

6. Oates, John F., David F. Preusser, Richard D. Blomberg, and Marlene S. Orban, Laboratory Testing of Alcohol Safety Interlock Systems. Vol. I: Procedures and Preliminary Analyses, Dunlap and Associates, Darien, CN, Jun. 1975.

7. Oates, John F., David F. Preusser, and Richard D. Blomberg, Laboratory Testing of Alcohol Safety Interlock Systems, Phase II, Dunlap and Associates, Darien, CN, Dec. 1975.

APPENDIX

This study is being conducted for the U.S. Department of Transportation to test a device designed to tell a driver that he may be too drunk to drive. Your participation in this project is important because you will be contributing to the improvement of measures for reducing alcohol related accidents. This phase of the study is aimed at helping us finalize our training strategy for the task.

You will use the device to perform a steering control task. We will provide you with an opportunity to become familiar with the operation of the device during three separate sessions. Each session will last approximately three hours.

The steering control task will be performed on a unit which is mounted on the side of the steering wheel in a specially modified car. The device consists of a meter (a display and a needle), a start switch, and a small computer in the trunk of the car. When the task begins, the needle is centered in the green area of the display. As the task proceeds, the needle begins to wander either to the left or the right. It is your job to keep the needle centered in the green area by moving the steering wheel in the direction you want the needle to move. It will become increasingly difficult to keep the needle centered.

The device will automatically set a passing level for your performance. This pass level is based on learning rates established in previous research. Trials will be conducted in groups of four. If you pass one trial in a group of four, you will receive a cash reward. If you have already passed one trial, it is important to continue to do your best on the remaining trials, since the task will continue to increase in difficulty, and your ability to pass future trials will be affected. The level of difficulty will change after each set of four trials. It will be necessay for the experimenter to reset the system periodically.

We expect that you will do poorly on the first few trials because you are completely unfamiliar with the task. You should not become discouraged; we know that your performance will improve with training.

You will be paid $3/hour plus a bonus of 75 cents per group passed for each session. You can expect to average $12-14 per session/

Your pay will be divided up into 2 parts — each day you will take home your hourly wages, thus on an average day you take home about $9.00. Your bonus money is held in a holding account, to be paid to you on your last day of experiments (approximately $13.50 in addition to your hourly pay).

Should you drop out sometime during the experiment, your holding account money will still be owed to you, but it can't be collected until the completion of the experiments, and you will have to collect in person.