METHODOLOGY IN THE ASSESSMENT OF COMPLEX HUMAN PERFORMANCE: THE EFFECTS OF SIGNAL RATE ON MONITORING A DYNAMIC PROCESS

W. Dean Chiles, Ph.D.
Cheryl Beveridge Bruni, B.S.
Robert A. Lewis, B.S.

Approved by

J. Robert Dille, M.D.
Chief, Civil Aeromedical Institute

Released by

P. V. Siegel, M.D.
Federal Air Surgeon

April 1969

Department of Transportation
FEDERAL AVIATION ADMINISTRATION
Office of Aviation Medicine

METHODOLOGY IN THE ASSESSMENT OF COMPLEX HUMAN PERFORMANCE: THE EFFECTS OF SIGNAL RATE ON MONITORING A DYNAMIC PROCESS

I. Introduction.

A number of the parameters of monitoring performance have been investigated extensively; a recent review of this literature was published by Jerison and Picket.[4] However, in the typical such study, the subject was concerned with only one task—the monitoring task being examined. This is rarely the situation that is obtained in the real world of work. Thus, such data are of limited direct usefulness in pursuing the long-term applied goal of research on human performance, viz, the development of techniques and data that will permit generalization to the "real world" domain of work behavior. Chapanis[1] has discussed this problem succinctly in the context of the design of equipment.

A recent paper by Chiles, Alluisi and Adams[2] identified and provided a rationale for the selection of a basic core of psychological functions held to be involved in complex work situations. Although these functions were selected in the context of aviation operations, they apply equally well to many ground environment situations. The functions enumerated are: monitoring, information processing, visual discrimination, and procedural performance. The tasks that were devised to assess these functions have been used in a variety of studies concerned with (a) confinement (in a general sense), (b) work schedules, (c) sleep loss, and more recently (d) infectious diseases as reported by Thurmond, Alluisi and Coates.[6] The findings of these studies give strong support to the proposition that subjects, through careful orientation, can be led to accept these tasks as being in the domain of work as opposed to the domains of experimentation or play.

Although the tasks have been found to be essentially orthogonal with respect to the behavioral skills required, they are not experimentally independent. For example, mental arithmetic scores are generally better when that task is performed concurrently with only monitoring tasks as compared to performance with a group problem solving task in addition to the monitoring tasks. An analogous finding holds for a visual discrimination task (target identification). In the case of the monitoring tasks, performance varies significantly as a function of the load imposed upon the man in the form of variations in the number of concurrently performed tasks.[3]

Essentially all of the studies conducted with this battery of tasks since 1962 have used virtually the same stimulus programs for the monitoring tasks; the average intersignal interval has been the same. Thus, although information has been obtained on the effects of workload on monitoring performance, information is lacking as to the way in which performance varies as a function of the parameters of the monitoring tasks themselves.

Thus, two of the purposes of the present experiment derived from this lack of information. The first purpose was to determine the effects on signal detection of variations in the intersignal interval of a task that involved the monitoring of a dynamic process; the task was called probability monitoring. The second purpose was to examine the interaction between the signal rate on probability monitoring and the other concurrently performed tasks.

The mean intersignal interval on probability monitoring in previous studies was 400 seconds—a signal rate of nine signals per hour. Because of the fact the major previous studies involved at least 8 hours of performance per day for periods of at least 12 days, the number of observations per subject per day (the typical unit of analysis) was sufficient to provide satisfactory stability of the data. In an experiment involving a 2-hour period of observation, only 18 signals would be presented; and, hence, stability might become a problem. Because of the amount of training required for mastery of the full battery of tasks, increasing the number of subjects is not a practical way of achieving greater stability. Thus, the third purpose of the experiment was to examine the feasibility of increasing the signal rate on the probability monitoring task as a technique for assuring stable measures with relatively short periods of performance.

In summary, the following questions were being asked in the experiment:

1. What is the effect of signal rate on a task involving the monitoring of a dynamic process when that task is performed as a part of a task complex?

2. What are the effects of the above variable on the other tasks of the task complex?

3. Does increased signal rate result in increased stability of the performance measures for short (2-hour) periods of observation?

II. Method.

Subjects. The 10 subjects were male students from the University of Oklahoma. They were paid for their services at the rate of $9.00 per 4-hour session. In addition, the subjects were told that if their performance for a given session was maintained at or above a level that was specified only in a general sense, they would receive a $2.00 bonus for that session. Payment of the bonus and knowledge as to whether or not it had been earned were withheld until the end of the experiment. Subjects were required as a condition of participation to agree that they would be paid only if they completed the entire series of experiments of which this study was the first; they would receive nothing if they dropped out prior to completion. Two subjects did drop out because of the "pressures of school work." These two subjects were replaced with new subjects, also from the University of Oklahoma.

The subjects were divided into two five-man crews; and, shortly before the end of the training phase, the best performing subject in each crew was designated as the "crew commander." (His rather simple duties will be outlined under the descriptions of the target identification and code-lock tasks.)

Apparatus. The specific tasks are described in detail in a previous report by Chiles, Alluisi and Adams[2] (pps. 152–158); and, therefore, only the probability monitoring task will be fully described here. This task required the subject to monitor four displays consisting of microammeters. The indications on each of the meters were programmed independently to yield a normal distribution of pointer positions for each meter ranging from -75 to $+75$ with a mean of 0; (the maximum scale values were ± 100). Each 0.5 second the meters were given a new instruction selected from 11 possible pointer positions with damping used so that a smooth pointer movement resulted. The critical signal to which the subject was to respond was a shift in the mean of the pointer positions from a normal mean of zero to either plus or minus 25, a shift of one standard deviation in the basic distribution. The shift was produced by adding a constant positive or negative voltage so that the range remained the same.

The subject was instructed to scan the meters systematically and, when he felt a particular meter to be biased, to watch that meter more closely until he had decided that it was or was not biased. The subject was also told that he should take action only if he was reasonably certain that a bias was present. He tested his hypothesis that a bias was present by throwing a three-position, center-off, lever-type switch in the direction he suspected the bias to lie. The meter would then automatically come to rest on its "true" mean, thus giving him immediate feedback as to the accuracy of his decision. If there was a bias present and the subject threw the switch in the correct direction, the bias was automatically removed when the subject released the switch. If it were biased in the direction opposite to his choice (but this very rarely happened), the subject had to move the switch in the correct direction to remove the bias. Any time that a subject moved the switch for any meter in either direction, a counter on the experimenter's console for that subject incremented by one count; if the direction were correct for the meter involved (e.g. thrown to the left when a negative bias was present on that meter) a correct counter incremented by one count. At the instant that a bias was introduced, a counter for each subject began incrementing at a rate of one count per second; a correct response stopped the counter.

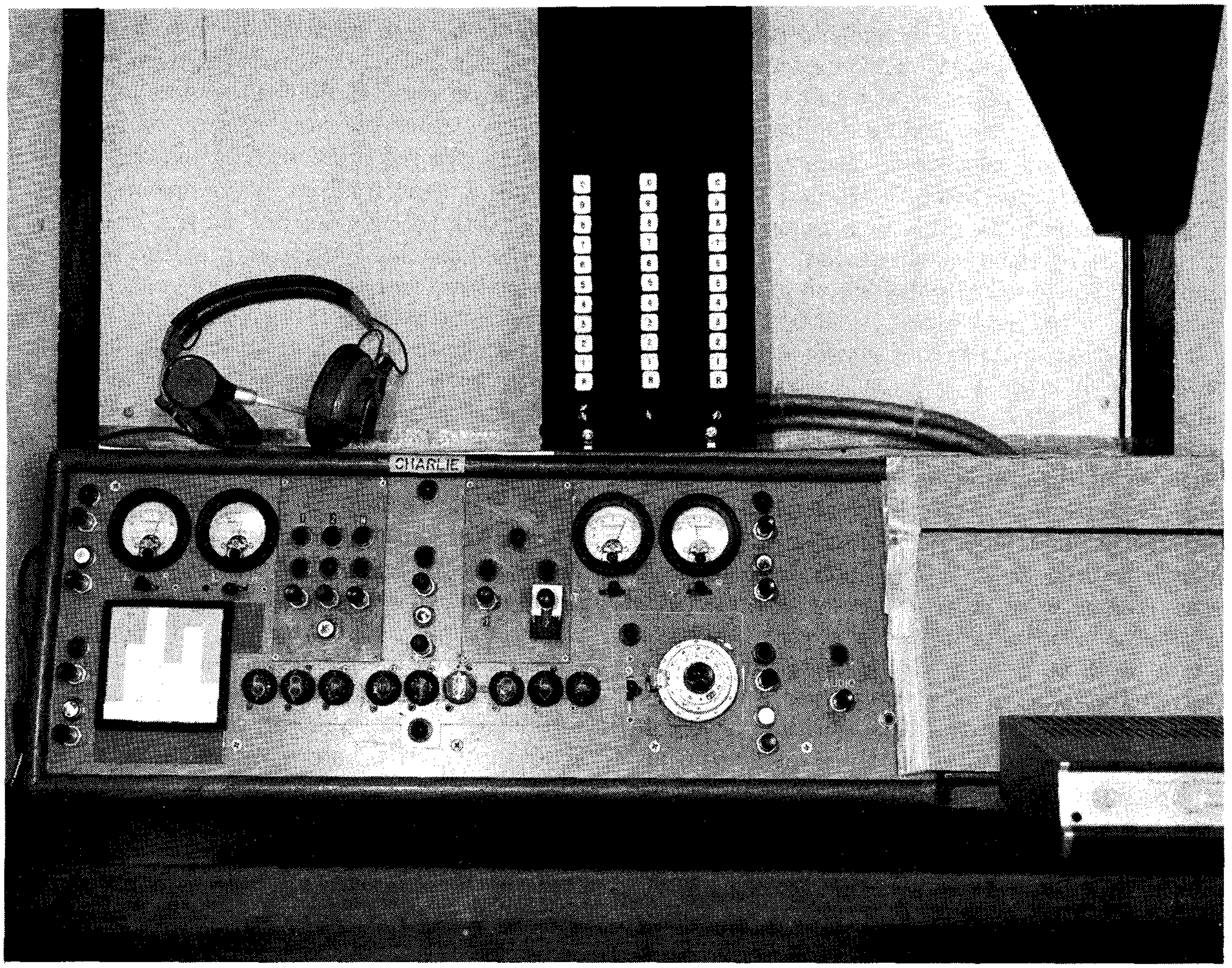

FIGURE 1. The performance panel.

One of the four remaining tasks also involved monitoring but, in this case, of a static process. This task, called warning lights monitoring, consisted of two aspects. The first aspect was represented by five red lights—one at each corner and one in the middle of the panel. The subject was instructed to push the button below a given red light any time that light was illuminated. The second aspect consisted of green lights paired with the above mentioned red lights. With this task, the subject was required to push a button below a given green light to re-illuminate that light if it should go off. For each of these tasks, response times were measured in tenths of a second from the onset (or offset) of a light until the subject returned the light to its normal condition by pushing the appropriate button.

The third task, mental arithmetic, required the subject to sum two three-digit numbers and subtract a third three-digit number from that sum. He entered his answer by appropriate manipulation of a set of three decade pushbuttons for the first three digits and a three-position, center-off switch to indicate whether the fourth digit was a "0" or a "1".

Problems were presented at a rate of three per minute; performance was scored in terms of the percentage of problems correct.

The fourth task was called target identification. This task required the subject to view a standard "target image" and then decide whether the first, second, or neither of two comparison images was the same as the standard image. The task was made somewhat more difficult by random distortion of the comparison images. Problems were presented at a fixed rate of two per minute. The presentations of the standard and two comparison images took place during the first 13 seconds of the problem period, and the response could be made any time after the second comparison image appeared, i.e., any time during the final 19 seconds of the problem period. The subject indicated his answer by depression of the appropriate one of three buttons marked "1", "2", and "N" (Neither). After the subject designated as the crew commander had made his response, an auxiliary display was illuminated at his crew position to indicate to him the choices of each member of the crew (including his own). Making whatever use of this information he wished, he then made a final selection of an answer "on behalf of the crew." The choice made as the crew response by the commander was then displayed on a three-light display at each individual crew position. Only the crew commander received feedback as to which answer was correct, and it was his option as to whether or not to pass this information on to the remainder of the crew. Performance on the task, both individual and crew, was measured in terms of the number of correct responses.

The fifth task, called code-lock, involved group performance. This task required the subjects to find the correct sequence in which each subject should push the code-lock button located on each subject's panel; entering the correct sequence illuminated a green light. The onset of a red light on each subject's panel indicated to the subjects that a problem was present. When the subject who was number one for that problem pushed his button, the red light would go out and remain out. If the next subject to push was any other than the number two subject in the sequence, the red light would come back on, thus providing error feedback to the group and indicating that the number one subject would again have to push his button so that the search for the number two subject could continue. The process continued in this manner until each subject had pushed his button once and only once in the correct sequence at which time the green light was illuminated. After a 30-second delay, the same problem was presented again and the subjects were to enter the previously obtained solution as rapidly as possible without error. Then, after another 30-second delay, a new problem was presented. The subjects used an earphone intercom system to coordinate both their initial search for the correct sequence and the entering of their second solution. The crew commander was responsible for assuring that a coordinated attack was followed in the search procedure and in entering the second solutions. Measures were made of the time required to reach the first solution, the time required to enter the second solution, the total number of errors and the total number of responses for the first and for the second solutions.

Procedure. Performance was measured on the combinations of tasks shown in Table 1; each test session lasted 4 hours and consisted of two of these programs in succession. Although no break was provided in the 4-hour schedule, subjects could, with permission, leave their duty

TABLE 1.—BASIC 2-HOUR TASK-PERFORMANCE SCHEDULE

Task	Minutes 00–15	15–30	30–45	45–60	60–75	75–90	90–105	105–120
Monitoring	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx
Arithmetic Computation		xxx	xxx					
Code-lock Solving			xxx	xxx	xxx	xxx		
Target Identification						xxx	xxx	
Level of Demand on Performance	low	med	hi	med	med	hi	med	low

NOTE: Each "x" in the table represents 5 minutes.

stations one at a time for purposes of going to the rest room or getting a drink of water during the period beginning at 105 minutes and ending at 130 minutes. The experiment consisted of three sessions: first, a baseline session in which signals were introduced on the probability monitoring task at the rate used during training; second, a session in which signals were introduced at a decreased rate; and third, a session in which signals were introduced at a rate faster than the one used during training. The characteristics of the three distributions of signals are shown in Table 2. The resultant signal rates were: Training—15.5 signals/hour; Slow—9.4 signals/hour; and Fast—20.6 signals/hour; the corresponding mean intersignal intervals were: 232 seconds, 381 seconds and 175 seconds.

TABLE 2.—INTERSIGNAL DISTRIBUTIONS FOR THE THREE SIGNAL RATES.

Frequency	Signals / hr.		
	Slow (9.4)	Medium (15.5)	Fast (20.6)
4	260	156	104
3	286	182	130
3	338	208	156
2	390	234	182
2	442	260	208
1	494	286	234
1	546	338	260
1	598	364	286
1	650	416	312
Mean Interval	381	232	175
Range	390	260	208

Prior to the beginning of the experiment, subjects were given extended training on the task complex. The specific amounts of training received by the two groups were as follows: four of the subjects in Baker crew received 48 hours of training (12 sessions) and one subject, who missed one session, received 44 hours of training; in Able crew three subjects received 44 hours of training, one received 24 hours and one 16 hours. Both of the subjects in Able crew who received the lesser amounts of training were "good subjects" and were judged to be adequately trained by the beginning of the experiment. In this respect, the remaining subjects were undoubtedly over-trained.

The five subjects in Able crew were routinely tested on Monday, Wednesday and Friday mornings from 8:00 to 12:00. Those in Baker crew were tested on Tuesday, Wednesday, and Thursday afternoons from 1:00 to 5:00. These were also the days and times at which the test sessions in this experiment occurred.

The only information that the subjects were given about the nature of the experiment was that it involved the work characteristics of the tasks. They were simply told that the training phase had ended and the test phase was beginning. At this time, the bonus procedure was also explained to them.

III. Results and Discussion.

Probability monitoring. The analysis of the probability monitoring data had to take into account the fact that a signal on this task remained until (a) the subject detected and eliminated the signal, or (b) the stimulus program called for a new signal to be introduced. As a result, changes in signal rate (mean intersignal interval) could directly affect the proportion of signals detected. The only assumption required to support such a prediction is the very reasonable one that the probability of detecting a given signal is proportional to the length of time for which the signal is present. A detection time measure based on only correct responses could be affected in an opposite manner. These possible direct effects were precluded by adopting an after-the-fact cutoff time for purposes of the analysis with no change in the task as viewed by the subject. Specifically, any response time greater than 104 seconds (the shortest intersignal interval of the fast signal-rate condition) was treated as a missed response in the primary analysis.

Separate nonparametric analyses (Friedman two-way analysis of variance by ranks) were carried out on the detection time measure (with the cutoff score) for each of the six conditions under which probability monitoring was performed as well as for the first 2 hours, the second 2 hours and the entire 4 hours of the sessions. The mean values for each condition and the results of the analyses are presented in Table 3.

The only performance condition under which the value of Chi square r reached the .05 level of confidence was that in which probability monitoring was performed concurrently with the visual discrimination task (target identification) and the group problem solving task (code-lock). The detection times for the fast-signal rate were approximately 38% shorter than for the slow-signal rate and 28% shorter than for the intermediate-signal training rate. However, the detection times for the slow and the intermediate rates were respectively 48% and 40% longer than the corresponding 4-hour means. In contrast to this, the typical finding in previous studies was a difference between 4-hour mean and target-identification-with-code-lock mean of 15% or less (with the longer response times being associated with the target identification condition). This leads to the speculation that whatever happened with this combination of tasks is perhaps not properly attributable to signal rate on the probability monitoring task. More will be said on that point subsequently.

No effects were found for the percent correct detections measure; and, therefore, only the 4-hour means will be given; they were 87.6%, 87.6% and 88.9% for the training, slow, and fast rates of signal presentation, respectively. As has been found in previous short-term studies with these tasks, the majority of the subjects (in this case 7 of 10) missed virtually no signals. The 2-hour means for each of the remaining three subjects ranged from 59% to 92%, 62% to 73%, and 42% to 72%. The high and low values for these three subjects were not systematically related to the signal rate conditions.

Warning lights monitoring. The mean response times to the onset of red lights are shown in Table 4 for each of the three signal rates on the probability monitoring task. Means and Chi square r's are shown for performance under each of the six task combinations, for the first 2 hours

TABLE 3.—DETECTION TIMES IN SECONDS FOR PROBABILITY MONITORING PERFORMANCE.

Task	Signal Rate			Chi Square r	P
	Training	Slow	Fast		
Monitoring only	21.4	24.9	21.6	3.2	
Arithmetic*	24.5	24.9	23.4	2.6	
Arithmetic and Code-lock	27.9	31.2	31.4	1.4	
Code-lock	24.0	22.6	22.8	4.2	<.20
Target Identification and Code-lock	35.5	40.0	24.5	7.4	<.05
Target Identification	26.6	26.5	24.6	1.4	
1st 2 hours	24.2	27.4	24.6	2.6	
2nd 2 hours	26.4	26.7	23.7	2.4	
4 hours	25.3	27.0	24.3	2.6	

* NOTE: The monitoring tasks were performed with all task combinations.

TABLE 4.—RESPONSE TIMES IN SECONDS TO RED LIGHTS.

Task	Signal Rate			Chi Square r	P
	Training	Slow	Fast		
Monitoring only	.9	.9	1.6	3.35	<.20
Arithmetic*	1.6	1.2	1.0	4.65	<.10
Arithmetic and Code-lock	1.0	1.1	1.2	4.85	<.10
Code-lock	1.0	1.1	1.0	.95	
Target Identification and Code-lock	2.2	1.0	1.4	8.15	<.02
Target Identification	.9	1.0	1.1	3.75	<.20
1st 2 hours	1.4	1.0	1.4	.65	
2nd 2 hours	1.0	1.1	1.0	4.55	<.20
4 hours	1.2	1.0	1.2	1.4	

* NOTE: The monitoring tasks were performed with all task combinations.

of the session, the second 2 hours and the entire 4 hours. The only instance in which response time was significantly affected by the differences in signal rate on probability monitoring was again in the case of the condition in which target identification was performed in combination with the code-lock task. However, in contrast to the finding for the probability detection time measure, the slow signal rate on probability was associated with faster response times to the onset of red lights.

In none of the previous studies has there been any indication that good performance on either the red lights or probability monitoring tasks is achieved at the expense of the other. And a rational analysis of the requirements of the two tasks would not lead to any such prediction. This again leads to the speculation that something unexpected happened with this particular task combination.

Data are shown in Table 5 for the offset of green lights. The only significant variation for this task was found for the data combined over the first 2 hours of the session. However, the mean values for the combined data for the second 2 hours showed a numerical—though not significant—reversal of the relation between the slow probability rate condition and the other two conditions. The data for the first 2 hours for each of the three conditions were further broken down in two ways—by 15-minute intervals as shown in Table 1 and by whether performance was or was not concurrent with code-lock. The first 15-minute period was the only one for which significant differences were found (Chi square r = 8.55; p is less than .02). However, the mean values for offset of green lights did not correspond to those for the 2-hour data; the means were 2.4, 2.7, and 1.1 seconds for the training, slow- and fast-probability rate conditions, respectively. Considering the second breakdown, the combined data for the four periods during which the code-lock task was not performed showed no effect of the probability rate variable; the mean values were 3.7, 2.1, and 4.9. For the four periods when code-lock was performed, the effect was significant (Chi square r = 7.8; p = .02). The mean values in seconds were 5.3, 2.4, and 2.8 for the training, slow, and fast probability conditions. When the data for the with-code-lock condition are further broken down by groups, it was found that the changes in the signal rate on the probability monitoring task affected the Able crew significantly (Chi square r = 7.6; p = .024) but not the Baker crew (Chi square r = 2.8).

Mental arithmetic and target identification. The data summarizing the performance of these two tasks, each of them with and without concurrent code-lock performance, are shown in Table 6 for the first and second 2-hour periods and for the entire 4-hour session. The only significant effect of the probability monitoring signal-rate variable was for the target identification task during the second 2 hours when performance was with the code-lock task. Inspection of Table 6 shows that this was the only 2-hour period in which performance of target identification under the slow probability rate condition was inferior to the other signal rate

TABLE 5.—RESPONSE TIMES IN SECONDS TO GREEN LIGHTS.

Task	Signal Rate			Chi Square r	P
	Training	Slow	Fast		
Monitoring only	2.5	1.8	2.2	1.25	
Arithmetic*	3.0	4.7	3.2	1.8	
Arithmetic and Code-lock	7.3	4.8	3.1	2.6	<.20
Code-lock	4.4	3.1	2.0	3.8	<.20
Target Identification and Code-lock	4.2	3.7	4.9	.8	
Target Identification	4.3	2.7	6.3	1.8	
1st 2 hours	4.7	2.4	3.3	11.40	<.01
2nd 2 hours	3.5	3.8	3.5	.35	
4 hours	4.1	3.1	3.4	.08	

* NOTE: The monitoring tasks were performed with all task combinations.

conditions; performance under this condition was intermediate during both 2-hour periods of the "without-code-lock" performance and was superior to the other two probability rate conditions during the first 2 hours of the "with-code-lock" performance. When the "with-code-lock" data were broken down by crews a puzzling result was obtained. In no case was the probability signal-rate effect significant with respect to target identification performance for the Able crew. For the Baker crew the effect was significant for the first 2 hours ($F = 6.81$; d.f. = 1,6; p is less than .05) with best performance associated with the slow probability rate condition. During the second 2 hours, the effect was also significant (Chi square $r = 8.4$; $p = .0085$), but best performance in this case was associated with the fast probability rate condition.

The task complex. The capriciousness with which the probability monitoring signal-rate variable appeared to affect performance on the battery as a whole is strongly suggestive of some uncontrolled group factor combined with the primary experimental conditions. There is a "built-in" possible provision for such a factor in the structuring of the groups as "crews" with a "crew commander." The two tasks on which the commander might be expected to exert influence are target identification and code-lock. In the case of target identification, the auxiliary display used by the commander in making his final crew response on this task provided him with complete information as to how each subject was doing on a problem by problem basis. Previous research has shown that the commander can and does develop expectations with respect to the capabilities of the individual subjects in performing this task. All that remains is for the commander to provide feedback to the individual subjects with some sort of commentary as to whether the subject is or is not doing as well as the commander thinks he should be doing. In this regard, the commander of Baker crew commented on the performance of the other crew members much more frequently than did the commander of Able crew.

The manner in which code-lock performance might be affected by the commander is straight forward. Because of the structure of the task and the search procedure, the commander is able to determine via the intercom how effectively each subject is contributing to the solution of the code-lock problems. Again, as with target iden-

TABLE 6.—PER CENT CORRECT FOR ARITHMETIC AND TARGET IDENTIFICATION TASKS.

Task	Period	Signal Rate			Chi Square r	P
		Training	Slow	Fast		
Arithmetic	1st 2hrs.	90.7	91.8	91.8	.20	
	2nd 2hrs.	90.7	90.9	90.7	.65	
	4-hour session	90.7	91.3	90.7	.20	
Arithmetic with Code-lock	1st 2hrs.	81.1	83.6	86.9	4.85	<.10
	2nd 2hrs.	85.1	84.9	84.9	.65	
	4-hour session	83.1	83.3	85.9	1.05	
Target Identification	1st 2hrs.	83.7	82.3	81.3	.45	
	2nd 2hrs.	82.0	83.0	85.0	.80	
	4 hour session	82.8	82.6	83.1	3.35	<.20
Target Identification with Code-lock	1st 2hrs.	67.3	75.0	66.3	4.85	<.10
	2nd 2hrs.	71.3	67.7	76.0	6.65	<.05
	4 hour session	69.3	71.3	71.1	.15	

tification, the commander of Baker crew was apparently much more willing to impose his thoughts and suggestions on the remainder of the crew than was the commander of Able crew.

Unfortunately, both crews began to place greater emphasis on speed in solving the code-lock problems at about the beginning of the study. Baker crew gave evidence of this in their within-crew conversations one or two sessions before the end of the training phase, and Able crew began emphasizing speed during the first session of the study proper. Thus, it is entirely possible that the obtained effects of the probability monitoring signal-rate variable only appeared to be effects of signal rate and that the actual processes underlying the changes in performance were crew or crew-commander effects. Alternately, it could very reasonably be argued that finding four comparisons to be significant at the .05 level or better out of a total of 39 comparisons would best be interpreted as a result of chance. From this point of view, then, it would be argued that the experimental design had insufficient power to reveal the effects of signal rate. Or, it could be argued that differences in signal rate, within the range of values used in this experiment, have little or no effect when the monitoring task in question is performed as a part of a task complex over a relatively long period of performance. The only real reason to doubt that the results were chance is the fact that the significant effects were consistently associated with the target identification/code-lock task combinations.

Good performance on the probability monitoring task is primarily dependent on the development of efficient scanning habits in time-sharing this task with the remaining tasks of the battery. If the training resulted in the development of a scan frequency that was high enough, then neither increasing nor decreasing the signal rate should have much of an effect on detection times. Obviously this argument rests on some sort of assumption to the effect that the requirements of the battery as a whole are sufficiently demanding that the subjects do not develop expectancies with respect to when the next probability monitoring signal will occur—or at least if they do develop them, the expectancies are weak enough that the subject's scanning habits are not altered. The better performers on the battery as a whole (and to some extent all subjects) do in fact develop regular scan patterns based largely on the "gaps" in the on-going stimulus program of the task complex. For example, after a suitable amount of training, the average subject can complete an arithmetic problem in about 10 seconds or less. With practically all subjects, at least a portion of the remaining 10 seconds before the next arithmetic problem is presented is spent in scanning the displays of the monitoring tasks and the probability meters in particular. Thus, the typical trained subject scans the meters once and perhaps twice every 10 to 15 seconds while performing the arithmetic task. The scan frequency is slightly lower during the time when the target identification task is presented, and it is much less regular when target identification is presented concurrently with code-lock and when arithmetic is presented with code-lock. However, even under the heaviest workload condition, the scan frequency is high enough that changes in signal rate should not affect the probability of detecting biases on the meters and should not have much of an effect on detection time. Unfortunately, these comments about scanning behavior are not based on quantitative data. They are based only on observations of and discussions with the subjects.

Stability of measures. The best estimate of the stability of the response-time measure on the probability monitoring task is provided by the reliability coefficient based on the shortest performance interval of interest. For the purposes of most of the projected research with this battery of tasks, the shortest period over which performance will be measured is 2 hours. Therefore, reliability coefficients were computed based on the first 2 hours of a session versus the second 2 hours for each of the probability monitoring signal-rate conditions. These coefficients are presented in Table 7. Both the Spearman r and the product-moment r are shown for each condition. (The mean values corresponding to the data used in computing these coefficients are shown in Table 3.) Table 8 presents the analogous coefficients based on the response-time measure without application of a cutoff score.

TABLE 7.—TWO-HOUR TEST-RETEST RELIABILITIES: PROBABILITY MONITORING DETECTION TIMES WITH CUTOFF SCORE.

	Training	Slow	Fast
Spearman r	.770	.742	.924
Product-Moment r	.877	.527	.964

TABLE 8.—TWO-HOUR TEST-RETEST RELIABILITIES: PROBABILITY MONITORING DETECTION TIMES WITHOUT CUTOFF SCORE.

	Training	Slow	Fast
Spearman r	.879	.855	.967
Product-Moment r	.868	.755	.951

The first point to be noted is that the reliability coefficients are quite respectable in all but one case considering the fact that they are based on an N of 10. The exception is seen in the product-moment correlation for the slow signal rate shown in Table 7, and this low value relative to the other coefficients is produced by the curtailment of the distribution resulting from the application of the 104-second cutoff score, a procedure that is peculiar to making comparisons among signal rates.

The data of primary interest are those of Table 8 which did not involve a cutoff score; and, of these coefficients, the product-moment correlations are of more interest in that they can legitimately be used in developing a regression equation. All three of these coefficients are significant at the .05 level or better as evaluated by t-tests based on a hyperbolic tangent transformation of the coefficients.[5] Using the same transformation, the difference between the reliability coefficient for the slow-signal rate (.755) and the fast-signal rate (.951) does not quite reach significance. The value of t (.10) is 1.895 and t (.05) is 2.365, each for seven degrees of freedom; the t value for the obtained difference was 2.254.

If it is assumed that increasing the signal rate to 21/hour does not affect the performance of this task other than to decrease the time required to present a given number of signals, then the data from this experiment can be compared with the data from previous research. Specifically, in a previous experiment using what is here called the slow-signal rate, a correlation of .76 was obtained between successive 2-hour periods of performance (i.e. 19 signals) based on an N of 20 subjects. Correlations were also computed on these same 20 subjects for 9 hours of performance (85 signals); in this case the reliability coefficient was .945, and the difference between these two correlations is highly significant. Thus, at some point in the vicinity of 40 signals, the reliability of this task appears to approach a maximum.

IV. Summary and Conclusions.

This experiment examined the performance of two five-man crews of highly trained subjects on a task requiring the monitoring of a dynamic process. The subject was to determine whether or not the mean position of any of four continuously fluctuating pointers had deviated from zero. This task was performed in a time-shared manner as a part of a task complex which also involved the monitoring of static processes, mental arithmetic, visual discrimination, and group problem solving.

The variable manipulated was the rate at which signals were presented on the dynamic monitoring task. Three rates were used: 9.4 signals per hour; 15.5 signals per hour (also the rate used during training); and 20.6 signals per hour. Performance on the dynamic monitoring task varied significantly over the three signal rate conditions with only one combination of tasks—when performance was concurrent with the visual discrimination and the group problem solving tasks. Poorer performance was associated with the slowest signal rate.

One of the two subtasks involving the monitoring of static processes (reaction time to the onset of a red light) showed significant variation with the same combination of tasks; however, in this instance, better performance was associated with the slowest signal rate. The other subtask (reaction time to the offset of a green light) showed significant variation in average performance for the first 2 hours. Better performance was associated with the slow-signal rate on the dynamic monitoring task; but, upon further breaking down the data, this finding held only for the average of those task combinations involving the group problem solving task.

The other task which varied significantly as a function of the signal rate on the dynamic monitoring task was the visual discrimination task. When that task was performed concurrently with the group problem solving task, performance was poorer during the slow-signal rate condition, but only for the second 2 hours of the session. This finding is beclouded by the fact that, when the data were further broken down by crews, the effect was significant for only one of the crews. During the first 2 hours, the differences were significant for that crew with the slow-signal rate condition yielding the best performance, whereas during the second 2 hours the effect was significant but with the fast-signal rate being associated with better performance.

The tentative explanation offered to account for the obtained pattern of results is that (1) the null hypothesis applies to this situation with respect to the effects of variations in the probability monitoring signal rate on the performance of that task as well as the other tasks of the battery, and (2) the obtained significant effects are real and should be attributed to the group problem solving task and possibly to the group component of the visual discrimination task.

Thus, with respect to the major purpose of the experiment, it is concluded that it is feasible to achieve increased stability of the performance measures on a dynamic monitoring task by increasing the signal rate on the task. This can be done without materially affecting either the basic properties of the task or its relation to the remaining tasks of the battery. This increased stability was evidenced by the substantial (though with these data, not statistically significant) increase in the 2-hour test-retest reliability coefficient on the task.

REFERENCES

1. Chapanis, A.: The relevance of laboratory studies to practical situations. *Ergonomics*, 10: 557–577, 1967.
2. Chiles, W. D., Alluisi, E. A., and Adams, O. S.: Work schedules and performance during confinement. *Human Factors*, 10: 143–196, 1968.
3. Hall, J. T., Passey, G. E., and Meighan, T. W.: Performance of vigilance and monitoring tasks as a function of workload. AMRL–TR–65–22 (March), Wright-Patterson AFB, Ohio, 1965.
4. Jerison, H. J., and Pickett, R. M.: Vigilance: A review and re-evaluation. *Human Factors*, 5: 211–238, 1963.
5. Smith, J. G., and Duncan, A. J.: *Sampling Statistics and Applications*. New York, McGraw-Hill, 1945.
6. Thurmond, J. B., Alluisi, E. A., and Coates, G. D.: An extended study of the behavioral effects of respiratory *Pasteurella Tularensis* in man. ITR–68–8 (February), Performance Research Laboratory, Univ. of Louisville, Louisville, Kentucky, 1968.