Contrails

FOREWORD

This research was supported by the Behavioral Sciences Laboratory, 6570th Aerospace Medical Research Laboratories, through the loan of the basic vigilance apparatus. The research supports Project 7183, "Psychological Research on Human Performance," Task 718305, "Behavioral Effects of Environmental Stress." Primary support for the research was provided by a grant from the Engineering Experiment Station of the Ohio State University. The essentials of this report have also been submitted as a doctoral dissertation to the Ohio State University.

The author is indebted to his joint advisers, Dr. G. E. Briggs and Dr. Daniel Howland. Thanks are also due to the staff of the Systems Research Group and the Staff of the Laboratory of Aviation Psychology, Ohio State University, and Dr. Harry J. Jerison, Antioch College. Dr. W. Dean Chiles, Environmental Stress Section, Training Research Branch, Behavioral Sciences Laboratory, was instrumental in providing the loan of the vigilance equipment.

Approved for Public Release

Approved for Public Release

Contrails

ABSTRACT

An experiment was performed to determine the transfer effect of knowledge of results and signal rate on performance in a Mackworth-type vigilance task. Subjects were run the first day under feedback conditions of zero, partial, and full knowledge of results, and 16, 32, and 48 signals during a 48-minute run. On the second day all subjects were run under the conditions of zero knowledge of results and 32 signals. The results showed that the two experimental variables differentiated subjects on both the initial exposure and the transfer condition. In terms of percentage of signals detected, groups initially trained with knowledge of results and high signal rates showed superior performance on both days. The usual decrement in performance over time was noted on both days. With respect to commissive errors, large individual differences contaminated the results. The data suggest that partial knowledge of results may encourage the operator to make more commissive errors than either zero or full feedback. The findings appear to recommend training with full knowledge of results and high signal rates when an operator must be placed in a situation with no knowledge of results and low signal rates.

PUBLICATION REVIEW

This technical documentary report has been reviewed and is approved.

Walter F. Grether
WALTER F. GRETHER
Technical Director
Behavioral Sciences Laboratory

Approved for Public Release

Contrails

TABLE OF CONTENTS

Approved for Public Release

Contrails

LIST OF ILLUSTRATIONS

Approved for Public Release

Contrails

LIST OF TABLES

Approved for Public Release

Contrails

THE MONITORING PROBLEM

A single day in 1959 serves as a tragic testimonial to the importance of the human monitor in man-machine systems. Late in the afternoon of February 3, during the first week of America's entry into the commercial turbojet market, a Pan American Airways Boeing 707 inbound to Gander, Newfoundland, at 35,000 feet entered an uncontrolled descending spiral. About 6,000 feet above the ocean, the aircraft was recovered and flown to Gander with only minor injuries to the occupants and slight damage to the airframe. The synopsis of the Civil Aeronautics Board's accident report (ref. 10) reads in part:

> The autopilot disengaged and the aircraft smoothly and slowly entered a steep descending spiral. The copilot was not properly monitoring the aircraft's instruments or the progress of the flight and was unaware of the actions of the aircraft until considerable speed had been gained and altitude lost.*

With this inauspicious beginning, the United States entered the commercial jet age. But the day was not over yet. Five hours later an American Airlines Lockheed Electra, during the first month of operation of an American-built turboprop, and again under autopilot control, crashed into the East River, a mile short of the runway at LaGuardia Airport, this time with a loss of 65 lives. This marked the first of the string of highly publicized Electra accidents. The CAB report on this accident, while recognizing a number of contributing factors, states (ref. 11):

> The Board determines the probable cause of this accident was premature descent below landing minimums which was the result of preoccupation of the crew on particular aspects of the aircraft and its environment to the neglect of essential flight instrument references for attitude and height above the surface.

These two accidents provide a perhaps overdramatic introduction to the problem of the human monitor, and point out

---

*As a result of this accident the carrier directed that during autopilot operation one pilot give continuous attention to the attitude and flight of the aircraft.

Approved for Public Release

Contrails

that the need to understand and improve man's performance of the monitoring task is considerably more than an academic question. This section traces the historical development of the need for and interest in the human monitor, points out a number of applications of the monitoring role, and discusses possible solutions to this difficult problem.

The origin of the problem may be traced to about 1760, which marks the approximate beginning of a movement which has come to be known as the Industrial Revolution. Beginning in England and spreading slowly to most of the world, the Industrial Revolution has had vast economic and political consequences. But forgetting these for the moment and concentrating only on the aspect of man and machine, this can be said: human and animal muscle power was replaced by steam and combustive power. Man was no longer needed as a power producer or transmitter, but he was by no means eliminated from the system. His role was simply changed to that of a machine minder, tender, or operator, and this is a role which man has maintained on the industrial front until very recently.

Now we find ourselves in the midst of another Industrial Revolution, this one fully as important as the first. This revolution began in the United States around 1930, primarily in the chemical and petroleum industry. The prime mover this time was the automatic control, or self-regulating device which allows complex machines to "run themselves." What of man's role now that he is no longer needed as a machine tender? The answer is once again that he is not, as many have feared, removed from the system altogether, but his role is changed, this time to that of a monitor. Now the operator's role is neither to supply power nor to regulate its use, but to act as an overseer, monitor, or standby controller to take over in case of system breakdown. One may ask why we need a monitor when the system is adequately controlled by an automatic device. The answer is that the human brings into the industrial or military scene those remarkable human attributes that no machine possesses. He is on hand to take over when the automatic device fails (as in the case of the auto-pilot already mentioned), to make decisions when rare, unusual circumstances dictate, or hopefully to detect and remedy a situation before trouble occurs. A paper by Hick (ref. 19) supports this view. Hick listed three functions for which the human is particularly well suited (compared to a machine or computer), and these three apply especially to the monitoring situation. These functions or subtasks are:

1. To compensate for failures
2. To take emergency action
3. To anticipate future states of the system

Approved for Public Release

This brings up the inevitable question of the optimal allocation of system-monitoring functions to the man or the machine. Many monitoring functions can be performed by either man or machine, leaving the system designer to decide in any given case which can perform the control function most efficiently. In certain cases, the function will be shared, as mentioned previously, with the machine providing the "minute-by-minute" control needed to keep the system in steady-state, and the human standing by to spot gross deviations or trends, preferably before they exceed tolerance limits.

Such a case is the modern chemical or petroleum plant which is highly automatic, but which also requires constant monitoring by the human. Modern military systems are also rapidly moving in this direction. In manned aircraft and especially in space vehicles, the operator is relieved of his duties as a "stick and rudder man" and assigned more passive functions as a monitor and decision-maker.

In order for the systems designer to make rational decisions on assignment of monitoring subtasks to man and machine, he must have dependable information on the performance characteristics of the human in this role.

There is every evidence at the present that the monitoring role is one which man finds difficult to accept. The task of "waiting for nothing to happen," as Mackworth (ref. 32) has put it, is disconcertingly passive, monotonous, and unrewarding. While no one would argue the importance of vigilance, it seems to be an attribute which is unrewarded in a society which prides itself on "action." As monotonous as a routine machine-tending task might seem, it is probably infinitely more interesting than simply watching for an event which the monitor may see only once a week, once a year, or once in a lifetime on the job. One may think of the example of the Ballistic Missile Early Warning System (BMEWS) operator who spends hours every day watching his scope for a possible rocket launch.

Another example of the monitoring task, one of unparalleled importance, is that of control of disarmament agreements. In the event that the nuclear powers of the world can some day agree to a test ban, the success of this agreement will depend to a large degree on a positive means of detecting surreptitious tests and violations of the pact. This will require the continuous monitoring of various radiological and seismic-sensing devices. The monitoring task will be especially difficult as any violator will attempt to conceal his detonation, or mask its report with noise.

Approved for Public Release

Contrails

Development of the Research Problem

The term vigilance, which is still preferred by most workers in the field, was first used by Head in 1926 (ref. 18) to describe a physiological or psychological readiness to respond. Previously the term attention had been in favor. The term vigilance is generally used as an intervening variable between signal and response. Operationally, vigilance may be defined as the probability of response to a low-probability, near-threshold event occurring in the monitor's environment. Though the monitoring task (or subtask) has never been adequately defined, it would seem wise to define it in similar terms:

> The monitoring task is that of maintaining a watch for low-probability, near-threshold, or highly noise-cluttered events. Monitoring may be the sole task of the operator, or it may be a subtask time shared with other related activities *

A practical interest in vigilance first arose in the field of quality control in the British munitions industry. In 1932 Wyatt and Langdon published a report (ref. 41) dealing with an investigation of visual inspection of cartridge casings. They reported that the probability of detection of flaws dropped markedly during a 4-hour work period, with the most rapid drop occurring during the first 30 to 45 minutes. This phenomenon, the decline of probability of detection over time, has come to be known as "vigilance decrement" and has been the focal point of investigations in the area.

Interest in the vigilance problem remained dormant until World War II, at least as far as the literature reflects. The exigencies of war, a war fought largely on "sensory margin" as S. S. Stevens has pointed out (ref. 40), revived the investigation into man's ability to serve as a monitor.

The problem was first encountered over the English Channel in 1943 when the RAF Coastal Command, flying long anti-submarine patrols with airborne radar, noted a phenomenon similar to that reported by Wyatt and Langdon. First, the detection rate of the radar operators was rather low. Frequently, other crew members who happened to pass the radar operator's station pointed out

---

*In this sense almost any man-machine operator task may contain a monitoring subtask. Take for example driving a car, loading bombs into a plane, running a loom, or performing surgery.

Approved for Public Release

targets on the scope that had eluded his attention. It appeared that the man in the best position to detect the targets was the least likely to do so. Further, it was found that a vigilance decrement appeared to be operating, as the occurrence of non-detections increased over the length of watch.

At this point the RAF sought the aid of the Applied Psychology Unit of the Medical Research Council at Cambridge in an effort to solve the operational problem of how long an operator should be left on duty before a radar scope. These studies, conducted by Dr. N. H. Mackworth, mark the beginning of a vast research effort dealing with man's ability to serve as a monitor.

The question of optimal watch time is still unsolved. Research by Jerison and Wallis (refs. 21, 22) indicates that the greatest part of the vigilance decrement occurs even faster than suspected, possibly during the first 10 to 15 minutes. Thus, the practicality of relieving the operator of his duties before his vigilance declines seems quite doubtful, and other methods of improving performance must be resorted to.

The research of Mackworth (refs. 30, 31, and 32) confirmed the RAF's findings of a rapid decrement. Figure 1 shows the typical Mackworth function, which has been subsequently confirmed by many workers in the field.

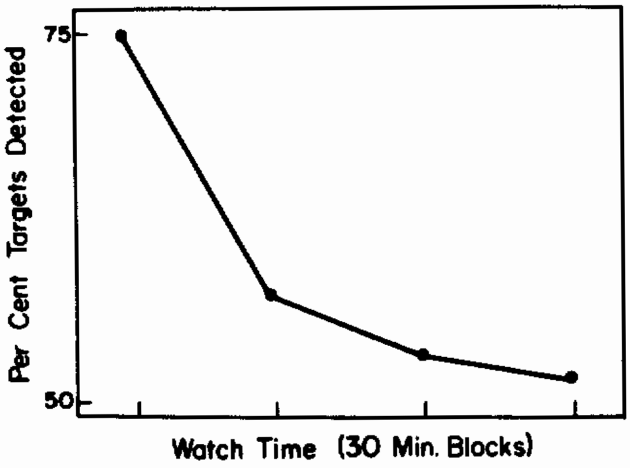

Figure 1. Percentage of Detected Signals as a Function of Time

Judging from figure 1, the systems designer has two problems

Approved for Public Release

Contrails

to be alarmed about: (1) the rapid decay in vigilance and (2) an initial level that is quite unsatisfactory. The actual figures for probability of detection and the curvature of the decremental function vary considerably from one experimental situation to another. But generally the empirical evidence gathered in the field of vigilance has supported Mackworth's early findings, though some have found various means of preventing vigilance decrement. However, it is a safe generalization that, where vigilance has remained at some initial level, this level is still unacceptably low. The vigilance problem is not simply one of preventing the decrement, but also one of elevating the entire response function.

Thus, we see that there are three important lines of thought which converge to emphasize the importance of studying man as a monitor:

1. An increase in the reliance on automatic control devices

2. A general increase in system complexity, accompanied by increasing costs for mistakes and oversights

3. A belief, supported by experimental and real-world evidence, that man is a poor monitor

SIGNAL RATE AND ARTIFICIAL SIGNALS

In the previous section a monitoring or vigilance problem was defined in part as one of guarding against a low-probability event, although no actual numerical limits on this have ever been specified. The early field observations and laboratory work indicated that it was this very aspect of signal infrequency that led to the poor performance of the monitor. The question for research was to determine just what effect probability of signal occurrence exerted on probability of detection.

Signal Rate

Much of the empirical work in vigilance has been directed toward the question of the effect of signal rate, and there has been an unusual amount of agreement in the findings. Six such studies are summarized in table I.

Approved for Public Release

Contrails

TABLE I

SIX EXPERIMENTS IN WHICH SIGNAL RATE WAS A MAIN VARIABLE. RATE IS EXPRESSED IN SIGNALS PER HOUR, REGARDLESS OF ACTUAL SESSION LENGTH.

	Authors	Display	Signal rate (per hour)	Findings
1.	Nicely and Miller (ref. 37)	Cathode ray tube	72, 12. The 72 rate was displayed in one quadrant, 12 in the remaining three.	Probability of detection significantly higher in 72/hr.quad. No decrement in this quad, decrement in others.
2.	Deese and Ormond (ref. 13)	Cathode ray tube	10. 20. 30. 40	Prob. detection increased with sig. rate.
3.	Jenkins (ref. 20)	Voltmeter	7,5, 30. 60. 480	Less decrement as rate increased. Prob. of detection roughly linear function of log. sig. rate.
4.	Kappauf and Powe (ref. 24)	Audio-Visual discrepancy checking	8, 20, 40, 80	Prob. of detection increased, decrement decreased with increasing sig.rate.
5.	Harabedian, McGrath, and Buckner (ref. 17)	Aural and visual sonar simulator	6, 30	Same
6.	Pollack and Knaff (ref. 38)	Light; meter	3 to 682	Same

These studies strongly indicate that vigilance performance depends to a great extent on the number of signals appearing during the vigil, at least within the range of signal rates presented here. It is worthwhile to note, in view of what has been said of the extremely low target rates (approaching zero in many cases) that may be expected in "real world" systems, that the lowest signal rate considered was 3/hour. If performance is relatively poor at the low end of the signal rate variable in these experiments, it is very alarming to extrapolate it down to the extremely low signal rates that we expect in these systems.

Artificial Signals

Just such reasoning and empirical evidence has led investigators in the field of vigilance to discuss the feasibility of inserting dummy or artificial signals into the monitor's display to elevate synthetically the signal rate and, hopefully, the probability of detection of the "real" signals.

Approved for Public Release

Contrails

Of all of the recommendations for increasing vigilance performance (including such things as drugs, punishment, monetary incentives, selection of monitors, and multi-man parallel circuits), artificial signals appear to be the most promising. The use of artificial signals presents serious practical problems. For example, how are the dummies to be introduced, and more important, how are they to be removed? How can the system be made "fail-safe" such that a commissive error cannot result from a dummy signal? What is the long-range effect of artificial signals on operators' motivation and task satisfaction?

As for the signals themselves, there is the question of whether or not they must be identical to the real signals to be effective. If they are identical, this is just a special case of the signal rate problem, and the monitor will not know, unless a feedback circuit is installed, whether he has detected a real or artificial signal. If the artificial signals are distinguishable from the real signal, there is the question of whether an increase in detection rate of the real signal will actually result. And finally, there is the question of the benefit of providing knowledge of results with these artificial signals, a question which will be explored in some detail in the next section.

Fortunately, there has been at least a beginning in investigating this intriguing matter of artificial signals.

Garvey, Taylor, and Newlin (ref. 16) investigated the effect of dummy signals, both identical and distinguishable, in a dial monitoring task. Using reaction time as a performance measure, the authors report that performance improved with the inclusion of either type of signal. While the evidence offered here, especially regarding the distinguishably different signals, is encouraging, it must be pointed out that the signals were fairly persistent, remaining on the display for 30 seconds unless previously detected.

Baker (ref. 5) added what he termed artificial signals in a visual monitoring task, where the signal persisted for 0.6 second. However, the "artificial" signals were identical to the real signals, and this experiment lends little support to the notion of inserting artificial signals beyond the evidence already presented in table I.

Thus, the empirical evidence to date indicates strongly that probability of detection of a signal is dependent on the rate of signal presentation, and that system performance may be enhanced by the insertion of dummy signals to artificially increase the apparent signal rate. The effect of nonidentical artificial signals on human performance, especially over a protracted period,

Approved for Public Release

Contrails

has not been investigated sufficiently at this time.

An interesting application of what actually amounts to artificial signals was suggested in a paper by Bailey (ref. 2). The author was concerned with a safety problem which has plagued aviation since the invention of the retractable landing gear: inadvertent gear-up landings. Many of these oversights escape the attention not only of the pilot, but the tower and runway control personnel as well, until it is too late. Bailey suggests that Navy operations sections schedule a number of deliberate gear-up approaches (not landings) at airfields in order to boost the probability of detection of such an event when it occurs inadvertently.

Again on the practical side, Merk (ref. 35) suggests including artificial signals in radar systems, with a senior monitor to receive and filter the responses, and furnish knowledge of results. However, neither Bailey nor Merk furnish empirical evidence to support their recommendations.

The actual manner in which artificial signals can be used, either in a training device or "on-the-job," brings up the central question of this dissertation: the effect of knowledge of results on the monitor's performance. The following section will explore the general question of knowledge of results in monitoring tasks, and present the experimental results to date.

KNOWLEDGE OF RESULTS

The topic of knowledge of results, or information feedback, has attracted considerable interest recently in engineering psychology. A large research effort, particularly in the field of motor skills, has gone into assessing the value of furnishing the operator information on the level or accuracy of his performance.

Knowledge of results is generally viewed as a feedback signal, either closing an open loop (in the case where the operator has no information about his performance), or augmenting the information already available in a closed loop. Most studies in the area have demonstrated that there can be little or no improvement in performance without some form of information feedback. In studies where supplemental feedback has been added to a closed loop, this information has resulted in a gain in performance.

While these results may be intuitively obvious in

Approved for Public Release

Contrails

acquisition of motor skills and performance of active tasks, the contribution of such feedback in vigilance tasks is somewhat less clear, due to the very nature of the task. More will be said of this shortly.

The Nature of Knowledge of Results

Any attempt to partial out the motivational and informational components of knowledge of results would be as fruitless as the long-fought controversy of the separate roles of heredity and environment. The only reasonable course is to recognize that both exist, and attempt to evaluate the role of knowledge of results in performance, whatever the task.

To date there have been almost as many attempts to classify knowledge of results as there have been authors in the field. A recent dissertation by Kinkade (ref. 25) reviews these, and summarizes the various classifications. Some of the attempts have divided knowledge of results into motivation-information (Brown, ref. 9), intrinsic-extrinsic (Annett and Kay, ref. 1), action-learning (Miller, ref. 36), as well as supplemental, augmented, psychological, fundamental, and high-achievement feedback.

Kinkade's own three-way classification seems to be satisfactory, and it recognizes differential levels of motivation and information components in each category. These categories are:

1. Fundamental feedback. The feedback is the only knowledge of performance furnished to the subject: he is unable to view the outcome of his action, as, for example, in long-range artillary firing. This form of feedback has a rather large component of information, since the operator could not possibly improve his performance without this knowledge. To be sure, there is also a motivational element. In servo terms, this information closes an open loop.

2. Summary feedback. Quantitative supplemental information summarizing the operator's performance is given after a period of work. In this form, knowledge of results follows a series of responses, and gives the subject some information regarding his over-all performance, and not the effectiveness of any particular response. This form no doubt has a large motivational component, and furnishes only general information which can guide future responding.

3. Augmented feedback. Here supplemental information about the effectiveness of the operator's responses is fed back during responding. A classic example is the on-target auditory signal

Approved for Public Release

Contrails

that a subject receives on a pursuit rotor. In the case of augmented feedback, the subject is already working in a closed-loop task, and information about performance is available even without the augmented feedback circuit. The augmented feedback, in Kinkade's terminology, serves as an "instructor surrogate" providing both information and commendation for high achievement.

Knowledge of Results in Tracking and Monitoring Tasks

Let us look at two types of tasks. In the tracking task the operator is continuously active, and immediate knowledge of results contains error information, aiding him in minimizing error, or guiding him to the target. In monitoring, the response is discrete, intermittent, and essentially binary: the subject either responds or does not respond. Thus, knowledge of results comes (as always) after the fact, but can hardly be thought of as providing any guidance toward a correct response to the next signal.

That the two types of tasks are fundamentally different is evidenced by the effect of practice. In any task involving tracking or continuous motor performance, ability, as measured by almost any performance measure, increases with practice until it becomes asymptotic. In monitoring tasks, performance not only declines during a single watch as already pointed out, but also shows no improvement over several days of practice. Vigilance data in no way resemble data found in other experiments, and therefore we might not expect knowledge of results to affect performance in the same manner.

The foregoing would indicate that the informational component of knowledge of results in vigilance task is fairly low, although at least one theory (Baker, refs. 3, 6, and 7) would take exception to this. Baker's theoretical and empirical work on knowledge of results will be reviewed later.

But even before attacking the question of just what it is about knowledge of results that affects vigilance performance, there is a more pressing need for empirical evidence to indicate what the effect is, and how knowledge of results may be employed.

Methodological Problems

The discussion of knowledge of results in monitoring has wandered considerably from the question that brought it up, signal

Approved for Public Release

rates and artificial signals. Now an attempt will be made to link these together.

When one discusses the matter of knowledge of results feedback in monitoring tasks he must face a question of real-world relevance. Specifically, he must answer the obvious criticism that, in most systems, knowledge of results is little but an academic issue. If knowledge of results can be supplied, then some person or device must have cognizance of the signal and the correctness of the monitor's response or failure to respond, so why is the monitor needed?

The real-world relevance of knowledge of results in monitoring is twofold. First, we return to the matter of artificial signals. If dummy signals can be programmed into the monitor's display, a very simple feedback circuit to furnish knowledge of results to these signals can easily be added. We then have the problem of whether or not to supply knowledge of results about artificial signals when they are employed in conjunction with the monitoring task.

Second, there is the most neglected area of vigilance research: training. Although much has been written about identification and selection of good monitors, little has been said to date about how a monitor is to be trained for his job once selected. This is perhaps due to the fact just mentioned that monitoring studies have shown no learning effect from session to session. But this very fact, rather than disparaging a training approach, accentuates the need for finding methods of improving monitoring performance beyond the initial level which the operator brings into the situation. Here again the question of knowledge of results enters the picture. Specifically, it would seem wise to know whether or not training under conditions where knowledge of results is furnished has any persistent effect on performance when feedback is no longer available. Put in another way, is there any advantage to initial training with knowledge of results if the monitor must later be transferred to a situation where he must perform his task without such knowledge? In the future, training of monitors may be performed with system simulators, into which it would be a simple matter to program any given signal distribution as well as knowledge of results. For this reason alone it is considerably more than an academic question to assess the effect of knowledge of results on the acquisition of monitoring skill, and even more important, on the maintenance and retention of this skill after initial training. If feedback of this information proves to be a worthwhile training aid, further research must be directed toward finding the most effective method by which the monitor can be "weaned away" from his training device and transferred to the real system.

Approved for Public Release

Contrails

At the time of this writing the author knows of no experimenter who has examined this question of removing knowledge of results in a vigilance task. The experiment reported by McCormack (ref. 33) could have shed light on the transfer problem, but the necessary comparison was not included in his analysis of variance.*

A Brief Review of the Literature

The empirical work to date in knowledge of results in monitoring indicates a beneficial effect, especially in preventing time decrement. In Mackworth's study (ref. 32) subjects were run under his standard conditions (two hours monitoring a Mackworth clock), with one group receiving verbal knowledge of results for both responses and omissive errors. The no-knowledge group showed the typical Mackworth decrement function, while the knowledge of results group remained at the initial level throughout the watch. McGrath, Harabedian, and Buckner (ref. 34) state that this experiment did not clearly make a case for knowledge of results, since the communication was verbal. They remark (p. 42): "The findings may be confounded with motivational variables associated with personal communication." Though knowledge of results is always confounded with motivation, their objection is especially appropriate in vigilance studies, since, as they state, the verbal feedback indicates to the subject that his performance is being continuously observed by the experimenter. In the light of Fraser's finding (ref. 15) that the mere presence of the experimenter in the room will prevent decrement, it seems advisable to isolate social and motivational factors from knowledge of results to the greatest possible degree by presenting the knowledge of results by nonverbal means, keeping the experimenter as remote from the subject as experimental conditions will allow.

An interesting, but perhaps somewhat limited, theory of knowledge of results has emerged from Baker's studies of stimulus regularity (refs. 4, 6). This view is best explained by quoting one of Baker's theoretical papers (ref. 3):

> When knowledge of results is given during a vigilance task, i.e., informing the observer when signals are correctly detected and when

---

* With a tracking task, Kinkade (ref. 25) has shown a persistent superiority of feedback groups even after removal of the information feedback.

Approved for Public Release

Contrails

> they are missed, the general level of performance will be higher than when such knowledge is not given. Further, knowledge of results can be discontinued after the observer has learned the temporal nature of the series, with little decrement in performance. This prediction is based on the proposition that knowledge of results serves to establish perception of the true sequential nature of the series, i.e., to increase the probability of expectancy confirmation. Once established the procedure can be dropped.

In a recent experiment, Baker (ref 4, Experiment 3) ran three groups: (1) no knowledge, (2) knowledge of results (commissive errors, omissive errors, and detections), and (3) feedback (no knowledge, but signals repeated at 5-second intervals until detected). The knowledge of results was furnished nonverbally via a visual display. The results showed a significant decrement for the no-knowledge group as expected, and no decrement for the experimental groups.

In a recent study mentioned in the previous section, Baker (ref. 5) demonstrated the effect of combining artificial signals with knowledge of results. The task involved detecting a 2-millimeter dot of light in a 4-inch square screen; the signal persisted for 0.6 second. Under the control condition, subjects ran for 1-1/2 hours with a total of 36 signals per watch period. The experimental group had the same schedule of "real" signals, plus identical "artificial" signals inserted between the real ones, bringing the total up to 58-63 signals per watch. No knowledge of results for either real or artificial signals was presented to the control group. The experimental group received verbal knowledge of results ("correct," "missed one," or "false") for detections, omissive and commissive errors on the "artificial" signals, and for correct responses only on the real signals.

The results showed that the experimental group detected significantly more real signals than the control, and had a shallower decrement function. In this study knowledge of results is confounded with signal rate, which in itself has been shown to provide these experimental results. As Baker states (ref. 5, p. 337),

> The purpose of this study was not to determine the relative merits of increased signal frequency using artificial signals, versus knowledge of results in maintaining the level of vigilance, but rather to demonstrate that when both factors are employed monitoring

Approved for Public Release

Contrails

performance is superior to the case when neither is employed.

Unfortunately, this is a weak case for knowledge of results, and provides the systems designer with little guidance.

An experiment by McCormack (ref. 33) showed that knowledge of results could result in superior performance even when the knowledge does not relate the temporal sequence of the events. In his study, the feedback took the form of notifying the subject whether his response time was longer or shorter than on the previous signal. The dependent variable was median response time, and since no mention is made of either omissive or commissive errors it is assumed that the signal was easily detected. Subjects were run for 50-minute watches on two successive days: one group received knowledge of results on the first day but not the second, and vice versa. Both groups showed an increase of response time over a single watch, but the increase was significantly greater for the no-knowledge group. The time periods, feedback conditions, and periods-by-conditions interactions were all significant.

As mentioned previously, McCormack unfortunately combined his knowledge and no-knowledge conditions irrespective of their order of presentation, so that no information was available on transfer effects.

McCormack's finding is further supported in a study by Loeb and Schmidt (ref. 27) in which the subjects were given verbal "pseudo-feedback" following each response to an auditory signal. The subject was informed after each response whether the response time was shorter or longer than on the previous signal, but the information was random and not related to actual performance. Under one of the two intensity conditions, the pseudo-feedback group showed significantly shorter reaction times and a higher detection rate. This study showed that the feedback which contains false information could still produce a beneficial effect, indicating that factors other than information content were responsible. In a followup study (ref. 28) these authors essentially repeated the earlier experiment, this time adding a true feedback condition in which subjects were correctly informed of their response time relative to the previous response time. Another experimental group received after each response an indication simply acknowledging the response--the indication carried no information about the subject's performance. A control group received no information or acknowledgment. Both the true and pseudo-feedback groups showed less performance decrement (in terms of median response time) than the control and acknowledge groups. Further, the group

Approved for Public Release

with true knowledge of results was superior to the pseudo-feedback group.

Pollack and Knaff (ref. 39) investigated the effect of a loud auditory signal in maintaining alertness to visual signals. The subjects were required to report large excursions of a fluctuating needle; the signal rate was 48/hour, and a watch session was 80 minutes. Three experimental conditions were employed: neutral, reward (monetary), and punishment. The punishment was a half-second-115-db. blast of a truck horn 18 inches in front of the subject each time a signal was missed. The reward was provided at the end of a session, and thus did not provide knowledge of results until after the experiment was completed, but the punishment condition can be considered immediate feedback. In terms of percentage detection, the experimental conditions differed significantly, with the punishment group showing superior performance. There was a significant time decrement for each group. Although the authors did not include the conditions-by-time interaction in their list of significant results, it appears from the graph presented that the punishment group showed an extremely small decrement.

This experiment constitutes a somewhat special case of knowledge of results due to the high energy level of the information feedback. No doubt the same information could have been provided to the subjects in a less unpleasant manner. As the authors report (p. 1014):

> Gross qualitative changes in the observer's behavior to the truck horn were noted. The observers sat straight on the edge of their chairs, (they shook themselves to maintain alertness); they urged their neighbors to be 'on the ball' so that the group would not be exposed to the acoustic leak-over of the individual horns; and reaction times were apparently greatly shortened.

The authors' use of the term "punishment" rather than "information" supports this view. Furthermore, under the punishment condition there was a slight increase in commissive errors. As the authors point out, this increase was not great enough to account for the drop in omissive errors shown by this group.

The studies of Garvey, Taylor, and Newlin (ref. 16), already discussed in connection with artificial signals, examined the effect of two types of nonverbal knowledge of results on response latency. Subjects were furnished knowledge of results of two types. The first type was what the authors called

Contrails

"feedback of a purely informative nature," supplied by a red light under the appropriate meter if the subject failed to respond within 3.2 seconds after an artificial signal appeared. Pushing the response button would turn off the light and reset the pointer to the "safe" position. The other type of feedback was called "punishment," the addition of a 120-db. white noise tone through a speaker near the subject's head during the time that the red light was on.* The nature of the task supplied some feedback, for correct responses re-centered the needle.

The results showed a slight but significant improvement (in terms of median response time) when the light was presented, and a similarly small but significant improvement over this condition with the addition of the tone. However, the authors state that the performance gain due to the feedback of results was meager compared to the gain associated with employment of artificial signals. They summarize (p. 7), "Weighing equipment complexity against the expected benefits to human performance, it is concluded that although the insertion of artificial signals might be worthwhile in a practical situation, the inclusion of informational and motivational feedback circuits might not be worthwhile."

In summary, the evidence presented demonstrates that knowledge of results favorably affects performance in a vigilance task. But the empirical work to date should only rouse mild enthusiasm about the use of knowledge of results. The case for this form of feedback has been weakened by the fact that in most experiments it has been presented in such a way as to be confounded with other effects such as social interaction between subject and experimenter, and signal rate.

Focusing primarily on the informational aspects of knowledge of results, the strongest case can be made under the following conditions:

1. Nonverbal presentation
2. Presentation at an energy level sufficient to be well above threshold, but not high enough to be punishing or noxious
3. If artificial signals are employed with knowledge of results, an equal number should be presented to the

---

* Cf. Pollack and Knaff's punishment condition (ref. 39).

Approved for Public Release

Contrails

control (no-knowledge) group to keep signal rate balanced.

A TRANSFER OF TRAINING EXPERIMENT

Introduction

The previous sections have outlined the present "state of the art" in the monitoring field with respect to two factors, signal rate and knowledge of results. This experiment was conducted (1) to re-examine the effects of these two variables during initial exposure to a standard vigilance task and (2) to determine the transfer effects of these variables on a second exposure to the task when all subjects are treated uniformly. Specifically, the purpose of the experiment was to determine whether knowledge of results furnished during one session would continue to result in superior performance once this form of feedback was withdrawn.

Method

Apparatus: The task was the Jerison adaptation (ref. 22) of the standard Mackworth clock task (ref. 29). The display was a Standard Laboratory Timer with a 1 rps sweep hand, the instrument face being reversed so that behind the sweephand was a uniform white surface. Each clock was placed in a wooden box with a plexiglas window, and shock-mounted inside the box with foam rubber. This display is illustrated in figure 2.

The movement of the sweep hand was controlled by a series of relays and timers, so that a normal pulse delivered to the clock advanced the hand 1/32 of a revolution, or 11.25 degrees once each second. The signal for which the subject was monitoring was a jump of 1/18 of a revolution, or 20 degrees. For convenience, this will be referred to as a "double jump," though it is actually only 1.78 times the normal jump. This ratio was determined experimentally in a preliminary study. The customary ratio of 2.00 was found to be too easy a task even without knowledge of results.

The occurrence of a double jump was determined by a schedule punched into standard 5-channel teletype tape, stepped through a Western Union Model 24-B tape reader. The apparatus allowed three subjects to be run at once independently; all clocks received signals at the same time.

Approved for Public Release

Contrails

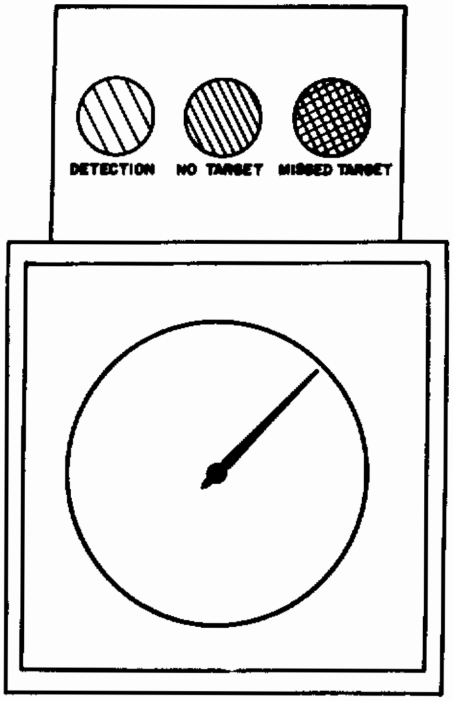

Figure 2. Mackworth Clock and Knowledge of Results Display

The knowledge of results was presented automatically by means of a device designed specifically for the experiment. The occurrence of a signal started a timer in the knowledge of results apparatus. If a response was made during the allotted time (3.5 seconds), the left light (figure 2) remained illuminated (green) as long as the subject held his switch closed. If a response was made when the timer was not timing out the 3.5 seconds, it was a commissive error and the center light illuminated (red) as long as the switch was held. If no response were made to a signal during the allotted time, at the end of the interval the right-hand light illuminated (amber) for 5 seconds. This interval was regulated by a time-delay relay. Information feedback for each subject was independent of that of the other subjects. The knowledge of results display was mounted in a green metal box above the clock box, as shown in figure 2.

Each signal and each subject's responses were automatically recorded on an Esterline-Angus-20-pen model AW recorder. One pen was assigned to the signal, and one to each subject.

The apparatus was fully automatic, in that the single and

Approved for Public Release

double jumps were programmed by the teletype, and the signals and responses were recorded on Esterline-Angus chart paper. The experimenter played only a monitoring role once the experimental session was under way.

Each subject sat in a booth at a desk which held the clock and knowledge of results display. The subject responded with a hand-held Packard-Bell SW-141-K silent squeeze-type switch. The subjects were isolated from each other by the walls of the booth, the three booths being separated from the rest of the room by a black curtain. Auditory isolation was achieved by having the subjects wear earphones playing white noise at 75 db, as measured at the earpiece by a Hermon Hosmer Scott Type 410 sound-level meter. Subjects were not able to detect the occurrence of a signal or a response from other subjects from apparatus sounds. The master control which pulsed the clocks and the knowledge of results apparatus were placed in soundproof celotex boxes.

Signal Schedules: Three signal schedules were used in this experiment: 16, 32, and 48 signals per 48-minute session. The schedules were made up by determining the intervals between signals with a table of random numbers from a uniform distribution. The following restrictions were imposed on the random assignments:

1. The signals were equally assigned to four 12-minute blocks within the 48-minute session. For example, in the 16-signal schedule, there were 4 signals each 12 minutes.
2. No inter-signal interval was shorter than 0.3 minute.
3. No signal occurred during the first minute of a session.

The actual signal schedules are shown in Appendix II. The 32-signal schedule used for all subjects on the second day was the first-day 32 schedule run in reverse order.

Subjects: Subjects were 96 female undergraduate students at The Ohio State University. They were recruited for the experiment by various means, including newspaper ads and posters. None had previously served in any type of vigilance experiment. Subjects were paid two dollars for their participation. The data from six subjects were not used in the analysis for the following reasons:

1. Apparatus failed on two runs (3 subjects).
2. Subject was observed reading a book during the

Approved for Public Release

second session.

3. Subject removed earphones and went to sleep during second session.

4. Subject made over 450 commissive errors during the first session. After the session she volunteered the information that she felt she did not understand the instructions.

Experimental Design: The experimental design for each session involved 3 signal rates, 3 knowledge of results conditions, and four 12-minute time blocks, with 10 subjects nested in each of the 9 signal rate-knowledge of results combination groups, running under all 4 time blocks, for a total of 90 subjects. The design is familiar to many as a Lindquist Type III (ref. 26, p. 281), illustrated in figure 3. In Lindquist's terms, signal rate and knowledge of results conditions and their interaction were "between subjects" effects and all comparisons involving the time dimension were "within subjects" effects.

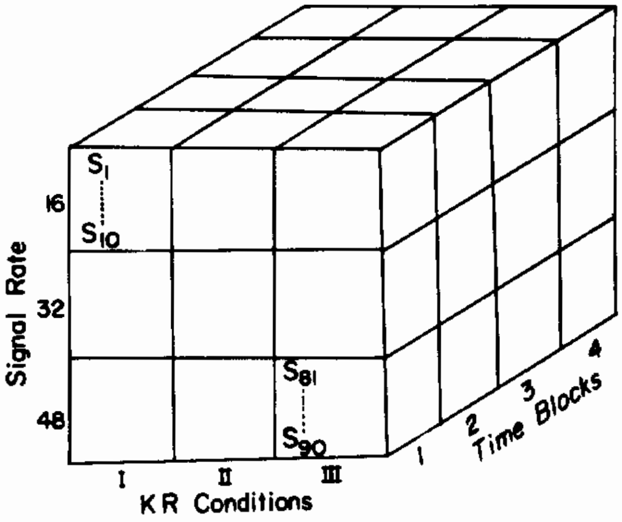

Figure 3. Experimental Design

The three knowledge of results conditions are identified by Roman numerals. They are:

I. No knowledge of results

Approved for Public Release

Contrails

II. Knowledge furnished only upon the subject's response (i.e., correct detections and commissive errors).

III. Correct detections, commissive errors, and omissive errors.

The full knowledge of results condition (III) was included as an ideal which could be attained in a training device. The partial knowledge of results condition (II) stimulates many real-world systems where immediate or slightly delayed knowledge of results can be obtained when a response is made, but knowledge of a missed signal may be impossible to obtain, or at best highly delayed.

On the second day, all subjects were run under the no-knowledge of results condition with 32 signals. This experimental condition is identical to that of Group I-32 on Day 1, except that the actual signal schedule was reversed. For purposes of analysis of the transfer effect, subjects on Day 2 were identified by their Day 1 experimental group, and an identical analysis of variance was performed on the data from Day 2.

Procedure: In scheduling subjects for the experiment, every attempt was made to run three at a time; however, many doubles and singles were also run. Subjects were randomly assigned to the 9 groups by means of a random number table, with the restriction on complete randomness that the groups were kept roughly balanced in size as the experiment progressed.

Subjects were given standard instructions appropriate to their group (see Appendix I). Groups run under knowledge of results Condition I were told nothing about knowledge of results, and the display was not present. On the second day the knowledge of results display was removed and groups run under knowledge of results Conditions II and III were simply told that the session would be as before, with the exception that no knowledge of results would be presented. At no time on either day was signal rate discussed, except that all groups were told that the signals would be very infrequent.

Watches were removed from the subjects before each session so that they could not keep time. They knew only that the session would last about an hour. Following the second session, the experiment was fully explained to the subjects and all questions answered, with the exception that the actual signal rates were never revealed. This was to avoid having the information available to the subject prior to his service in the study.

As the instructions in Appendix I reveal, the subject

Approved for Public Release

was cautioned to be alert for all double jumps, and at the same time was warned that "false alarms" would count against her. No attempt was made to set relative weights on the importance of these errors, but only to impress upon the subject the great importance of both. No subject asked about relative weights or penalties for "guessing."

The subjects were given a certain amount of familiarization with the task, as indicated by the instructions. In this familiarization, which amounted to about 10 signals appearing in rapid order, the subjects sat together in a booth. Both their responses and those of the experimenter furnished some knowledge of results. So it is not entirely accurate to say that subjects in Condition I had received no training with knowledge of results. However, this small amount of feedback is necessary for the conduct of any experiment where the task is unfamiliar to the subject. Subjects were warned that the signals occurring during their actual session would be far less frequent than during the instructional period.

Results

The raw data analyzed in this experiment were the percentage of detected signals* and the number of commissive errors, during each 12-minute time block.

Detected Signals: The raw data in terms of percentage signals detected did not meet the assumptions of analysis of variance, and were consequently transformed to radians by the arcsine transformation (refs. 8, 23). This transformation took the form:

$$Y = 2 \sin^{-1} \sqrt{X}, \text{ where } 0 \leq X \leq 1.0$$

$$0 \leq Y \leq \pi$$

The analyses of variance were run on an IBM 704 computer making possible comparisons of significance levels before and after the application of the arcsine transformation. All conclusions reported will be based on the transformed

---

* For purposes of analysis, this performance measure was used rather than the negative form, omissive errors.

Approved for Public Release

Contrails

data.*

All figures in this section are reported in terms of raw percentage detections.

Figure 4 and figure 5 show the percentage detection as a function of time on Day 1 and Day 2. These separate figures report the same data: in figure 4 the parameter of the curves is the three knowledge of results conditions averaged across the three signal rates. Figure 5 is just the opposite. This method of presentation was preferred over the cumbersome alternative of presenting all nine groups separately for each day. Figure 6 shows the over-all detection rate as a function of time for each day, summed over the nine groups. The means of each group and time block are tabulated in Appendix III.

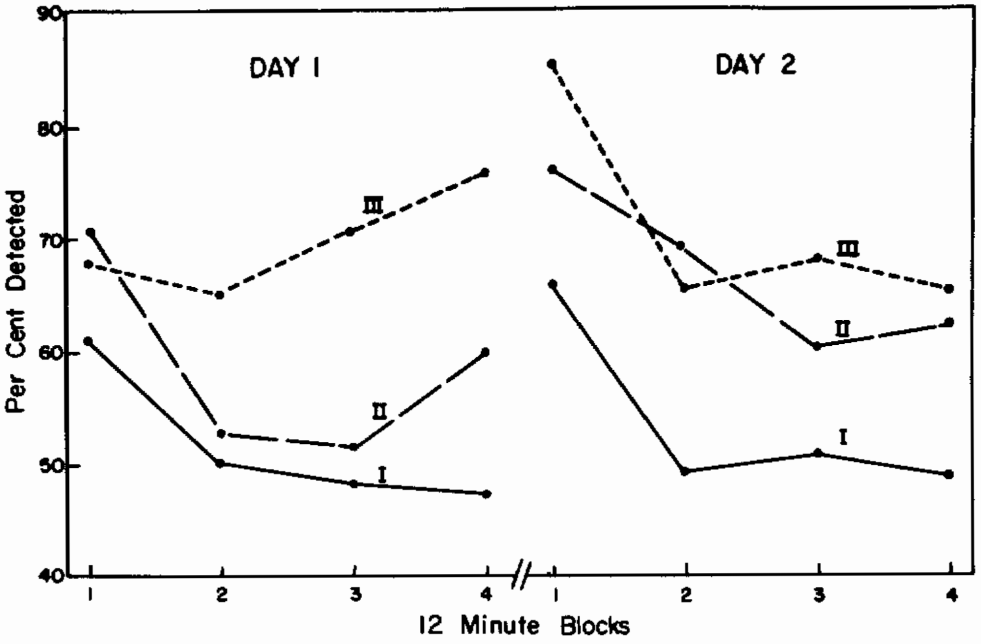

Figure 4. Per Cent Detection as a Function of Length of Watch. Knowledge of Results Groups Are Averaged Over the Three Signal Rates.

---

* In only one comparison, a second-order interaction, did the conclusions from the transformed data fail to agree with the raw percentage data. In this case an F test previously significant at the .05 level dropped slightly below the critical F for .05 significance. This information is included only as an incidental fact, giving further testimony to much recently reported evidence of the "robustness" of parametric tests.

Approved for Public Release

Contrails

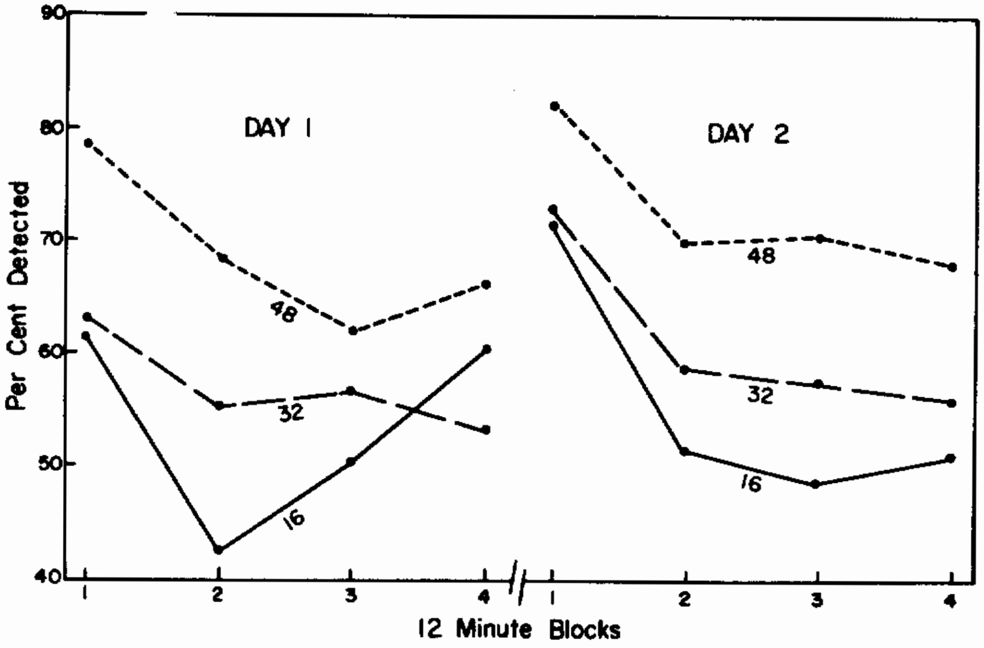

Figure 5. Per Cent Detection as a Function of Length of Watch. Signal Rate Groups Are Averaged Over the Three Knowledge of Results Conditions.

The analysis of variance for each day separately is presented in tables II and III. Probability levels greater than .05 are listed as nonsignificant. Table II indicates that on Day 1 all main effects (time, knowledge of results, and signal rate) were statistically significant as were both two-way interactions involving time. Table III indicates that on Day 2 only the three main effects were significant.

Although differences from day-to-day were not central to this investigation, a two-way analysis of variance was performed on the differences between days and time blocks, as shown in figure 6. For this analysis, the ten subjects in each knowledge of results signal rate group were averaged for each time block. Thus, nine scores were available for each time period of each day for the analysis presented in table IV. The time blocks were significantly different, but the difference between days and the day-by-block interaction were not.

To explore further the decremental function, a Duncan Multiple Range Test (ref. 14) was performed on the means for each day (figure 6). On both days, the mean of the first time

Approved for Public Release

Contrails

block was significantly different from all others, and none of the remaining three time block means produced significant differences.* This further confirms the view of Jerison and Wallis (refs. 21, 22), that vigilance decrement occurs early in the watch period.

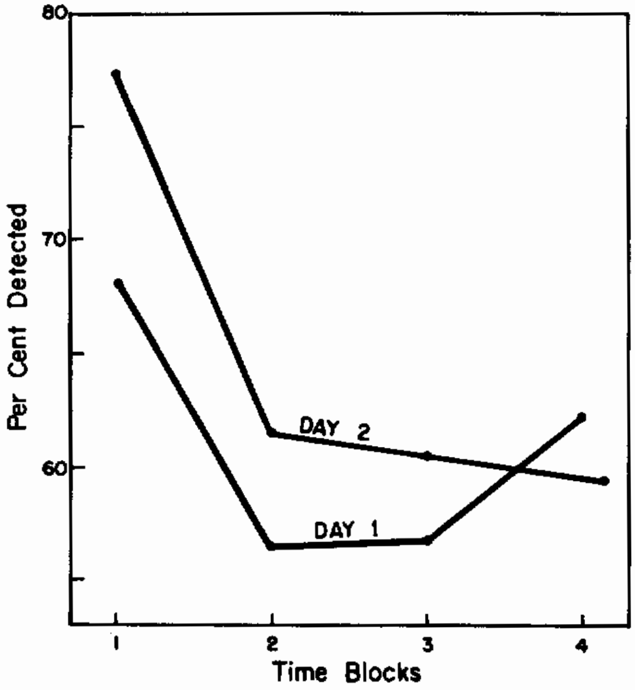

Figure 6. Per Cent Detection as a Function of Length of Watch, All Groups Combined

---

* All these differences were significant at the .01 level except Day 1, Block 1 vs. Block 4, which was significant at the 0.5 level.

Approved for Public Release

Contrails

TABLE II

ANALYSIS OF VARIANCE OF PERCENTAGE OF DETECTIONS ON DAY 1

SOURCE	df	MS	F	p
Between Subjects	89			
Feedback (F)	2	6.54	6.14	< .01
Signal Rate (S)	2	4.74	4.44	< .025
FXS	4	1.49	1.40	NS
Error (between)	81	1.07		
Within Subjects	270			
Time Periods (T)	3	1.77	5.70	<.001
TXF	6	.91	2.94	<.01
TXS	6	.77	2.48	<.025
TXFXS	12	.20	-	
Error (within)	243	.31		
Total	359			

TABLE III

ANALYSIS OF VARIANCE OF PERCENTAGE OF DETECTIONS ON DAY 2

SOURCE	df	MS	F	p
Between Subjects	89			
Feedback (F)	2	7.35	7.83	< .001
Signal Rate (S)	2	6.25	6.65	< .005
FXS	4	2.19	2.33	NS
Error (between)	81	.94		
Within Subjects	270			
Time Periods (T)	3	4.95	24.65	< .001
TXF	6	.38	1.88	NS
TXS	6	.09	-	
TXFXS	12	.13	-	
Error (within)	243	.20		
Total	359			

Approved for Public Release

Contrails

TABLE IV

ANALYSIS OF VARIANCE OF PERCENTAGE DETECTED SIGNALS AS A FUNCTION OF TIME, DAY 1 AND DAY 2

SOURCE	df	MS	F	p
Time Periods (T)	3	5833	4.59	< .01
Days (D)	1	2415	1.90	NS
DXT	3	880	-	
Residual	64	1272		
Total	71			

To determine the day-to-day reliability of the subjects' scores, the Pearson product-moment correlation of each subject's detection rate (averaged over the four time periods) between Day 1 and Day 2 was computed. This was found to be r = +.67, significant at the .001 level for n = 90.

Commissive Errors: A commissive error was defined as any response occurring more than 3.5 seconds after the previous signal. Thus, the subject was virtually unbounded in the number of commissive errors she could make. The commissive error data are shown in tables V and VI. The raw data revealed an extreme positive skewness. Summing across the four time blocks for each subject, it was found that the majority of subjects made 3 or fewer commissive errors, but the mean number of such errors on Day 1 was 12.9, due to a vast number made by a few subjects. The percentage of the total 1158 commissive errors on Day 1 attributed to the eight worst offenders is shown in figure 7. We see from this plot that 5 out of 90 subjects (about 6 per cent) accounted for 49 per cent of the commissive errors on the first day.

TABLE V

COMMISSIVE ERRORS BY SIGNAL RATE AND KNOWLEDGE OF RESULTS CONDITION, DAY 1

		KR CONDITION			
		I	II	III	Total
SIGNAL RATE	16	186	104	18	308
	32	108	267	41	416
	48	95	274	65	434
	Total	389	645	124	1158

Approved for Public Release

Contrails

TABLE VI

COMMISSIVE ERRORS BY SIGNAL RATE AND KNOWLEDGE OF RESULTS CONDITION, DAY 2

KR CONDITION

		I	II	III	Total
SIGNAL	16	54	39	9	102
RATE	32	44	99	31	174
	48	58	103	52	213
	Total	156	241	92	489

The extreme skewness of the data called for some special handling. First, an attempt was made to fit a Poisson distribution to the data, in the hopes of justifying a square-root transformation. This was abandoned after it was shown that the data were by no means Poisson distributed. Again, this was due to the fact that, while the mean was large, the scores were generally clustered near zero. The large value of the mean caused the expected values of the Poisson to be very small in the region near zero, resulting in significant deviations in a chi-square goodness-of-fit test.

Any attempt to perform an analysis of variance on these data was abandoned, and a nonparametric method was adopted. A preliminary Kruskal-Wallis test on the four time blocks revealed no significant differences due to time. Therefore, further analyses were conducted on the total number of commissive errors made by each subject during the 48-minute run, the 90 subjects being ranked with respect to this measure.

Since no acceptable two-way nonparametric test was available, two separate Kruskal-Wallis tests were performed on the data, and these summarized in table VII, one to test the signal rate effect and one to test the KR conditions.

Tables V and VI show that the number of commissive errors increased with signal rate on both days, but the difference was significant only on Day 2. The effect of knowledge of results conditions was significant on both days. This effect was interesting, in that condition II, partial knowledge of results, was accompanied by a large increase in commissive errors, while condition III, full knowledge of results, showed far fewer errors than either of the other conditions.

Tables giving commissive errors by experimental conditions in terms of sum of the ranks, rather than number of errors, are

Approved for Public Release

Contrails

included in Appendix IV. These tables give the same information provided in tables V and VI, but tend to damp out the effect of the few subjects who contributed an inordinate number of commissive errors.

The Day 1-Day 2 correlation of the subjects' performances was computed by Spearman's nonparametric method. This analysis yielded a rho of +.66, significant at the .001 level for n=90.

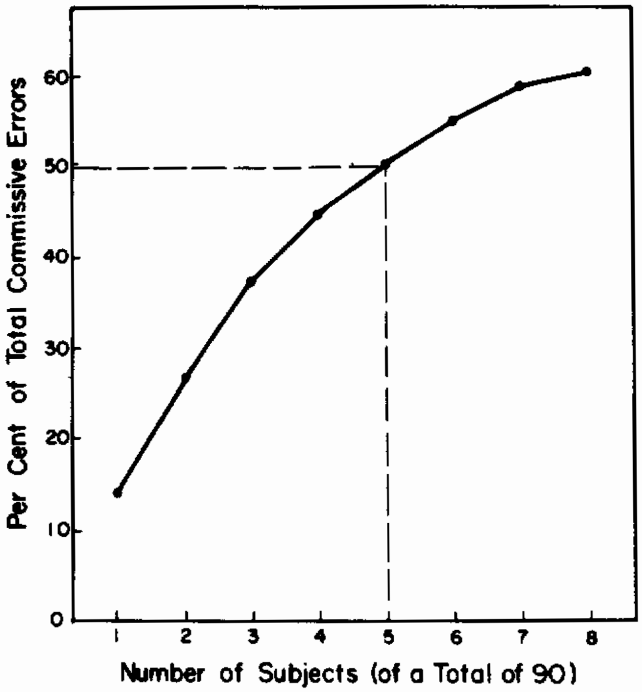

Figure 7. The Fraction of the Total 1158 Commissive Errors on Day 1 Contributed by 8 of 90 Subjects. The Dotted Line Illustrates that 5 Subjects (6 Per Cent of Total) Contributed Almost Half of the Commissive Errors.

Approved for Public Release

Contrails

TABLE VII

CHI-SQUARE VALUES OF THE KRUSKAL-WALLIS TESTS FOR COMMISSIVE ERRORS

	Day 1	Day 2	df
Signal Rate	1.8	7.0*	2
Knowledge of Results	12.4**	6.6*	2

$*p < .05$
$**p < .01$

Subsidiary Analyses: Two analyses were conducted on the data from Day 1 merely as a check on the experimental procedures. These were chi-square tests of goodness-of-fit to determine whether signal detection was affected by (1) the three stations used by the subjects or (2) the size of the groups that were run. In each case, the expected number of detected signals, based on the number of subjects run at each station or in each size group, was compared to the obtained frequency. Both analyses resulted in nonsignificant chi-squares, indicating no departure from internal consistency in the experimental procedure.*

Discussion

Detected Signals: The major results of this experiment support the use of knowledge of results as a training aid. The results of Day 1 support what has already been demonstrated by Mackworth, Baker, and others, namely that knowledge of results affects vigilance performance. But more interesting are the results from Day 2, which show that even when knowledge of results is withdrawn, the groups initially exposed to this feedback continue to perform at a level which is superior to that of the control group. This divergence between groups, as shown in figure 4, is encouraging from a training point of view. Whether this difference would persevere through long periods without knowledge of results is a question that only further research can answer.

---

* χ_2^2 = 2.28 for stations; χ_2^2 = 0.07 for group size.

Approved for Public Release

It is interesting to note the rather dramatic change in the relative position of groups first exposed to condition II, partial knowledge of results, from Day 1 to Day 2. On Day 1 those groups more closely resembled the control groups than the full knowledge of results groups. However, on Day 2, the knowledge of results groups tended to coincide, and were clearly different from the control. This may indicate that partial knowledge of results, which is furnished only when the operator responds, is just as effective as the full knowledge of results condition which requires furnishing information about missed signals.

The signal rate effect further supports the unanimous finding that higher signal rates lead to higher detection rates. The exception noted in these data is the rather peculiar behavior exhibited by the 16-signal curve on Day 1. The interesting point is again the transfer effect. When all groups were transferred to a 32-signal schedule for the second day, the groups trained with higher signal rates were clearly superior. Further, the effect of time is uniform over the signal rates on the second day. The marked time-by-signal rate interaction found on Day 1 was absent on the second day, when all three curves exhibited the "typical" vigilance decrement.

The data presented strongly suggest training monitors with schedules of high signal rate even when they are to be transferred to low signal rate environments. Just how rapid a signal rate is desirable for training should be the subject of further research, but the clear superiority of the 48-signal groups would suggest a large number of signals during initial training.

There are several results which are disappointing from a training point of view. First there is the absence on either day of a signal rate-by-knowledge of results interaction. Since high levels of both of these conditions were found to be beneficial, both during the initial training and the transfer session, one would expect that the effects would be interactive. Since the amount of feedback is dependent on the appearance of signals (with the possible exception of commissive error feedback), certainly it is reasonable to expect that higher signal rates would enhance the knowledge of results effect, but this was not supported by the data.

Some comment on the shape of the knowledge of results functions is necessary. The data presented in figure 4 show that the control groups behaved in the usual manner for a no-feedback condition, while the full-feedback condition (III) generaly displayed an increase in performance over initial level. Condition II showed a decremental function not unusual in vigilance work, where performance dropped markedly at first,

Approved for Public Release

then showed a decelerating drop, ending with a climb in performance. Unfortunately the number of subjects is too small to take any single period-to-period change too seriously. But there are indications that the knowledge of results conditions vary considerably in their time effects on Day 1 (as supported by the significant time-knowledge of results interaction). These differences are not present on Day 2, indicating that it is a general level of vigilance, not a resistance to time decrement, that transfers.

As the full-knowledge of results group improves its performance, detecting more signals, it denies itself the missed-signal feedback which distinguishes it from condition II. That is to say, missed signal feedback, if effective, is self-eliminating. On this basis one might predict that if knowledge of results can produce a steady increase in performance, conditions II and III will converge in information content, and possibly in performance. There is still, however, the difference that subjects in condition III know that they will be informed if they miss a signal, for whatever this is worth at high performance levels.

Comparison of the two days lends further support to previous studies which report little or no "learning" effect from session to session. Furthermore, the nonsignificance of the day-by-block interaction demonstrates that the subjects did not acquire a resistance to vigilance decrement with practice. These comparisons substantiate the view that vigilance tasks are fundamentally different from active tasks, and practice alone is not sufficient to elevate performance.

Commissive Errors: The highly skewed distribution of commissive errors makes generalization very difficult. Indeed, most authors in the field of vigilance have either ignored commissive errors entirely in their reports, or have passed them off with the statement that very few false reports were made.

The unusual finding presented here is that partial knowledge of results leads to marked increase in commissive errors over the control condition, while full knowledge of results leads to a reduced number of such errors. It is easy, and often misleading, to make ex post facto speculations on the reason for unexpected results. However, such an attempt to explain the large number of commissive errors under condition II seems worthwhile. Under the control condition, the subject received no information regardless of what action or inaction she chose to take. However, under condition III the subject received full information regardless of her decision. If the subject responded she was able to determine immediately whether that response was right or

Approved for Public Release

Contrails

wrong. If she did not respond, the missed signal light or its absence gave the same information. But in condition II, the subject received information only by responding. Thus, if the subject was doubtful about whether a presentation was signal or noise, she could resolve this ambiguity by responding. Thus it seems reasonable that the large number of commissive errors found under this condition may be accounted for by the subject's desire for information, which could be satisfied only by responding. What is more, this tendency transferred to Day 2 even when no information was available. One might say that partial knowledge of results as presented here (as well as in practical situations) trains an observer to make false responses. Although it was pointed out previously that the two knowledge of results conditions performed equally with respect to detections following transfer to a no-knowledge of results condition, the difference in probability of false response recommends the addition of missed signal information during early training.

The effect of signal rate was in the direction of higher commissive error rates for higher signal rates, as shown in tables V and VI, but the difference was significant only on Day 2.

Theoretical Implications: Although this study has not explored the various theoretical positions in the vigilance area, the results presented bear some relevance to the expectancy theory of Deese (ref. 12) and Baker (refs. 3, 5, 6), mentioned briefly in the previous section. Deese's theory states, in brief, that the likelihood that the subject will respond to a randomly occurring signal depends on his expectancy about the appearance of a signal. This expectancy is built up as a kind of "averaging process" based on past signals. The Baker extension of this theory is based on the notion of the observer perceiving the true temporal nature of the signal schedule.

These theories explain the effect of signal rate outlined in the section on "Signal Rate and Artificial Signals," by stating that with higher signal rates, (1) the subject has more experience upon which to base his perception of the signal structure; and (2) the subject can estimate the shorter intersignal intervals more accurately, making more precise his expectancy of the next signal. Knowledge of results is explained (by Baker) as furnishing better information upon which to base one's expectancy.

The expectancy theory does not attack the problem of transfer of training directly, but it is not difficult to make predictions about transfer consistant with the general theory. As for signal rate, one would predict that the higher the signal

Approved for Public Release

rate during training, the higher the expectancy during transfer, and thus the better the performance. This is supported by the data. However, there is a slight inconsistency here. The expectancy theory would predict also that the groups transferred to the identical signal rate schedule would have the most accurate perception of the intersignal interval, and thus the theory would predict that the 32-signal groups would perform better during transfer than the other two groups.

The first view seems more plausible. It emphasizes a general alertness to signals built in by a high presentation rate; the second view emphasizes an accurate perception of the true intersignal interval. Under the transfer condition on Day 2, the 48-signal groups were presented with fewer signals than on Day 1, but they carried their expectancy with them into the second day, accounting for their high performance, even on a different signal schedule. The 16-signal groups were at a disadvantage from either point of view. They suffered from both low expectancy and a change of signal schedule when transferred to the 32-signal condition.

Though the expectancy theory does not specifically deal with commissive errors, it again seems consistent with the theory to predict that, under transfer, the higher signal rate groups would commit more false reports. One might say that, under transfer, a large number of commissive errors is the price of building in a high expectancy. This view is supported by the data, which show a nonsignificant signal rate effect on Day 1, and a significant effect in the predicted direction on Day 2.

Applications: Some implications of this research for practical application have already been discussed. Primarily they consist of:

1. Initially training monitors with full knowledge of results

2. Initially training monitors with high signal rates. Just how high this should be cannot be answered by this study.

Only further research can determine the permanence of such training. However, if these findings are substantiated by further research, the question of permanence need not be the deciding factor in determining feasibility of such training. It is entirely possible to retrain monitors at regular intervals after they have been on the job without knowledge of results. In this manner it may be possible to "refresh" the monitor with

Approved for Public Release

Contrails

periodic reexposure to a monitoring training device with full knowledge of results feedback.

Summary

An experiment was performed to determine the transfer effect of knowledge of results and signal rate on performance in a Mackworth-type vigilance task. Subjects were run for the first day under feedback conditions of zero, partial, and full knowledge of results, and 16, 32, and 48 signals during a 48-minute run. On the second day all subjects were run under the conditions of zero knowledge of results and 32 signals.

The results showed that the two experimental variables differentiated subjects on both the initial exposure and the transfer condition. In terms of per cent signals detected, knowledge of results and high signal rates resulted in superior performance on both days. The usual decrement in performance over time was noted on both days.

With respect to commissive errors, large individual differences contaminated the results. The data suggest that partial knowledge of results may encourage the operator to make more commissive errors than either zero or full feedback.

The implications of the experimental results for training are discussed. The results appear to recommend training under full knowledge of results conditions and high signal rates when an operator must be placed in a situation with no knowledge of results and low signal rates.

Approved for Public Release

Contrails

APPENDIX I

INSTRUCTIONS TO SUBJECTS

This experiment deals with your ability to detect infrequent signals. The job is similar to that of a radar operator or flight engineer.

Your task is to monitor that dial. The hand on the dial steps around the circle at very regular intervals; at irregular intervals, it will take a larger step than normal, almost twice the usual step. When you detect one of these "double steps," report by squeezing this switch. I will now demonstrate the single and double steps.*

Remember that it is important to be on your toes for all these targets. However, you must also guard against a false alarm, reporting a target when there was none. This will count against you also.

Do not try to beat the game by counting or timing. It will not help you a bit. You are being run together merely as a matter of convenience. You are neither working for or against each other.

After your second session I will be happy to answer any questions you may have about the experiment.

---

* At this point groups run under conditions II and III received an explanation and demonstration of the knowledge of results display.

Approved for Public Release

Contrails

APPENDIX II

SIGNAL SCHEDULES

The schedules are described in minutes between signals. The first number is the time from the start of a session until the first signal. The last number is the time from the last signal to the end of the session.

16 Signals

3.6
3.2
1.1
4.0
2.5
3.1
1.7
4.7
0.9
2.2
4.6
4.4
2.6
2.5
3.2
3.5
0.2

32 Signals

1.3
1.3
1.7
1.5
1.9
2.0
1.1
1.1
1.3
1.6
0.7
1.9
2.5
1.1
1.2
1.8
1.0
0.3
0.7
1.9
2.7
2.0
2.2
1.3
1.4
0.7
1.3
2.4
1.5
2.7
0.7
1.2
0.1

48 Signals

1.3
0.3
1.0
0.6
1.8
0.9
0.5
0.9
1.3
0.5
0.7
2.1
1.1
1.9
1.4
0.4
0.5
0.5
0.9
1.0
0.9
0.5
0.8
2.1
1.5
0.8
0.7
1.6
0.7
1.2
0.4
1.2
0.9
0.4
1.7
1.0
0.5
1.1
0.6
1.1
0.9
1.5
0.7
1.1
1.3
0.7
1.4
0.6
0.5

Approved for Public Release

Contrails

APPENDIX III

TABLE VIII

MEAN PERCENTAGE OF DETECTED SIGNALS BY SIGNAL RATE, KNOWLEDGE OF RESULTS CONDITION, AND TIME BLOCK, DAY 1

	TIME BLOCK				
SIGNAL RATE	1	2	3	4	
16	62.5	45.0	52.0	50.0	KR-I
32	51.2	42.4	36.1	34.9	
48	77.5	64.1	52.8	56.8	
16	70.0	35.0	40.0	65.0	KR-II
32	61.3	56.3	56.4	49.9	
48	81.8	66.6	56.7	65.8	
16	52.0	47.5	60.0	70.0	KR-III
32	76.4	71.5	80.2	79.1	
48	78.2	76.7	76.7	80.1	

TABLE IX

MEAN PERCENTAGE OF DETECTED SIGNALS BY SIGNAL RATE, KNOWLEDGE OF RESULTS CONDITION, AND TIME BLOCK, DAY 2

	TIME BLOCK				
SIGNAL RATE	1	2	3	4	
16	69.0	35.9	36.2	43.9	KR-I
32	57.7	40.0	37.5	37.5	
48	75.2	69.0	62.7	62.7	
16	81.5	71.3	61.4	64.0	KR-II
32	67.6	67.8	53.8	61.4	
48	82.7	70.1	66.5	66.3	
16	67.6	51.2	48.6	49.8	KR-III
32	93.9	70.3	78.9	72.7	
48	92.7	73.8	76.4	76.3	

Approved for Public Release

Contrails

APPENDIX IV

TABLE X

SUM OF RANKS OF SUBJECTS WITH RESPECT TO COMMISSIVE ERRORS ON DAY 1

KNOWLEDGE OF RESULTS CONDITION

SIGNAL RATE	I	II	III	Total
16	388.5	491.5	282.0	1162.0
32	408.0	609.5	406.5	1424.0
48	291.0	683.0	527.0	1501.0
Total	1087.5	1784.0	1215.5	

TABLE XI

SUM OF RANKS OF SUBJECTS WITH RESPECT TO COMMISSIVE ERRORS ON DAY 2

KNOWLEDGE OF RESULTS CONDITION

SIGNAL RATE	I	II	III	Total
16	381.0	422.0	294.0	1097.0
32	427.0	574.0	374.0	1375.0
48	347.0	675.0	600.0	1622.0
Total	1155.0	1671.0	1268.0	

Approved for Public Release

Contrails

LIST OF REFERENCES

1. Annett, J., and H. Kay, "Knowledge of Results and Skilled Performance," Occup. Psychol., Vol 31, pp 69-79, 1957.

2. Bailey, A.W., A Suggested Procedure for Reducing the Incidence of Involuntary Wheels-up Landings of Naval Aircraft, Naval Research Laboratory Letter No. 5122-30, Washington 25, D.C., July 1957.

3. Baker, C.H., Vigilance: Two Tentative Theoretical Approaches, Commonwealth Advisory Committee on Defense Science, Memorandum of the Canadian Delegation, CWS(P)/P(58)18, Toronto, 1958.

4. Baker,C.H., Three Minor Studies of Vigilance, Defense Research Board, Defense Research Medical Lab. Rep. 234-2, Toronto, April 1959.

5. Baker, C.H., "Maintaining the Level of Vigilance by Means of Artificial Signals," J. appl. Psychol., Vol 44, pp 336-338, 1960.

6. Baker, C.H., "Toward a Theory of Vigilance," In A. Morris and E. P. Horne (Ed.), Visual Search Techniques, National Research Council, Washington, 1960.

7. Baker, C.H., Further toward a Theory of Vigilance, Defense Research Medical Lab. Rep. No. 234-10, Toronto, Paper read at Symposium on Vigilance, Santa Barbara, Calif., January 1961.

8. Bartlett, M.S., "The Use of Transformations," Biometrics, Vol 3, pp 39-52, 1947.

9. Brown, J.S., A Proposed Program of Research on Psychological Feedback (Knowledge of Results) in the Performance of Psychomotor Tasks, Human Resource Research Council Conference Report 49-2, pp. 81-87, Lackland Air Force Base, Texas, 1949.

10. Civil Aeronautics Board, Aircraft Accident Report No. 1-0001, 3 November 1959.

11. Civil Aeronautics Board, Aircraft Accident Report No. 1-0038, 10 January 1960.

12. Deese, J., "Some Problems in the Theory of Vigilance," Psychol. Rev., Vol 5, pp 359-368, 1955.

Approved for Public Release

13. Deese, J., and E. Ormond, Studies of Detectability during Continuous Visual Search, Wright Air Development Center Technical Report 53-8, Wright-Patterson Air Force Base, Ohio, 1953.

14. Duncan, D.B., "Multiple Range and Multiple F Tests," Biometrics, Vol 11, pp 1-42, 1955.

15. Fraser, D.C., "The Relation of an Environmental Variable to Performance in a Prolonged Visual Task," Quart. J. exp. Psychol., Vol 5, pp 31-32, 1953.

16. Garvey, W.D., F.V. Taylor, and E.P. Newlin, The Use of "Artificial Signals" to Enhance Monitoring Performance, Naval Research Laboratory Report 5269, Washington 25, D.C., 1959.

17. Harabedian, A., J.J. McGrath, and D. Buckner, "The Probability of Signal Detection in a Vigilance Task as a Function of Intersignal Interval," Human Factor Problems in Anti-Submarine Warfare, Tech. Rep. 3, Human Factors Research, Inc., Los Angeles, February 1960.

18. Head, H., Aphasia, Cambridge University Press, Cambridge, 1926.

19. Hick, W.E., "Why the Human Operator?" Transactions Soc. Instru. Technol., Vol 2, pp 67-77, 1952.

20. Jenkins, H.M., "The Effect of Signal Rate on Performance in Visual Monitoring," Am. J. Psychol., Vol 71, pp 647-6661, 1958.

21. Jerison, H.J., and R.A. Wallis, Experiments on Vigilance, II: One-Clock and Three-Clock Monitoring, Wright Air Development Center Technical Report 57-206, Wright-Patterson Air Force Base, Ohio, 1957.

22. Jerison, H.J., and R.A. Wallis, Experiments on Vigilance, III: Performance on a Simple Vigilance Task in Noise and Quiet, Wright Air Development Center Technical Report 57-318, Wright-Patterson Air Force Base, Ohio, 1957.

23. Kappauf, W.E., The Use of Angular Transformation in the Statistical Treatment of Error Frequencies, U.S. Army Air Force Air Materiel Command Report TSEAA-694-1J (ATI 110578), Wright-Patterson Air Force Base, Ohio, September 1947.

Approved for Public Release

Contrails

24. Kappauf, W.E., and W.E. Powe, "Performance Decrement at an Audio-Visual Checking Task," J. exp. Psychol., Vol 57, pp 49-56, 1959.

25. Kinkade, R.G., The Effect of Augmented Feedback on the Acquisition and Transfer of a Tracking Skill, Doctoral Dissertation, Ohio State University, Columbus, 1960.

26. Lindquist, E.F., Design and Analysis of Experiments in Psychology and Education, Houghton Mifflin, Boston, 1953.

27. Loeb, M., and E.A. Schmidt, Influence of Time on Task and False Information on Efficiency of Responding to Pure Tones, U.S. Army Medical Research Laboratory Report 426, Ft. Knox, Kentucky, 1960.

28. Loeb, M., and E.A. Schmidt, A Comparison of the Effect of Different Kinds of Information in Maintaining Efficiency on an Auditory Monitoring Task, U.S. Army Medical Research Laboratory, Report 453, Ft. Knox, Kentucky, 1960.

29. Mackworth, N.H., Notes on the Clock Test--A New Approach to the Length of Watch for Radar Operators, Applied Psychology Unit Report No. 1, Medical Research Council, Cambridge, 1944.

30. Mackworth, N.H., Faulty Perception Caused by Blank Spells without Signals during Experiments on Prolonged Visual Search, Applied Psychology Unit Report No. 17/45, Medical Research Council, Cambridge, 1945.

31. Mackworth, N.H., "The Breakdown of Vigil during Prolonged Visual Search," Quart. J. exp. Psychol., Vol 1, pp 6-21, 1948.

32. Mackworth, N.H., Researches on the Measurement of Human Performance, Medical Research Council Report, Series No. 268, Cambridge, 1950.

33. McCormack, P.D., "Performance in a Vigilance Task with and without Knowledge of Results," Canad. J. Psychol., Vol 13, pp 68-71, 1959.

34. McGrath, J.J., A. Harabedian, and D. Buckner, "Review and Critique of the Literature on Vigilance Performance," Human Factor Problems in Anti-Submarine Warfare, Tech. Report 1, Human Factors Research, Inc., Los Angeles, 1959.

Approved for Public Release

35. Merk, R.V.R., A Field Method for Monitoring Radar Operator Performance, Air Force Cambridge Research Center Technical Report 54-106, Cambridge, August 1954.

36. Miller, R.B., Handbook on Training and Training Equipment Design, Wright Air Development Center Technical Report 53-136, Wright-Patterson Air Force Base, Ohio, June 1953.

37. Nicely, P.E., and G. A. Miller, "Some Effects of Unequal Spatial Distribution on Detectability of Radar Targets," J. exp. Psychol., Vol 53, pp 195-198, 1958.

38. Pollack, I., and P.R. Knaff, "Effect of Rate of Target Presentation on Target Detection Probability," Paper presented at 1958 meeting of Amer. Psychol. Assn., Abstract in Amer. Psychol., Vol 13, p 414, 1958.

39. Pollack, I., and P.R. Knaff, "Maintenance of Alertness by a Loud Auditory Signal," J. acoust. Soc. Amer., Vol 30, pp 1013-1016, 1958.

40. Stevens, S.S., "Machines Cannot Fight Alone," Amer. Scientist, Vol 34, pp 389-400, 1946.

41. Wyatt, S., and J.N. Langdon, "Inspection Processes in Industry," Industrial Health Research Board Reports, Vol 63, London, 1932.

Approved for Public Release