

*Contrails*

WADC TECHNICAL REPORT 55-395

Cleared: April 17th, 1980

Clearing Authority: Air Force Wright Aeronautical Laboratories

## **EXPERIMENTAL STUDIES OF PROLONGED WAKEFULNESS**

*WALTER D. CHILES, 1ST LT, USAF*

*AERO MEDICAL LABORATORY*

*DECEMBER 1955*

PROJECT No. 7193

TASK No. 71612

**WRIGHT AIR DEVELOPMENT CENTER  
AIR RESEARCH AND DEVELOPMENT COMMAND  
UNITED STATES AIR FORCE  
WRIGHT-PATTERSON AIR FORCE BASE, OHIO**

# Contrails

## FOREWORD

These experiments were carried out under Project No. 7193, Task No. 71612, "Changes in Performance as a Function of Increased Fatigue," with 1/Lt. Walter D. Chiles acting as project engineer and task scientist. The assistance of Mr. C. A. Dempsey, Capt. N. R. Burch, Capt. T. H. Greiner and of all the other participating Laboratory personnel is gratefully acknowledged.

WADC TR 55-395

Two experiments were carried out concerning the effects of prolonged wakefulness and of fatigue on performance of psychomotor and psychological tasks. The first experiment involved four subjects each of whom sat individually in an aircraft cockpit for 56 hours. During this time, measures of reaction time and alertness were taken, and, at the end of the experiment, two of the subjects flew ILAS passes in a Link trainer. The reaction time and the alertness measures reflected considerable variability in the attentiveness of the subjects. The performance in the Link trainer was judged to be within the limits of acceptability.

The second experiment involved the assessment of the effects of 30 hours of wakefulness with continuous work (painting a barracks) on the performance of 15 subjects on a psychological and a psychomotor task. In addition, the effects of two different drugs, d-amphetamine sulphate and cortisone, were compared with those of a placebo. The psychological test involved the summing of rows of figures, and the psychomotor test involved arm-hand steadiness. Each of these tests distinguished, to some extent, the two drug groups from the placebo group. It was concluded that the two drugs improve performance on the addition test, whereas they tend to decrease arm-hand steadiness as a function of time.

PUBLICATION REVIEW

This report has been reviewed and is approved.

FOR THE COMMANDER:



JACK BOLLERUD  
Colonel, USAF (MC)  
Chief, Aero Medical Laboratory  
Directorate of Research

*Control*  
TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTION.....	1
EXPERIMENT I.....	1
PROCEDURE.....	2
Apparatus.....	2
Program.....	2
RESULTS.....	3
DISCUSSION.....	5
The Tasks.....	5
EXPERIMENT II.....	7
EXPERIMENTAL PROCEDURE.....	7
Subjects.....	7
Tests.....	7
Experimental Program.....	8
RESULTS.....	8
DISCUSSION.....	12
SUMMARY AND CONCLUSIONS.....	14
BIBLIOGRAPHY.....	16

LIST OF TABLES

<u>Table</u>		<u>Page</u>
I	Analysis of Variance, Steadiness Aiming Test.....	9
II	Analysis of Covariance - Addition Test.....	11

LIST OF ILLUSTRATIONS

<u>Figure</u>		<u>Page</u>
1	The experimental situation.....	3

*Continuity*  
LIST OF ILLUSTRATIONS

<u>Figure</u>		<u>Page</u>
2	Relative frequency distribution of the reaction time measures. The measures taken on all subjects were combined for the purposes of this graph. The total range of the responses was from .6 to 90 seconds.....	4
3	A plot of the mean and median reaction times for the period of continuous performance. The data for each point are the combined measures for the four subjects with eight responses per subject..	4
4	A plot of the clock test measures for the initial, continuous performance, and final periods of the test. Each point in the initial period is based on 15 minutes performance combined over two subjects; during the continuous performance period each point is based on one hour of performance combined over two subjects; and, during the final period, each point is based on 15 minutes performance again combined over two subjects.....	4
5	A plot of the average number of errors (contacts) made by each group on the steadiness aiming test as a function of time. Each point in each of the three curves is based on two 30 second trials per subject, these two scores being averaged and then the average taken over the five subjects in the respective groups....	8
6	A plot of the average number of correct responses on the addition test as a function of time. Each point in each of the three curves represents the average of the number of rows of figures correctly handled by the five subjects for each of the groups in a three-minute period.....	10
7	A scatter plot of the control vs the experimental scores for each subject on the addition test. The control scores are the mean of the performance on the first two periods for each individual subject. The experimental scores are the mean of the performance on the final four periods for each individual subject.....	10
8	The data presented in Figure 6 for the last four periods adjusted by the analysis of covariance regression equation.....	12
9	A plot of the total number of errors for each group on the addition test as a function of time.....	12

## INTRODUCTION

Past psychological studies which are related to the area of prolonged wakefulness and "fatigue" have dealt with a wide variety of both tests and experimental situations. One of the most frequently studied of such experimental conditions has been the loss of sleep complicated by various amounts and kinds of work or activity. The tests have been aimed at assessing the influence of these "fatiguing situations" on almost all of the psychological functions which have been isolated in man. The conditions have been varied from eight hours of driving a truck to 112 hours of wakefulness, which included forced marches of up to 65 miles (3,4).

The results of these tests are in fairly good agreement. Most psychological functions which have been measured show no effect of fatigue or wakefulness when the tests involve a relatively short period of performance. For example, a reaction time test lasting two minutes showed no change after 36 hours of sleep deprivation, but, when the test was increased to about eight minutes, an increase in response time was produced (4). In general, then, tests which involve 15 minutes or less of moderately paced performance show no change with fatigue; when this period is increased, tests which require considerable concentration or sharp focusing of attention usually are performed less proficiently.

Many investigators have indicated that they feel that the qualitative observations of the behavior of the individual are more revealing than the quantitative measures which have been thus far utilized. For example, wakefulness for periods exceeding 36 hours leads to a change in the ability of most subjects to express themselves; they tend to slur their speech; they may at times speak incoherently (2). A subject may even have hallucinations, and there are frequently transient personality changes, e.g., the subject may become belligerent or euphoric.

The individual's ability to concentrate on any complicated problem for any length of time has been reported to be deleteriously effected after 36 hours of wakefulness (4). It is this inability to sustain vigilance that is the most frequently reported product of wakefulness for long periods of time.

A series of experiments has been initiated to assess further the effects of fatigue and wakefulness on the ability of the individual to maintain an alert attitude for a relatively long period of time.

### EXPERIMENT I<sup>1</sup>

This investigation concerned itself with the ability of an individual to maintain a particular gross physical orientation for a relatively long period of time - up to 56 hours. During the first 18 hours, the subject was to be

---

1. This report concerns a portion of a study carried out by the Aero Medical Laboratory under the supervision of Mr. C. A. Dempsey.

allowed to relax, except that he was required to be capable of returning to an alert state when requested to do so. During the next 19 hours, the subject was to be required to remain alert at all times. During the last 18 hours, the subject would again be permitted to relax except when requested to alert himself. And for the last half hour of this final period the subject was to remain alert. The experimental-design problem was primarily that of devising a test which would afford an indication as to the "mental alacrity" of the individual with respect to his responding to alterations in his physical environment.

## PROCEDURE

### Apparatus:

The subjects were given a reaction time task to perform. Whenever a red light (situated on the lower left hand portion of the instrument panel, Figure 1) came on, the subject pushed a button which turned out that light. This yielded a measure of reaction time which included a vigilance component; i.e., the subject had to see or recognize that the light was on before he could turn it off.

The subjects were provided with a second task which consisted of a pointer that rotated at the rate of one revolution per minute on a circular display. The subject was required to depress a toggle switch when he thought that the pointer was directly aligned with a mark at the top of the display. (See Figure 1.)

The reaction time was recorded to the nearest one-hundredth of a second. The accuracy of the pointer alignment was scored in terms of per cent direct "hits", and there were two scoring positions immediately "before" and two immediately "after" the center, "hit", position. Each of the five scoring positions was 2 and 1/2 degrees wide which resulted in the requirement that the response, to be "perfectly" accurate, must occur within a particular 1/2 second period in the 60 second cycle of the pointer.

Two of the subjects, who were rated pilots, were removed from the cockpit at the end of the 56 hours; they were permitted to don street clothes and were then placed on a stretcher and carried to another building. Here, they were helped into a Link trainer and were required to fly three ILAS passes.

### Program:

Each subject, upon entering the test situation, was required to perform the clock test for the first one hour period. The reaction time light was turned on eight times, at randomly selected time intervals, during this period. The subjects were then required to perform the clock test - with one reaction to the light - for fifteen minutes during each hour until they decided to go to sleep. Beginning with the 19th hour, the subject performed the clock task continuously through the 37th hour. Eight reaction time measures per hour were taken during this period, again at random time intervals. From the 37th hour through the 55th hour, the subject was permitted to relax. And then the subject

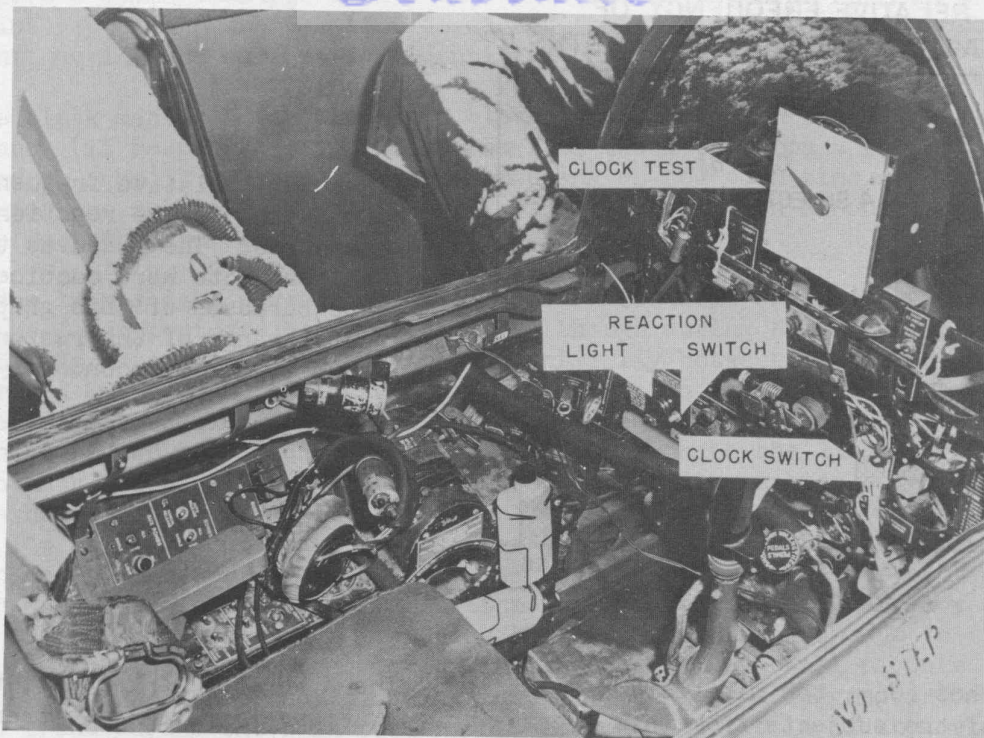


Figure 1: The experimental situation.

performed the clock test and had one reaction time test during the final fifteen minutes of the 56th hour. The two subjects, for whom the clock apparatus was in working order, performed this task 15 minutes per hour from the 47th hour through the end of the test.

## RESULTS

Complete data were obtained on all four subjects on the reaction time task. However, because of mechanical failure, data were obtained on only two of the subjects on the clock test.

Figure 2 is a relative frequency distribution of the reaction time data from all subjects. The data were grouped, according to the time taken to make the reaction, into .6 second time intervals (represented on the abscissa). Each point in the curve represents the per cent of the reactions which fell into the corresponding time class identified on the abscissa. From this figure it can be seen that there was a great range of reaction times. The median reaction took 1.5 seconds; 25 per cent of the reactions were longer than 2.3 seconds ( $Q_3$ , third quartile) and 10 per cent of the reactions were longer than



*Contrails*

## RELATIVE FREQUENCY OF REACTIONS OF VARIOUS LENGTHS

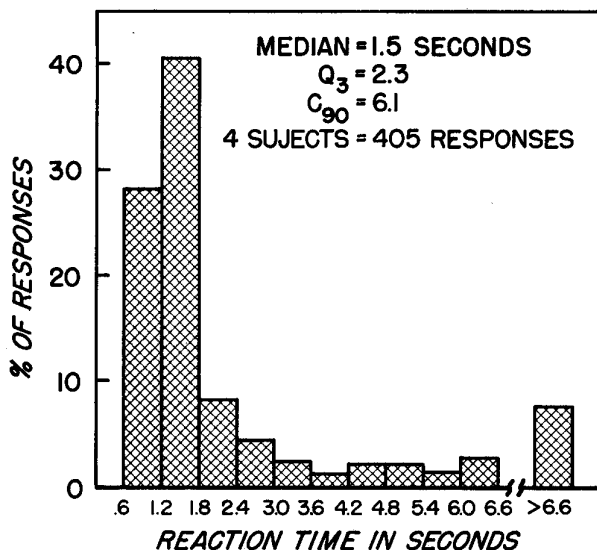


Figure 2: Relative frequency distribution of the reaction time measures. The measures taken on all subjects were combined for the purposes of this graph. The total range of the responses was from .6 to 90 seconds.

6.1 seconds ( $C_{90}$ , the ninetieth percentile). There were no marked deviations of any of the subjects from this average curve. Figure 3 reveals no trend in either the mean or median reaction time during the 19 hour period.

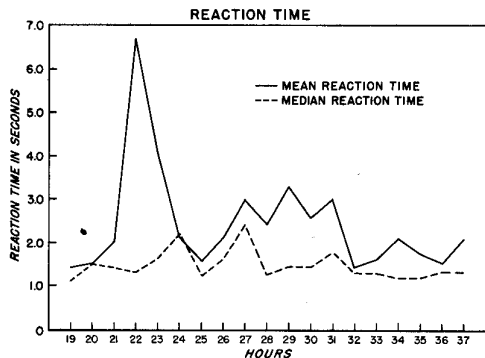


Figure 3: A plot of the mean and median reaction times for the period of continuous performance. The data for each point are the combined measures for the four subjects with eight responses per subject.

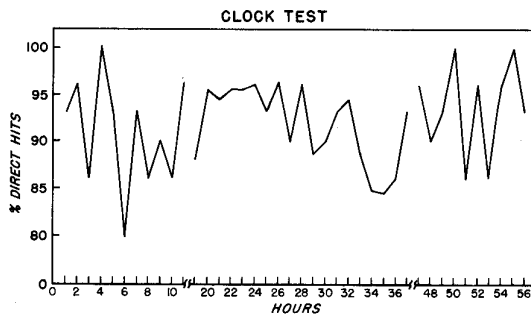


Figure 4: A plot of the clock test measures for the initial, continuous performance, and final periods of the test. Each point in the initial period is based on 15 minutes performance combined over two subjects; during the continuous performance period each point is based on one hour of performance combined over two subjects; and, during the final period, each point is based on 15 minutes performance again combined over two subjects.

Figure 4 summarizes the data from two of the subjects on the clock test. Each point represents the per cent of direct "hits" with respect to pointer alignment, during the time period identified on the abscissa.

The clock measures for the first eleven hours are each based on 15 minutes performance (15 responses). The measures from the 19th through the 37th hours are based on from 53 to 60 responses, depending on how much time was taken for other functions required of the subject. And from the 47th through the 56th hour the measures are again based on 15 responses.

There is a slight suggestion of a downward trend in the per cent hits during the first period - hours 1 to 11. Although during the 4th hour 100 per cent hits were made, the scores range primarily between 96 and 86 per cent, one score of 80 per cent having been noted in the 6th hour. The high performance period started out with a score of 88 per cent hits during the 19th hour, and then jumped up to 90 to 96 per cent from the 20th to the 23rd hours. From the 24th hour the performance fluctuated considerably, reaching a low point at the 29th hour, climbing back up to 95 per cent during the 32nd hour and reaching the low point of this part of the test (85 per cent) at from the 34th to the 36th hours. The performance then again came back up to 94 per cent during the 37th hour. The performance during the final ten hours was rather variable, ranging from a low of 86 per cent during the 51st and 53rd hours up to a high of 100 per cent during the next to the last hour. The over-all picture shows the slight downward trend noted during the first period, a suggestion of a downward trend during the high performance period, and little if any detectable trend during the final period. The salient feature seems to be the great variability coupled with a possible periodicity, although with only two subjects this cannot be taken very seriously.

Comparison of the mean and median reaction times with the clock test data (Figure 3 vs Figure 4) gives no indication of any consistent relationship between the subjects' performances of the two tasks.

The two rated pilots performed satisfactorily in the Link trainer. Each of the three passes which each of the subjects made, using the ILAS, were acceptable in terms of their arriving at the end of the runway at an appropriate altitude, bearing, and air speed, as well as having stayed within the glide path and localizer beams reasonably well. It should be pointed out that the performance was also "satisfactory" when each of these subjects flew one pass with the rough air simulator set on one-half rough.

## DISCUSSION

### The Tasks:

What is the significance of the data summarized in Figure 2? As was pointed out above, 10 per cent of the responses to the reaction time light took longer than 6.1 seconds. To evaluate what this might mean in a practical sense, let us consider a hypothetical situation. Suppose that on board an aircraft there is an electronic Gadget which is quite sensitive to temperature, i.e., there is

*Control*

danger of its exploding if it gets too hot. Let us further suppose that if the Gadget automatic flutter valve sticks (requiring manual override) the critical temperature will be reached in 6 seconds, i.e., the operator must see and respond to the Gadget automatic flutter valve malfunction signal in something less than 6 seconds. Clearly, then, the probability of mission abortion, because of a stuck Gadget automatic flutter valve, would be a function of the probability of occurrence of such a malfunction and the probability obtained in this experiment, .1, that a signal light would go undetected for 6 seconds or longer. (Some other rather obvious contingencies are omitted for the sake of simplicity.) On the basis of previous work in the area, it can be said that the obtained reaction time data are in line with what would be expected in a situation such as this.

The clock test may be thought of as measuring the ability of the subject to pace himself with respect to periodically focusing his attention on a particular aspect of his physical environment. Though the subject may permit his mind to wander as he pleases most of the time, he must be able to return to the consideration of his immediate task in response to a temporal stimulus which he himself must provide. In a sense, then, it could be said that the test measures the ability of the subject to generate accurately such stimuli; and this task would presumably require a fairly high level of alertness.

Figure 4 suggests that there is an alteration in the subject's ability to remain alert. These data are compatible with those obtained from the reaction time task, except that there is a suggestion of a trend, here, which was not apparent in the reaction time data. In combination, these two tasks indicate that the subjects were not by any means alert at all times during the period in which they were required to be so.

Any positive assertion concerning practical situations from tests such as these would have to be made with due consideration of a very potent but essentially unknown factor. This factor is the strength of the motivation of the subjects in the experimental situation as compared to their expected motivation in an analogous operational situation. At the present time the best guess would be that the motivation in the operational situation would be considerably higher than in the test situation. Performance on either of these tests would be improved with increased motivation, but there is no way to say how much better an alert state would be maintained under the pressures of the "real thing."

In this same vein, it could be argued that the highest level of motivation during the test would be induced by the challenge presented to the two rated pilots when required to fly ILAS approaches in the Link trainer. This represented a task, the successful performance of which was quite meaningful to them. Though the argument would hardly be carried so far as to suggest that the motivation in the Link would approximate that of the actual flight situation, nevertheless, the greater "meaningfulness" as compared to the other tests cannot be ignored.

It is the performance of these two subjects in the Link trainer which suggests that, from the point of view of psychomotor performance, further investigation in operational test situations is in order, keeping in mind the cautionary

notes inserted by the variability in alertness indicated by the results on the simpler psychological tests.

## EXPERIMENT II

This investigation<sup>1</sup> was designed to assess the effects of 30 hours of wakefulness, coupled with continuous work, on the performance of a psychological and of a psychomotor test with certain types of pharmacological assistance.

### EXPERIMENTAL PROCEDURE

#### Subjects:

The fifteen subjects in this experiment were volunteers from among airmen stationed at Wright-Patterson Air Force Base. Nine of the subjects had regular duties as painters, and six had regular duties as hospital corpsmen. The incentive for participating in the experiment was a three day pass to be awarded on the week end following the completion of the study. For purposes of controlled drug administration, the subjects were divided unsystematically into three groups - identified as Group I, Group II, and Group III - with five subjects in each group.

#### Tests:

Each subject performed on the SAM Steadiness Aiming Test at approximately 4-hour intervals. The performance consisted of two 30-second trials separated by 30 seconds of rest. For purposes of analysis, these two 30-second trials were averaged for each session. This test, described in detail elsewhere (1), required the subject to hold a pivoted stylus on a rest point while attempting to prevent the tip of the stylus from making contact with the sides of a ring in which the tip was located. Scoring was provided through the completion of an electrical circuit each time the stylus touched the target ring; this circuit completion activated a Veeder Counter. Time of contact measures were also taken to provide a check as to whether or not the subject deliberately held the stylus against the edge of the ring to reduce the total number of contacts. In addition to this test, the subjects were assembled as a group at approximately 4 hour intervals and were given an arithmetic test. Each subject was given a sheet of paper on which were printed rows of digits. The test required the subject to begin summing the series of digits presented in a given row, and, when the cumulative sum equalled a figure presented in the left margin next to that row, the subject marked the final digit necessary to reach the indicated figure and went on to the next row. The subject worked for a total of 3 minutes on this task at each test. Scoring was provided by a count of the total number of rows correctly handled as well as by a count of the number of errors made. Eight equivalent forms were used, one for each session, in order to prevent the occurrence of any direct learning effects.

1. This report concerns a portion of a study carried out by the Biophysics Branch, Aero Medical Laboratory, under the supervision of Captain N. R. Burch and Captain T. H. Greiner. Their assistance is gratefully acknowledged.

Two drugs were selected for study, both of which are considered to be central excitants. These drugs were dextro-amphetamine sulphate and cortisone. The dexadrene was administered in 5 mg tablets and the cortisone was administered in 25 mg capsules. The control subjects were given a 10 mg placebo. At the time of administration, each subject was given his pill and a cup of water, and an attempt was made by observation of the subject to be certain that he swallowed the pill as instructed.

Experimental Program:

The experiment officially began at 0800 hours on a Wednesday. At this time the subjects began to paint the interior of a barracks. At 1600 on the first day each of the subjects received a pill; those in Group I each received 5 mg of dexadrene; those in Group II each received 25 mg of cortisone; and those in Group III each received a placebo. These dosages were repeated at 1900 and 2300, and at 0100, 0400, 0700 and 1100 the next day. At 12 noon of the first day, the first cycle of testing began; during the course of the experiment the subjects were tested at approximately 4 hour intervals, each period of testing requiring a total of approximately 1 hour for the whole group.

The subjects ate meals at their normal meal times and in addition ate a meal at 2300 hours. No foods or liquids, other than water, were permitted in the intervening periods.

The painting continued without other than the above mentioned interruptions until about 1000 hours of the second day, at which time the painting job was finished. From that time until the end of the experiment, the subjects were permitted to do as they chose except that sleeping was prohibited.

The final testing period, which concluded the experiment, was completed at 1400 hours on the second day after a total of 30 hours of wakefulness.

RESULTS

In general the subjects seemed to accept with interest the challenge presented to them by the steadiness aiming test. As can be seen in Figure 5, the

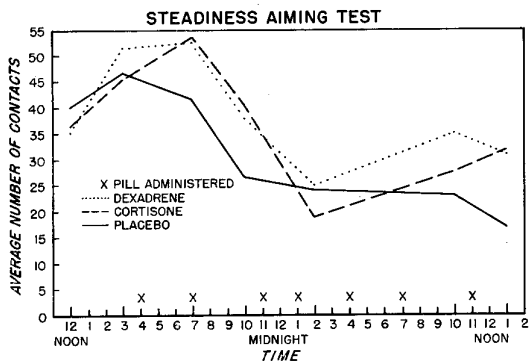


Figure 5: A plot of the average number of errors (contacts) made by each group on the steadiness aiming test as a function of time. Each point in each of the three curves is based on two 30 second trials per subject, these two scores being averaged and then the average taken over the five subjects in the respective groups.

performance on this test after a slight initial deterioration, improved throughout the course of the experiment, and this was true of all three of the groups. The statistical analysis which was carried out on the data from the last four sessions (Table I) proved the improvement trend to be significant at better than the .001 level of confidence. Although there were no significant differences

TABLE I

Analysis of Variance, Steadiness Aiming Test

Source	d.f.	SS	MS	F
Between Subjects	14	8,862.2	633.0	
Drugs	2	1,032.8	516.4	
Error (between)	12	7,829.4	652.2	
Within Subjects	45	3,617.7	80.4	
Time	3	1,246.7	415.5	9.17**
Time x Drugs	6	741.2	123.5	2.73*
Error (within)	36	1,629.8	45.3	
Total	59	12,479.9		

\*\* significant at .001 level of confidence.

\* significant at .05 level of confidence.

Summary table of the analysis of variance applied to the data from the Steadiness Aiming Test. The analysis is based on the mean of two 30-second trials per subject for each of 15 subjects - five subjects per group.

(F ratio is less than 1.00) among the groups, there is found to be a significant interaction between the effects of the drugs and the passage of time. In this respect, the Placebo Group shows a fairly steady improvement in performance, whereas the two experimental groups show considerable fluctuation in their levels of performance.

As is indicated in Figure 6, there was a fairly large initial separation among the three groups with respect to the number of rows correctly handled on the addition test. By the fifth test period, the groups had fairly well stabilized with respect to their ordering, the dexadrene group doing the best, the placebo the next best and the cortisone group the poorest. Since each subject's performance was quite consistent from one test period to the next, it was decided

# Contrails

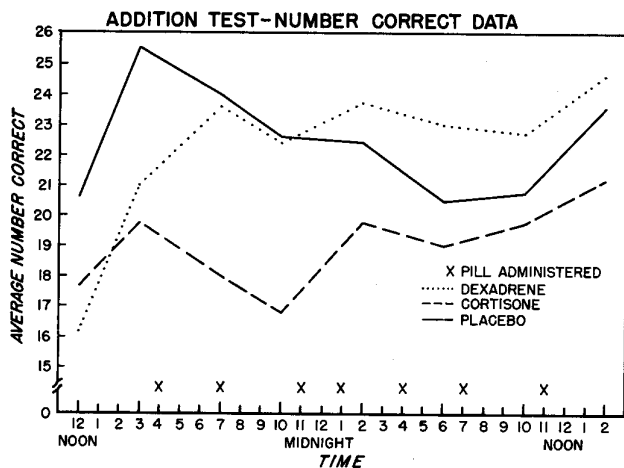


Figure 6: A plot of the average number of correct responses on the addition test as a function of time. Each point in each of the three curves represents the average of the number of rows of figures correctly handled by the five subjects for each of the groups in a three-minute period.

to attempt a statistical analysis that would take account of this consistency and, hence, correct for the initial level of performance of the individual, as well as that of each group considered as a whole. Although the number of subjects was smaller than is generally recommended for the application of the analysis of covariance, the scatter plot of initial score vs the mean of the last four scores for each subject (see Fig. 7) suggested that such an analysis

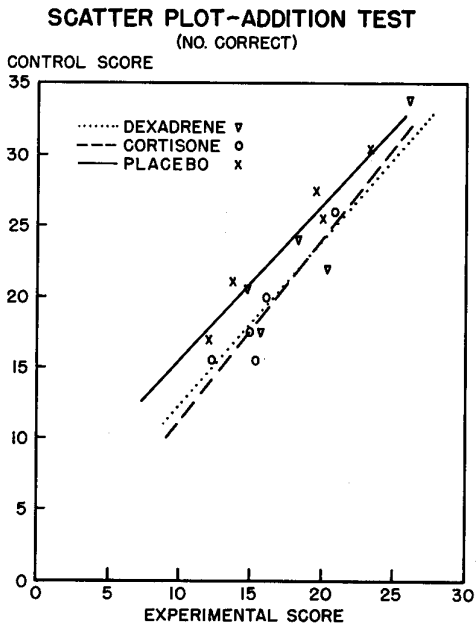


Figure 7: A scatter plot of the control vs the experimental scores for each subject on the addition test. The control scores are the mean of the performance on the first two periods for each individual subject. The experimental scores are the mean of the performance on the final four periods for each individual subject.

might be well worth the effort. Therefore, the first two periods were averaged and used as the control variable; the mean of the last four periods was used as the criterion variable. The initial test periods were dropped since the time course of the cortisone is uncertain, and it was desired to treat only the sessions in which the drugs were known to be active. The results of this analysis are summarized in Table II. The test of the significance of the difference between groups is found to be significant at only slightly better than .10 level

TABLE II

Analysis of Covariance - Addition Test

Source	d.f.	SS	MS	F
Between Groups	2	13	6.5	2.98
Within Groups	11	24	2.2	
Total	13	37		

"t" tests

Dexadrene vs Placebo  $.02 < P < .05$

Cortisone vs Placebo  $.05 < P < .10$

Summary table of analysis of covariance applied to the addition test data. The analysis is based on: for the control measures, the mean of the first two performances for each subject, and for the experimental measures the mean of the last four performances for each subject. N=15 with five subjects per group.

of confidence. However, since the primary concern of the experiment was the effect of the drugs as compared to the effect of the placebo, the t-tests, the results of which are presented in Table II, were carried out. The test for the difference between the dexadrene and placebo groups is significant at almost the .02 level of confidence. The difference between the cortisone and the placebo groups is not quite significant at the .05 level of confidence. Each of these t-tests was performed on the scores as adjusted by the analysis of covariance regression equation (Fig. 8). The correlation coefficients for total, between groups and within groups obtained from this analysis were .92, .83 and .94 respectively. The effect of taking account of these high correlations is demonstrated in the very gross shifts in the relative performances of the three groups as seen when Figure 8 is compared to Figure 6.

The error data from the addition test are summarized graphically in Figure 9. Reference to this figure indicates that the general envelope of the plot of



ADJUSTED MEANS FOR ADDITION TEST  
NUMBER CORRECT

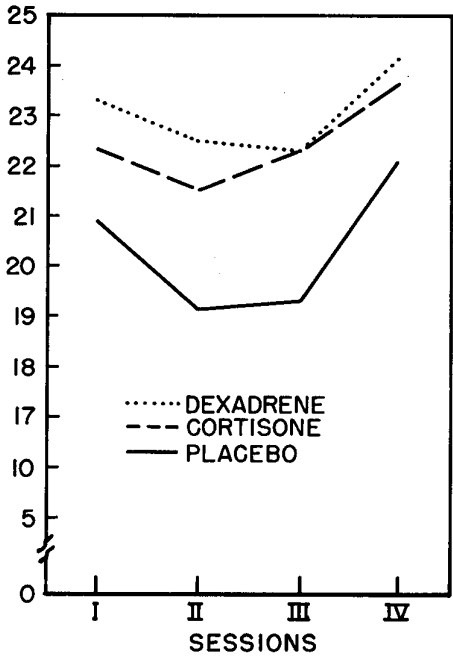
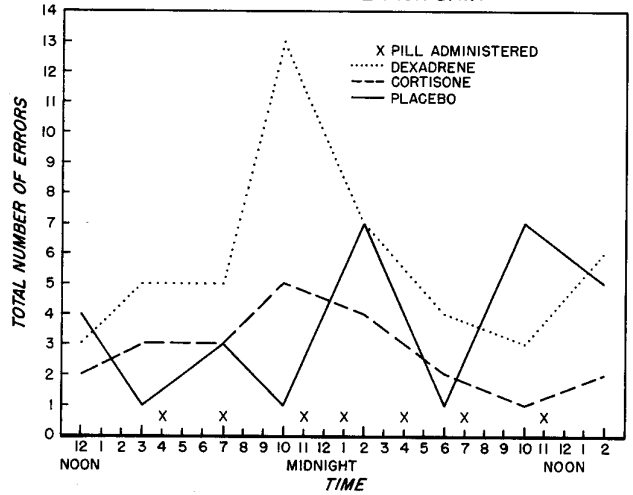


Figure 8: The data presented in Figure 6 for the last four periods adjusted by the analysis of covariance regression equation.

ADDITION TEST-RAW ERROR DATA



The over-all improvement of the subjects on the steadiness aiming test over time, though not really surprising, does pose problems of interpretation. The first source of explanation lies in the direction of increased familiarity with the test and testing procedure. A second explanation might evolve from the notion that, with increased fatigue, the subjects were better able to relax in performing the test; and relaxation seems to be a keynote of good performance on this test. A third explanation might be that the subjects were merely learning how to perform the task more efficiently. These explanations might be abbreviated as adaptation, relaxation and learning respectively. The adaptation and learning factors would be expected, on the basis of information available about this test, to run their courses in a matter of a few trials, which fact would leave the continued improvement over the final trials unaccounted for. The possibility that increased relaxation accounted for the results, though attractive, cannot be accepted without additional information. One further consideration is the fact that the subjects were painting during most of the experiment, presumably primarily with the preferred hand, and they performed the test with the preferred hand. This suggests that a check of the present results, with respect to the effects of fatigue, could be carried out by requiring subjects to manipulate a hand and arm dynamometer for various lengths of time and then measuring their performance of this test. Using only one or two test periods under these conditions should permit ruling out any effects of learning or of adaptation on performance on the steadiness aiming test.

The lack of a significant difference among the groups could be attributed in part to the small number of subjects in each group, but, in the light of the actual magnitude of the F ratio (less than 1), it would seem reasonable to accept the hypothesis of no differences at face value. However, there is the significant interaction between drugs and time to be considered, .05 level of confidence. This significant interaction can be interpreted to mean that the differences among the performances of the three groups varied significantly as a function of the test periods (time). In this respect, the placebo group remained at a rather consistently higher level of performance, i.e., were more steady, than the other two groups. The performance of the placebo group was also more stable on the average as compared to the two experimental groups which fluctuated over time. This fluctuation on the part of the experimental groups might have resulted from some peak effectiveness phenomena with respect to the time of administration of the drugs, and this peak effect might be assumed to result in more "nervousness" on the part of the experimental subjects. As the drug effects wore off, the experimental subjects would become more "like" the placebo subjects. Although some of the known subjective effects of such central excitants are consistent with such an interpretation, the time schedule for the administration of the test does not permit any firm conclusions; i.e., although the data are plotted as points with respect to the time axis, the testing of each subject individually required that the testing be spread out over a period of approximately one hour for each session thus making any good estimate of time between a test session and the last drug administration impossible. Obviously, an experimental schedule could be arranged which would provide satisfactory measures of this time variable.

The raw data from the addition test present an unexpected picture in that, although the two drugs employed are presumably similar in effect, the cortisone

group was inferior to the placebo group in performance, whereas the dexadrene group was superior to the placebo group. However, as the analysis of covariance demonstrated, this was merely an example of the possible vagaries of random sampling with small numbers of subjects. This sampling artifact is clearly seen in the comparison of Figure 8 with Figure 6.

The failure of the F ratio to achieve a customary level of significance is not particularly surprising in view of the similarity between the treatments of the two experimental groups and of the small numbers of subjects per group.

The results of the "t" tests applied to the adjusted means of the dexadrene and the cortisone groups comparing each of them with the placebo group are more or less in line with what would be expected in this type of experimentation. From the statistical point of view, it is of interest to note that cortisone group is seen to have improved as a result of the administration of the drug. It should be pointed out that, although the adjusted scores place the cortisone group above the placebo group (not quite significant at the .05 level of confidence), the placebo group actually performed better as seen in the raw data.<sup>2</sup> The failure of the error data from the addition test to display any consistent trend is probably a result of the fact that this test measures a greatly over-learned ability; therefore, the resultant frequency of occurrence of errors was too small to reveal a trend even if there were one inherent in the situation. The similarity between the curves for the two drug groups cannot be explained satisfactorily within the context of the present experimental structure, although it was undoubtedly related in some way to the time course of the effects of the drugs.

## SUMMARY AND CONCLUSIONS

### Experiment I.

1. Each of 4 subjects individually sat in an aircraft cockpit for 56 hours.
2. Measures of reaction time and alertness reflected considerable variability in the attentiveness of the subjects during the course of the experiment.
3. Performance in a Link trainer was judged to be within the limits of acceptability.
4. It is concluded that further study of the problem in operational test situations is in order.

---

2. This apparent incongruity is introduced by the fact that the analysis of covariance technique, in effect, compares the performance of the subjects with what would be expected of them on the basis of the correlation between the control and the criterion scores; the shift in the relative positions of the two groups merely means that the cortisone group did better than would be expected on the basis of the control measures, whereas the placebo group did not do as well as would be expected.

# Contrails

## Experiment II.

1. Fifteen subjects remained awake and worked for a 30 hour period.
2. The subjects were divided into three groups of 5 subjects each. Each subject periodically received a pill; Group I received 5 mg pills of d-amphetamine sulphate; Group II received 25 mg capsules of cortisone; and Group III received 10 mg placebos.
3. The subjects were tested at approximately 4 hour intervals on a hand-arm steadiness test and on an addition test.
4. The two tests distinguished to some extent the two drug groups from the placebo group.
5. It is concluded that relative to the effects of a placebo the two drugs improve performance on an addition test and under the described experimental conditions tend to decrease arm-hand steadiness as a function of time.

*Contrails*  
BIBLIOGRAPHY

1. Apparatus Tests, Edited by A. W. Melton, Report No. 4, Army Air Forces Aviation Psychology Program Research Reports, U.S. Government Printing Office, Washington, D. C., 1947, p. 501-557.
2. Edwards, A. S. Effects of the loss of one hundred hours of sleep. Amer. Jr. Psychol., Vol 54, 1941, p. 80-91.
3. Finan, J. L., Finan, S. C., & Hartson, L. D. A review of representative tests used for the quantitative measurements of behavior-decrement under conditions related to aircraft flight. USAF Technical Report No. 5830, July 1949.
4. Tyler, D. B. The effect of amphetamine sulfate and some barbiturates on the fatigue produced by prolonged wakefulness. Amer. Jr. of Physiol., Vol 150, 1947, p. 253-262.