### EFFECTIVENESS OF RESTRAINT EQUIPMENT IN ENCLOSED AREAS

**Author(s):**
D. L. Lowrey, E. D. Langston, W. Reed, and John J. Swearingen, D.Av.T.

**Performing Organization Name and Address:**
FAA Civil Aeromedical Institute
P. O. Box 25082
Oklahoma City, Oklahoma 73125

**Sponsoring Agency Name and Address:**
Office of Aviation Medicine
Federal Aviation Administration
800 Independence Avenue, S. W.
Washington, D. C. 20590

Research leading to preparation of this report was performed under Project AM-A-71-PRS-37.

**Abstract:**
A series of 20-g decelerations of a crash sled was conducted to determine the magnitude of head impact decelerations while wearing various types of restraint equipment in small confined areas. Restraint webbing loads and head impact decelerations are presented for three directions of impact (straight forward, and 90° to left and right). Restraint webbing undoubtedly reduces head impact velocities, especially in the forward direction. However, this study shows that, in most instances, head strikes may be expected even while using upper and lower torso restraint because of the close proximity of surrounding structure in general aviation aircraft. Introduction of upper torso restraint along with lap belts in general aviation aircraft will not relieve the need for lethally surrounding structures.

**Key Words:**
Crash injury, aircraft design, aviation safety, body flailing during deceleration, shoulder harness design, injury potential

**Distribution Statement:**
Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia, 22151, for sale to the public.

**Security Classification:**
Unclassified
EFFECTIVENESS OF RESTRAINT EQUIPMENT IN ENCLOSED AREAS

I. Introduction.

Numerous research scientists have investigated the effectiveness of various restraint devices during deceleration utilizing instrumented dummies, volunteer human subjects, and various primates. However, in all cases, tests were conducted with a seat mounted openly on a crash sled in such a manner as to allow freedom for kinematic motions in all directions, limited only by the restraint equipment being tested. Since man, as an operator of a vehicle, spends most of his time during transportation in a small enclosed area sitting next to rigid structure on his left, within 2 to 4 feet of structure on his right, and within 18 to 20 inches from rigid structure in front of him, this study was conducted to determine head impact forces on surrounding structures and body kinematics while wearing seven different designs of restraint equipment in an enclosed area. For economy reasons, sections of automobile bodies instead of aircraft fuselages were utilized in these tests. Decelerations were measured in the forward and in both lateral directions.

II. Test Equipment and Procedures.

Four automobile bodies (a 1961 Chevrolet, a 1961 Ford, a 1965 Mercury, and a 1965 Plymouth) sectioned just anterior to the firewall and just aft of the back of the front seat, were purchased from a local salvage yard and rigged for mounting on the CAMI crash sled. Sled impact velocity was programmed for approximately 29 miles/hour (42.53 ft./sec.) to produce a peak deceleration of 20 g's. Actual test results indicated a peak terminal sled velocity of 42.3±1 ft./sec. and all impact tests with the exception of test runs Numbers 3 (17 g's) and 20 (23 g's) had peak decelerations of 20 g's. Photographic coverage of the event was provided to record both top and side views at 24, 400, and 2000 frames/second.

Three 250-g CEC Model 4-202-0001 strain gage accelerometers were mounted tri-axially in the head of an Alderson F-50 anthropometric dummy to measure forward, lateral, and vertical head impact decelerations. Belt load transducers manufactured in the CAMI machine shops were calibrated on a 5000-pound capacity Dillon dynamometer prior to test procedures. Output signals for the belt load transducers and accelerometers were recorded on a CEC Model 5-124A oscillographic recorder in conjunction with a Sanborn 550M signal conditioning system.

A total of 24 tests was conducted to evaluate the effectiveness of whole body restraint systems for head protection. Seven different types of restraint equipment were used for these tests in the forward and in both lateral test positions. For comparison, one test was made in each of the three directions with the dummy unrestrained. These seven different designs of restraint equipment are shown diagrammatically in Figures 2 through 8, 10 through 16, and 18 through 24. The restraint equipment tested consisted of (1) a single diagonal chest strap without seat belt (as used in some foreign cars) (Figure 2), (2) a single diagonal chest strap with a separate seat belt (as installed on over 10,000,000 late-model American cars) (Figure 3), (3) the so-called three-point restraint system—diagonal chest strap with lower end attached to one-half of the seat belt and upper end attached to the "B" post (Figure 4), (4) a three-point harness identical to that described above with the exception that the upper end of the chest strap goes over the shoulder and seat back and is attached to the floor structure behind the seat (Figure 5), (5) a double chest strap arrangement with the lower ends sewn to the lap belt, near the seat back and pan intersection, and the upper ends joined behind the neck to form a single strap which, in turn, passes over the seat back and attaches to the floor structure behind the seat (Figure 6), (6) a restraint system the
same as number 5 with the exception that the shoulder straps are crossed behind the neck and pass over the seat back to separate floor attachments (Figure 7), and (7) the Pacific Scientific quick-release harness in which shoulder straps and lap belts plug into a quick-release buckle. In these tests, the common upper torso strap passed over the seat back to a fixed floor attachment instead of going into an inertial reel (Figure 8).

III. Results.

Head impact forces with various structures and restraint webbing loadings, as well as vertical forces on the neck from centrifugal forces, along with time readings of each peak force occurrence from the time the crash sled began deceleration, are shown diagrammatically in Figures 1 through 24.

Forward Decelerations: Figure 1 demonstrates, as other researchers have shown repeatedly, that the unrestrained human body slides forward in an erect sitting posture until the knees contact the lower instrument panel, at which time the upper torso flexes forward into the windshield and/or control column. In the case depicted, the head penetrated the windshield (1965 Plymouth) at about the same time that the neck contacted the upper steering wheel rim.

In studying Figures 2 through 8, the following observations are made:

All harness configurations allowed the driver's head to impact the steering wheel rim and/or hub.

Deceleration impact force was proportional to the length of shoulder strap webbing in designs anchored to the "B" post and roof structure; i.e., the longer the webbing, the greater the impact force with increase in stretch distance.

In designs where shoulder straps traversed the top of the seat and anchored to the floor structure behind the seat, head impact force varied with yield characteristics of the seat back structure—the top of the seat flexed forward 6 to 8 inches.

Head impact force with steering structures varied between 45 and 107 g's (forward and lateral head forces were vectored since there was usually some degree of head rotation). These forces could be tolerated without serious injury on well-designed structures, but would cause serious facial fractures and head injuries in most current transportation vehicles.

A vertical peak force of 20 to 40 g's was recorded in all forward deceleration tests and, since this peak force occurred approximately 0.02 seconds before head impact, it may be assumed that the stress resulted from centrifugal force, producing considerable stretching of the neck along the radius of motion.

In general, as would be expected, the fewer the number of straps restraining the body, the higher the webbing loads. For example, when only a single upper torso strap was used without a seat belt (Figure 2), the total strap load exceeded 2200 pounds and would have been higher, but lack of a seat belt allowed the knees to impact the lower dash panel, taking some of the deceleration load off of the chest strap. With the body totally restrained by using both chest strap or straps and lap belt, total strap loading was approximately 3000 pounds. Attaching shoulder straps to the lap belt as in Figures 4, 5, 6 and 7 substantially increases (nearly doubles) the strap loading on that portion of the lap belt that serves as a common attachment for both a shoulder strap and a portion of the seat belt.

Lateral Impact—Occupants thrown to the left: In comparing Figures 10 through 16 with that of the unrestrained dummy in Figure 9, it is clear that none of the seven different restraint systems offers any appreciable protection against head impact in this direction (as attested to by low readings on the restraint webbing) since the occupant is seated in such close proximity to side structures on his left. The following observations are, however, worthy of note:

a. Impact of the side of the head may occur against the "B" post or against the door glass, depending on slight variations in the angle of impact. Head impact against the "B" post in this study ranged from 110 g's to 158 g's; these levels would probably be fatal since the loads were concentrated on such a small area of the head due to rigid, nonyielding construction of the post. Head impact forces to break the door glass were 100 and 122 g's for pre-1965 laminated glass as compared with only 44 and 60 g's for tempered glass used in 1965 and later-model vehicles. According to Lissner,9 none of these
forces against yielding glass is sufficient to produce skull fracture.

b. As shown in Figure 10, the lap belt is an absolute necessity in this type of impact to prevent ejection. Use of the single shoulder strap without a seat belt in Figure 10 allowed the buttocks to be ejected out the door, and permitted the strap to catch under the chin, putting an 800-pound load on the webbing. In such situations, the body slides down the strap and a knife-action is produced on the neck which has been reported to decapitate the occupant.\(^5\)

c. In Figures 14 and 15 (double shoulder straps), it will be noted that the strap forces are considerably higher, especially on the strap over the left shoulder. A double shoulder strap system running through a strong integrated seat could offer considerable protection against left side impacts.

20-g Right Side Impact: In the unrestrained test, the dummy (Figure 17) slid across the seat in a sitting position in less than 0.3 seconds and the right side of its head impacted on the right “B” post with a force of 150 g’s. Body impact caused the door to open and the dummy was ejected.

In all instances of single strap restraint over the left shoulder, the upper torso slipped side-ways out of the shoulder strap and the body folded to the right, over the seat belt, with the head striking the seat cushion (Figures 18, 19, 20 and 21). These lateral head impacts with the seat cushion, varying between 20 and 50 g’s, are insignificant; however, lap belt loads between 1675 and 2850 pounds cutting into the side of the abdomen would probably cause some internal injuries.

Double strap shoulder restraint over the seat back in conjunction with a seat belt allowed some lateral motion of the trunk, but held it almost upright (Figures 22, 23 and 24). In these tests the side of the head hit the top of the seat back with insignificant forces varying from 49 to 73 g’s and abdominal loads from seat belts were significantly reduced (975 to 1925 lb.).

[V. Discussion.

Caution should be used when the results of the cinematics of the body along with head impact forces presented in this study are applied to general aviation aircraft because of the difference of internal measurements of the occupant spaces. The front seat of an automobile is approximately 5 feet wide, while that of popular general aviation aircraft varies from 3 to 4 feet, with a large majority of seats only about 3.5 feet wide. A previous study\(^1\) has shown a side displacement of the head from the centerline of the body (in the sitting position, with seat belt restraint) in excess of 36 inches during application of a one-g force. Since the centerline of the body is approximately 12 inches from the left side of the aircraft, only 2 to 2\(\frac{1}{2}\) feet of clearance are available to the right side for prevention of head impact with side structures. Hence, the kinematic motions allowed to the right side (without head impact) with the types of restraint equipment shown in Figures 18 through 21 would be expected to allow head impact with the right side of the cabin in most general aviation aircraft. Only the double shoulder strap designs shown in Figures 22 through 24 sufficiently restrained side motions of the body to protect the head in small aircraft.

In the forward deceleration tests, the maximum motion of the head was approximately 14 inches in the horizontal plane and 14 inches in the vertical plane. This compares favorably with Chandler’s study\(^2\) of human subjects in which he reports a forward horizontal motion of the head of 1.03 feet during a 12-g deceleration. Young\(^3\) has reported 22 to 24 inches of horizontal displacement of a dummy’s head with a vertical drop of 16 to 18 inches, while the dummy was restrained by a single strap shoulder harness during decelerations from velocities between 40.6 and 41.3 feet/second. In the same report Young showed that double shoulder strap harnesses restrict the forward motion of the head to a range of 10 to 18 inches and the vertical motion to 16 to 20 inches. It is probable, therefore, that forward head motions in the tests presented in this report were limited by impact with the steering assembly and that, when these measurements are corrected and applied to general aviation aircraft, occupants of the front seat can be expected to experience head impact with the upper instrument panel and control wheel, even while wearing shoulder harness restraint, if they are exposed to decelerations of the magnitude tested in these studies.
V. Conclusions and Recommendations.

As demonstrated by the test dummy, restraint of the human body by the use of belts during crash decelerations is difficult. In the small cockpit area of general aviation aircraft, a single diagonal chest strap used in conjunction with a lap belt may reduce, but will not prevent, head impacts during forward and side decelerations. A double shoulder strap-lap belt design of restraint equipment can further reduce crush injuries, especially in crashes involving deceleration forces which will throw the pilot to the right side of the cockpit, provided that it is designed as an integral part of a strong seat or fastened to a strong aircraft structure near the seat. The findings given in this report indicate that even with the upper torso and lap belt restraints, front seat occupants will suffer injuries from forward and side crash impact forces under the acceleration conditions of the test; therefore, engineers should stress design of structures in the cockpit area to minimize injuries from head impact.

REFERENCES


Figure 1. Dummy unrestrained, forward deceleration.
Figure 2. Dummy restrained by single diagonal chest strap, no seat belt, forward deceleration.
Figure 3. Single diagonal chest strap and a separate seat belt, forward deceleration.
Figure 4. Three-point restraint system (diagonal chest strap with lower end attached to one-half of the seat belt and upper end attached to the "B" post), forward deceleration.
Figure 5. Three-point restraint system with chest strap going over seat back and anchored to the floor behind the seat, forward deceleration.
Figure 6. Double chest strap with lower ends sewn to lap belt near the seat intersection and upper ends joined behind the neck to form a single strap which, in turn, passes over the seat back to a floor attachment, forward deceleration.
Figure 7. Double chest straps sewn to seat belt, crossed behind seat, and extending to separate floor attachments, forward deceleration.
Figure 8. Pacific Scientific quick-release restraint equipment. Double seat belts converge to one behind the neck and attach to floor structure behind the seat, forward deceleration.
Figure 11. Single diagonal chest strap and a separate seat belt, left side impact.
Figure 12. Three-point restraint system (diagonal chest strap with lower end attached to one-half of the seat belt and upper end attached to the "B" post) left side impact.
Figure 13. Three-point restraint system with chest strap going over seat back and anchored to the floor behind the seat, left side impact.
Figure 14. Double chest strap with lower ends sewn to lap belt near the seat intersection and upper ends joined behind the neck to form a single strap which hooks to a door attachment. Left side impact.
Figure 15. Double chest straps sewn to seat belt, crossed behind seat and extending to separate floor attachments, left side impact.
Figure 16. Pacific Scientific quick-release restraint equipment. Double chest straps converge to one behind the neck and attach to floor structure behind the seat. Left side impact.
Figure 17. Dummy unrestrained, right side impact.
Figure 19. Single diagonal chest strap and a separate seat belt, right side impact.
Figure 20. Three-point restraint system (diagonal chest strap with lower end attached to one-half of the seat belt and upper end attached to the "B" pillar) right side impact.
Figure 21. Three-point restraint system with chest strap going over seat back and anchored to the floor behind the seat, right side impact.
Figure 22. Double chest strap with lower ends sewn to lap belt near the seat intersection and upper ends joined behind the neck to form a single strap on the seat back to a floor attachment. Right side impact.
Figure 23. Double chest straps sewn to seat belt, crossed behind seat and extending to separate floor attachments, right side impact.
**Title and Subtitle**
EVALUATION OF A FIBERGLASS INSTRUMENT GLARE SHIELD FOR PROTECTION AGAINST HEAD INJURY

**Author(s)**
E. D. Langston  
John J. Swearingen, D. Av. T.

**Performing Organization Name and Address**
FAA Civil Aeromedical Institute  
P. O. Box 25082  
Oklahoma City, Oklahoma 73125

**Supplementary Notes**
Research leading to preparation of this report was performed under Project NO. AM-A-71-PRS-47.

**Abstract**
An all fiberglass prototype glare shield has been evaluated in terms of head injury protection. In 30-ft./sec. head impacts, the 9-1/2-inch protrusion folded down over the heavy instruments, offering significant improvement in head injury protection when compared to current aircraft instrument panel design. However, in this particular design the fiberglass broke, allowing the forehead to contact a thin, sharp edge with sufficient force to produce fatal head injuries. Design changes to eliminate this fracture point and incorporation of fiberglass glare shields of similar design in future general aviation aircraft could lead to a significant reduction of head injuries during crash decelerations.

**Key Words**
Crash safety, engineering design, crashworthiness

**Distribution Statement**
Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia, 22151, for sale to the public.
EVALUATION OF A FIBERGLASS INSTRUMENT GLARE SHIELD FOR PROTECTION AGAINST HEAD INJURY

I. Introduction.

In view of the large number of serious or fatal head injuries resulting from head impact against upper instrument panels during crashes of general aviation aircraft, the Protection and Survival Laboratory at CAMI maintains interest in new design concepts which may offer protection against head trauma.

II. Test Equipment and Procedures.

Two fiberglass instrument panels (without instruments), along with their integrated glare shields, were submitted to the laboratory for evaluation. These were mounted at the end of the CAMI crash decelerator (Figure 1). The position of the glare shield was adjusted to conform with measurements provided by the manufacturer. An instrumented dummy head, taken from an Alderson F-50 anthropometric dummy, was rigidly attached to a weighted arm which was free to swing about a pivot on the decelerator sled. The sled was braked prior to impact, allowing the head to swing forward in an arc, impacting the protruding edge of the glare shield in the orientation shown in Figure 2. Yellow chalk applied to this protruding edge allowed determination of maximum area of initial face contact and strips of adhesive tape indicated depth of head penetration. Sled decelerations were calibrated to produce head impact velocities of 30 and 35 ft./sec. since these head impact velocities will occur in crashes of approximately 7-G magnitude as measured on the floor structure under the seat. Head impact deceleration was measured by a single 250-G CEC Model 1-202-0001 strain gage accelerometer mounted in the dummy head with data recorded using a Sanborn 550 M signal conditioning system and a CEC Model 5-124A oscillographic recorder. Time lines on the oscillograph paper, a sweep-second clock and time marks on all high-speed motion picture film allowed synchronization of deceleration peaks on the tracing with structural collapse of the fiberglass glare shield. Motion picture coverage was provided by four cameras, one operating at 24 fps, two at 400 fps, and one at 2,000 fps.

III. Description of Glare Shield.

The glare shield shown in Figure 3 consisted of a basic structure of a thin fiberglass layer covered with a ¼-inch thick layer of Ensolite and a thin layer of plastic vinyl. It extended 9½ inches from the instrument panel toward the pilot and was elevated about 13 degrees above the horizontal. On the protruding edge nearest the pilot the shield was rolled down and under with an inside radius of curvature of approximately ¼-inch. After impact testing had been completed, the covering was removed from a portion of the shield so that the thickness of the fiberglass could be measured. The first five inches (nearest the pilot) were of a uniform thickness of .020 inches, increasing to .025 inches at six inches, to .050 inches at seven inches, and .065 at eight. At nine inches the thickness abruptly increased to 0.200 inches due to the lap with the instrument panel.

It should be noted at this point that the glare shield protruded only in front of the left pilot and that the right side dropped back such that the head impact area for the right seat occupant consisted of a layer of the thin padding over a rigid structure at the top of the instrument panel (Figure 1).

IV. Results.

Figure 4 presents a pictorial sequence of events correlated with deceleration measured during Test One (30 ft./sec. head impact velocity). The initial head contact with the edge of the glare shield was with the maxilla just below the
nose producing an initial deceleration peak of 10 G's followed by a second peak of 30 G's as the shield began to collapse. Chalk deposited on the head form from the edge of the glare shield during this phase of the impact was lifted by means of masking tape and the area of facial contact determined to be 3.75 square inches. Based on previous research, it appears unlikely that major facial fractures would occur during this initial head impact. From 0.025 seconds to 0.04 seconds, the glare shield folded down over the instruments and prevented head contact with them. However, beginning at about 0.04 seconds, the fiberglass broke about 6 1/2 inches from the edge of the glare shield. The thin edge of the shield, supported by the instrument panel flange penetrated the 1/4-inch padding and cut the rubber head form covering, as shown in Figures 5 through 8. A peak head form deceleration of 60 G's was produced as the rigid head form contacted this thin edge. In Reference 7 it is stated that the forehead can tolerate 80 G's on one square inch without bone fracture. If we assume that the forehead contacted a four-inch length of this 0.035-inch-thick fiberglass edge, we can calculate a contact area of 0.14 square inch and any impact deceleration in excess of 12 G's would be expected to produce injuries. The impact of 60 G's in this test would produce severe lacerations (Figure 8) and extensive fracture of the anterior cranium.

The results of this test were confirmed by a second test which produced almost identical failure patterns and head form decelerations.

V. Conclusions.

Since the fiberglass glare shield folded down over the heavy instruments and sharp knobs and edges and produced a maximum deceleration force on the head of only 60 G's while distributing the load over large facial areas, as compared to 300-G forces produced on small areas of the head in similar impact tests of conventional light aircraft instrument panels without the glare shield (Reference 2), it must be concluded that a glare shield constructed of fiberglass or of some similar material could substantially reduce the large number of severe head injuries now occurring in general aviation crashes. The failure of the glare shield in these tests exposing a relatively rigid sharp edge to the head could cause fatal injuries, even in the relatively low impact velocities represented by these tests. Correction of this fault by a minor design modification could lead to a significant improvement in crashworthiness.

REFERENCES

Figure 1. Fiberglass instrument panel and glare shield mounted at the end of the crash sled.
Head Impact Test of Fiberglass Instrument Panel (Velocity = 30ft/sec)

Top View
A B C D E

Side View

Deceleration in G
0 20 40 60

Time in Seconds
0 .01 .02 .03 .04 .05 .06 .07

Figure 4. Results of a 30-ft./sec. head impact against the fiberglass glare shield.
Figure 7. Fiberglass edge at break (see arrow). Dime is shown for thickness reference.
Figure 8: Insertions (of expanse) in rubber covering on dummy head forehead from impacting broken fiberglass.
Thirty-nine human subjects were exposed to repetitive backscatter light stimulation (off a white wall or fog) from a Grimes capacitance discharge airplane anticollision light flashing at 1.27 Hertz. Both tonic (light stimulus absent) and phasic (light stimulus present) stimulus-bound occipital EEG, heart rate, respiration, skin potentials, and eyeblinks were recorded. In the first experiment, response decrement (habituation) to the flashing light occurred only with one out of five response measures (skin potential) over a 40-trial session indicating that the flashing light was a potent stimulus. None of the subjects demonstrated photic driving, seizure activity, or theta wave activity in his EEG. Most subjects reported the light as noxious, none became nauseated, but many became drowsy over the course of the experiment and several reported dark deadaptation from the flashing light. In Experiment II, eyeblink and skin potential measures differentiated between an instrument-rated pilot group and an age-matched control group of non-pilot professional men.

The results suggest that, although the flashing anticollision light induces changes in physiological measures which are resistant to habituation, these changes do not extend to the induction of nausea.
ACKNOWLEDGMENTS

We wish to thank the FAA Civil Aeromedical Institute in Oklahoma City, Oklahoma, for the use of its facilities during our data collection for Experiment II and Drs. P. P. Iampietro and C. E. Melton for their assistance and advice. We also thank Dr. E. A. Higgins, who manned the fog generator in Experiment II, and Gerald Vardiman, who worked hard and ably as our assistant.
EFFECTS OF BACKSCATTER OF BRIEF HIGH INTENSITY LIGHT ON
PHYSIOLOGICAL RESPONSES OF INSTRUMENT-RATED
PILOTS AND NON-PILOTS

The view is fairly common that some of the immediate effects of intermittent photic stimulation may result in altered states of consciousness, vertigo, fear and visceral disturbances. This has wide implications for people exposed to flickering light stimulation. Of most concern is the effect of intermittent light stimulation on airplane pilots since impairment of ability to function can have the gravest consequences in the event of an emergency.

The nature of the evidence for the view that intermittent photic stimulation can cause impairment of ability to function comes from several sources. First, much of the evidence is of an anecdotal sort and as such presents problems of valuation. The second form of evidence comes from clinical EEG laboratories where intermittent photic stimulation is used as a tool in assessing epilepsy. The third type of evidence comes from experimental research laboratories. Since there are wide variations in research methodology, the intensity of the photic stimuli used, the site at which the light pulses were presented, and in the subject populations employed, there is room for contradictory claims. Discussion of different aspects of the problem as well as good views of the literature are abundant.

Most of the earlier experiments dealing with the effects of intermittent light stimulation on physiological indices and performance have employed frequencies higher than the flash rates of most airplane anticollision lights and have used electroencephalograms almost exclusively as the index of physiological response. These studies do not bear as directly on the problem confronting pilots flying through clouds or fog as would studies which use flash frequencies in the range countered by pilots. It was the purpose of the experiment to determine under controlled conditions the effects of brief high intensity light pulses from an actual airplane anticollision light flashing at 1.27 Hertz on a variety of physiological measures in a normal non-pilot population. Post-experimental verbal reports of the experience were also obtained from the subjects.

EXPERIMENT I

Methods

Subjects. A group of 13 young adult males (age range 17-25 years) comprised of medical students from the University of Oklahoma Medical School and introductory psychology classes from Oklahoma City University with no history of epilepsy participated in the experiment. Each experiment lasted for about one hour. Subjects were paid for participation in the experiment.

Procedure. Upon arriving at the laboratory, subjects were told that they would be part of a visual stimulation experiment and that a variety of physiological responses would be monitored. Skin surfaces where recording electrodes were to be attached were cleaned with acetone. Grass silver cup electrodes were used to record EEG and eye blinks. The occipital EEG lead was attached 1 cm above and 1 cm to the right of the occiput. This lead was referenced to the ipsilateral earlobe. For eye blinks, leads were attached above and below the left eye. Plate EKG leads were attached on the right shoulder and on the chest below the heart. Beckman skin potential leads were attached to the second finger tip of the left hand and to the volar surface of the left arm. A common ground lead was attached to the volar surface of the left arm above the reference skin potential lead. A strain gauge respiration lead was attached across the chest at the bottom of the rib cage. Physiological responses were recorded on a Beckman Type RM rectilinear inkwriting dynograph running at a paper speed of 10 mm/sec.
Design. The experiment was set up to provide 40 trials, each of which was 10 seconds in duration, of high intensity condenser discharge light stimulation. The light source was a Grimes model no. 30-360-1 high intensity condenser discharge airplane anticollision light. Flash rate was constant at 1.27 flashes per second. The intertrial interval between each 10 seconds of flash stimulation was 30–90 seconds, varying randomly in 15-sec steps as determined by a table of random numbers. A Gerbrands tape programmer turned on a Hunter 111C timer for 10-second periods at the randomly prescribed intervals. The Hunter timer, in turn, controlled the power source to the Grimes light. Preliminary tests indicated that the sound of the capacitor charging in the Grimes light was a confounding cue which would trigger physiological responses not due to the light flashes. To circumvent this problem, a Grason-Stadler 901B white noise generator was employed in conjunction with a Koss Pro 4A headphone set. Further testing revealed that an 80 db re: .0002 dyne/cm² white noise effectively masked the capacitor charging sound and other equipment sounds.

Response quantification. Occipital EEG was recorded with a 0.3-second time constant and at a gain of 50 μ volts/cm. To be scored as alpha wave activity, two or more consecutive brain waves of 8–12 Hertz had to be present in the EEG. Visually, alpha stands out from the background EEG in the awake subject by being of larger amplitude than the background activity and by exhibiting synchrony (i.e., when alpha occurs, it is generally present in “bursts” of several waves in a row). The percent alpha in the 10 seconds preceding each stimulus was scored as well as the percent alpha in the 10 seconds during flashing light stimulation.

Eye blinks were recorded at a gain of 0.2 mv/cm with a 0.1-sec time constant. Since spontaneous blinks occur with a frequency greater than zero and the spontaneous blink rate varies among individuals, blinks were scored both during the 10 sec preceding each stimulus and during the 10 sec of each stimulus.

Skin potential responses were recorded with a 1.0-sec time constant at a gain of 0.5 millivolts/cm. Response amplitude was measured in millivolts from response onset to peak of response. Response amplitudes were summed separately for the 10-sec pre-stimulus periods and 10 sec during light stimulation periods to yield baseline and stimulus-elicited responding scores, respectively.

Respiration rate was recorded with a 1.0-sec time constant. A flexible tube, which change its resistance when stretched, was taped to the subject's chest. This changing resistance was part of one arm of a Wheatstone bridge circuit. The distance between respiration peaks was measured and converted to rate/min. Rate in the 10 sec preceding a stimulus and the first complete respiration cycle during the stimulus were recorded for each trial.

Heart rate was recorded by a cardiograph coupled. It was qualified in terms of mean beat per minute change from the mean of the 10-sec period before stimulation compared to the 10 sec during stimulation.

Results

All trials effect comparisons were based on nonparametric Friedman analyses of variance pre-during stimulus comparisons and specific trial comparisons were based on correlated difference between means t-tests. The p<.05 twotailed rejection region was adopted throughout.

![Figure 1](image)

**Figure 1.** Mean number of eye blinks in 10 trial block pre- and during light flash stimulation.

Figure 1 summarizes the results for the blink data. Both the mean pre-stimulus blink rate and the mean rate of blinking during phasic stimulation are presented in 10 trial blocks. Several results are evident from this figure. First, collapsing the data across trials, there

---

2
a highly significant increase in blinking to the light flashes when compared with the pre-stimulus blink rate \( (t=8.30, \, df=12, \, p<.001) \). Second, the blinking rate to the flashing lights did not habituate as indicated by a nonsignificant Friedman nonparametric analysis of variance for trend across trials. Third, no significant changes in basal blink rate occurred over trials as indicated by a second Friedman analysis of variance.

![Figure 2. Mean magnitude of skin potential responses in millivolts for 10 trial blocks pre- and during-light flash stimulation.](image)

**Figure 2.** Mean magnitude of skin potential responses in millivolts for 10 trial blocks pre- and during-light flash stimulation.

Figure 2 summarizes the results for the skin potential data. A Friedman nonparametric analysis of variance indicated that the decrement in responding across trial blocks was significant \( \chi^2=12.15, \, df=3, \, p<.01 \). Further testing indicated that a significant decrement occurred between the first and second trial blocks \( (t=2.51, \, t^*<9, \, p<.05) \). None of the other comparisons was significant. Thus, most of the decrement occurred in the first 10 trials with no consistent change occurring thereafter.

![Figure 3. Mean rate of respiration in 10 trial blocks pre- and during-light flash stimulation.](image)

**Figure 3.** Mean rate of respiration in 10 trial blocks pre- and during-light flash stimulation.

Figure 3 summarizes the results for respiration data in 10 trial blocks. Collapsing across trials, significant difference in respiration rate was found with rate increasing during photic stimulation when compared to the rate just preceding atic stimulation \( (t=2.83, \, df=9, \, p<.02) \). The respiration changes across trial blocks were not significant for either the pre- or during-stimulus light.

![Figure 4. Mean heart rate over 10 trial blocks pre- and during-light flash stimulation.](image)

**Figure 4.** Mean heart rate over 10 trial blocks pre- and during-light flash stimulation.

Figure 4 summarizes the results for heart rate data over trial blocks. Collapsing across trials, a small but highly reliable heart rate deceleration was found when comparing pre- with during-stimulus heart rate \( (t=5.65, \, df=12, \, p<.001) \). The decrease in basal heart rate as measured over trial blocks in the pre-stimulus interval was not significant \( (C=4.88, \, df=3, \, p<.05) \). However, the heart rate decrease across trial blocks during light stimulation was significant \( (C=13.3, \, df=3, \, p<.01) \).
**EXPERIMENT II**

Since the subjects in the first experiment were not pilots and the possibility exists that pilots may respond differently to flashing lights than non-pilots, the first experiment was repeated, but it utilized instrument-rated pilots as subjects as well as an age-matched control group of non-pilots. Secondly, the actual back-scatter conditions were simulated more closely by using back scatter from man-made fog as the stimulus.

**Methods**

*Subjects.* A group of 12 instrument-rated pilots from the Oklahoma City area and a group of 14 non-pilot professional men from the University of Oklahoma Medical School served as subjects in the experiment. The average age of the subjects was 40 years. Pilots and non-pilot did not differ significantly on average age. Each session lasted about 80 minutes. At the end of the experiment subjects were administered post-experimental questionnaire. Post-experimental verbal reports and ratings were also obtained from the subjects. Subjects were paid for participation in the experiment.

*Procedure.* Upon arriving at the laboratory, subjects were told that they would be a part of visual stimulation experiment and that a variety of physiological responses would be monitored. Electrode placements and data quantification were the same as in Experiment I except that the pulse train duration in Experiment II was extended to 40 seconds for pre- and during stimulus measures. For eye blinks, skin potentials responses, and alpha EEG activity, data were quantified over 40-second intervals instead of the 10-second intervals of Experiment I. If heart rate data were scored over the 10 seconds preceding a stimulus and the first tenth of a second during the stimulus because preliminary data indicated that this is where the change was taking place. Respiration was not scored in the second experiment because of recording difficulties with the transducer.

Experiment II was performed in the chamber of the Civil Aeromedical Institute the FAA Center in Oklahoma City. After being instrumented, the subject, clad in a surgical gown, was placed in a semi-reclining position with a plastic lawn chair in the fog chamber. The temperature was $73^\circ$ F.±$2^\circ$ and the humid...
was 100%. A flashing Grimes airplane anti-collision strobe light was placed below and behind the subject. The fog density was such that neither the experimenters nor the subject could see more than five feet in the fog chamber. On two occasions the fog generator iced up and the data from the subjects were not used because the fog density thinned visibly.

**Design.** The experiment was designed to provide 20 trials, each 40 seconds in duration, of repetitive high intensity condenser discharge light stimulation. The light source was a Grimes model no. 30-360-1 high intensity condenser discharge airplane anticollision light. Flash rate was constant at 1.27 flashes per second and flash duration was ≃ 10 μsec. The intertrial interval was 30-90 seconds, varying randomly in 15-sec steps as determined by a table of random numbers. A Gerbrands tape programmer operated a Hunter 111C timer for 40-second periods at the randomly prescribed intervals. The Hunter timer, in turn, controlled the power source to the Grimes light. Sound masking was provided by a Grason-Stadler 901B white noise generator via a Koss ko-44A headphone set. The white noise generator was set to provide 80 db re: 0.0002 dyne/n² white noise at the earphones.

Both instrument-rated pilots and the control group of non-pilots received the same sequence of 20 trials of flashing light stimulation. For the pilot group a 100-watt overhead light was on in the fog chamber at all times during the experiment.

The non-pilot matched control subjects received the flashing light under two conditions. One-half of the control subjects the overhead light was on for the first half of the experiment and was turned off for the second half of the experiment. For the second half of the control subjects, these conditions were reversed. This procedure tested the hypothesis that the physiological and subjective effects of the flashing light stimulus are determined, in part, by contrast or amount of light change.

For each response measure the data were analyzed by 3-factor repeated measures (Case 1) analyses of variance. Factor one was the instrument-rated pilot versus non-pilot dimension. Factor two was the pre-stimulus basal measure compared with the during-stimulus measure. Factor three was repetition, i.e., response change over 4 trial blocks. In all cases two-tailed tests were used and the criterion for rejection of the null hypothesis was set up at p<.05 or better.

**Results**

Figure 6 summarizes the results for the eyeblink data. Both the mean pre-stimulus blink rate and the mean rate during light flash stimulation are presented for the four trial blocks. The instrument-rated pilots did not differ significantly in overall blink rate from the age-matched non-pilot group (F<1, df=1, 22, p>.05) indicating that the blink rate between groups was nondifferential. There was a significant groups-by-stimulus interaction (F=11.46, df=1, 22, p<.01), with the instrument-rated pilot group having a higher pre-stimulus blink rate than stimulus rate, whereas the group of non-pilot showed increased blinking during the flashing light stimulus over their pre-stimulus rate.

![Graph showing mean number of eye blinks in 5 trial blocks pre-and during light flash stimulation.](image)

There was a significant trials effect (F=4.64, df=3,66, p<.01) indicating that blink rate increased over the course of the experiment. The groups-x-trials interaction was not significant, suggesting that both pilots and non-pilots contributed about equally to the significant trials effect.
Figure 7. Mean magnitude of skin potential responses in millivolts for 5 trial blocks pre- and during-light flash stimulation.

Figure 7 presents the skin potential response data in summary form. The three-factor analysis of variance for the skin potential data indicates that (1) pilot and non-pilot groups did not differ in their overall level of skin potential responding \( F = 1, df = 1.22, p > .05 \); (2) there was a significant groups-x-trials interaction \( F = 3.48, df = 3.66, p < .05 \). The interaction was due to the fact that pilots did not differ in their pre-stimulus or during-stimulus responses nor did they change their responding over trials. The non-pilot group, on the other hand, had a stable pre-stimulus response rate over trials but their response to the flashing light stimulus was largest initially and showed steady decrement (habituation) over trials.

Figure 8 summarizes the heart rate data. The analysis of variance of the heart rate data did not disclose any significant differences between groups in either basal heart rate or in pre-during stimulus heart rate changes, although there was a non-significant decrease in rate from pre-stimulus base to stimulus for both groups. There was a significant trials effect \( F = 4.5, df = 3.66, p < .01 \) indicating that for both groups heart rate decreased over trial blocks.

Figure 9 summarizes the EEG data. Figure 9, in conjunction with a three-factor analysis of variance, indicates that neither groups, nor stimulus, nor trials had any significant effect on percent alpha rhythm.

In the control group there was suggestive evidence with the physiological measures that the degree of change or contrast determines the amount of physiological responding to the flashing light. The control subjects who received the flashing lights with the overhead light off responded with larger responses than did subjects who received the flashing light with the overhead lights on. The effect was not statistical significant because there was one reversal at the sample size was too small to tolerate such reversal.
The post-experimental questionnaire results indicated that none of the subjects in either pilot or non-pilot groups experienced any subjective feelings of nausea. Eight out of 12 pilots and even out of 12 non-pilots developed severe rowsness over the course of the experiment. Some felt that had the experiment continued longer they could not have stayed awake. There was no difference in the noisiness rating assigned to the flashing light by pilots or non-pilots. The non-pilot controls who received the flashing light under overhead light-on and overhead light-off conditions, rated the flashing light significantly more noisier when the overhead light was off ($t=5.05$, df = 11, $p < .001$).

**DISCUSSION**

In Experiment I, 13 young adult males ranging from 17–25 years of age were exposed to very intense reflected intermittent light flashes from a Grimes condenser discharge airplane anticollision light. Their EEG, eye blinks, skin potential responses, heart rate, and respiration were monitored. Over a 40-trial session, habituation (response decrement) occurred only with the skin potential response.

In Experiment II, 26 adult males ranging from 31 to 51 years of age were exposed to very tense backscatter off man-made fog produced by a Grimes condenser discharge airplane anticollision light. Brain wave activity, eye blinks, in potential responses and heart rate were continuously monitored over the course of twenty 60-second trains of light pulses. Heart rate and brain wave activity did not differentiate between groups, but significant differentiation occurred in eye blinks and skin potential measures.

In our sample of subjects neither photic driving nor EEG seizure activity occurred. These findings are in line with the results of previous experiments which indicate that for both photic driving and EEG seizure activity, the adequate stimulus is a light flashing at faster than 4 Hz.

Data of the present experiment readily fit into a framework of the habituation literature or finding that the skin potential response habituates sooner than does the occipital alpha frequency responses to intermittent light stimulation supports the findings of Sokolov and his workers.

The fact the eye blinks to the anticollision light flashes showed no decrement over trial blocks in either Experiment I or II suggests that the reflected light may have been intense enough to elicit a protective reflex. None of the commercially made photic stimulators on the market put out a light even approaching the intensity of the Grimes anticollision light. Thus, there is a scarcity of data with which to compare our findings on this point.

In Experiment II there was actually a significant increase in blinking in the pilot and non-pilot groups over trials. Since the increase occurred not only to the light stimulus but also during the pre-stimulus base level periods, it was probably not due to the effects of the light becoming more noisier over repeated presentations. A more parsimonious explanation would ascribe the increase to the humid environment. This interpretation is supported by statements of subjects from the post-experimental questionnaire.

The fact that pilots blinked less to the flashing light than during the pre-stimulus period can be ascribed to a combination of greater familiarity and possible adaptation to flashing light and to greater defensiveness by the pilots than non-pilots. Support for the latter interpretation comes from the post-experimental questionnaire. Several pilots volunteered the information that they actively resisted the effects of the flashing lights by concentrating on their great toe or on an imaginary instrument panel. Non-pilot subjects did not report engaging in such activities.

The skin potential data also offer support for the view that instrument-rated pilots were more defensive than non-pilot subjects. Whereas non-pilot subjects showed habituation of responding to the flashing light as had college students in Experiment I, the pilots by concentrating on something else while the light was flashing, were able to suppress their responses so that there was no difference in their pre-stimulus base versus during-stimulus response measures.

In our first experiment in which a younger, more responsive college student population was utilized for subjects, significant physiological changes to backscatter off a white wall from the Grimes light occurred in all measures (occipital EEG, skin potential, heart rate, respiration, eye blinks). In the second experiment, responses with occipital EEG and heart rate measures were
not significantly different although they were in the same direction as had been the case in the first experiment. A plausible explanation for this discrepancy is attributable to either (1) the fact that the intensity of backscatter from fog in the second experiment was much less than the backscatter from the white wall in the first experiment, or (2) the fact that young college-age subjects are more responsive than are older subjects.

Our finding, that none of the subjects in either experiment became nauseated, does not mean that under actual flying conditions pilots may not be affected. A potentially significant factor which was not present in our experiment but which is present under actual flying conditions is vestibular input and its interaction with visual phenomena.

Our finding that there was a consistent decrease in heart rate to the flashing light which did not habituate cannot be accounted for as an epi-phenomenon or by-product of respiration change. Since the first complete respiration cycle during light stimulation showed a consistent increase over the pre-stimulus respiration rate, then, if anything, heart rate during the light should have increased if it were positively related to respiration. Graham and Clifton in reviewing the literature on heart rate change to a variety of stimuli have concluded that the orienting response to stimuli is heart rate deceleration, a conclusion in line with our findings.

Subjects in other experiments as well as subjects in our experiments have noted that the experimental chamber appears appreciably darker just after light offset. This finding suggests that a marked degree of dark de-adaptation occurs which may lead to a brief temporary interference with seeing the instrument pane under conditions of backscatter from clouds or fog.

Conclusion

The purpose of the experiments was to determine, under controlled conditions, the effects of short, high-intensity light pulses from an actual aircraft anticollision light flashing at 1.27 Hertz on a variety of physiological measures. Post-experimental verbal reports of the subjective experience were also obtained from the participants. Three groups of subjects composed of instrument-rated pilots, age matched non-pilot medical students without flying experience were used for experiments in the laboratory of the University of Oklahoma Medical School and in the fog chamber of the FAA Civil Aeronautical Institute in Oklahoma City. The light source was a Grimes model No. 30-360-1 high-intensity condenser discharge flash tube. Heart, respiration and blink rates, EEG, and galvanic skin response were recorded. In general, heart rate and skin potential were slightly elevated in the early part but decreased in the later part of the experiments. The blink-rate responses were different for the three groups: the pilots showed decrease, and the non-pilots and medical student showed an increase in eye blinks in the course of exposure. None of the subjects demonstrated photic driving, seizure or theta-wave activity in their EEG, or nausea but experienced the backscattered light as noxious, and many became drowsy. The results bear directly on the problem of anticollision lights used during flight under backscatter conditions.
REFERENCES


1. Report No. FAA-AM-72-9
2. Government Accession No.
3. Recipient's Catalog No.
4. Title and Subtitle PILOT TRACKING PERFORMANCE DURING SUCCESSIVE IN-FLIGHT SIMULATED INSTRUMENT APPROACHES
5. Report Date February 1972
6. Performing Organization Code
7. Author(s) Paul G. Rasmussen and A. Howard Hasbrook
9. Performing Organization Name and Address FAA Civil Aeromedical Institute P. O. Box 25082 Oklahoma City, Oklahoma 73125
10. Work Unit No.
11. Contract or Grant No.
12. Sponsoring Agency Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D.C. 20591
13. Type of Report and Period Covered OAM Report

15. Supplementary Notes

16. Abstract

Eight instrument rated pilots with flying experience ranging from 600 to 12,271 hours each flew 10 simulated ILS instrument approaches in a single engine, general aviation aircraft equipped with a primary flight display arranged in a conventional "top" configuration. Continuous glide slope and localizer tracking performance were recorded during each approach. Approaches were flown consecutively at approximately ten-minute intervals, with a one-minute in-flight rest period prior to each approach.

The principal finding was that there were no statistically significant changes in glide slope tracking with respect to accuracy of tracking or consistency of performance as a result of the practice afforded by the ten approaches. Some performance measures did yield statistically significant changes in localizer tracking performance, but these changes were of a small order of magnitude. Results are discussed in terms of the factors operating to limit the precision with which the glide slope and localizer can be tracked.

17. Key Words Aviation safety, Instrument flight, Flight training

18. Distribution Statement Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.

19. Security Classif. (of this report) Unclassified
20. Security Classif. (of this page) Unclassified
21. No. of Pages 14
22. Price $3.00

Form DOT F 1700.7 (8-69)
Acknowledgment

The authors wish to extend their thanks to Mr. Richard Speck for his design and technical support on the data acquisition system.
PILOT TRACKING PERFORMANCE DURING SUCCESSIVE IN-FLIGHT SIMULATED INSTRUMENT APPROACHES

I. Introduction.

Making an instrument approach is a complex, multi-element task that requires extensive and rapid information acquisition and processing, decision making, and highly skilled psychomotor inputs to maneuver the aircraft within the constantly narrowing electronic "beam" of the Instrument Landing System (ILS) approach. It is generally accepted that such a task requires a high degree of instrument flying proficiency and that practice and currency of experience are important factors in achieving the necessary flying skills. In spite of the critical requirements on human capability for making an ILS approach, there appears to be a paucity of quantitative data defining the limits of human performance in relation to individual components of the total system requirement.

Most evaluations of pilot performance during actual or simulated instrument approaches (e.g., pilot proficiency check flights) are generally subjective and summary in nature. Except for extreme cases, it is practically impossible to rate one performance against another, or against an absolute standard, with the degree of precision or discrimination necessary to characterize subtle but real differences among individual performances. This is not to say that observational judgments are necessarily inadequate for the purpose of determining if a performance meets certain established criteria. In fact, specified flight maneuvers can be rated with a high degree of reliability between independent observers. However, such subjective judgments frequently fail to identify the extent to which the component tasks contribute to, or subtract from, the performance of the overall task. Quantitative data, such as is obtained by continuous, automated recording of selected elements of pilot performance or of flight parameters, are required for most of the research work on pilot performance, whether it is for the purpose of evaluating training techniques or designing a more effective man-machine interface.

The Control group from an earlier study provided an opportunity to study the glide slope and localizer tracking behavior of a group of instrument-rated pilots and to investigate the effects of practice afforded by an extended series of successive approaches.

The data to be reported here are not presented as being definitive since they cover only a limited number of performance parameters in one type of aircraft and for a relatively small number of subjects. It is believed, however, that the data suggest several potentially productive lines of investigation which could lead to the establishment of objective performance criteria with special reference to training, proficiency, and currency requirements, compatible control-display relationships, and the composition and structuring of complex, multi-element flying tasks.

II. Equipment and Methodology

Research Device. A four-place, single engine general aviation aircraft was used for the study. Figure 1 shows the aircraft instrument panel. It is representative of the "T" configuration frequently found in this type of aircraft except that the position of the vertical speed indicator has been moved to the right to provide a more centrally located space for the glide slope/localizer indicator. Quick-removal slats installed in the windshield area, and to the left of the subject, were used to simulate an instrument environment without obstructing the outside view of the safety pilot.

A portable, battery-powered tape recorder automatically recorded the deviations from glide slope and localizer centerlines by taking its input directly from the ILS receiver output. Event information such as passage over geographical "fixes", calibration checks, and narrative comments were supplied to the recorder by the observer who operated the equipment.
safety pilot placed the aircraft on the centerline of the glide slope/localizer “beam” at a fixed geographical point. Speed, power, gear, and flaps were set in “approach configuration” before turning control of the aircraft over to the subject. In this manner all subjects started their approaches in identical approach geometry from the same point in space. Upon completion of each approach at the middle marker, the safety pilot raised the landing gear and flaps, and instructed the subject to “go-around”. After go-around, the subject climbed back to initial approach altitude. He was then given a one-minute rest period while the safety pilot made a 180° turn in-bound to the glide slope/localizer centerline. The same procedure was repeated for each approach and the flight was terminated after the tenth approach.

III. Results.

The performance data were evaluated by several criteria that emphasize different aspects of the ILS tracking task. Only quantitative data are presented. No specific claims are made that any performance measure, or combination of measures, accurately reflect the overall “quality” of an ILS approach. The practical relevance of any given measure must be judged on the basis of how well the criteria for the individual measures reflect actual demands of the tracking task portion of the ILS approach.

All values are given in terms of the number of “dots” of deviation from the centerlines of the glide slope and localizer paths as indicated on the crosspointer instrument, except for the composite weighted scores which are expressed as percentages. Angular deviation (expressed in “dots” of instrument indication) is used, rather than linear distance, because this is the information that is displayed to the pilot and on which he must act without translation into linear distances. The relationship of indicated “dot” deviations to linear distances is also a changing ratio as a function of the distance between the aircraft and point of origin of the two signals. This makes it impractical to attempt to use linear deviation as a means of comparing performance at varying distances along the approach.

Position of Aircraft at the Middle Marker

This measure considers only the position of the aircraft with respect to the glide slope and localizer centerlines at the time the aircraft passed over the middle marker. It does not reflect performance prior to that point and gives no indication as to how well the pilot handled the aircraft with respect to such things as attitude, or rate of change with respect to the glide slope and localizer centerlines.

(1) Glide Slope Data. Figure 3 presents the arithmetic means of the deviations from glide slope centerline at the middle marker for each of the ten approaches. These values express only the magnitude of the deviations, not the direction above or below centerline. The means range from 0.76 to 2.36 “dots”. Individual performances range from “on centerline” to “off scale” (more than 5.0 “dots”). Eight of the 80 individual approaches resulted in deviations at the middle marker of more than 3.0 “dots”. Two of these larger deviations occurred as late as approach #9 in the sequence of ten approaches.

The algebraic means for the same glide slope data all fell within one dot of the centerline. Individual deviations range from 3.4 “dots” below, to off scale (more than 5.0 “dots”) above centerline. There are no significant differences among the means for the ten approaches as tested by analyses of variance.

Consistency of performance was examined by comparing the arithmetic differences between deviations on successive approaches for each subject, i.e., between approaches 1 and 2, 2 and 3, etc. The means of the differences between successive approaches range from 0.98 to 1.73 “dots” and differences by individual subjects range from zero to 4.6 “dots”. There are no significant differences among the difference values for the successive approaches as tested by analysis of variance.

(2) Localizer Data. The results for localizer deviations from centerline are presented in Figure 4. The arithmetic means range from 0.7 to 1.49 “dots” and individual deviations range from zero to 4.6 “dots” for the ten approaches. There are no significant differences among approaches as tested by analysis of variance.

Consistency of performance was examined in the same method as for the glide slope data. The means of the differences range from 0.59 to 1.09 “dots” and differences by individual subjects range from zero to 4.1 “dots”. There are no significant differences among the different values for the successive approaches as tested by analysis of variance.
Figure 3. Arithmetic means and maximum deviations from glide slope centerline at the middle marker.

Figure 4. Arithmetic means and maximum deviations from localizer centerline at the middle marker.
Range of Glide Slope Deviation Changes During the Last Thirty Seconds of Approach

This measure is a numerical expression of the range of the aircraft's vertical excursions relative to the glide slope centerline during the last half-minute of the approach prior to reaching the middle marker. It is derived by determining the number of equivalent “dots” between the highest and lowest position of the aircraft relative to the glide slope centerline during the specified period. The resulting value can be interpreted as a simple indication of the vertical stability of the aircraft’s track. The score does not reflect accuracy in tracking the glide slope centerline because a small value can indicate both large and small, but relatively constant deviations from centerline. The frequency and appropriateness of the indicated changes in deviations do not enter into the resulting values.

The means and maximum values of these changes in glide slope deviations during the last thirty seconds of the approach are given in Figure 5. With the exception of one approach, the means of the changes all fall between two and three “dots”. Individual values range from 0.7 to 7.4 “dots”. There are no significant differences among approaches with respect to this performance measure as tested by analysis of variance.

Consistency of performance was examined by taking the differences in magnitude of the range of glide slope deviation on successive approaches by the same subject. The means of the differences tend to be slightly more than the equivalent of one “dot”. The range of individual values falls between zero and 4.7 “dots”. There are no significant differences among the difference values for successive comparisons as tested by analysis of variance.

Figure 5. Means and maximum glide slope deviation changes during the last 30 seconds of the approaches.
Composite Weighted Tracking Scores

A composite weighted score was developed to provide a single overall tracking score for each of the two elements of the ILS system. These scores express the level of tracking performance or the whole approach between the outer and middle markers. The scores were derived by assigning an arbitrary value of 16 to deviations equal to, or less than, one “dot” (equivalent to taying within the “bulls-eye” or “donut”), and successively smaller values to larger deviations according to a fixed ratio as follows:

<table>
<thead>
<tr>
<th>Range of Deviations From Centerline</th>
<th>Point Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ne “dot” or less</td>
<td>16</td>
</tr>
<tr>
<td>reater than 1 but not more than 2 “dots”</td>
<td>8</td>
</tr>
<tr>
<td>reater than 2 but not more than 3 “dots”</td>
<td>4</td>
</tr>
<tr>
<td>reater than 3 but not more than 4 “dots”</td>
<td>2</td>
</tr>
<tr>
<td>reater than 4 but not more than 5 “dots”</td>
<td>1</td>
</tr>
<tr>
<td>ore than 5 “dots”</td>
<td>(-16)</td>
</tr>
</tbody>
</table>

These values were then multiplied by the number of seconds the indicated deviations fell within the respective ranges. The sum of these products was the raw score. The raw score was converted to a percentage of the maximum possible score for that approach based on the lapsed time between the outer and middle markers. Maximum performance is represented by 100% and is achieved by keeping the needle within the “bulls-eye” continuously between the outer and middle markers.

1) Glide Slope Data. The means and ranges of the composite glide slope tracking scores are presented in Figure 6. The means range from 77.6% to 89.5%. The range of individual scores is from 49.7% to 100.0%.

There are no significant differences among approaches on the basis of this criterion as tested by analysis of variance. The slight upward trend of scores during the first few approaches is significant only at a low level of confidence (p<.10).

Consistency of performance was examined by taking the difference between scores on successive

![Figure 6. Means and ranges of weighted glide slope tracking scores.](image_url)
approaches by the same subject. The means of these differences range from 2.2 to 17.3 percentage points. Individual difference scores range from zero to 36.2 percentage points. There are no significant differences among the comparison pairs as tested by analysis of variance but the trend of the scores is significant (p<.05).

(2) Localizer Data. The means and ranges of composite localizer tracking scores are presented in Figure 7. The means range from 82.1 to 94.4%. Individual scores range from 58.8 to 100.0%. The differences among approaches are statistically significant (p<.01) and the trend is essentially linear (p<.005). The analysis is summarized in Table 1.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>7</td>
<td>384.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approaches</td>
<td>(9)</td>
<td>140.41</td>
<td>2.96</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Linear</td>
<td>1</td>
<td>822.87</td>
<td>16.28</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>Quadratic</td>
<td>1</td>
<td>167.60</td>
<td>3.31</td>
<td>&lt;.10</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
<td>50.60</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>63</td>
<td>50.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consistency of performance was examined by the same method as the other scores. The means and maxima of the difference scores are presented in Figure 8. The means range from 2.8 to 13.3 percentage points. Individual difference scores range from zero to 30.1 percentage points. Differences among successive comparisons are significant (p<.01) and the trend is essentially linear. The analysis of variance is summarized in Table 2.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td>7</td>
<td>191.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approach Pairs</td>
<td>(8)</td>
<td>112.68</td>
<td>3.26</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Linear</td>
<td>1</td>
<td>490.22</td>
<td>11.83</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Quadratic</td>
<td>1</td>
<td>3.31</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>81.48</td>
<td>2.36</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Error</td>
<td>56</td>
<td>34.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Maximum Deviations from Centerlines Within Successive Segments of Approach

This performance measure examines the relative accuracy of glide slope and localizer tracking at various stages of the approach. Each approach was divided into ten equal time segments based on the transition time between the outer and middle markers. The magnitudes of the maximum deviations from centerline during each segment were used as the performance measure. Unlike the procedure for the weighted tracking scores, the duration of the deviation was not taken into consideration. Consistency by individual subjects was not evaluated for this performance measure.

(1) Glide Slope Data. The means of the maximum deviations from glide slope centerline during successive segments of the approach, together with the ranges of the means for the individual approaches, are presented in Figure 9. Data for all approaches have been combined because the individual means all fall within a relatively narrow range. The means of the maximum deviations within segments tend to be at the equivalent of one “dot” during the first half of each approach and then gradually increase to approximately twice that value as the middle marker is approached. With a few notable exceptions, individual deviations by subjects tend to be fairly consistent. There were four instances in which deviations exceeded four “dots” and one instance of an “off scale” (more than five “dots” deviation). There are no significant differences among approaches with respect to this performance measure. The differences among segments, however, are significant (p<.001). The analysis is summarized in Table 3.

(2) Localizer Data. The means of the maximum deviations from localizer centerline during successive segments of the approaches, together with the ranges of means for all approaches, are presented in Figure 10. As with the glide slope data, results for all approaches have been combined. The means tend to be less than the equivalent of one “dot” throughout most of the approach, but exhibit a gradual increase during the last three segments before reaching the middle marker. Deviations by individual subjects of 3.0 “dots” or more occurred during 800 segments. A disproportionate num
Figure 7. Means and ranges of weighted localizer tracking scores.

Table 3—Maximum Deviations from Glide Slope Centerline Within Successive Segments of Approach (Analysis of Variance)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approaches</td>
<td>9</td>
<td>59.23</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>A—Linear</td>
<td>1</td>
<td>104.50</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>A—Quadratic</td>
<td>1</td>
<td>184.28</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td>A—Other</td>
<td>7</td>
<td>34.92</td>
<td>.37</td>
<td></td>
</tr>
<tr>
<td>Segments</td>
<td>9</td>
<td>1,032.26</td>
<td>17.18</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>B—Linear</td>
<td>1</td>
<td>7,011.21</td>
<td>116.68</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>B—Quadratic</td>
<td>1</td>
<td>2,140.20</td>
<td>35.62</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>B—Other</td>
<td>7</td>
<td>15.56</td>
<td>.26</td>
<td></td>
</tr>
<tr>
<td>Subjects</td>
<td>7</td>
<td>563.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A x B</td>
<td>81</td>
<td>27.18</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>A x C</td>
<td>63</td>
<td>94.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B x C</td>
<td>63</td>
<td>60.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A x B x C</td>
<td>567</td>
<td>25.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>799</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These larger deviations were contributed during isolated instances where a single subject introduced a single large deviation over a period of several segments of a single approach. Differences among approaches are significant (<.01) as are differences among segments (p<.001). There are no significant interactions between approaches and segments. The trend of both variables has a definite linear component. The analysis of variance is summarized in Table 4.

Table 4—Maximum Deviations from Localizer Centerline Within Successive Segments of Approach (Analysis of Variance)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>A—Approaches</td>
<td>9</td>
<td>236.87</td>
<td>2.89</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>A—Linear</td>
<td>1</td>
<td>1,104.14</td>
<td>13.48</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>A—Quadratic</td>
<td>1</td>
<td>456.86</td>
<td>5.58</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>A—Other</td>
<td>7</td>
<td>80.26</td>
<td>.98</td>
<td></td>
</tr>
<tr>
<td>B—Segments</td>
<td>9</td>
<td>493.01</td>
<td>7.51</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>B—Linear</td>
<td>1</td>
<td>2,269.81</td>
<td>34.58</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>B—Quadratic</td>
<td>1</td>
<td>1,821.69</td>
<td>27.76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>B—Other</td>
<td>7</td>
<td>49.37</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td>C—Subjects</td>
<td>7</td>
<td>729.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A x B</td>
<td>81</td>
<td>24.48</td>
<td>.79</td>
<td></td>
</tr>
<tr>
<td>A x C</td>
<td>63</td>
<td>81.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B x C</td>
<td>63</td>
<td>65.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A x B x C</td>
<td>567</td>
<td>30.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>799</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IV. Discussion.

Probably the most significant finding resulting from this study is that there was practically no change in glide slope tracking performance as a function of the practice gained by the ten successive approaches. There was a small, but statistically significant, improvement in localizer tracking on two of the measures. This improvement may be attributed to increasing familiarity with the prevailing wind conditions, which turn made possible an earlier and more accurate heading correction and thus tended to minimize the effects of incipient localizer deviations. The result of this input would contribute more raising performance scores than to improving basic tracking ability.

The lack of overall improvement was expressed not only in terms of unchanged tracking accuracy
also by an undiminished variability in performance and the continuation of the instability of the track during the last stages of the approach. This evaluation does not deny the statistical significance of the differences reported in some of the localizer performance scores, but rather evaluates the magnitude of the improvement as having little, if any, practical implications from the standpoint of actual operational applications.

There are several significant factors that may have been operating to limit the effects of practice during the ten approaches. These may include, but are not necessarily limited to, one or more of the following: (1) the subjects initially had a high degree of tracking proficiency, (2) left little room for improvement in terms of increased overall accuracy; (3) system inability to introduce external variables such as meteorological conditions; (4) lack of instructional input to the subject coupled with a limited wedge of results; (5) monotony or lack of efficacious motivation; (6) the subjects may have adopted lower performance standards than they are capable of attaining.

Considering the performance data as a whole, it appears that, as a group, the subjects performed at a relatively high degree of proficiency permitted only a very limited opportunity for increased accuracy. This is perhaps best assessed by the tracking accuracy reflected by the means of the maximum deviations from centerline illustrated in Figure 9 and Figure 10. This level of tracking performance appears to be somewhat better than that generally found during actual ILS approaches by pilots of comparable experience. Unfortunately, quantitative data are not available from actual instrument weather ILS approaches for comparison purposes. If the apparent differences are accepted as real, they are readily explained by the operational differences between the approaches flown as part of this study and those encountered during actual ILS approaches, i.e., the subjects were not required to read approach plates, conduct ATC communications, or to monitor power plant performance. The initiation of the approach by the safety pilot, and the subject's knowledge that the safety pilot had constant visual contact outside the aircraft, may also have had some effect on the subject's subsequent performance. This unburdening of the pilot permitted concentrated attention to the primary tracking task with an adequate margin of residual attention for airspeed and attitude control, which, subjectively judged, were adequately accomplished.

As indicated by the ranges for the respective performance measures, the extreme values that comprise the worst performance by any subject during a given approach, are well below the group mean. Though a disproportionate number of these data points were contributed by rela-
tively few subjects, the latter by no means account for all of them. This aspect of individual performance is emphasized by the data comparing the differences on successive approaches by the same subject. The means of the differences are frequently equal to or greater than, the means of the absolute deviations from glide slope and localizer centerlines. This makes individual variability on the same order of magnitude as the absolute deviations. This variability would seem to indicate that a few instances of demonstrated proficiency may not be adequate to determine an individual’s practical performance level. It is, after all, the occasional extreme deviation from an otherwise acceptable performance level that poses the greatest threat to the safety of the pilot and the aircraft.

The fact that the group as a whole did perform at a high level of proficiency in spite of the differences in total flying experience, would seem to call for a closer examination of the present standards and requirements of proficiency and currency. Such an evaluation cannot be made on the basis of the limited data reported here. However, considering that the present standards apparently have been established on less substantive evidence, it would appear that certain economies might be realized in initial and recurrent training if more accurate data were available to define the practical limits of performance imposed by the whole man-machine system. If, for example, the level of tracking accuracy exhibited in the present study can be accepted as an approximation of an optimal level, then training and performance standards should probably be more heavily weighted toward the management of the whole approach task, rather than toward increasing accuracy beyond a practical level. It may be that improvement in overall performance should be sought by somehow achieving a better degree of repeatability or predictability of successive performances.

Of particular interest, and not unexpectedly, there is a significant increase in tracking error, as measured by the maximum angular deviation from centerlines, during successive segments of the approach. This increase in tracking error occurs primarily during the last half of the approach between the outer and middle markers. This increase is only slightly larger for glide slope tracking error than for localizer error in terms of “dots” of deviation. This small difference in indicated tracking error is particular noteworthy in that the ratio of the beam width is such as to make the glide slope indication roughly four times more sensitive with respect to linear deviations of the same magnitude at any given point along the approach. Absolute linear accuracy may actually improve even though the angular deviation increases as the aircraft approaches the middle marker. However, the pilot is required to react more rapidly to the angular deviations indicated by the crosspointer instrument in order to achieve the great degree of absolute linear accuracy required during the last stage of the approach. Since the ratio of indicated angular error to absolute linear deviation from centerlines is a constantly changing value as a function of the aircraft’s progress along the approach, it is impossible for the pilot to translate angular error to a linear deviation with any degree of practical usefulness. T

deterioration in tracking accuracy as measured by angular deviation errors is, therefore, the result of the increasing requirement for absolute accuracy as measured by angular deviation. Such demands impair the pilot’s ability to maintain consistent track within the framework of the man-machine system dynamics. This situation is further aggravated by the pacing stress imposed on the pilot. His control inputs must be continuously more rapid, but of diminishing magnitude, to correct indicated deviations of a comparable magnitude from centerlines during successive segments of the approach.

Perhaps of the greatest practical significance is the tracking behavior exhibited during the period immediately prior to reaching the middle marker. When we consider the instantaneous deviations from centerlines registered at the middle marker, we have a rough indication of the proximity of the aircraft to the desired flight path at a critical point in the approach, but such values do not reflect any indication of the ability of the track at that point in time. Examining the period immediately prior to reaching the middle marker (in this case arbitarily set at thirty seconds) it is found that there are often large excursions in glide slope deviation. These excursions must be attributed to unre

excessive, or inappropriate control inputs tended to correct what were actually relatively minor tracking errors in terms of linear displacement. The excursions frequently appear as a sequen-
of oscillatory reversals that can best be interpreted as indicating the aircraft was "getting ahead of the pilot". Such a condition occurs when proper control actions lag far enough behind the appropriate point in time to be no longer applicable and to become potentially inappropriate. Though the mean values of these excursions are equivalent to less than three "dots", a number of individual instances of more than five dots" (and one spanning more than seven "dots") were recorded. Many of these occurred over a period of only a few seconds, indicating rapid but excessive attempts at correction. Though such control actions never caused a serious threat to adequate control of the aircraft, they are indicative of an inadequate control-display relationship with respect to the dynamic characteristics of the system and the ability of the human operator to respond reliably with a high degree of precision. It is not improbable to speculate that it would be possible for the pilot of a high performance aircraft to introduce such excessive oscillatory control inputs that the aircraft might be forced outside its acceptable performance envelope with potentially serious consequences.

Considering that performance measures were effectively limited to glide slope and localizer tracking, it is possible that less obvious practice effects did in fact occur, but were not recorded otherwise observed. One possibility is that these were reflected as a performance range in a relatively consistent and controlled situation, the added practice on the tracking task may have resulted in an increase in residual attention. Such an increase could make it possible for the pilot to deal more effectively with the ancillary sub-tasks that comprise the total ILS approach or to handle the added demands of an emergency condition. There are no direct performance data to support this possibility, but art-rate data on the subjects which have been reported elsewhere may be suggestive of this. The mean heart rate for the group of subjects declined significantly as a function of the number approaches flown. This has been interpreted indicate a reduction of anxiety-induced stress. However, if a relatively low degree of residual attention in the performance of a task can be accepted as being stress-inducing, and conversely, an increase in residual attention is, therefore, likely to result in a reduction in stress, it is possible that the reduction in heart rate may in part be due to an increasing residual attention resulting from extended practice. If further investigation were to verify this assumption, it would support the desirability, stated earlier, for re-evaluating the criteria for proficiency and currency.

V. Summary and Conclusions.

There were few significant changes in tracking behavior as a result of the practice on ten successive approaches flown under the conditions of this study. Whether this overall lack of change was due to an artificial performance ceiling attributable to an initially high proficiency level, or whether the practice effects were limited by system characteristics, cannot be determined by the present data. It is suggested that a number of complex and interrelated factors may be operating.

It appears that the difficulties encountered in performing an ILS approach with a high degree of precision are related only in part to the difficulty of the tracking task involved in following the glide slope and localizer centerlines to the middle marker. Under optimized conditions, this task can apparently be accomplished with a relatively high degree of precision even though certain undesirable tracking characteristics remain operative. Under actual ILS conditions, the main difficulty may be associated with the pacing stress introduced by ancillary activities that must be performed, as well as by the psychological stress experienced by the pilot. This would suggest that once an acceptable level of tracking proficiency has been attained, further improvement of the whole task should perhaps be sought in exercising the management and allocation of attention to the component tasks.

Though the conventional crosspointer instrument can be considered generally adequate for most of the approach there appears to be an inherent deficiency in the control display ratio and relationship with respect to the requirements for rapid and precise control adjustments at the most critical stage of the approach at the middle marker.

Individual variability by subjects on successive approaches raises the question of what constitutes an adequate criterion for an acceptable level of performance. Further study will be required to determine suitable methods for increasing consistency of performance.
REFERENCES


1. **Report No.**
   FAA-AM-72-10

2. **Government Accession No.**

3. **Recipient's Catalog No.**

4. **Title and Subtitle**
   PHYSIOLOGICAL EVALUATION OF A MODIFIED JET TRANSPORT PASSENGER OXYGEN MASK

5. **Report Date**
   March 1972

6. **Performing Organization Code**

7. **Author(s)**
   Ernest B. McPadden, M.S.

8. **Performing Organization Report No.**

9. **Performing Organization Name and Address**
   FAA Civil Aeromedical Institute
   P. O. Box 25082
   Oklahoma City, Oklahoma 73125

10. **Work Unit No.**

11. **Contract or Grant No.**

12. **Sponsoring Agency Name and Address**
    Office of Aviation Medicine
    Federal Aviation Administration
    800 Independence Avenue, S.W.
    Washington, D.C. 20591

13. **Type of Report and Period Covered**
    OAM Report

14. **Sponsoring Agency Code**

5. **Supplementary Notes**
   Research leading to preparation of this report was performed under Project No. AM-A-69-PRS-13.

6. **Abstract**
   This paper describes altitude chamber experiments conducted with human subjects using a new continuous-flow disposable passenger mask applicable for emergency use to maximum altitudes of 41,000 feet. This mask design differs in configuration from the previous omni-directional masks designed to meet the requirements of National Aerospace Standard 1179 and FAA Technical Standard Order C-64 in that the inner face lap or seal has been eliminated and the cylindrical shape reduced to a modified one. The primary goal of this study was to determine if design modification of the mask induced an increase or decrease in physiological efficiency. Of paramount concern was the possibility that modification of the configuration and facial seal might increase the leakage rate of ambient air into the mask, and thereby compromise its ability to provide the level of protection required at the maximum altitude of the aircraft. Subjects were instrumented to obtain a variety of physiological data. This included EKG impedance pneumograph, ear oximetry, and expiratory nitrogen, and minute volume. The continuous flow of oxygen delivered to the passenger mask was uncontrolled and precisely measured. The mask was evaluated at altitudes of 14,000, 25,000, 29,000, 35,000 and 40,000 feet. The average inspired tracheal oxygen partial pressure remained above 83.8 mm Hg under all conditions of rest and exercise at all altitudes except for the third and last minute of exercise at 40,000 feet. If a higher degree of physical activity is to be maintained for more than a few minutes, oxygen flow rates should be increased commensurate with the anticipated level of activity. Larger subjects with increased body surface areas exhibit lower blood oxygen saturations and tracheal oxygen partial pressures at a standardized oxygen flow due to increased oxygen consumption.

7. **Key Words**
   Oxygen masks, decompression, hypoxia, altitude, passengers

8. **Availability Statement**
   Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.

9. **Security Classification (of this report)**
   Unclassified

10. **Security Classification (of this page)**
    Unclassified

11. **No. of Pages**
    11

12. **Price**
    $3.00

---

m DOT F 1700.7 (8-59)
INTRODUCTION

This report describes altitude chamber experiments conducted with human subjects using new Sierra Engineering Series 289–601 disposable passenger oxygen masks. These masks, applicable for emergency use to 40,000-foot altitudes, differ in configuration from the previous Sierra mask (although both are designed to meet requirements of NAS 1179 and FAA TSO C–64) in that the inner face flap or seal has been eliminated and the cylindrical shape reduced to a modified cone.

The unique characteristics of continuous-flow oxygen masks, in terms of human respiration, are frequently not well understood. Although less costly than crew masks and deceptively simple in appearance, continuous-flow passenger masks have physiological performance characteristics that are relatively complex.

Continuous-flow, phase-dilution type masks, signed to the NAS 1179 and FAA TSO C–64 requirements, employ a unique design in which the reservoir is interposed between the delivery valve and the mask. The reservoir is separated from the mask by a sensitive check valve. The continuous flow of oxygen fills the reservoir bag during the respiratory pause and exhalation. The flow also continues at the same rate during expiration. The mask-wearer receives 100% oxygen from the reservoir during an entire inspiration unless the bag is emptied; if this occurs, the spring-loaded valve in the mask opens, and ambient air is introduced to provide sufficient pressure to meet the remainder of the inspiration. The flow of 100% oxygen is provided at the most advantageous point in the respiratory cycle, the beginning of inspiration. For example, a human subject’s tidal volume is 650 cc, and the reservoir contains only 500 cc at the beginning of inspiration, the 500 cc of 100% oxygen will be inspired first and delivered to the active gas of the lungs. The ambient air valve will then open and deliver 150 cc of ambient air, which will enter the mouth, trachea, and other “dead” or inactive spaces of the respiratory system. Upon expiration, this dead-space air is the first to exit through the exhalation valve. This process is repeated with each respiratory cycle. In normal practice, the reservoir bags are capable of containing a maximum of 1,100 cc which provides for increased tidal and minute volumes.

In summary, continuous-flow, phase dilution masks offer the following advantages:

1. Oxygen economy is afforded by use of a reservoir bag that fills and retains the oxygen flow during both the respiratory pause and exhalation.

2. At lower altitudes, reduced oxygen flow rates may be employed and the air-oxygen dilution controlled in a manner more reliable than other methods employed in constant-flow oxygen masks.

3. Oxygen concentrations approaching 100%, required at 35,000 to 40,000 feet, may be obtained with moderate and reasonable flow rates.

One basic disadvantage of all continuous-flow oxygen systems is their inability to adjust automatically to the respiratory changes associated with changes in emotional and physical activity of the wearer. A healthy young male breathing air at rest normally exhibits a (volume/breath) tidal volume of about 550 cc and a minute volume (volume/minute) of 7,700 cc. Emotional and/or physical activity may cause these values to increase greatly. Concern with this problem is reflected in the Federal Aviation Regulations Part 25 (formerly Part 4B), 25.1443, which requires maintenance of a mean tracheal oxygen partial pressure of 83.8 mm Hg at a tidal volume of 1,100 cc, and a 50-liter BTBS (body temperature pressure saturated) minute volume for altitudes of 18,500 to 40,000 feet.
With the introduction of jet-transport passenger aircraft certified to operate at high altitudes, new oxygen systems and masks were designed and evaluated. Subsequently, standards for passenger oxygen masks were compiled and published. The National Aerospace Standard 1179 and Federal Aviation Administration Technical Standard Order (TSO) C-64 set forth manufacturing, material, and testing standards for passenger oxygen masks.

An excellent description of the basic physiological rationale of oxygen equipment design for aircraft has been prepared by the SAE A-10, Aircraft Oxygen Equipment Committee. An additional report delineates the basic criteria and design philosophy of jet transport passenger oxygen systems.

II. Methods

Procedure. The altitude chamber flight profile is shown in Figure 1. Six subjects who had received altitude indoctrination within the past two years were employed. Instrumentation is shown in Figures 2 and 3. The mask was no donned until after air-breathing baselines were established at 10,000 and 14,000 feet. A safety observer accompanied each subject.

A Waters ear oximeter, Model XE 350, was affixed to the antihelix of the subject’s ear 10 to 15 minutes prior to the chamber flight, to allow stabilization. The output of the earpiece was fed into an XE-350 oximeter and continuously recorded on a Physiograph. The latter was also used to record signals from EKG electrodes at the output from impedance pneumograph electrodes attached to the subjects. The impedance pneumograph was included in the experiments to attempt to determine if changes in the respiratory activity baseline occurred during subsequent ascent to altitude. At the present time there is no satisfactory method of measuring respiratory volumes and activity while wearing a passenger mask without compromising the performance of the mask.

After a preliminary assurance of the subject’s capability to equalize ear pressures, the subject rested quietly at 10,000 feet until the ear oximeter reading indicated a stabilization of blood oxygen saturation. Then the subject was depressurized to 14,000 feet to establish similar resting at altitude. When it appeared the blood saturation was stabilized at 14,000 feet, the subject donned crew-type demand oxygen mask and commenced breathing 100% oxygen. Immediately followi crew-mask donning, exercise on a bicycle ergometer (modified so that it could be operated each subject while seated in a passenger type aircraft seat) was initiated. The exercise level was set at 45 rpm (speed) and 45 watts (load) to obtain the desired respiratory activity (approximately 25 to 30 liters/minute). This regarded as a light to moderate work load approximately equivalent to walking at 3.0 to 4.0 mph.

Exercise was continued until the desired minute volume, as indicated by a dry gas meter, was obtained and stabilized. The subjects became denitrogenated during this period, thereby ducing the possibility of bends with exercise at the subsequent higher altitudes.

After completion of nitrogen washout at 14,000 feet, the subject removed the crew mask.
Figure 2. Instrumentation of subjects prior to nitrogen washout and determination of minute volume at an altitude of 14,000 feet. Subject wearing a crew-type mask.

...asily donned the Sierra passenger mask (Figure 3). Each subject wore a new mask; thus six masks were used in this evaluation. The flow of oxygen to the mask was regulated by an altitude-sensitive regulator of the type used in multi-passenger oxygen systems of transport aircraft. The flow from this regulator, instead of being transmitted directly to the subject, was first routed outside the chamber through flowmeter and needle-valve arrangement in order to obtain precise measurement and control of the flow (Figure 3). Excess oxygen provided by the regulator was vented outboard outside the chamber.

The subject exercised for three minutes at each altitude and rested during the climb to each successive higher altitude level. The chamber as leveled off at 14,000, 21,500, 29,000, 35,000 and 40,000 feet and readings were taken pre-exercise, during the first minute and last minute of exercise followed by post-exercise recordings...
Figure 3. Instrumentation of subjects following completion of nitrogen washout and determination of minute volumes. Subject has completed passenger mask-donning. Oxygen flow to the mask was routed from a pressure control regulator within the chamber through the passenger flow orifices outside the chamber where was reduced to standard conditions and returned to the mask. Excess oxygen from the regulator was dumped overboard.

In the evaluation of oxygen mask efficiency, one of the most important measurements is that of the partial pressures of inspired gases. A continuous flow mask, while worn by a human subject, defies direct measurement of these parameters (without compromise of mask performance) due to the rapid changes in gas composition at the facepiece. The percentage of gases in the facepiece of the mask may be averaged by an integrating reservoir; this average indicates a trend, but is influenced by the inactive gases of the facepiece and anatomical dead space. In order to estimate the composition of inspired gases an indirect approach was used; this technique is based on the assumption that the end-expiratory gases are completely mixed and have equilibrated with the blood in the alveoli.

Nitrogen is not involved in metabolic exchange. If the absorption of O₂ and the production of CO₂ were exactly the same, the amount of nitrogen inspired would equal the amount expired; i.e., nitrogen molecules inspired = nitrogen molecules expired. The metabolic respiratory quotient (R.Q.) would equal one. However, the metabolic R.Q. = \( \frac{CO_2 \text{ produced}}{O_2 \text{ consumed}} \) and is normally equal to one, or unity. Under the conditions there may be a relative difference in the nitrogen composition of inspired and expired gases. The metabolic R.Q. depends upon the predominance of carbohydrates (1.0), protein (0.82), or fat (0.71) being metabolized and usually about 0.83. The respiratory R.Q. may vary temporarily from the metabolic R.Q. during unstable states such as hyperventilation. The increased lung ventilation produces a blow off of CO₂ from the blood with an apparent,
misleading, increase in CO₂ production and an R.Q. greater than 1.0. Conversely, hypoventilation and retention of CO₂ indicate an apparent, but misleading, decrease in CO₂ production resulting in a decreased R.Q. which may be less than 0.7.

One must keep in mind that in a steady state condition the unequal exchange of oxygen and carbon dioxide involves only that portion of the gases consumed and produced. For example, if during a one-minute period at rest 0.3 liters of O₂ were consumed and 0.25 liters of CO₂ were produced (R.Q. = 0.83), the resultant volumetric difference of 0.05 liters in the seven or eight liters passing through the lungs during this one minute would be relatively small, and the error (only a few per cent) would be well within the experimental error of the determinations of end expiratory nitrogen.

All of the nitrogen which dilutes the inspired gas originates from air by mask leakage or dilution-valve activation with the exception of that derived from the tissues; after six to eight minutes of breathing oxygen under a steady state condition this constitutes less than one per cent of the lung volume.

Using calculations suggested by Laft, the admixture of air can be determined; i.e.:

\[
\frac{\text{Admixture of air}}{100} = \frac{\text{Inspired nitrogen fraction}}{\text{Nitrogen fraction of air}}
\]

By substituting end expiratory nitrogen for inspired nitrogen:

\[
\text{Admixture of air} = \frac{\text{End expiratory nitrogen}}{\text{Nitrogen fraction of air}} \times 100
\]

Using these formulas, the percentage of dilution, supply oxygen, oxygen from the ambient air, and total oxygen may be derived according to the following calculations:

\[
\text{Per Cent Dilution} = \frac{\text{End expired N₂} \times 100}{\text{N₂ of air} (79.03)}
\]

\[
\% \text{ Oxygen from Supply} = 100\% - \% \text{ dilution}
\]

\[
\text{Oxygen from Ambient} = \text{Per cent dilution} \times \text{oxygen contained in ambient air}
\]

\[
\text{Total oxygen} = \text{oxygen from supply} + \text{oxygen from ambient air}
\]

\[
\text{Calculated inspired oxygen partial pressure} = (P_o - 47) \times \frac{\text{Per cent}}{\text{total O₂}}
\]

Where: \(P_o\) = Total pressure in mm Hg at ambient altitude.

47 = Pressure, in mm Hg, of saturated water vapor at body temperature.

III. Results

The oxygen flow of the passenger mask NTPD (normal temperature pressure dry, 70° F. -760 mm-dry) and BTPS (body temperature pressure saturated, 37° C. -ambient-saturated) is shown in Figure 4. The flow to Subject 5 was increased at 29,000 and 35,000 feet during exercise due to a precipitous fall in blood oxygen saturation. Flows to the remaining subjects were adjusted to the minimal values, according to FAR Part 25.

![Figure 4. Oxygen flow rate.](image)

Heart Rate. Predictably, the mean heart rates increased with exercise, this being more pronounced in the third minute of exercise. There was also an over-all increase in mean heart rate with increase in altitude (Figure 5). These latter increases at similar exercise levels may have been partially due to subject apprehension with increase in altitude; however, the main stimulus appears to be hypoxic stress as indicated by lower blood oxygen saturations and lower oxygen partial pressures at the higher altitudes (Figures 6 and 7). At 40,000 feet, the average tracheal oxygen partial pressure dropped from 86.4 mm Hg at the end of the first minute to 77.2 at the end of the third minute (Figure 7). A typical nitralyzer record is shown in Figure 8. The oximeter earpiece used on Subject
No. 1 appeared erratic and unstable with readings, for example, of 90% blood saturation while breathing 100% oxygen at 14,000 feet. The earpiece and amplifier were replaced. The average blood oxygen saturation dropped to its minimum value at the end of three minutes of exercise at 40,000 feet (83.6%); however, re-saturation to an average of 90.8% occurred within a short period of time (30–45 seconds) following cessation of exercise (Figure 6).

**Blood Oxygen Saturation.** The National Aerospace Standard states that blood oxygen saturation baselines established at 10,000 and 14,000 feet should be obtained with the subject engaged at the same level of activity as during the evaluation at higher altitudes. Comparison as the mean blood oxygen saturation of the subjects at rest breathing air at 14,000 feet (93.8%) to that at the last minute of exercise (89.9%) indicates a difference of 3.9% which agrees with results obtained in a previous study.9 Average blood oxygen saturation levels prior to exercise exceeded the resting baseline (93.8%) at all altitudes. Post-exercise saturations were 1.3% to 4.6% lower than the above baselines.

**Figure 5.** Average heart and respiratory rates. Note the increase in heart rate from the first to the last minute of exercise at altitude.

**Figure 6.** Average blood oxygen saturation as determined by ear oximetry. Note the decrease in blood oxygen saturation from the first to the third minute of exercise.

**Figure 7.** Average inspired oxygen partial pressure. The air breathing baseline at 10,000 and 14,000 feet is merely a function of barometric pressure; i.e., (Ps–47) x 20.94% and is identical irrespective of physical activity. The combination of a crew mask and breathing 100% oxygen at 14,000 feet causes the inspired tracheal oxygen partial pressure to drop off scale to a calculated value of approximately 400 mm Hg.
The 8. Nitralyzer record of Subject No. 4 at 29,000 feet. Slope of line indicates increased dilution and reduced $P_{O_2}$ due to exercise. Numbers 0-1.0 equal 0-100% nitrogen subject to calibration correction. Nitrogen should not be confused with flow, volume or mechanics of respiration. (1) Inspiration of oxygen from reservoir. (2) Dilution valve activated. (3) Continuing. (4) Oxygen, inspired primarily at (1) above, exhaled. End expiratory nitrogen concentration.

**Oxygen Partial Pressure.** The tracheal oxygen partial pressure remained above 83.8 mm Hg in all subjects under all conditions of altitude for the first minute of exercise except for Subject 5, who experienced a drop in $P_{O_2}$ 71 mm Hg at the end of the first minute of exercise at 40,000 feet. In the third minute of exercise, the $P_{O_2}$ of this subject dropped to 70, 71, and 64 mm Hg at altitudes 39, 35, and 40 thousand feet, respectively. To the precipitous drop in $P_{O_2}$ and blood oxygen saturation at 35,000 feet, the flow was raised from 2.58 lpm to 3.10 lpm; the postexercise $P_{O_2}$ climbed to 93 mm Hg and 99%. The recovery after 1½ minutes of exercise, in which the $P_{O_2}$ and blood oxygen saturation dropped to 64 mm Hg and 80%, was accomplished at 40,000 feet by merely discontinuing exercise; this produced a rapid climb in $P_{O_2}$ blood oxygen saturation to 93 mm Hg at.

**Expiration.** Subject No. 5 was the oldest (age and the largest (225 lbs) of the subjects used in this series of experiments. His larger body surface area undoubtedly increased his oxygen consumption in comparison to the other subjects. He exhibited marked increases in respiratory rate from 14 respirations per minute at 14,000 feet (at which time the baseline minute volume was determined) to 24, 25 and 24, respectively, at 29, 35 and 40 thousand feet. A crude estimate of this effect may be made by assuming that the baseline tidal volume (determined during exercise at 14,000 feet) remained constant (1,857 cc) during subsequent ascent to higher altitudes. The indicated increase in respiratory rate would, therefore, increase the baseline minute volume from 26 lpm (BTSPS) at 14,000 feet to 44.6 lpm at 29,000 and 40,000 feet, and 46.4 lpm at 35,000 feet, an overall minute volume increase of 48% to 54% above the specified lpm. With these conditions prevailing, the basic design criterion of the mask is exceeded by approximately 50%.

Subject No. 2 exhibited satisfactory $P_{O_2}$ and blood oxygen saturation levels at the end of the first minute of exercise; however, oximetry indi-
Exercise time at altitude was held to a minimum in order to minimize the probability bends and to reduce fatigue.

Previously, high altitude evaluations of passenger masks have been carried out on resting subjects. In some evaluations, a brief episode of voluntary hyperventilation took place to increase minute volumes to 30 liters/minute. This procedure is recommended in NAS 1179, but is practically impossible for a sedentary subject to maintain this level of respiration for more than 30 to 45 seconds without experiencing severe symptoms of hypoxemia (dizziness, paraesthesia, muscular cramps, etc.).

In addition, voluntary hyperventilation reduces the alveolar and blood carbon dioxide partial pressures by the washout effect. By mere physical reduction of the carbon dioxide partial pressure, an increase in alveolar oxygen partial pressure is induced. As pointed out previously, this physiologically unsteady state may be maintained for only brief periods. Changes in blood chemistry and cerebral blood flow induced by voluntary hyperventilation detract from its usefulness in mask evaluation.

A controlled and measured work load was used in these experiments to stimulate respiration to the 30 liters/minute standard with imposing severe changes in respiratory blood-gas composition.

Admittedly, increased work load produces an increase in oxygen consumption. The level of work load used in these experiments should produce an increase in oxygen consumption of approximately 35 to 50 liters/minute above resting value.19

Neither voluntary hyperventilation nor exercise is entirely a satisfactory means of elevating the minute volume; voluntary hyperventilation tends to over-estimate the physiological effect of a mask, and exercise tends to underestimate this value. The use of exercise, however, increases the margin of safety and assures that an adequate blood oxygen saturation be maintained during periods of increased ventilation.

One disadvantage of using exercise in evaluations at altitude is the increased susceptibility to the development of bends. The depth of denitrogenation, altitude profile, and exercise time must be carefully considered in relation to the use of exercise.

IV. Discussion

The National Aerospace Standard (NAS) recognizes gas analysis and blood-oxygen-saturation determination as the two principal alternate methods to be used in altitude-chamber evaluations of passenger masks.

In this study, the experiments were designed to measure both of these variables simultaneously.
The increased minute and tidal volumes developed during exercise impose efficiency requirements of mask performance in excess of similar valuations conducted on the sedentary, resting subject.

At 40,000 feet, 8.1 liters/minute NTPD equals 8.4 liters/minute BTPS. A subject breathing 9 liters/minute will empty the reservoir bag and raw in air through the ambient air valve. The average composition of inspired gas is shown in Figure 10.

If, however, there are significant and uncontrolled openings around the periphery of the mask, ambient air may be drawn through these during peak inspiration rather than through the check valve of the reservoir bag.

The ear-oximeter determinations exhibit some degree of variability, depending upon physical activity and subject movement. Fluctuation of the ear oximeter was pronounced during resting and air breathing at 14,000 feet, became more stable with 100% oxygen and exercise at 14,000 feet.

<table>
<thead>
<tr>
<th>Altitude (ft)</th>
<th>Activity</th>
<th>RP</th>
<th>IE</th>
<th>LE</th>
<th>RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>21,500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 10.** Histogram showing dilution of inspired gas with ambient air as calculated from the end expiratory nitrogen concentrations.
feet, and again fluctuated to a marked degree at 40,000 feet on oxygen. It appeared that this "hunting" phenomenon was more marked when some degree of hypoxia was present. In general, the ear-oximeter readings appeared to be more stable during exercise than at rest.

Maintenance of an adequate blood-oxygen saturation is the desired end result. Instrumentation artifacts and variations in the physiological response of the mask-wearer may result in considerable variation in the ear-oximetry indications of blood-oxygen saturation.

The function of the mask is to deliver sufficient oxygen to produce an adequate tracheal oxygen partial pressure. Since pressure breathing is not involved in passenger systems, the mask cannot provide partial pressures in excess of those provided by a 100% concentration of oxygen. A hypothetical leak-free passenger mask, capable of delivering undiluted 100% oxygen to the alveoli throughout inspiration would therefore exhibit maximum efficiency. The resulting oxygen partial pressure, therefore, becomes merely a function of the ambient barometric pressure minus the vapor pressure of saturated gas inspired, i.e.: $P_a - P_g$.

It should be pointed out that a discrepancy exists between the performance of any oxygen equipment evaluated at constant altitude, as presented in this report, and the dynamic changes which occur during rapid decompression. Experiments have been conducted in this area by Bryan and Donaldson, but are beyond the scope of this report.

V. Conclusions

1. The average inspired tracheal oxygen partial pressure remained above 83.8 mm Hg under all conditions of rest and exercise at all altitudes except for the third (last) minute of exercise at 40,000 feet.

2. The average blood oxygen saturation during exercise remained above the 14,000-foot air breathing baseline (90.8%) during the first minute of exercise except at an altitude of 40,000 feet where it dropped to 89.6%.

3. Following three minutes of exercise, mean blood oxygen saturation dropped below the 14,000-foot air breathing control baselines at altitudes by differences of 0.7, 4.8, 5.9, and 6 per cent at 21.5, 29, 35, and 40 thousand feet respectively.

4. The average respiratory rate at the maximum altitude increased 20% during the first minute of exercise, and 37% during the last minute of exercise when compared to baseline minute volumes determined at 14,000 feet breathing 100% oxygen. Providing tidal volumes remained constant, the mean volumes were approximate 28.8 to 32.8 liters/minute. The large individual variations, however, with the 50 lpm minute volume requirements of T5 C-64 being exceeded by as much as 50%.

5. As in previous evaluations larger subjects with increased body surface area exhibit lower blood oxygen saturations and tracheal partial pressures at a standardized oxygen flow to increased oxygen consumption.

6. An ideal technique for increasing the respiratory minute volume to the levels specified NAS-1179 and TSO C-64 is not readily attainable. Originally formulated by general consensus of a number of physiologists, physicians, and engineers, these levels are representative of anticipated minute volume that a naive passenger would exhibit during the excitement of aircraft rapid decompression.

7. Voluntary hyperventilation leads to overestimation of the efficiency of a mask design whereas evaluation during exercise underrates mask efficiency within the framework of intent of NAS 1179 and the Federal Aviation Regulations.

8. If a moderate degree of physical activity is to be maintained for more than a few minutes oxygen flow rates should be increased commensurate with the anticipated level of activ
REFERENCES


EFFECTS OF ALCOHOL ON A PROBLEM SOLVING TASK

Introduction.
In a previous study (Chiles and Jennings, 70) we found alcohol (an average blood alcohol level of 105 mg%) to produce significant decrements in skills important to flying. Specifically, both monitoring and two-dimensional tracking performance suffered, although decrements were not found in the performance of a simple mental arithmetic task with subjects who were given substantial practice before testing. Such tasks were performed in different combinations which required time sharing, and we concluded that the subjects probably tended to place higher priority on the arithmetic task and thus affected their performance of that task.

Anson, Jones, Simpson, and Vega (1971) reported decrements in performance on a variety of tasks with (Breathalyzer) blood alcohol levels of 80 mg%; included were tests of short-term memory (WAIS Digit Span and dichotic presentation of digits), simple and choice reaction time, and abstracting and conceptual lity (Shipley-Hartford). However, with practiced subjects, alcohol effects were noted only on one of these tasks—3-choice reaction time.

Similarly, Carpenter and Ross (1959) reported decrements in short-term memory with an alcohol of 1 ml/kg of body weight. (Blood alcohol levels were not specified.) Again, this effect was greatly mitigated by increased practice.

The present study examines the effect of a relatively high level of blood alcohol on a simple problem-solving task. Basically, the subject required the subject to discover the correct sequence in which to push five buttons in order to turn on a green light; a red light provided feedback to the subject. The solution had to be remembered for 10 seconds and re-entered. Thus, performance of the task was primarily dependent on short-term memory and perceptual-motor speed, both of which are important elements in aviation operations.

II. Method.
A. Subjects. Thirty-one paid volunteer male subjects were tested; they were college students in their twenties who described themselves as "moderate social drinkers." The subjects served concurrently in an experimental study that involved rotational stimulation of vestibular mechanisms. (The testing schedule in that other experiment was such that the subjects could serve in the problem-solving study without additional subject costs.)

B. Apparatus. The apparatus consisted of a small metal box (approximately 11 x 18 x 9 cm) on which five push buttons were mounted in a row. The buttons were 8 mm in diameter and were spaced with 2.4 cm between edges; they were labeled with the numbers one through five. The subjects were required to make all responses with the index finger of the preferred hand. Three indicator lights (amber, red, and green) were mounted above the box on a panel in front of the subjects. The lights provided information in the following manner. At the beginning of a problem the red light was illuminated to indicate to the subject that he should begin to search for the problem solution. Whenever any button was pushed, the amber light would come on and remain on until the button was released; this light indicated that the response had been registered by the programming and scoring equipment. Coincident with the onset of the amber light, the red light would go out. When the subject released the button, the red light would come back on if the response was incorrect, but it would remain out if the response was correct. Thus, the number-one button for a given problem was the button that would keep the red light out when the button was released. If the next button pushed after finding the first correct button was the correct second button, the red light would remain out; if it were not the correct second button, the red light would be re-illuminated and...
the correct first button would have to be pushed again in order to continue the search for the second button. Similarly, if at any time during the search the red light came on, the subject had to re-enter the partial solution he had already discovered before he could continue the search for the next button in the problem sequence. Since the correct solution used each button only once, those buttons which were included in a partial solution were to be eliminated from the search. The onset of the green light indicated that the problem had been solved. Thus, if the solution for a given problem were 3, 5, 1, 4, 2, a subject who used the search sequence would enter the following sequence of button pushes (an R representing the onset of the red light): R, 1, R, 2, R, 3, 4, R, 3, 5, 1, 2, R, 3, 5, 1, 4, 2, Green Light. After a 10-second delay, the green light would go out, the red light would come on, and the same problem would be presented a second time. Thus, each problem was solved twice—once by discovering the sequence by a trial-and-error process through application of the search procedure and once by re-entering the already obtained sequence; the error light functioned the same during the second solution as it did for the first solution. The green light was also illuminated for a 10-second interval following the entry of the second solution. Thus, the subject also had to remember whether the onset of the red light meant that he was to repeat the preceding problem solution or that a new problem was present.

The subjects were required to search for the problem solution by trying buttons in a left-to-right sequence; for this purpose, the left hand button was to be considered as the extreme right hand button. This procedure was chosen because it provided an easily-learned way of simplifying and standardizing the memory load during search. It also provided a means of standardizing error detection during search; specifically, the number of errors (red lights) for a given problem can be predicted exactly for a given problem by relating the problem solution back to the search sequence. For example, the sample problem shown in the preceding paragraph should be solved with only four “errors” if the search sequence is followed properly; of course, it is possible to make fewer errors by chance deviations from the search sequence at the risk of forgetting what one has done and thus, making redundant responses.

Responses were recorded on punch tape in means of an automatic scoring system. Time at which each event occurred with respect to the beginning of the experimental session was recorded to the nearest 1/100 sec. The button pushed, whether or not it was in the correct sequence, and whether it was associated with an initial or re-entered solution were also recorded. The punch tapes were then compiled and analyzed by computer.

Four time measures were taken separately for the first and second solution phases. These were mean time per response, mean time per correct response, mean time per incorrect response, a mean time per solution. In addition, a mean time per problem with first and second solution data combined was computed.

An error measure for first solution performance was derived by taking the mean difference between the number of incorrect responses and the number of incorrect responses which would be expected if the search procedure were followed correctly. Since the subject should make any errors during the second solution, error measure for second solution was simply the number of incorrect responses.

C. Procedure. The 31 subjects were randomly divided into three groups; the control group had N of 11 and two experimental groups each had N of 10 each. One experimental group received alcohol on the first day of testing and the other received alcohol on the second day of testing. The experimental subjects were given alcohol in the amount of 2.5 ml of 100-proof Wild Turkey bourbon or Smirnoff vodka per kg of body weight. This is equivalent to slightly over absolute alcohol per kg of body weight and calculated to give a peak blood alcohol level about 100 mg%. The vodka or bourbon was diluted with orange juice to a total volume of about 30 fluid ounces. Control subjects received a placebo of equal volume with a few drops of rum extract added to simulate the addition of alcohol. The subjects were given 30 minutes which to drink the beverage. Six of the alcoholic first subjects were given bourbon and four given vodka; in the alcohol-second group, the subjects were given bourbon and six were given vodka.
Blood alcohol level (BAL) determinations were made on all subjects before they were given the alcohol or placebo. In addition, the experimental subjects had blood samples drawn for BAL determinations at one and two hours after they were given the alcohol. Testing on the problem-solving task occurred about halfway between the first and second hour BAL determination. Results of these BAL determinations are summarized in Table 1. The mean BAL cross subjects was 94.5 mg% which approximated the target figure of 100 mg%.

Table 1.—Blood Alcohol Levels in mg%.

<table>
<thead>
<tr>
<th>Group</th>
<th>1 hour after ingestion</th>
<th>2 hours after ingestion</th>
<th>Mean of 1 and 2 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>99</td>
<td>95</td>
<td>97</td>
</tr>
<tr>
<td>S</td>
<td>25</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>B</td>
<td>96</td>
<td>87</td>
<td>92</td>
</tr>
<tr>
<td>S</td>
<td>22</td>
<td>18</td>
<td>18</td>
</tr>
</tbody>
</table>

The subjects were trained and tested individually with all training and testing sessions following essentially the same procedure. During training, the task was explained to the subject and the experimenter demonstrated the solution of one complete problem (first and second solution). Then the experimenter "talked a subject through" one other problem. Additional instruction was given if necessary until the subject solved one complete problem by following the correct procedure with essentially no distance from the experimenter. The subject then allowed to work five practice problems while the experimenter observed through a half-opaque mirror from the adjoining room. Whenever necessary, the subject was reminded, via intercom, to follow the proper procedure.

The testing sessions consisted of 18 problems (six problems each). The duration of each session was dependent on the length of time it took each subject to complete the 18 problems. Prior to each testing session, the subject worked one complete practice problem and the experimenter corrected any errors in procedure. The subject was then reminded to work as quickly as possible while trying to avoid making necessary errors. The experimenter then entered the adjoining room and momentarily turned the task off while activating the scoring circuitry. The subject was alerted via the intercom system, and the task device was turned on to start the testing session. Voice contact with the subjects during the testing period was kept to a minimum.

The testing schedule for the three groups is summarized below.

1. Pre-Alcohol Day. Only the alcohol-second group was tested on a pre-alcohol day, and these subjects were not exposed to vestibular stimulation on this day. The subjects were trained on the task and, immediately following training, were tested on 18 problems, with each problem being solved twice.

2. Alcohol Day. First, subjects in all groups had pre-drinking BAL blood samples drawn, immediately after which Ss in the control group and the alcohol-first group were trained on the problem solving task. Then they were given their drink, either alcohol or placebo. One hour after ingestion, blood samples were drawn on Ss who had received alcohol and all subjects then experienced a period of rotational vestibular stimulation. Following this, (about 1½ hours after receiving their drink) all subjects were tested on the problem solving task; another blood sample was drawn after code-lock testing was completed, i.e., at two hours after they received alcohol.

3. Post-Alcohol Day. Following a period of rotational vestibular stimulation, control and alcohol-first subjects were again tested on code-lock.

All subjects were tested on the same set of 18 problems on the alcohol day; a second set of 18 problems was used on the post-alcohol sessions; the alcohol-second group was given a third set for their pre-alcohol testing.

III. Results.

The data from the experimental groups were analyzed using a Lindquist Type II design in which day of receipt of alcohol was the Latin square factor. The sources of variance in each of the 11 analyses were: day of testing (first vs. second day); alcohol (vs. no alcohol); group (order of testing with respect to the alcohol/no alcohol conditions); and subjects. The results of these analyses are summarized in Table...
2. For each measure, the means for the two levels of each of the variables and the "F" value for the test of the effect are shown.

The order of testing with respect to the day of administration of alcohol is reflected in the difference between groups in this design. In no case was that difference significant.

The practice effect was significant for all but three of the 11 measures. With the exception of those three measures, performance on day 2 was significantly better than on day 1. Two of the measures that did not show a significant difference between day 1 and day 2 were error measures, namely, errors per solution for first solution and errors per solution for second solution. The third measure for which no difference was found was the time per incorrect response for the second solution. However, this latter measure is based on a relatively small number of responses and a reduced number of subjects. Since seven of the 20 subjects made no incorrect responses during the second solutions on one or both testing sessions, the four subjects in the alcohol-first group and the three subjects in the alcohol-second group who made no errors were dropped from the analysis, along with one subject selected at random from the alcohol-second group. For the remaining six subjects per group the time per incorrect response for the second solution was based on an average of only six responses per session. For all of the other time measures, the difference between day 1 and day 2 was significant with better (i.e., faster) performance occurring on day 2. Specifically, time per correct response, time per incorrect response, and time per solution for both first- and second-solution phases of performance treated separately differed significantly for the two days of the study; time per incorrect response differed significantly for the first solution only; and time per problem differed significantly for the two solutions combined.

The main effect of alcohol was significant for only one measure taken during the first-solution phase of testing; that measure was the time per correct response. The alcohol effect was significant for three second-solution measures—time per response, time per solution, and errors per solution.

### Table 2.—Means and F Ratios for Analysis of Variance.

<table>
<thead>
<tr>
<th></th>
<th>Time Per Correct Response</th>
<th>Time Per Incorrect Response</th>
<th>Time Per Solution</th>
<th>Errors Per Solution</th>
<th>Time Per Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Solution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Day 1</td>
<td>1.35</td>
<td>1.26</td>
<td>1.46</td>
<td>23.3</td>
<td>.88</td>
</tr>
<tr>
<td>Mean Day 2</td>
<td>1.14</td>
<td>1.11</td>
<td>1.20</td>
<td>18.1</td>
<td>.50</td>
</tr>
<tr>
<td>F (df=1/18)</td>
<td><strong>15.53</strong></td>
<td><strong>17.23</strong>*</td>
<td><strong>23.83</strong>*</td>
<td><strong>16.22</strong>*</td>
<td>1.19</td>
</tr>
<tr>
<td>Mean with Alcohol</td>
<td>1.27</td>
<td>1.23</td>
<td>1.36</td>
<td>21.6</td>
<td>.97</td>
</tr>
<tr>
<td>Mean without Alcohol</td>
<td>1.23</td>
<td>1.14</td>
<td>1.30</td>
<td>19.7</td>
<td>.51</td>
</tr>
<tr>
<td>F (df=1/18)</td>
<td><strong>1.07</strong></td>
<td><strong>5.09</strong>*</td>
<td><strong>20.7</strong></td>
<td><strong>2.51</strong></td>
<td>5.09*</td>
</tr>
<tr>
<td>Mean Group A</td>
<td>1.27</td>
<td>1.23</td>
<td>1.36</td>
<td>20.7</td>
<td>.74</td>
</tr>
<tr>
<td>Mean Group B</td>
<td>1.21</td>
<td>1.04</td>
<td>1.30</td>
<td>20.6</td>
<td>.64</td>
</tr>
<tr>
<td>F (df=1/18)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Second Solution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Day 1</td>
<td>.92</td>
<td>.90</td>
<td>1.18</td>
<td>5.5</td>
<td>.46</td>
</tr>
<tr>
<td>Mean Day 2</td>
<td>.77</td>
<td>.77</td>
<td>.88</td>
<td>4.3</td>
<td>.27</td>
</tr>
<tr>
<td>F (df=1/18)</td>
<td><strong>17.53</strong>*</td>
<td><strong>15.52</strong>*</td>
<td><strong>3.84</strong></td>
<td><strong>5.95</strong></td>
<td>1.29</td>
</tr>
<tr>
<td>Mean with Alcohol</td>
<td>.89</td>
<td>.88</td>
<td>1.08</td>
<td>5.5</td>
<td>.55</td>
</tr>
<tr>
<td>Mean without Alcohol</td>
<td>.89</td>
<td>.89</td>
<td>.99</td>
<td>4.3</td>
<td>.18</td>
</tr>
<tr>
<td>F (df=1/18)</td>
<td><strong>5.91</strong>*</td>
<td>3.40</td>
<td>X</td>
<td><strong>6.39</strong></td>
<td><strong>5.43</strong></td>
</tr>
<tr>
<td>Mean Group A</td>
<td>.91</td>
<td>.89</td>
<td>1.06</td>
<td>5.4</td>
<td>.47</td>
</tr>
<tr>
<td>Mean Group B</td>
<td>.78</td>
<td>.77</td>
<td>1.00</td>
<td>4.4</td>
<td>.26</td>
</tr>
<tr>
<td>F (df=1/18)</td>
<td>1.07</td>
<td>1.99</td>
<td>X</td>
<td>1.78</td>
<td>2.83</td>
</tr>
</tbody>
</table>

* p < .05  ** p < .01  *** p < .001  X  F < 1.0
solution. Time per problem (first and second solutions combined) also showed a significant effect.

The individual cell means for the five measures for which a significant alcohol effect was found are shown in Tables 3 through 7 along with the day-1 and day-2 means for the control group. Individual "t" tests were applied to the differences between days for each of the alcohol groups treated separately. The performance of the alcohol-first group was significantly better on day 2 than on day 1 (p<.05) for five of the measures (Tables 3, 4, 5, and 7); only errors on the second solution did not differ significantly across days (Table 6). The only significant difference found across days for the alcohol-second group was in the case of the time per problem measure (both solutions combined). It can be seen in Table 7 that, for this measure, performance on day 2 (the alcohol day) was significantly superior to that on day 1 (p<.05); note also that the direction of the effect for this test is opposite that found with the overall "F" test for the effect of alcohol (Table 2). For two of the measures, the performance of the alcohol-second group was faster on their alcohol day (day 2) than on day 1, but the differences were not significant; the measures were time per correct response (first solution) shown in Table 3 and time per response (second solution) shown in Table 4. For time per solution (second solution), the mean performance on the two days was the same for the alcohol-second group (Table 7). Only the error measure on the second solution (Table 6) was in the direction of better performance on day 1 than on day 2 for the alcohol-second group, but that difference was not significant.

**Table 3.**—Cell Means for Time Per Correct Response—First Solution.

<table>
<thead>
<tr>
<th></th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Alcohol First</td>
<td>1.35*</td>
</tr>
<tr>
<td>Alcohol Second</td>
<td>1.17</td>
</tr>
<tr>
<td>Control</td>
<td>1.20</td>
</tr>
</tbody>
</table>

*The value underlined is for the session on which the group received alcohol.

**Table 4.**—Cell Means for Time Per Response—Second Solution.

<table>
<thead>
<tr>
<th></th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Alcohol First</td>
<td>1.03*</td>
</tr>
<tr>
<td>Alcohol Second</td>
<td>.81</td>
</tr>
<tr>
<td>Control</td>
<td>.91</td>
</tr>
</tbody>
</table>

*The value underlined is for the session on which the group received alcohol.

**Table 5.**—Cell Means for Time Per Solution—Second Solution.

<table>
<thead>
<tr>
<th></th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Alcohol First</td>
<td>6.6*</td>
</tr>
<tr>
<td>Alcohol Second</td>
<td>4.4</td>
</tr>
<tr>
<td>Control</td>
<td>5.4</td>
</tr>
</tbody>
</table>

*The value underlined is for the session on which the group received alcohol.

**Table 6.**—Cell Means for Errors Per Solution—Second Solution.

<table>
<thead>
<tr>
<th></th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Alcohol First</td>
<td>.66*</td>
</tr>
<tr>
<td>Alcohol Second</td>
<td>.16</td>
</tr>
<tr>
<td>Control</td>
<td>.44</td>
</tr>
</tbody>
</table>

*The value underlined is for the session on which the group received alcohol.

**Table 7.**—Cell Means for Time Per Problem—First Plus Second Solution.

<table>
<thead>
<tr>
<th></th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Alcohol First</td>
<td>32.0*</td>
</tr>
<tr>
<td>Alcohol Second</td>
<td>26.8</td>
</tr>
<tr>
<td>Control</td>
<td>29.4</td>
</tr>
</tbody>
</table>

*The value underlined is for the session on which the group received alcohol.
In addition to the above "across-days tests," individual "t" tests were applied to the differences between groups for each day treated separately. For the day-1 comparisons, significant differences between groups were found only in the case of the time per response measure (second solution). For this measure, shown in Table 4, the performance of the alcohol-first group was significantly poorer than that of the alcohol-second group (p<.05) and the control groups (p<.05). And the performance of the control group was significantly poorer than that of the alcohol-second group (p<.05). (Note that the control group underwent vestibular stimulation immediately prior to testing whereas the alcohol-second group did not). For the day-2 comparisons, no significant differences were found. However, for the errors per solution (second solution), the differences between the alcohol-second group and the other two groups approached significance (.10>p>0.05) with the performance of the alcohol-second group being inferior to that of the other two groups (Table 6). Since the control group and the alcohol-first group did not differ on this measure, these two groups were combined and an "F" test was applied to the difference between the combined groups and the alcohol-second group. This difference was significant at the .05 level of confidence.

IV. Discussion.

The significant practice effect for time but not error measures suggests that subjects followed a strategy of trying to minimize errors by holding their speed of performance to a value that would permit achievement of that goal. The subjects apparently followed the search procedure rather carefully on the first solution as seen by the very small number of surplus errors when a correction is made for the expected errors for a given problem. And the errors were even less frequent for the second solution. The error-minimization strategy is also revealed by the fact that between-subject variance accounted for about 40% of the total variance for the error measures whereas it accounted for about 70% of the variance on the time measures. This finding that individual differences on this task are measured primarily by time measures is consistent with the findings of an earlier study by Chiles and Smith (1971) in which significant correlations with the Otis test of mental abilities were found for the time measures but not for the error measures.

Although a significant main effect of alcohol was found for five of the measures, the interpretation of the findings with respect to the effects of alcohol on problem-solving performance is rather complicated. First, the fact that the study here "attached to" the vestibular stimulation study creates a problem. Specifically, in the comparisons on day 1 between the alcohol first and the alcohol second groups, the effects of the two variables (alcohol and vestibular stimulation) are confounded. The alcohol-first group was exposed to vestibular stimulation prior to testing on day 1 whereas the alcohol-second group was not. Similarly, on day 2 there is partial confounding of the two variables in that the alcohol-second group was exposed to the vestibular stimulation for the first time on the day. Second, of the five significant F-ratios for the alcohol effect, in only one case was the performance of the alcohol-second group actually poorer with than without alcohol, and even in that case (errors on the second solution), performance was not significantly poorer under alcohol.

The simple effects of alcohol on day 1 for the time per response measure (second solution) are relatively clear-cut. For this analysis, there is both a significant effect of alcohol and a significant effect of the vestibular stimulation and/or the placebo. The alcohol-second group responded the fastest, the control group (placebo and vestibular stimulation) second fastest, and the alcohol-first group (their alcohol day) responded the slowest, all of these differences being significant. The possibility that it was the vestibular stimulation that produced the significant effect on the performance of the control group makes interpretation of the day-2 findings for second solution errors rather difficult. Although the alcohol-second group differed significantly from the combined control and alcohol-first group this was the first exposure of the alcohol-second group to the vestibular stimulation. Considerin the performance of the control group, the improved performance of those subjects on day is probably reflective of practice effects but may also be reflective of some kind of adaptability to the vestibular stimulation. Thus, the greater number of second-solution errors made on day by the alcohol-second group relative to the oth
two groups may have been in part a result of an alcohol effect, in part a residual of this having been their first exposure to vestibular stimulation, and in part an interaction of the two variables. The significant differences on this measure between the alcohol-first and the alcohol-second groups on day 1 is even less clear. The control group was also poorer than the alcohol-second group but better than the alcohol-first group, though not significantly so in either case. Thus, again, the apparent effect cannot be unambiguously attributed to any one of the experimental variables.

The effect of the order of testing with respect to the administration of alcohol could not be measured with any great degree of sensitivity in this study. Since this was a between-subjects comparison, 10 subjects per group does not afford much statistical power for the direct test of the order effect. However, there were indications in the individual tests of the simple effects of alcohol and the simple effects of practice that suggest that there was in fact an interaction between the two variables. For one measure, time per problem (first and second solutions combined), the alcohol-second group performed significantly faster on day 2 (with alcohol) than on day 1 despite the fact that the overall effect of alcohol was significant and in the other direction for this measure. This suggests the possibility that subjects who are motivated to do well on a task will adjust their speed to take account of what they perceive to be possible effects of alcohol on their performance. And it also suggests that the amount of adjustment (reduction in speed) will be greater on a relatively new task (alcohol-first group) than on a task with which the subject has become somewhat familiar (alcohol-second group). It may well be that this is an important aspect of the increased errors for the alcohol-second group on day 2. They incorrectly perceived the potential for disruption caused by the alcohol and, as a result, went too fast to avoid making second-solution errors. One is tempted to conclude that their judgment was impaired by the alcohol, but that conclusion would be an unwarranted extrapolation in view of the confounding noted previously.

There are three primary mechanisms in the performance of the code-lock task that might be affected by the blood alcohol levels found with these subjects, namely, perceptual-motor skill, short-term memory, and “memory-response” encoding. Although the first solution phase of the code-lock task is ostensibly a problem-solving situation, performance generally becomes a rather simple, repetitive process that is fairly automatic when the subject has a small amount of practice. The limiting factors on performance during this phase are motor skill in pressing buttons, speed of reaction to the error feedback light, short-term memory, and memory-response encoding. Short-term memory is clearly important in re-entering partial solutions following an error. It is also important in executing the search sequence, since any button that is a member of the already-discovered partial solution must be eliminated from the search sequence.

Short-term memory is perhaps even more important in the second solution phase; the subject must remember the correct sequence discovered in the first solution phase and, after a 10-second delay, re-enter that sequence. Speed of response to the feedback lights was generally not important since the subjects tended to enter the second solution without reference to the lights. Since the lights were separated from the pushbuttons by a vertical distance of about 10 cm (4 inches), it was possible to determine by observation where many of the subjects were focusing their attention. They seemed not to be looking at the feedback lights during the second-solution phase. This can also be inferred from the fact that it was not uncommon for a subject to enter a complete 5-button sequence even though intermediate responses were incorrect.

The pattern of significant effects on the code-lock measures suggests that the most likely mechanism for an effect of alcohol was interference with short-term memory and/or the memory-to-response encoding mechanism. During the first solution phase, only time per correct response and solution time were affected by alcohol. The number of correct responses during this phase of the task is constituted primarily of responses made in re-entering partial solutions, a process that depends heavily on short-term memory. Solution time for a given problem is a function of both speed of response and the number of surplus errors, but, because relatively few surplus errors were made (an average of less than one error per problem), speed of response in re-entering partial solutions is the dominant component. Hence, short-term mem-

7
ory and/or memory-to-response encoding are also implicated for this measure.

Efficient performance during the second solution is almost entirely dependent on short-term memory and the memory-to-response encoding process. However, the occurrence of errors during the second solution could also be the result of motor involvement. Although this study does not afford a means for clearly distinguishing between these two possibilities, the separation between buttons (2.4 cm) would seem to make "aiming" a relatively small factor. Thus, we are inclined to conclude that the obtained effects are in the informational rather than the motoric domains.

Several factors that are present in the aviation environment were not included in this experiment. Three of these are of direct relevance to the interpretation of the findings in relation to flying skills. First, the problem-solving task can be thought of as involving a homogeneous set of behaviors. In contrast, the pilot is required to exercise a variety of behaviors, and, of even greater importance, he is required to time share their performance. Thus, as suggested by our previous study (Chiles and Jennings, 1970), if time sharing were added to the situation, the detrimental effects of alcohol would be expected to be enhanced. Second, the nature of the problem solving task is such that the subject can adjust his speed of response as a way of avoiding errors if he interprets the potential effects of alcohol to be likely to lead to errors; this is in fact what the subjects in this study did. In the operation of an aircraft, the pilot has rather limited freedom in making such adjustments and, in emergencies, may have no room for adjustment of his "speed of response." And, third, the pilot is subjected to a variety of potentially disturbing motions of the aircraft which, as reported by Collins, Gilson, Schroeder, and Guedry (1971), tend to enhance the effects of alcohol on performance. Factors such as these are undoubtedly the reason that Billings, Wick, Gerke, and Chase (1971) found significant decrements in pilot performance at blood alcohol levels that were substantially lower than those used in the present study.

V. Summary and Conclusions.

Three groups of subjects were tested on two separate days on a simple problem-solving task which required them to discover the correct sequence in which to push five buttons in order to turn on a green light; the correct sequence was then re-entered after a 10-second delay. One group of subjects (N=10) received alcohol on the first day and nothing the second; a second group (N=10) received nothing the first day and alcohol the second; the third group (N=11) received a placebo the first day and nothing the second. The group that received alcohol the first day and the placebo group were also subjected to vestibular stimulation on both days; the alcohol-second group received vestibular stimulation only on the second day of testing. The average blood alcohol level for the experimental groups was about 94.5 mg%.

Alcohol as a main effect was significant for five of the 11 measures analyzed; three of the significant effects found were for second-solution performance, one was for first-solution performance, and one was for the solution time per problem which included both solutions. One of the measures that was affected by alcohol was errors on the second solution; the other four were time measures. Analysis of the simple effects showed that alcohol has a greater effect during the time when the task is still being mastered than after a day's practice (18 problems) has been given. However, there is suggestive evidence from the error measure, second solution performance, that the performance of practiced subjects may also be adversely affected by alcohol. There was also suggestive evidence of a residual effect of the vestibular stimulation on the problem-solving performance of the control group.

We conclude that alcohol does have an effect on problem solving of the sort used in the present study and that the effect is greater on subject when they are less practiced than when they are more experienced on the task. We also tentatively conclude that there may be a residual effect of vestibular stimulation per se on performance even though some time has elapsed since the rotational experience. Further research is required to clarify this point, but, if the finding does prove reliable on replication, then the implications with respect to recovery times in the case of pilots could be very important.
REFERENCES


A COMPARISON OF SERUM CHOLINESTERASE METHODS: II

Crane, Charles R., Ph.D., Sanders, Donald C., M.S., Abbott, John K., B.A.

FAA Civil Aeromedical Institute
P. O. Box 25082
Oklahoma City, Oklahoma 73125

This work was performed under Task No. AM-B-72-PHB-19.

Among aerial applicator personnel, the primary value of periodic blood cholinesterase (ChE) assays is the detection of pesticide poisoning indicated by a decrease in enzyme activity since the previous (or pre-season) assay. Comparison of these values is difficult if they are based on different methodologies and expressed in different units, which is frequently the case. This report provides an evaluation of four serum or plasma ChE methods which are currently in use and establishes the relationships for interconversion among their respective units. This was accomplished by performing simultaneous assays by each method on a series of samples whose activities covered the range from normal levels down to the very low level found in organophosphate-poisoned individuals. The resulting conversion data (regression equations) are also compatible with those described in Report No. AM-70-13, thereby providing direct interconversion capability among the units of seven commonly-used serum ChE methods.
A COMPARISON OF SERUM CHOLINESTERASE METHODS: II

. Introduction.

The utility of measurement of the activity of the acetylcholine-hydrolyzing enzyme present in pooled plasma as a diagnostic tool in organophosphate poisonings and for monitoring the healthfety aspects of aerial applicator operations has been previously discussed.¹

The activity of this enzyme in unexposed individuals varies widely, with the normal range extending above and below the population mean as much as 50%. Therefore, any such value obtained after potential contact with an anticholinesterase should be compared to the individual's normal, pre-exposure value to determine whether inhibition is present. Such a comparison can be made only if both values are expressed in the same units. Unfortunately for the purpose of comparison, the various methods of cholinesterase (ChE) assay produce dissimilar units, and no single "standard" method exists; consequently, a means of converting activity units obtained by one analytical method to those of another is required.

This paper represents the second in a series which will provide conversion data to transform the activity units obtained by one method into the corresponding units of other methods. Most of the methods developed primarily to measure the activity of the pseudocholinesterase present in a fluid portion of blood, including those compared in this study, customarily utilize serum for the determination. However, plasma can be used for the measurement, with identical results, provided that the presence of fibrinogen does not produce interference. There is a distinct advantage in performing the assay on plasma if the method permits, since it allows the measurement of erythrocyte enzyme activity on the same blood specimen. Therefore, plasma was utilized in this study.

The procedure used to develop these comparison data was to perform simultaneous assays by each method on aliquots of the same sample of plasma, using the pH-Stat method as a primary reference. Heparinized, whole blood was purchased from a local blood bank. Part of the separated plasma was heated in a water bath at 60°C until pH-Stat assay indicated a loss of about 95% activity (10-20 minutes). The inactivated plasma was then mixed in varying proportions with fresh plasma to obtain a gradient of enzyme activity.

II. pH-Stat Reference Method.

*. Materials and Method. Butyrylcholine iodide (Mann Research Laboratories, New York, N.Y.) was dried overnight in a vacuum desiccator and stored therein until used. An aqueous solution, 0.163 M (49 mg/ml), was prepared weekly and stored at 4°C. The sodium hydroxide titrant was prepared at a concentration of approximately 0.005 N using CO₂-free distilled water. The solution was stored in the titrant reservoir of the pH-Stat and protected from CO₂ absorption by a drying tube filled with Mallscorb, 30-50 mesh, indicating type CO₂ absorbent (Mallinckrodt). This solution was standardized daily by titration of a potassium hydrogen phthalate standard. Ringer's salts solution for diluting plasma samples was prepared as needed using 9.00 g NaCl, 0.20 g KCl, 0.26 g CaCl₂, and 0.20 g NaHCO₃, brought to a 1 liter final volume with distilled water.

A Radiometer pH-Stat (Radiometer A/S; Copenhagen, N.V., Denmark) was used for all constant-pH titrations, and consisted of the following units:

PHM26—expanded scale pH meter
TTT11b—electronic titration control unit
ABU1b—semi-micro automatic buret for titrant delivery, 2.5 ml capacity
TTA3—glass-jacketed reaction vessel, 3 ml, with plastic-coated magnetic stirring disc
VTI13c—constant-temperature water bath and circulator, set to maintain 25°C in reaction vessel
SBR2c—recorder for indicating volume of titrant delivered per unit time

¹
Aliquots of samples were diluted 1:10 with Ringer's salts solution. One ml of the diluted plasma was added to the reaction vessel, the magnetic stirrer was started and the solution was allowed to come to temperature equilibrium (25° C.). The pH was adjusted with the automatic titrator to the pre-set value of 8.1. One-half ml of butyrylcholine iodide substrate was added to the reaction vessel using a 1.0 ml tuberculin syringe and the recorder chart drive was turned on. The reaction was allowed to continue until the recorder had plotted a straight line (titrant volume vs time) for a minimum of 3 minutes, or until a minimum titrant volume of 0.5 ml had been added.*

Calculations:

All activities were expressed as micromoles of substrate hydrolyzed/minute/ml of plasma. Calculations were made as follows:

$$\text{Activity} = \frac{1000 \times (0.025) \times (N) \times (1000) \times (D.F.)}{S.D. \times t}$$

Where

- $S.D.$ = recorder scale divisions (linear portion)
- $t$ = time in minutes for linear portion of trace
- 1000 = μmoles/m mole
- 0.025 = ml of titrant per S.D.
- $N$ = normality of titrant
- D.F. = dilution factor (i.e., 10)

III. Michel Method.

The Michel method is an electrometric determination based on the measurement of the acid produced by the action of ChE on acetylcholine.² The acetic acid produced is measured in terms of the change in pH produced in a buffer-substrate mixture during a specific incubation period. The decrease in ChE activity with decreasing pH is taken into consideration by selecting a buffer whose capacity closely parallels the activity loss over the range from pH 8 to pH 6; the buffering effect of plasma is minimized by sufficient dilution (1:111). As performed in this laboratory, the basic method of Michel² as modified by Larson³ was used.

**Materials and Method.** Acetylcholine perchlorate (K & K Laboratories, Plainview, N.Y.) was dried overnight in a vacuum desiccator over Drierite and stored therein until used. The substrate was prepared by dissolving 0.405 g of the salt in a final volume of 10.0 ml of water. This solution is stable for up to four weeks if refrigerated.

A buffer stock solution was prepared by dissolving 4.124 g sodium barbital, 0.545 g KH₂PO₄ and 44.730 g KCl in distilled water and diluting to a final volume of 200 ml. Immediately before use, the plasma buffer was prepared by diluting 6.4 ml of the stock solution with 75 ml of distilled water; the pH was adjusted to 8.00 with 0.1N HCl and the solution adjusted to a final volume of 100 ml with distilled water.

Plasma aliquots of 20 μl were withdrawn using a Sahl pipet and expelled into 12-ml test tubes containing 1.0 ml of distilled water. One ml of buffer solution was added and the mixture was allowed to equilibrate in a 25° C. water bath for 10 minutes. After recording the pH to the nearest 0.01 pH unit, 0.2 ml of substrate was added and the tubes were allowed to incubate at 25° C. for one hour. The pH was read again and the change in pH/hr was calculated. Control tubes containing only distilled water; buffer and substrate were treated similarly and the change in pH in the control tubes was subtracted from the values obtained from plasma to correct for non-enzymic hydrolysis.

Calculated Michel units ($\Delta$ pH/hr) were plotted against pH-Stat units obtained for the same plasma samples assayed simultaneously. The linear regression equations and correlating coefficients were calculated, each point representing the mean of duplicate assays of the same sample by each method. The statistical regression of the pH-Stat units on Michel units graphically represented in Figure 1.

![Figure 1](https://via.placeholder.com/150)

**Figure 1.** Correlation between pH-Stat units and Michel units. *N*=20 points; correlation coefficient =0.986

---

*Since full scale on the recorder (100 scale divisions) represents 2.5 ml of titrant, 0.5 ml is represented by 20 scale divisions. The recorder trace can be read to 0.2 scale division; therefore, the precision in reading titrant volume is 1%. With a chart speed of 2 cm per min, the precision in reading elapsed time of 3-min trace is also 1%.
IV. Pfizer ChE-tel Method.

The Pfizer method is a colorimetric determination based on Garry and Routh's modification of the earlier work of Ellman. In this reaction, the artificial substrate, propionylthiocholine, is hydrolyzed by cholinesterase and the released thiocholine displaces the yellow anion of 5-thio-2-nitrobenzoic acid from 5,5'-dithiobis-2-nitrobenzoate (DTNB). The rate of color production is measured at 410 nm. Quinidine sulfate, an inhibitor, is used to stop the reaction at the end of the incubation period.

Materials and Method. All materials used for the assay were supplied in the Pfizer 5360 ChE-tel Test Set (Pfizer diagnostics, 300 West 3rd St., New York, N.Y. 10036) and the method outlined supplied with the kit was followed. Spectrophotometric measurements were performed using a Beckman DU monochromator attached to a Gilford Model 222 photometer and power supply.

Individual assays were performed by adding 0.4 ml of plasma to a tube containing 4.0 ml of 0.1% DTNB (0.05M Tris buffer, pH 7.40, with ionic strength increased to 0.153 with NaCl) previously equilibrated at 37° C. With the tube immersed in the incubation bath, 0.50 ml of substrate (propionylthiocholine diode, 0.028M) was added and mixed. After exactly 3 minutes incubation, 1.0 ml of quinidine sulfate (0.5% w/v) was added to stop the reaction and the absorbance was read at 410 nm. Zero absorbance calibration was provided by a blank prepared exactly as described for the sample except that the quinidine sulfate was added prior to the substrate.

ChE-tel units were calculated by multiplying the absorbance by 120, the conversion factor for 1-mm cuvettes. Calculated ChE-tel units over the range of 9.5 to 90 were plotted against pH at units obtained for the same plasma samples assayed simultaneously. The linear regression equations and correlation coefficients were calculated, each point representing the mean of duplicate assays of the same sample by each method. The regression of pH-Stat units on ChE-tel units is shown in Figure 2.

Caraway Method.

The Caraway method is a colorimetric procedure utilizing the acid-base indicator, phenol red, to measure the reduction in pH which accompanies the hydrolysis of acetylcholine. The phenol red absorbance is decreased as the acetic acid produced lowers the pH of the buffered reaction mixture. The ratio of the final to initial absorbance in each sample is compared to a standard curve, similarly prepared by substituting known quantities of acetic acid for the substrate. The method requires equipment ordinarily possessed by a clinical laboratory and the reagents are relatively stable.

Materials and Method. Acetylcholine perchlorate (K & K Laboratories, Plainview, N.Y.) was dried overnight in a vacuum desiccator; a 0.663M solution was prepared by dissolving 4.074 g of the reagent in distilled water and diluting to 25 ml. The buffer-indicator mixture was prepared by grinding 100 mg phenol red with 28 ml of 0.01N NaOH until dissolved; then 10 ml of the phenol red solution was added to a solution containing 6.85 g of anhydrous Na2HPO4 and 0.45 g of anhydrous KH2PO4, and diluted to a final volume of 1 liter with distilled water.

For each assay, duplicate tubes containing 0.5 ml of plasma and 5.0 ml of the buffer-indicator solution were allowed to come to temperature equilibrium for 10 minutes in a 25° C water bath. The substrate was added (0.5 ml), mixed,

![Figure 2. Correlation between pH-Stat units and Pfizer ChE-Tel units. N=25 points; correlation coefficient = 0.980](image)

3
and the absorbance was determined at 540 nm on an instrument zeroed with a distilled water blank. After 30 minutes of incubation, the absorbance was determined a second time and the ratio of the final absorbance (A₂) to the initial absorbance (A₁) was determined (i.e. A₂/A₁). Two control tubes, with 0.5 ml of distilled water substituted for the plasma, were prepared and incubated, then the A₂/A₁ ratio was obtained as before. The average value of A₂/A₁ for the control tubes was subtracted from the ratio for each test replicate to correct for non-enzymatic hydrolysis.

A standard curve was prepared by reading the absorbances of a series of eight tubes, each containing 5.0 ml of buffer-indicator solution, 0.5 ml of pooled plasma, and 0.5 ml of 0.01N-0.15N acetic acid (corresponding to 10-150 Caraway activity units/ml) and plotting the ratios of each of these absorbances to the absorbance of a tube similarly prepared but containing no acetic acid. These ratios were plotted on the log axis of semi-log paper against the corresponding Caraway units on the linear axis and the sample activity was read directly from the curve.

Calculated Caraway units were plotted against the pH-Stat units obtained for the same plasma samples assayed simultaneously. Regression equations and correlation coefficients were calculated, with each point representing the mean of duplicate assays by each method. The regression of pH-Stat units on Caraway units is depicted in Figure 3.

![Figure 3. Correlation between pH-Stat units and Caraway units. N=20 points; correlation coefficient = 0.987](image-url)

VI. Acholest Method.

The Acholest method is a simple "test-paper" procedure of the type used for field screening where multiple cases of suspected pesticide exposure have occurred. The method is based on the time required for the enzyme activity of a plasma sample to change the color of a piece of filter paper impregnated with buffer, substrate, and pH indicator to match a filter paper similarly treated but containing no substrate. Under stable temperature conditions, we have found the results from this method to correlate well with those obtained by more sophisticated techniques. The principal advantages are inexpensive materials with long storage life and very simple equipment requirements. These factors might induce some clinical laboratories having few requests for ChE assays to maintain this capability.

*Materials and Method.* All materials used for the assay were supplied in the Acholest Test Paper Kit (donated by E. Fougera & Co., Inc. Hicksville, New York, Distributor) and the modified procedure described by Braid and Nix was followed. These modifications consisted of cutting both test and control strips into 1/4-in. discs with a paper punch prior to use and utilizing an incandescent light source for viewing the color change against a dull-black background.

Individual plasma samples were assayed by placing six 7-μl drops on a clean microscope slide; three test discs and three control discs were lowered onto the drops with tweezers and the timer was activated. A second slide was lowered onto the first and pressed gently to assure saturation of the discs with the plasma. The time required for the color of the test discs to match the color of the control discs was recorded. All tests performed during the unit conversion study were at 25°C.

Reciprocal Acholest units (i.e., 1/minutes of color match) were plotted against pH-Stat units obtained by simultaneous assays of the same samples. Regression equations and correlation coefficients were calculated, each point representing the mean of duplicate assays of the same sample by each method. The regression of pH Stat units on Acholest reciprocal units is shown in Figure 4.
c. To convert Pfizer (ChE-tel) units to pH-Stat units:
PfH-Stat units = (0.08)(ChE-tel units) + 0.30

d. To convert pH-Stat units to Pfizer (ChE-tel) units:
ChE-tel units = (12.07)(pH-Stat units) - 1.83

e. To convert Caraway units to pH-Stat units:
pH-Stat units = (0.092)(Caraway units) + 0.306

f. To convert pH-Stat units to Caraway units:
Caraway units = (10.67)(pH-Stat units) - 3.15

g. To convert Acholest units (minutes to color match at 25°C.) to pH-Stat units:
pH-Stat units = (44.8/T) + 0.827 where T = minutes to color match by the Acholest method

h. To convert pH-Stat units to Acholest units:
Acholost units = (81.4/pH-Stat Units) - 5.19

Conversion equations for the relationships between the units of the Michel, Pfizer, Caraway, and Acholest methods may be readily derived from the preceding equations. Also, since the same pH-Stat primary method was used in an earlier paper to interconvert the units of the Sigma and Boehringer methods, these two papers provide the clinical worker with the necessary data to effect a conversion of units between any of seven commonly used serum (or plasma) ChE methods.

Thus, it is possible to directly compare pre-exposure ChE units with post-exposure units, even when these values were determined by different techniques. Such a comparison can frequently identify a case of low-level chronic ChE inhibition which would be undetectable by comparing a single ChE value with the broad normal range of values for a specific method. With direct interconvertibility of units, any one of these methods best suited to the resources and facilities of the laboratory can be used for the medical monitoring of aerial applicator personnel or for confirmatory diagnosis in acute organophosphate poisoning. Regardless of the method selected, however, each aerial applicator should have a pre-exposure test from which any subsequent depression of ChE activity can be measured.
REFERENCES


