PERFORMANCE TASKS FOR OPERATOR-SKILLS RESEARCH

Richard G. Pearson, Ph. D.

Approved by

J. ROBERT DILLE, M.D.
CHIEF, CIVIL AEROMEDICAL INSTITUTE

Released by

P. V. SIEGEL, M.D.
FIELD AIR SURGEON

June 1966

FEDERAL AVIATION AGENCY
Office of Aviation Medicine

FOREWORD

This report describes the characteristics and uses of several research tasks developed as part of a capability for conducting experiments involving skilled-operator performance. In this context, the success of the project scientist's efforts at times depends greatly upon the design and reliability of the equipment with which he and his subjects are involved. The author has been fortunate in this regard to have at his disposal commendable engineering, electronics, mechanical, and shop services, involving consultation, design, development, maintenance, and fabrication. Development of the tasks described was a product of these services and the individuals who rendered them deserve recognition: Marion C. Oviatt, Harlie W. Huffman, Francis E. Anderson, and Armand C. Schmidt, all from CAMI's Research Engineering Service; and Clifton E. Hunter and Don P. Harris, from the FAA Depot.

PERFORMANCE TASKS FOR OPERATOR-SKILLS RESEARCH

I. Introduction.

Engineering psychology as a discipline is concerned with the efficient and optimal combination of men and machines in both normal and unusual operator environments. In civil aviation, the focus of its interest is upon pilots, other airborne personnel, air traffic controllers, and other ground personnel—taking into account the capabilities and limitations of the skilled operator, the equipment with which he must interact, and the environment in which he must work. With the establishment of a new research facility and its associated mission, there existed a need for performance tasks that would permit study of problems associated with maintenance of operator proficiency, whether these problems reside in the operator, the task, or the work environment. This report describes the selection, development, and operation of several tasks that, in support of these needs, are now part of the engineering-psychology research capability of the Office of Aviation Medicine.

The tasks described are intended, collectively, to sample a broad spectrum of abilities required by complex operator systems and, individually, to be a challenge to the experimental subject. Some were developed to simulate closely real-life tasks, while others are of a more abstract nature. The tasks also vary to the extent with which demands are placed upon various levels of functioning of the human organism; e.g., attentional, sensory input, central processing and decision, and motor output.

II. Tracking and Monitoring.

The tracking and monitoring task (TMT–I) was developed to simulate, in part, some of the functions associated with flying skill. Figure 1 (all figures are in the appendix) shows the subject's station, which includes, in addition to the main tracking-monitoring task, a capability for assessing either choice reaction time or simple mental-arithmetic ability during rest periods.

These auxiliary tasks will be described in following sections. The principal (central) task involves two-dimensional, compensatory tracking of a randomly moving target appearing on a 5-inch diameter display. Specifically, the operator's task requires centering a target spot on cross-hairs by moving a small spring-loaded joystick. Concurrently with the tracking task, two subtasks, meter and warning-light monitoring, can be programmed to occur at irregular intervals during a given trial. The purpose of these subtasks is to introduce some complexity into the test situation by requiring the operator to time-share his visual tracking, scanning, and peripheral-vigilance activities.

Taken together with the auxiliary tasks, the attraction of the TMT–I system for operator-skills research is that a number of diverse human functions—simple motor responses, psychomotor coordination, monitoring, peripheral attention, and problem-solving—can be evaluated quickly and repeatedly at a single station. Consequently, a profile of performance effects can be obtained as a function of time on the task under the experimental treatments of interest.

A. *Display.* The tracking-task display is provided by a Hewlett-Packard Model 120B oscilloscope equipped with P^{31} phosphor and an internal graticule to eliminate parallax. For testing under dark conditions, the face of the CRT can be illuminated with black light. The table surface is $27\frac{1}{2}$ inches above the floor, and the center of the CRT is 13 inches above this. The display panel is situated 12 inches back from the front edge of the table.

B. *Control.* The joystick assembly is Raytheon Type FA–7553, obtained as surplus RBDE–5 equipment. Handle length is $4\frac{1}{2}$ inches and diameter is 1 inch. Maximum stick displacement is thru an arc of approximately 100°. As noted in Figure 1, the control is offset 3 inches to the right of a line through the center of the CRT. Potentiometers attached at the point of rotation of the X and Y axes of the joystick yield a voltage in-

dication of the control position coordinates. Initially, voltage to the control system was supplied by two 1.4-volt mercury cells, later superceded by a small regulated DC power supply.

C. *Subtasks.* The meter-monitoring subtask requires the operator to detect needle deflections of each of two panel meters, one located on each side of the CRT display. Tips of the needle pointers are located 5 inches from the center of the CRT display. If a movement of a needle from a null position on either meter is observed, the operator must respond with his left hand by pressing the appropriate push-button (Micro Switch Series 2) located at the lower left of the panel. Timely responses null the needle; if, however, the deflection goes undetected, the needle is automatically nulled within the time frame of the cam program described below. Detections within this time frame are "rewarded" by the onset of an amber lamp located above each meter.

The warning-light monitoring subtask requires the operator to detect the onset of a yellow visual indicator located at the extreme lower left-hand corner of the panel, 14 inches from the center of the CRT display. Operator response, which extinguishes the light, requires depression of a push-button located in the top of the control stick. As in the case of the meter subtask, the cam program automatically extinguishes the light after a period of time if the event goes undetected.

To provide for variety in the occurrence of subtask events, a combination of cam-timer and stepping-switch circuitry was designed involving 10 different subtask schedules, or test trials, as shown in Table 1. With each new trial, the advance of the stepping switch selects a different combination of cams to control the pattern of events occurring over time. The program shown provides for 3 meter events and either 0 or 1 warning-light events for a given trial. Trial events are programmed to occur at irregular intervals during a 2-minute task trial. Allowing for a 1-minute rest period between trials (which has been our practice) a specific pattern of subtask events does not repeat itself within 30 minutes.

D. *Task Load.* The task load provided by the subtasks is, of course, secondary to that of the forcing function input to the central tracking task. This latter input is provided by movement of two target function cams that vary the voltage input to the X and Y terminals of the oscilloscope by electromechanical means, thereby determining the position of the target on the CRT display. The two irregular cams functionally represent, respectively, the mixture of sine waves of 3, 5, and 8 cycles per minute and 3, 8, and 10 cycles per minute. The motors driving the cams operate at different speeds, however, so that the former cam pattern runs its course in 120 seconds, the latter in 160 seconds. Maximum target excursion on the display was set at 4 cm from center; limiters prevent the target from disappearing off the display in the event of an inappropriate control response.

E. *Scoring.* For the tracking task, voltages defining the position of the target on the display (i.e., input function minus joystick voltages) are fed into a PACE TR–20 analog computer, which integrates the absolute error over the task trial. Readout appears on a Non-Linear Systems Model 4700 digital voltmeter. Computer wiring is in accord with the recommendations of Frost.[1] Scoring of performance by the computer is initiated 10 seconds following the start of a trial (to allow for operator warm-up effects) and ceases 5 seconds before the end of a trial. Subtask responses are timed with Industrial Timer Corporation Model SC–100 stop clocks in 0.01-second units. Clock reset is push-button, solenoid-actuated.

F. *Typical Operation.* Figure 2 shows the experimenter's station, which includes a task-programming and scoring control console (left), the analog computer, and digital voltmeter and 48-volt DC power supply (right). A 20-cam Industrial Timer Corporation Model RC–4 timer located inside the control console accomplishes several functions: (a) Activation of buzzers that indicate "ready," "start," and "cease" tracking-task status, (b) control of analog computer "hold," "integrate," and "reset" modes; (c) blanking out of the CRT target during rest periods; (d) operation of tracking-target function cams; and (e) activation of the monitoring subtasks and associated stop clocks. Separate circuits permit recycling of the tracking-target function cams and resetting of the stepping switch that controls the subtask event program. Figure 3 is a block diagram depicting more clearly the circuitry involved in task control and scoring.

Checkout procedure prior to daily testing includes balancing of the analog computer and centering of the CRT target on the cross-hairs of the display with all inputs balanced. With extended testing of 2 to 10 hours, routine hourly calibration

Table 1. Eight-cam subtask program

Cam No.		1	2	3	4	5	6	7	8
Time of event (sec)		13	21	39	60	69	81	96	101
Stepper position	1	L+		W		L−		R−	
	2		R−		R+		W		L−
	3	L+			R+				L−
	4		R−			L−	W		L−
	5		R−	W		L−		R−	
	6	L+			R+		W		L−
	7		R−			L−		R−	
	8	L+		W	R+				L−
	9	L+			R+		W	R−	
	10		R−	W		L−		R−	

Code: L=left meter; R=right meter; W=warning light; (+)=needle moves upward; (−)=needle moves downward.

of the CRT target is practiced to compensate for oscilloscope amplifier drift.

Approximately 2½ hours of training appear to be required on the TMT task for subjects to reach a relatively stable baseline of performance and demonstrate subtask integration. Figure 4 indicates the performance means for a group of 30 Agency test pilots. On Day 1 they received 45 minutes practice, followed on Day 2 by 150 minutes of continuous practice with knowledge of performance given about every five trials (15 minutes). Following initial training, college students typically show small, but steady, improvement with 15 to 30 hours experience at the task.

Experience with the TMT–I indicated some desirable modifications that have been incorporated in a new model, the TMT–II, shown in Figure 5. The display in this case is provided by a Hewlett-Packard Model 130B oscilloscope equipped with P–7 phosphor. Vertical-movement, edgewise-type, panel meters are used, and detections of needle movements are acknowledged by illumination of the associated push-button that must be depressed. The TMT–II does not provide for the mental-arithmetic auxiliary task. A control has been added at the experimenter's station to allow selection of three variations in speed of the motors driving the tracking-target function cams. This provides two faster forcing functions for more stressful task loading.

Stop clocks in the new TMT–II scoring console (Figure 6) are Standard Electric Co. Model S–1 with manual, lever reset. Analog-computer circuitry is being assembled, at the time of this writing, from three Philbrick Model MP solid-state operational amplifier manifolds containing as a system ten P65AU and two P25A amplifiers, two PSQ Quadratic Transconductors, and power supply.

III. Choice Reaction Time.

This task, located at the subject's station of both the TMT–I and TMT–II, is the traditional two-choice reaction time. It requires the subject to respond as rapidly as possible to the onset of a red or green signal by hitting the telegraph key located immediately below the associated indicator light (see Figures 1 and 5). For a given trial, the subject must rest his right hand on a small metal plate imbedded (over a small, sensitive miscroswitch) in the console-table surface. This illuminates a "ready" indicator light at the experimenter's control console. The experimenter then gives a "ready" command (usually over a microphone-headset hookup) and, following a delay of 2, 3, or 4 seconds (selected at random by the experimenter), depresses a push-button switch that triggers the signal light and scoring circuitry. Circuitry is depicted in the block diagram (Figure 3). In the TMT–I, two measures are obtained directly, total RT, and decision RT, which is the time it takes for the subject's hand to leave the ready pad. This latter time may be presumed to reflect the time required for sensory reception, central identification, and neuromuscular command. A third measure, motor RT, is derived from the difference between total and decision RT's, and reflects the hand-movement time required to bridge the 9-inch gap from pad to response key. The stop clocks involved are those employed in the TMT subtasks. In the TMT–II, all three measures are obtained on separate clocks.

Three RT trials can normally be obtained during a 1-minute period. We have followed the practice of obtaining our RT data from two successive rest periods during a 15-minute period of performance (five trials) at the TMT apparatus, and use the median of six scores as the index of behavioral efficiency.

IV. Mental Arithmetic.

This task is located at only the TMT–I subject's station. It requires short-term memory store and involves addition of two 2-digit numbers and subtraction of a third 2-digit number (e.g., 22 + 19 − 17) that appear in succession on adjacent visual readouts (IEEE Series 10,000) on the panel (Figure 1). The subject "accepts" each of the first two temporally separate inputs by depressing the associated push-button as each readout is illuminated in succession. With the appearance of the third input, the subject is then required to punch his answer as rapidly as possible on the panel-mounted keyboard and indicate the certainty of his response by depressing a "check-off" push-button. As indicated by the block diagram of the circuitry (Figure 7), scoring begins with presentation of the third input and ends with the "check-off" response. Use of the check-off button avoids the problem encountered with many subjects who treat the task in the manner of a reaction-time problem and, accordingly, make frequent errors in punching the keys.

In practice, the experimenter advises the subject of the start of a series, usually giving four problems during a 1-minute "rest" period. A buzzer sounds with the appearance of each number and ceases with the subject's response. Both response errors and speed are scored.

The digital readouts for this application have a range of 12 numbers, 11 to 22, which provides a relatively simple task. With undergraduate college students, however, considerable individual differences have been observed, and so to date the sensitivity of the task to experimental treatment has not been demonstrated. Paper-and-pencil practice tests, incidentally, have confirmed this problem. For some subjects a strong practice (relearning) effect is evident.

V. Problem Solving.

Circuitry described by A. T. Welford [2] was followed in the construction of the apparatus shown in Figures 8 and 9. Basically, the task requires the subject to make a key-pressing response (from 1 of 12 keys) to a light appearing in a bank of 12 indicators. A correct response extinguishes the stimulus light, a new one appears, the process is repeated, etc., for any given number of lights up to 12. The plugboard (Figure 9) permits the experimenter to select and vary both the sequence in which the lights come on and their relationship to the response keys, so that an infinite number of series combinations are possible. For any given pattern that is wired, selection of both starting and ending points within a series can also be controlled by two 12-position switches.

We have used a serial learning paradigm in which the subject must find the sequence of keys that will cause six lights to change in series. Presentation of the first stimulus light is controlled by a switch on the experimenter's panel that also starts a stop clock. The final correct response in

a series stops the clock and is signalled by red indicator lamps on both the experimenter's console and the subject's display panel. Data are obtained from measures of incorrect responses and of time to attain error-free performance; i.e., to learn the "code." With the above paradigm, both age of the subject and degree of practice appear, in our experience, to be critical variables during early exposure to the task. Typically, 8 or 9 trials per problem are required to reach criterion with initial practice, while subjects who have had experience with 15 to 20 problems require only 2 or 3 trials per problem.

The flexibility of the apparatus has prompted its use by other investigators in studies of aging effects on skill, display-control relationships, short-term memory store, prefrontal brain injury, aircrew fatigue, and drug effects.[2-4]

VI. Vigilance.

Vigilance is defined here as a form of continuous, prolonged monitoring behavior in which the operator must detect a critical signal that occurs infrequently and unpredictably. The task may vary in terms of absolute signal rate, critical signal to noncritical signal ratio (or contrast), or signal to background contrast. According to McGrath, "The signal to be detected may be either a discrete stimulus added to or taken away from the environment, or a change in a continuously or intermittently presented stimulus."[5] Typically, detection efficiency can be demonstrated to decline in tasks of 30 minutes to 1 hour, and consequently frequent rest periods for observers are recommended.

A. *Bakan Task.* A tape-recorded copy of this auditory task has been made available to us by Bakan.[6] The task requires the subject to listen to a voice repeating a series of digits recorded at the rate of one per second. The critical signal is a set of three successive odd numbers, all of which are different; e.g., 5–3–7, 9–1–5, etc. Upon hearing such a set, the subject writes them down in a column on an answer sheet. During a 64-minute "watch," 40 critical signals occur. Data normally consist of the number of correct signal detections and of "false positives" (i.e., subject responds to a nonsignal) for a 16-minute subperiod.

Although we have used the Bakan task successfully in one experiment, certain considerations led us to develop the new auditory vigilance task described below. Since the Bakan task is a popular one, some subjects anticipate a relationship between their responses and the space provided on the answer sheet; i.e., they are familiar with the task or compare notes with others regarding the number of signals. Difficulty has been encountered with some older (age 45 to 65) noncollege subjects in indoctrinating them on the task; indeed, some showed no evidence of being able to detect a critical signal with 4 to 5 hours of training. Finally, we decided to use a task that involved a purer element of signal *detection* and that did not imply a requirement in it for memory; i.e., a Bakan subject's response requires both *recall* and recording of all three digits correctly.

B. *"Beep" Task.* This task was developed by recording 1,000 cps tones at 3-second intervals. Nonsignals are of 1.0 second duration, while signals are 1.4 seconds long. The difference in length between the two tones appears to avoid any problem involving the difference threshhold; experience indicates subjects have no problem identifying which of two tones, 1.0 and 1.4 seconds, is the longer. With prolonged listening to the tone, however, the traditional characteristics of the vigilance task become evident. Thirty-six signals occur during a 60-minute watch, 9 every 15 minutes.

In a typical experiment, the subject listens to the recording through a headphone set and indicates detections by depressing a push-button. This produces a stylus deflection on a 20-channel Esterline-Angus event recorder. With this equipment it is possible to obtain data on several subjects concurrently. A communications-systems laboratory, described in detail elsewhere by its developer,[7] has enabled us to test concurrently up to 20 subjects, each located in a separate soundproof booth. The subject station shown in Figure 10 includes Telex headphones with microphone, a volume control, lamp, a response key, and a communication link to the experimenter.

VII. Radar Air Traffic Control.

A simulator complex, described in considerable detail elsewhere,[8] provides a capability for assessing a number of operator functions that characterize the enroute radar air traffic controller's task. These include vigilance, problem solving, monitoring, and reaction time. Programming of display conditions is accomplished by punched tape and relay circuitry. Task events are thereby pre-

sented repeatedly, reliably, and concurrently to each of six experimental subjects seated at separate task consoles (Figure 11). Radar targets are presented by filmstrips made from an animated display. Hypothetical pilot communications and flightstrip information are played concurrently from prerecorded tapes. Responses to critical events are measured automatically with electronic counters, and periodically, the data are typed out by a Friden Flexowriter and punched on paper tape for computer analysis. Duration and frequency of subject radio transmissions are also scored. Monitoring of system operation and data collection requires only a single experimenter.

A unique photographic technique provides identical target workloads for each of several 1-hour tests; yet the subject perceives these as different task programs. This permits day-to-day comparison of performance under comparable environmental and task conditions while controlling for possible memory effects.

Initial experimentation indicates that approximately 8 hours of practice are required for subjects (ATC instructors) to integrate effectively all subtasks and to achieve a stable level of proficiency. A new flightstrip format has been developed to be commensurate with system design, and, currently, closed-circuit television is being installed for radar target display. Research applications include studies of drug effects on proficiency, fatigue, stress, aging, automated training, target and display characteristics, workload, and control techniques.

VIII. Assessment of Affective State.

To supplement and complement the performance data derived from the operator tasks described, two affective-state checklists have been used in our research. A fatigue checklist with two, 13-item, equivalent forms is described in detail elsewhere.[9, 10] An extension and elaboration of this, developed by the author in 1956 at the USAF School of Aviation Medicine, is identified as the SAM Affective Checklist, Form X, and includes 81 items as listed in Table 2. The clustering is based on item sorts made by psychologists working in stress, drug, and sensory-deprivation research. A unique feature of the items (in contrast to many mood checklists in use) is that they are understandable by low-vocabulary individuals. Screening trials included individuals with as low as seventh grade education and from wide geographical, racial, and ethnic backgrounds. Typical checklist format requires the subject to check those items that describe his mood or feelings as of a given moment.

TABLE 2.   SAM affective-checklist items

Cluster	Polarity	Items
1. Anxiety	(+)	Quiet; easy going; peaceful; calm, cool, and collected.
	(−)	Anxious; worried; disturbed; excited; uneasy; tense.
2. Alertness	(+)	Alert; careful; cautious; wide-awake; clear-headed.
	(−)	Careless; reckless.
3. General morale	(+)	Agreeable; sociable; witty; talkative.
	(−)	Irritable; unfriendly.
4. Elation-depression	(+)	Cheerful; carefree; happy; daredevil; happy-go-lucky.
	(−)	Moody; sad; depressed; serious.
5. Cooperation	(+)	Cooperative; eager.
	(−)	Indifferent.
6. Organic	(+)	Dreamy; light-headed.
	(−)	Drowsy; dizzy; headache; weak; feverish; heavy-headed; hot; cold; numb.
7. Work attitude	(+)	Energetic; industrious; strong; raring-to-go; fit-as-a-fiddle; ambitious.
	(−)	Lazy; dull; sluggish; stale; lifeless; run-down; slowed-down.
8. Relaxation	(+)	Relaxed; patient.
	(−)	Nervous; jittery; restless.
9. Persistence	(+)	Self-confident; on the ball; steady; tireless.
	(−)	Unsure; unsteady; confused; disorganized; useless, timid.
10. Boredom	(+)	Interested.
	(−)	Bored; uninterested.
11. Fatigue	(+)	Quite fresh; extremely peppy.
	(−)	Tuckered out; dead tired.

REFERENCES

1. FROST, G. G.: A review of scoring methods for tracking tasks. *USAF, Aerospace Medical Research Laboratories Memorandum* P–14, Wright-Patterson AFB, Ohio, October 1962.
2. WELFORD, A. T.: An apparatus for use in studying serial performance. *Amer. J. Psychol.*, 65: 91–97, 1952.
3. KAY, H.: The effects of position in a display upon problem solving. *Quart. J. Exp. Psychol.*, 6: 155–169, 1954.
4. KRUGMAN, A. D., ROSS, S., VICINO, F. L., and CLYDE, D. J.: A research note: effects of dextro-amphetamine and meprobamate on problem solving and mood of aged subjects. *J. Gerontal.*, 15: 419–420, 1960.
5. BUCKNER, D. N., and McGRATH, J. J., eds: *Vigilance: a symposium*, New York: McGraw-Hill, 1963.
6. BAKAN, P.: Extraversion-introversion and improve-ment in an auditory task. *Brit. J. Psychol.*, 50: 325–332, 1959.
7. O'CONNOR, W. F.: A twenty-station behavioral communications laboratory with flexible networks and on-line digital logging. *FAA Office of Aviation Medicine Report*, in press.
8. PEARSON, R. G., HUNTER, C. E., and NEAL, G. L.: Development and evaluation of a radar air traffic control research task. *FAA Office of Aviation Medicine Report*, in press.
9. PEARSON, R. G.: Scale analysis of a fatigue checklist. *J. Appl. Psychol.*, 41: 186–191, 1957.
10. PEARSON, R. G., and BYARS, G. E., Jr.: The development and validation of a checklist for measuring subjective fatigue. *School of Aviation Medicine, USAF, Report* No. 56–115, Randolph AFB, Texas, December 1956.
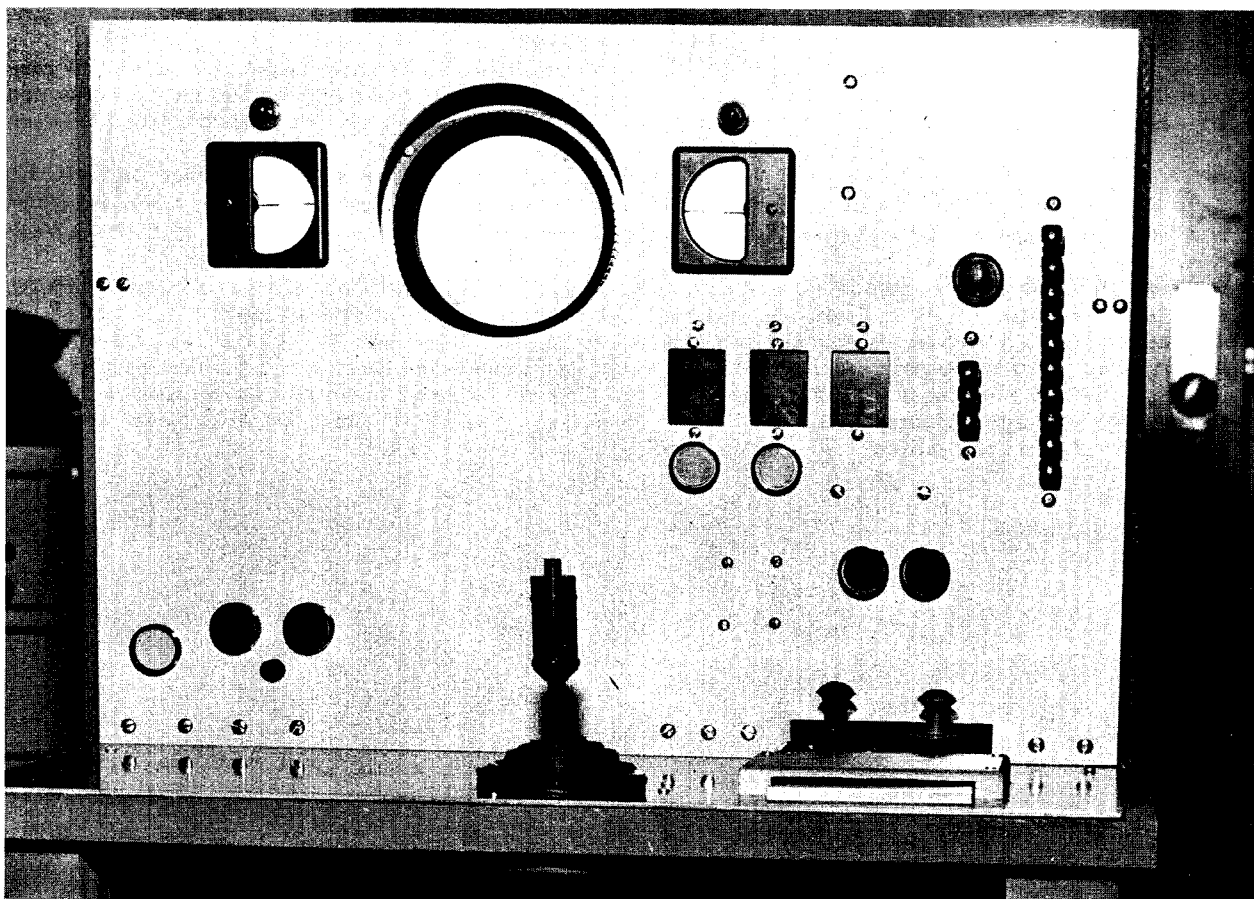
REFERENCES

1. FROST, G. G.: A review of scoring methods for tracking tasks. *USAF, Aerospace Medical Research Laboratories Memorandum* P-14, Wright-Patterson AFB, Ohio, October 1962.
2. WELFORD, A. T.: An apparatus for use in studying serial performance. *Amer. J. Psychol.*, 65: 91-97, 1952.
3. KAY, H.: The effects of position in a display upon problem solving. *Quart. J. Exp. Psychol.*, 6: 155-169, 1954.
4. KRUGMAN, A. D., ROSS, S., VICINO, F. L., and CLYDE, D. J.: A research note: effects of dextro-amphetamine and meprobamate on problem solving and mood of aged subjects. *J. Gerontal.*, 15: 419-420, 1960.
5. BUCKNER, D. N., and McGRATH, J. J., eds: *Vigilance: a symposium*, New York: McGraw-Hill, 1963.
6. BAKAN, P.: Extraversion-introversion and improve- ment in an auditory task. *Brit. J. Psychol.*, 50: 325-332, 1959.
7. O'CONNOR, W. F.: A twenty-station behavioral communications laboratory with flexible networks and on-line digital logging. *FAA Office of Aviation Medicine Report*, in press.
8. PEARSON, R. G., HUNTER, C. E., and NEAL, G. L.: Development and evaluation of a radar air traffic control research task. *FAA Office of Aviation Medicine Report*, in press.
9. PEARSON, R. G.: Scale analysis of a fatigue checklist. *J. Appl. Psychol.*, 41: 186-191, 1957.
10. PEARSON, R. G., and BYARS, G. E., Jr.: The development and validation of a checklist for measuring subjective fatigue. *School of Aviation Medicine, USAF, Report* No. 56-115, Randolph AFB, Texas, December 1956.
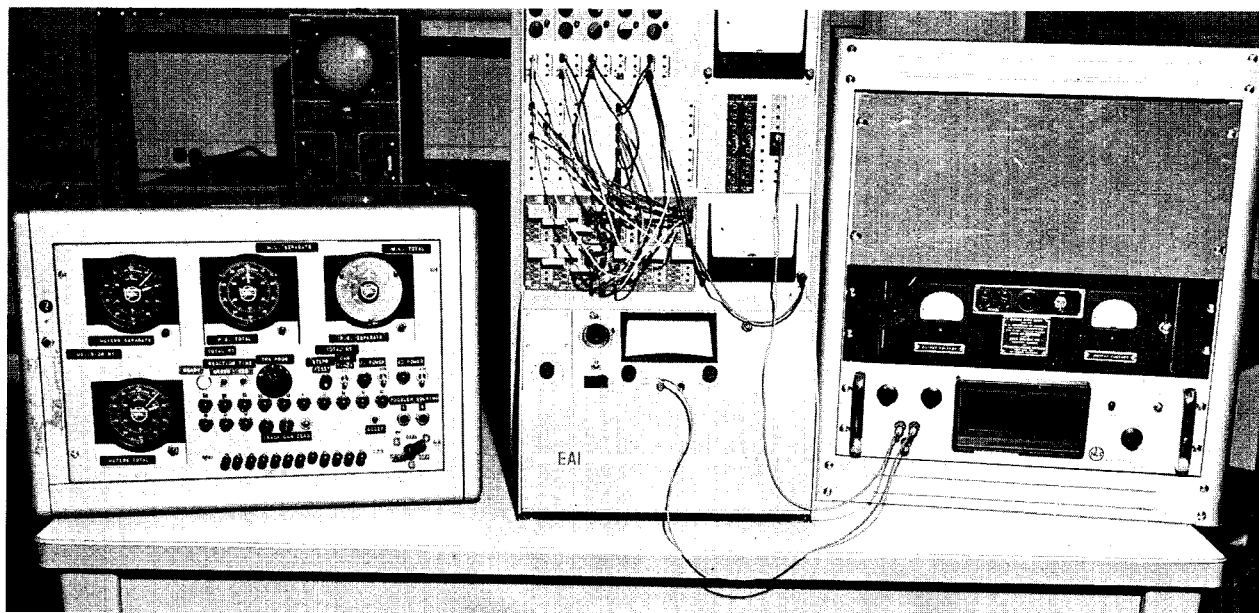
FIGURE 1.   Tracking and monitoring task, Model I.

FIGURE 2.   TMT–I programing and scoring complex.

FIGURE 3. TMT–I system block diagram.

FIGURE 4. TMT–I skill acquisition with practice.

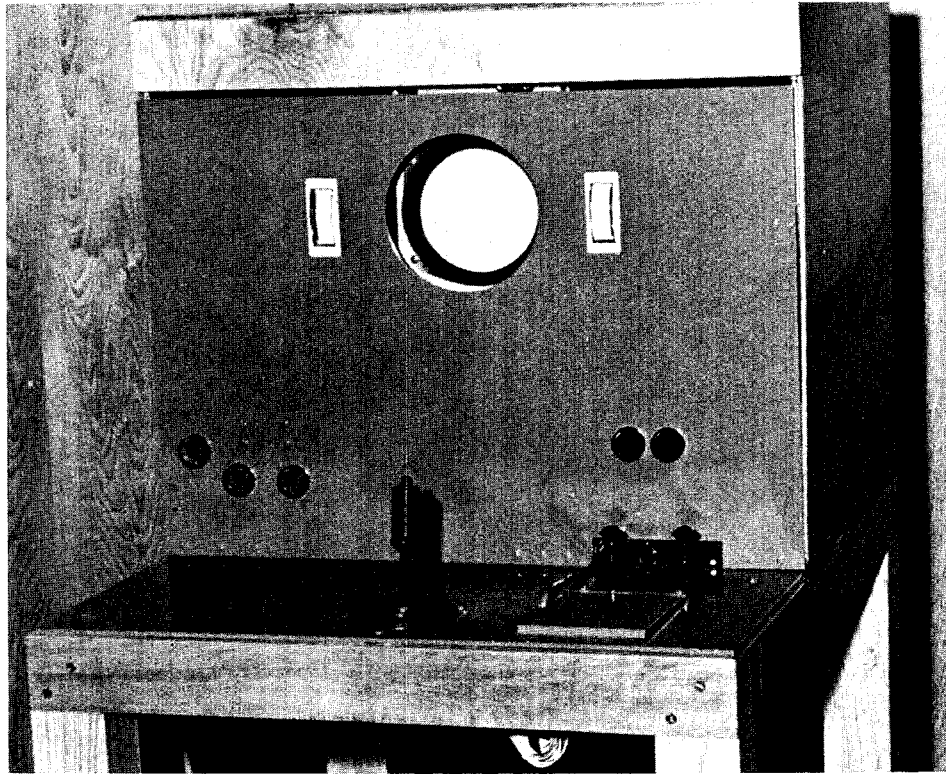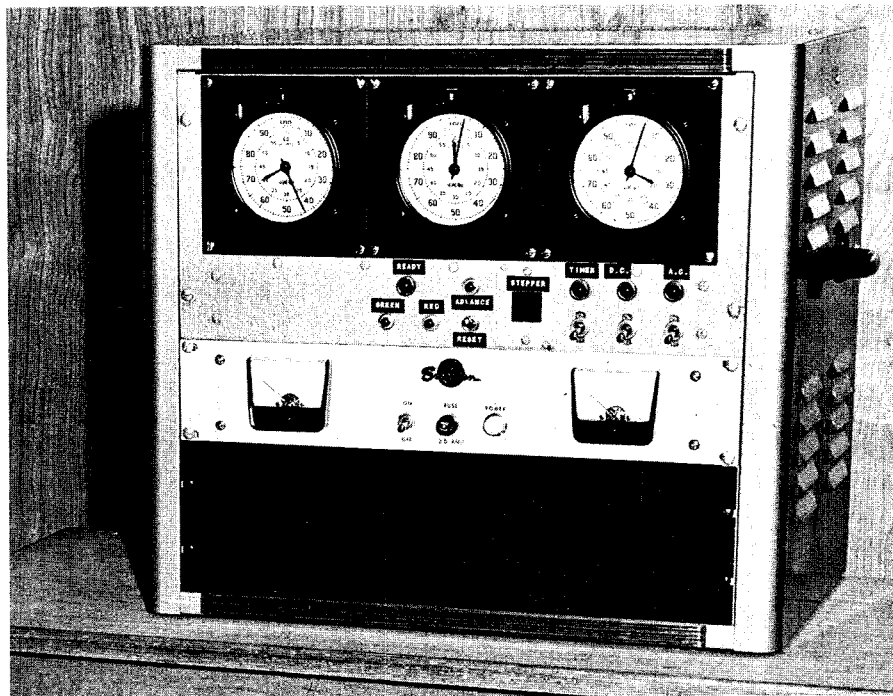FIGURE 5. Tracking and monitoring task, Model II.
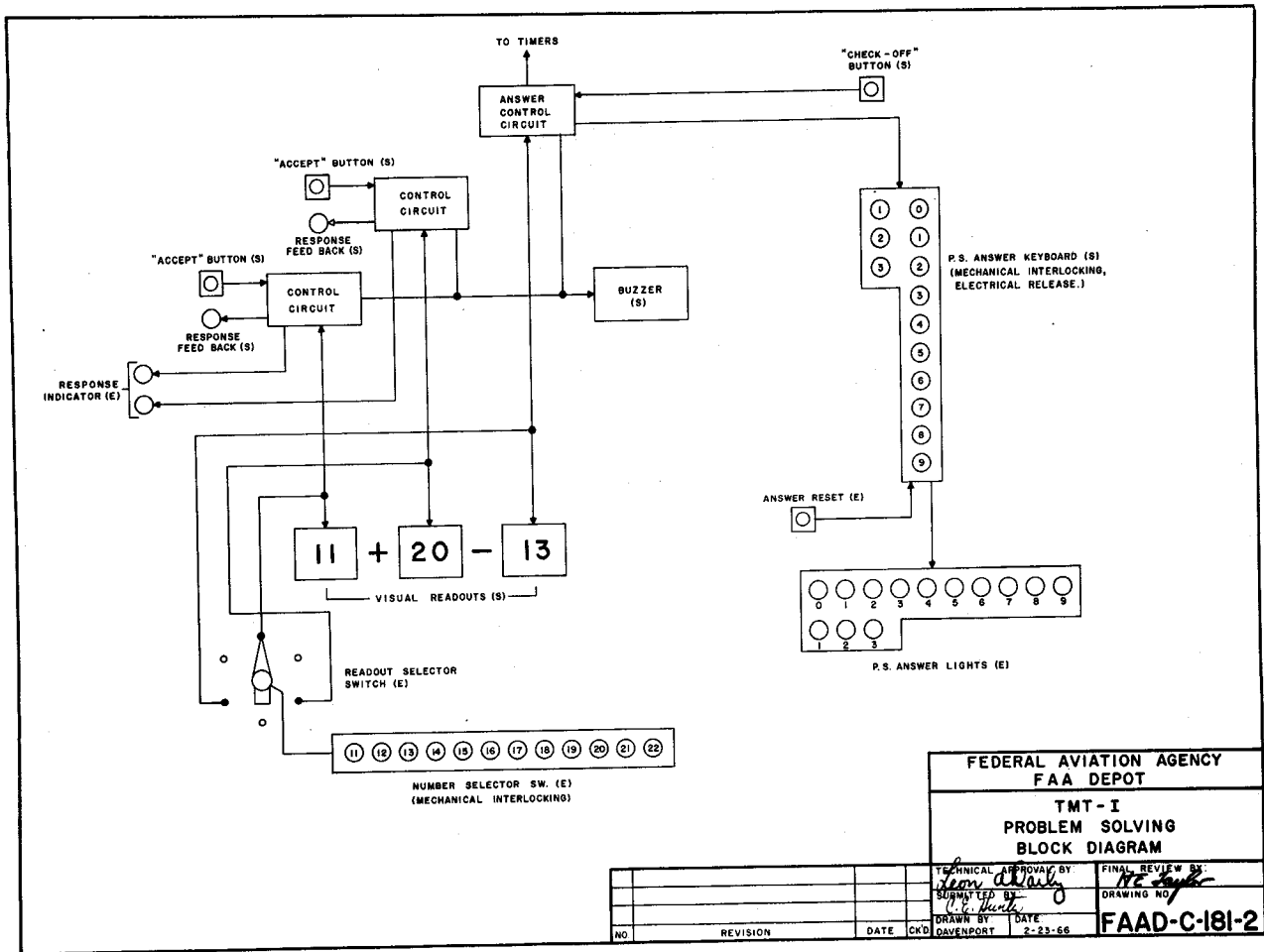

FIGURE 6. TMT–II programing and scoring unit.

FIGURE 7. TMT–I problem-solving task, block diagram.
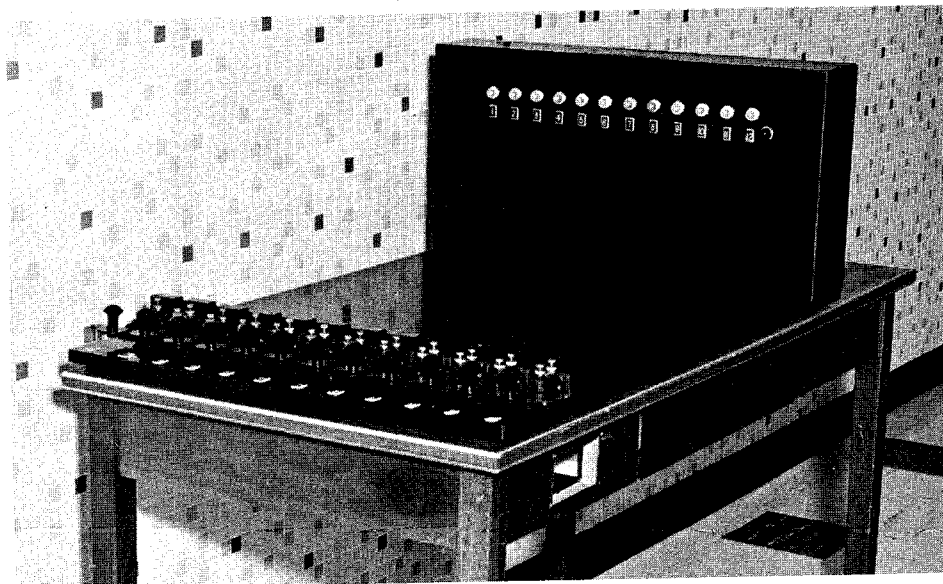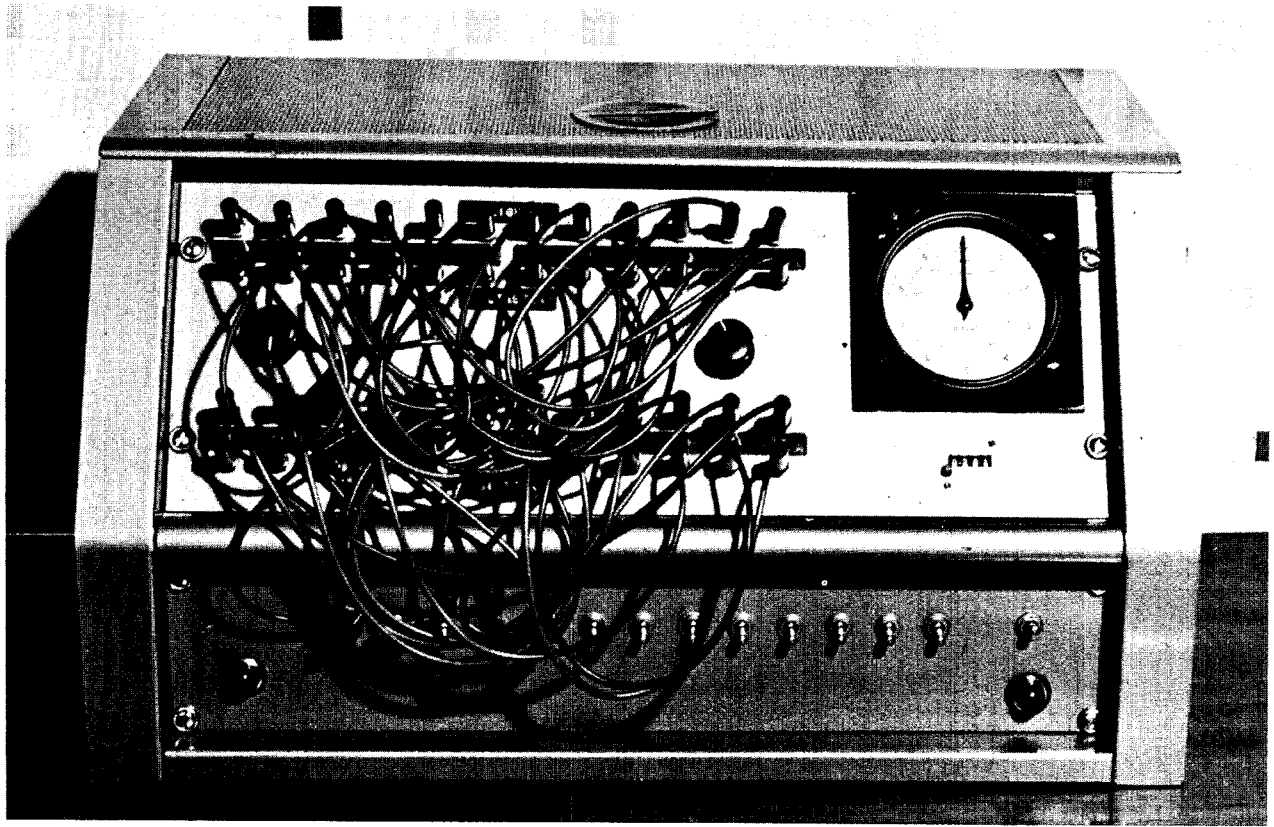

FIGURE 8. Serial-performance task.
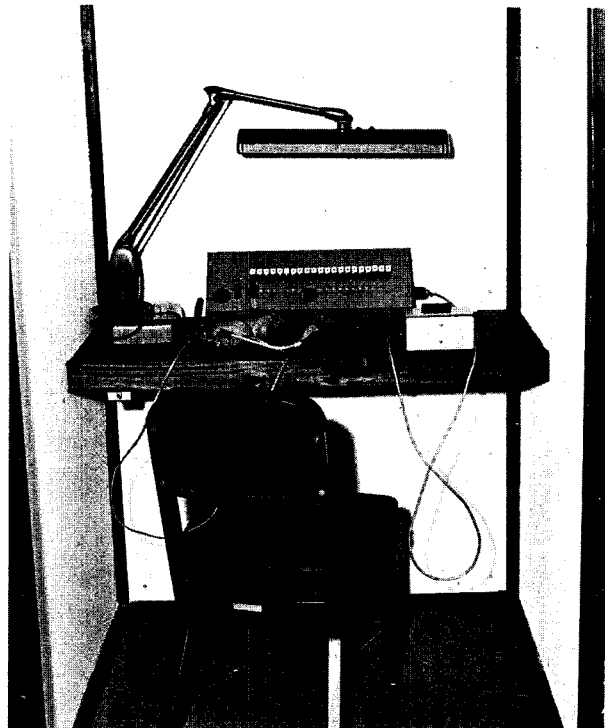
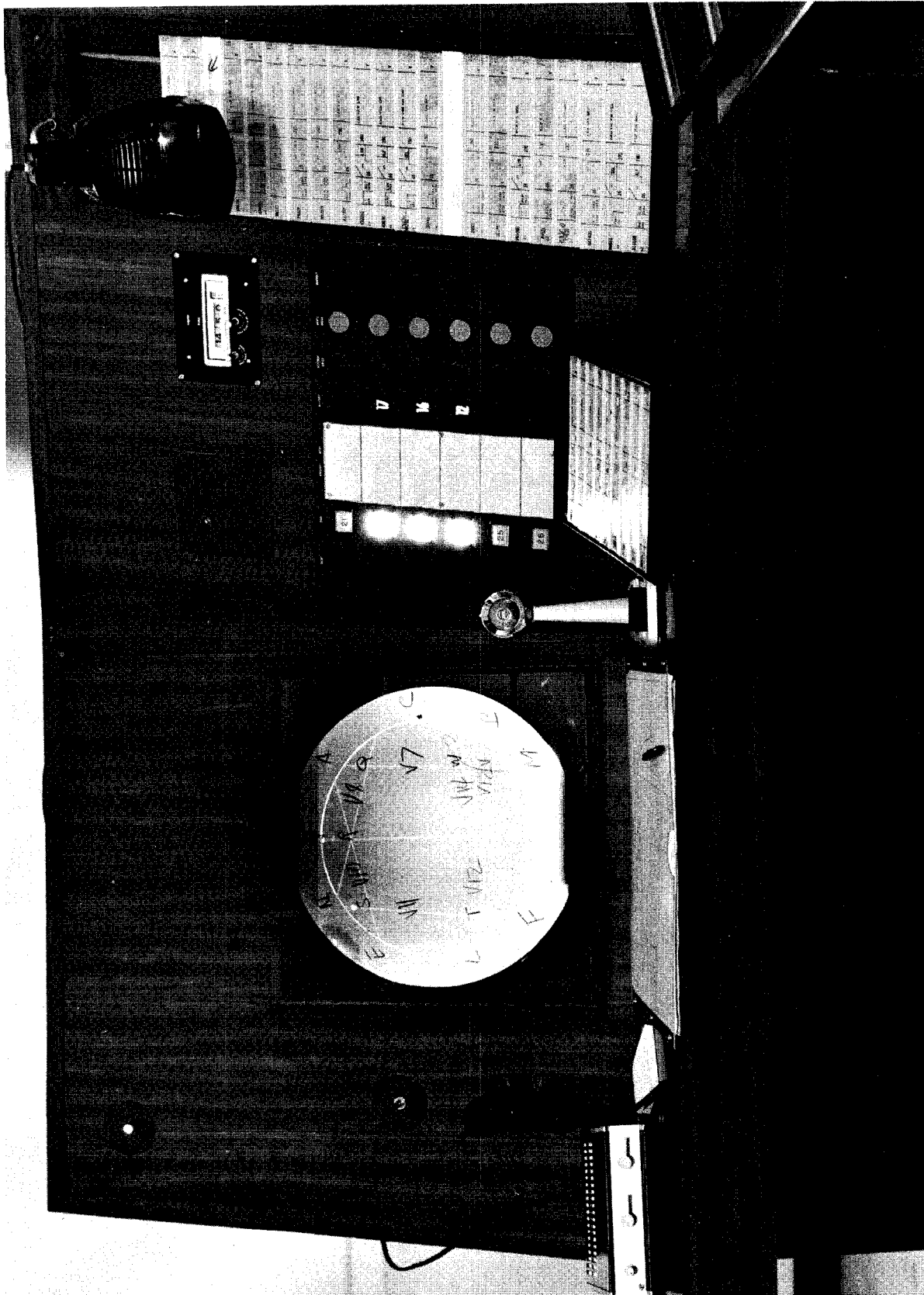FIGURE 9. Serial-performance programing unit.

FIGURE 10. Communication-laboratory subject station.

FIGURE 11. Radar ATC task console.