PILOT WORKLOAD MEASUREMENT AND EXPERIENCE ON SUPERSONIC CRUISE AIRCRAFT

Terrence W. Rezek
Dryden Flight Research Center

SUMMARY

The YF-12 aircraft is considered as representative of high workload supersonic cruise aircraft. A study was performed to determine which aircraft parameters and which physiological parameters would be most indicative of crew workload. This study is summarized and the recommendations formed a basis for a continuing study in which variations of the interval between heart beats is examined as a measure of nonphysical workload. Preliminary results of this work are presented. Current efforts in further defining this physiological measure are outlined.

INTRODUCTION

The need to understand the limits of human capability in all areas of aeronautical flight has spawned a large number of physiological studies of the human operator of flight vehicles. Each aeronautical system has a unique set of characteristics which impact the operator, but all modern systems exhibit steadily increasing operator workload demands. These demands take two forms; the physical, that is, extremes of heat, cold, noise, vibration, pressure and g forces; and the nonphysical, that is, personal risk, mission responsibility, and the large amount of information needing to be perceived, thought about, and acted upon in short time spans. The effects of the aeronautics environment on the operator are well documented (ref. 1) and some studies have demonstrated physiological effects in the absence of physical work (refs. 2 and 3). There have even been some attempts to quantify the effects of mental workload, albeit in a clinical atmosphere (refs. 4 to 6). From all this has come the realization that certain physiological phenomena accompany nonphysical workload. One objective of NASA's human factors research is the quantification of these nonphysical effects and their separation from physical effects.

In any mechanical-human interactive system, operation is achieved through an exchange of energies. This exchange and the resulting work can easily be measured quantitatively with regards to the mechanical system. This is not the case with the human system. The total workload imposed on the human organism by any given task has two sources, the physical and the nonphysical, which may interact over a range from 0 to 100 percent. The purely physical, such as the labor involved in ditch-digging, can be measured in terms of both cause (weight of dirt thrown through a given height against a force of 1*g*) and effect (oxygen consumed, calories burned, and so forth). The nonphysical workload is much more difficult to measure. But as anyone knows who has come home exhausted after a day during which the most strenuous activity was the sharpening of a pencil, nonphysical work is very real.

The physiological monitoring effort at Dryden Flight Research Center (DFRC) has produced many hours of heart rate information from a large number of subjects in a variety of aircraft. However, correlative data concerning pilot activity occurring at given heart rates were not available. At the time the heart rate data were acquired, the possibility of assessing nonphysical workload through physiological measures was not fully appreciated.

The addition of the YF-12 aircraft to the DFRC research vehicles offered an opportunity to study nonphysical workload in a high-demand system. The high performance capabilities of this aircraft and its unique handling characteristics suggested such high demands but also implied that using secondary tasks to augment ambient workload might be hazardous. To avoid any possibility of compromising mission performance or flight safety, the studies would have to be conducted on a noninterference basis; consequently, a three-part program was conceived. First, a contract was let with a company specializing in mission evaluation and task analysis. The results and recommendations from this study are summarized. Second, those parameters from both the aircraft and the pilot which showed sensitivity to workload and were amenable to recording were selected for long-term monitoring, and the resulting data were analyzed. Finally, a possible method was devised for separating the effects of physical from nonphysical workload on a select physiological parameter.

PILOT PERFORMANCE MEASUREMENT STUDY

The principal objective of the study contract was to investigate pilot performance measurement in theory and practice during high-performance aircraft research at DFRC. The best approach was felt to be a study of realistic pilot and vehicle tasks. The study involved: (1) the development of measurement sets for system and pilot performance, (2) the investigation of objective, subjective, and physiological pilot performance measures, (3) the development of measures for such applications as pilot workload and crew control display effectiveness, and (4) the development of practical and feasible automatic and semiautomatic data processing techniques.

Flight Crew Task Analysis

Information collected from the aircraft, engines, and avionics is listed in table 1. This list covers the conditions of vehicle control, attitudes, systems, and engine thought to be most indicative of the pilot's control and performance. This information was needed to permit the evaluation of operational pilot tasks and system performance in lieu of imposing a secondary task. The choice of these parameters resulted from an extensive task analysis specific to the YF-12 airplane and associated missions. A different set of parameters is likely to result when different aircraft, missions, or both are considered.

In this study it had to be recognized that the crew operations were already established. The aircraft, the test program, the crews, the missions, and the tests were given. The task at hand was to learn how the operations were conducted, and how to develop quantitative and qualitative measures of crew performance.

The Pilot's Duties

As aircraft commander, the pilot is ultimately responsible for flight planning and execution. He maintains the aircraft attitude and airspace position in a specified mission profile. He divides his attention between attitude and position instruments, outside visual references, and air traffic in the maintenance of attitude and position. To follow the mission profile precisely, he must monitor and adjust all aircraft systems for normal and emergency conditions. In the performance of a test maneuver, he establishes the pretest conditions for the aircraft systems, attitude, and position. During the tests, he alters aircraft systems, attitudes, and positions in the prescribed way. Upon recovery to a posttest condition, he reestablishes pretest conditions or proceeds to a point in the flight profile for the next test. The pilot also is required to monitor such housekeeping items as fuel, center of gravity, oxygen suit, and environment. In terms of communications, he is responsible for contact with the flight test engineer (FTE), control tower, ground radar, the chase pilot, the tanker boom operator, and other air-to-air communications. He shares responsibility with the FTE for communications with ground control, the tanker commander, air traffic centers, and other flight test supporting units on the ground.

Flight Test Engineer's Duties

Despite the ultimate responsibility of the pilot for the mission, a flight could not be successfully completed without the full-time participation of the flight test engineer. One of his primary duties is navigation. He is responsible for setting up, monitoring, and updating all navigation equipment en route. He provides the pilot with headings and distances to checkpoints and turns; he takes tactical air navigation (TACAN) fixes and communicates with air traffic centers and ground control for clearances and confirmation of position fixes.

The FTE assists the pilot in timing and maintaining the flight profile. He provides the information the pilot needs before, during, and after the test. He signals the start of the test and intermediate event points during and at the end of a test, or both. He records his own and the pilot's parametric observations before, during, and after each test.

From entry into the cockpit to leaving the cockpit he performs all checklist items of procedures; he also serves as a source of detailed information concerning subsystem operations. He monitors fuel consumption and computes the center of gravity to check the semiautomatic-to-automatic center of gravity computation during flight.

The FTE observes, communicates, and records all unusual events such as turbulence and unstarts. He shares responsibility with the pilot for communication with ground control, the tanker and commander, air traffic centers, and other flight test support units on the ground.

DEVELOPMENT OF MEASUREMENT DATA SETS

Mission and task analyses were made to establish measurement requirements. All subsystem functions and maneuvers were included, and critical tasks were analyzed in more detail, using timeline analysis techniques. Data collection procedures were examined, and a prototype data reduction and processing system was developed.

Human Performance Measurement

There are at least five ways to acquire information about human performance: system performance, secondary task workload, physiological activity, pilot control models, and statements of subjective opinion.

System performance measurements.—System performance measurement includes the comparison of all system and subsystem parameters with mission requirements. Such comparisons are often most relevant to the solution of system design problems, but are perhaps only obliquely related to human performance, since system performance reflects the combined performance of human and machine. System performance measurement may be a necessary part of the total measurement set, but it is not likely to be sufficient in and of itself.

Secondary task workload measurement.—A common measurement technique for the purpose of workload assessment calls for the measurement of performance on an added secondary task; the human operator is instructed to attend to the secondary task only to the degree that performance on the principal, or primary, task will permit. Measurement of the performance of only the secondary task indicates the level of performance for an unloaded operator; poorest performance on the secondary task (no attention given to it) may indicate the level of performance for a completely loaded operator. In this way, a workload scale can be constructed that indicates the percentage of loading of an operator with a given primary task. Unfortunately, the secondary task may interfere with the performance of the primary task, that is, the human operator may adopt a strategy for performing both tasks simultaneously in a way which is no longer relevant to the study of the primary task. Further, in such operational settings as the YF-12 airplane, a secondary task may compromise flight safety. Nevertheless, it may be possible to employ the secondary task concept. For example, in some circumstances the control of aircraft pitch attitude may be considered the primary task, and roll control may be considered the secondary task; under heavily loaded situations the pilot may be able to prevent pitch performance from deteriorating only if roll control is sacrificed (therefore, pitch control would indicate the presence of a high degree of operator loading). This and other adaptations of the secondary task measurement concept may be suitable for measurement in the YF-12 flight test program. As a rule, in an operational setting, measurement must be accomplished on subordinate tasks embedded in the normal task structure.

Physiological parameter measurement.—As the human operator can often maintain a fixed level of performance until the actual point of overload, it was felt that an earlier indication of loading might be obtained through physiological measurement. Many physiological parameters have been measured in relation to human work capacity and reserve. However, this study was confined to a consideration of electrocardiograph (ECG) signals. Recent literature has related heart rate variability to mental workload (refs. 4 to 6). Heart rate variability, often termed sinus arrhythmia, is a variation in the time interval between successive heart beats. The variability has been shown to decrease when a subject is given a mental task and to increase when the subject is not noticeably occupied; however, some variation is also attributable to dynamic and static physical workload, respiration rate, emotion, and age. The evidence is so encouraging that sinus arrhythmia measurement must be considered a candidate for measurement system development.

Pilot control model measurement.—A significant aspect of the pilot's duties is the direct manual control of the aircraft. The manner in which the pilot controls various parameters can be mathematically modeled in a manner consistent with total vehicle control analysis. One form of mathematical model can be derived from spectral analyses of the input information to pilot's display and pilot's control signals. Depending on the nature of the control task, such a model may have from one to four parameters. These parameters vary as pilot control performance levels change, and in particular as the nature of pilot control changes. For example, from such a model, the pilot's gain or sensitivity is apparent, and also the degree to which course changes and smoothing disturbances are anticipated. In short, the manner in which the pilot controls is quantified so that changes in control activity may be apparent even though the level of performance error in relation to mission requirement does not change. Such modeling should, therefore, be considered for part of the preliminary candidate measurement set.

Statement of subjective opinion.—The pilot is the only available source of some kinds of information, and information may be volunteered by the pilot or crew which the investigator would not have known to ask about. Subjective opinion is perhaps the easiest of all information to obtain, but it is difficult to obtain in quantitative form. Perhaps the most widely known technique for the quantification of pilot opinion of vehicle handling qualities is the Cooper-Harper scale.

The flight crew may be the best, and it is sometimes the only source of some types of information, such as details of the crew tasks, task performance criteria, the nature of performance tradeoffs, the dimensions of task difficulty, and unplanned or unmentioned flight events. Examples of such subjective data are shown in figure 1. At the suggestion of the crew, the events in any flight were classified into four workload categories: communications, vehicle control, time sharing, or busy (induced by flight stresses). The crew identified instances of most, routine, or least workload during each monitored flight. They also offered comments on specific events as they happened and the time when each occurred. Subjective measurement, then, must be considered in the attempt to create a comprehensive measurement set.

Audio and Time Recording

In the flight test setting, completed performance measures are meaningful only if the exact times and conditions of the events are known. Many events and their corresponding times can be gleaned only from adequately transcribed and timed communication recordings (flight test engineer's log, flight test engineer-to-pilot intercom, and crew-to-ground).

PHYSICAL AND NONPHYSICAL WORKLOAD CORRELATION TECHNIQUES

Flight Tests

This early study gave some indication of what was possible in the way of performance measurement of the crew of a high performance aircraft. Unfortunately, the major question was still unanswered: How can a measurement of nonphysical workload be extracted from ECG, the single available physiological parameter? As suggested above, as this study was concluding, new literature appeared which suggested that sinus arrythmia was a reliable indicator of nonphysical workload. Since this literature was the product of several independent researchers, the schemes devised to score the amount of level of arrythmia varied widely. When examined, the only measure which produced consistent results was standard deviation. NASA's programing efforts were therefore directed towards producing a computer program to determine the sensitivity of variation statistics, especially standard deviation, to nonphysical workload. Data reduction limitations at DFRC prevented the digitization of early YF-12 physiological data. While this problem is being resolved, the concepts were applied to data from some of DFRC's remotely piloted research vehicle programs.

The normal human ECG is a tracing of an electric potential function driving a muscular pump. As with any pump, the activity is cyclic and the driving function must be basically rhythmic. Left to itself, this rhythmic function would seem to originate in a free-running oscillator with a relaxed regulatory system. Under no other load than basic life sustenance, an average resting heart rate, on a minute-by-minute basis, remains fairly constant. If examined closely by measuring the interval between each beat and calculating the heart rate for that interval, a wide variation becomes evident. As either physical or nonphysical workload (or both) increases, this variation dramatically decreases, almost as if the oscillator becomes more stable under stress.

DFRC had developed a method for obtaining ECG under a wide variety of human activities. Originally intended to collect the ECG in analog form aboard high performance aircraft, the method and equipment have proved adaptable to the collection of physiological data in digital form. The data are transmitted directly to ground-based computers as in the remotely piloted vehicle program, or to in-flight recorders, as used on the YF-12 airplane. What happens to the heart's cyclic activity during a time of nonphysical flight stress is shown in figure 2. These data were taken during a remotely piloted flight of the 3/8-scale F-15 airplane. In this series of flights a scale model of the F-15 with full onboard avionics and a single forward-looking video camera for visual data was dropped from beneath another aircraft and flown through unpowered maneuvers to a lakebed landing. The pilot performing these maneuvers was on the ground in the remote piloting facility. The top graph shows heart rate as a 15-second average; that is, each point is an average of the instantaneous rate of every beat occurring in the preceding 15 seconds. There is the gradual climb to launch, a dramatic jump at launch, and a decline after droque chute deployment. The bottom graph shows one method of displaying heart rate variability. Over the same 15-second intervals discussed above, the instantaneous rates of each beat are compared to find the minimum and maximum in milliseconds. The difference between these two measures is found and plotted against the same time scale as used in the top graph. The variation in these numbers is wide throughout most of the flight, and the variation is very small during the portions most demanding of the pilot.

Clinical Tests

Since both physical and nonphysical workloads have stabilizing influences on the function generator, it is important to know how much influence each has. A series of clinical tests is being constructed to attempt to separate these factors. In these tests, the pilots at DFRC and some volunteers from the employee population are to be tested in the DFRC stress physiology laboratory. First, heart rate variability is to be measured under carefully quantified physical workloads. Then, on the assumption that nonphysical workload can be adequately simulated by increasing demands on the subject's information-processing capability, a simple decision task with minimal physical involvement is being devised. If an individual's saturation point under nonphysical demand can be successfully measured, the two tasks can be combined and the total effect, including any synergism, can be measured.

YF-12 Cold Wall Tests

As of this date, data handling techniques have evolved to a reliable stage. A physiological indicator of nonphysical workload has been developed but requires proof. The YF-12 test missions have recently included the type of scheduled partitioning of flight (cold wall experiments) which have shown the most pronounced ECG variability on other programs. During these flights, the aircraft had to be maneuvered to a precise point in space and had to maintain a specified altitude and airspeed through the experiment. At the moment when all conditions were met, a pod suspended below the aircraft was blown open, uncovering a cryogenically cooled cylinder. As might be expected, this greatly disturbed the airflow about the aircraft, which occasionally disturbed the pilots.

Some preliminary results from these flights are shown in figure 3. This started out as a sequential histogram, that is, the rate at which each heart beat occurred was plotted sequentially against time for a period which began just before the cold wall experiment and ended about 10 minutes after it. For the figure as shown, a line was drawn through the minimum and maximum points to create a graphic envelope of heart rate. Again, the variability narrows dramatically as the cold wall test approaches. Figure 4 shows the same data presented in a manner similar to the remotely piloted vehicle data discussed above.

ONGOING RESEARCH: THE YF-12 AIRCRAFT AS A RESEARCH VEHICLE

The initial study demonstrated a potential measurement system for workload comparisons and indicated the direction further research might take towards the development of a complete theory of flight workload. The YF-12 aircraft is nearly ideal for the study of pilot workload. The measurements that can be produced are not likely to be exceeded in a laboratory or simulated environment; in fact, the simulation of the YF-12 environment is not within the current state of the art. Therefore, the continued use of the workload measurement system during the conduct of the ongoing YF-12 program is recommended.

Electrocardiograph Recordings

Information to be collected from the aircraft, engines, and avionics is listed in table 1. This information supports the measurement of pilot tasks and system performance. The large number of recorded measurements is a direct result of the extensive task analysis made on this class of vehicles and the projected flight test missions. A different set of parameters is likely to result when different aircraft, missions, or both are considered.

Audio and Time Recordings

In the flight test setting, most applications of computed measures depend on knowing the times and the conditions corresponding to measured performance. Many events and their corresponding time can be gleaned from the flight test engineer's log, but some are not available from this source. The primary source of information is the set of communication recordings after careful transcription and timing.

Subjective Measures

Subjective measures depend on direct access to the crew's knowledge, opinions, and ratings through postflight interview sessions in addition to information obtained from flight briefings and debriefings.

Minimum Parameter Set

The minimum parameter set consists of the ECG recordings, the parameters listed in table 1, communications recordings, and crew interviews. Other information relevant to the conduct of an experiment is to be obtained from briefings, debriefings, and flight monitoring.

The YF-12 airplane is an excellent stressor for the human operator. This research with it may help to establish a quantifiable measure of workload, and, perhaps, a system saturation point. Overloaded systems suffer breakdowns; when the system includes human operators, the cost of breakdown may be measured in lives.

REFERENCES

1. Bioastronautics Data Book. NASA SP-3006, 1964.

2. Roman, James; Perry, John J.; Carpenter, Lewis R.; and Awni, Shaiban A.: Flight Research Program — VI. Heart Rate and Landing Error in Restricted Field of View Landings. Aerospace Medicine, vol. 38, no. 2, Feb. 1967, pp. 128-132.

3. Roman, James: Risk and Responsibility as Factors Affecting Heart Rate in Test Pilots. The Flight Research Program — II. Aerosp. Medicine, vol. 36, no. 6, June 1965, pp. 518-523.

4. Kalsbeck, J. W. H.: Sinus Arrhythmia and the Dual Task Method in Measuring Mental Load. Measurement of Man at Work; an Appraisal of Physiological and Psychological Criteria in Man-Machine Systems, W. T. Singleton, J. G. Fox, and D. Whitfield, eds., van Nostran Reinhold Co., 1971, pp. 101-113.

5. Rohmert, W.; Laurig, W.; Phillip, U.; and Luczak, H.: Heart Rate Variability and Work-Load Measurement. Ergonomics, 1973, vol. 16, no. 1, pp. 33-44.

6. Mulder, G.; Mulder, W. R. E. H.; and van der Meulen, Hajonides: Mental Load and the Measurement of Heart Rate Variability. Ergonomics, 1973, vol. 16, no. 1, pp. 69-83.

7. Muckler, Fred; Obermayer, Richard W.; and Nicklas, Douglass R.: Pilot Performance Measurement Study — Final Technical Report. NASA CR-143842, 1977.

TABLE 1.—AIRCRAFT PARAMETERS

Spike position
Forward bypass position
Aft bypass position
Engine
Coarse pitch rate
Fine pitch rate
Coarse roll rate
Fine roll rate
Coarse yaw rate
Fine yaw rate
Pitch attitude
Roll attitude
Center-of-gravity normal acceleration
Center-of-gravity lateral acceleration
Duct pressure ratio
Right rudder position
Longitudinal stick position
Lateral stick position
Event
Right power lever position
Left power lever position
Computed angle of attack
Computed angle of sideslip
Altitude
Inlet system condition
Spike tip total pressure
Time (hr, min, sec, msec)

FLIGHT 015 CREW POS. Pilot

BUSY, TIME-SHARING

Least/Most	Time	Nature of Event/Comments
Routine	102045-102400	TP #1
Most	102400-102800	TP #2 - Off schedule on [illegible]; missed sideslip points; four times as difficult
Routine	102900-103300	Item 10
Most	114850-115600	TP #5 - Busy (also an unstart at 115508)

VEHICLE/EQUIPMENT CONTROL DIFFICULTY

Least/Most	Time	Nature of Event/Comments
Most	102400-102800	TP #2 - (See explanation above).
Most	103328 - 103800 120310 - 120800	Items 11 & 27 - difficult to control
Least	104018 - 104435	TP #3 - Easy (used autopilot)
Most	121136 - 121155	TP #6 - Wasn't satisfied with 1st one (performed again); had trouble on 1st one with trim; had autopilot on and was cross-controlling against it.
Least	121520 - 122300	TP #7 - Used autopilot; went well (flt eng. did not turn tape on until 122137)
Least	122500 - 122900	Did another TP #2 - "Went real well"

Figure 1.—Subjective data.

FLIGHT 015 CREW POS. Pilot

COMMUNICATION

Least/Most	Time	Nature of Event/Comments
Most	121630 - 122900 (get more off tape)	Went well - no saturation except Trying to get in touch with NASA southbound off of second loop

STRESSFUL

Least/Most	Time	Nature of Event/Comments
Most	110206 - 110648	Item 15 & 16 - trying to find the tanker
Most	104556 - 104915	TP #4 - know you're going to get an unstart - but don't know when or how bad (unstart about 104840)
Most	115022 - 115600	TP #5 - Looking for unstarts (unstart 115508)

Figure 1.—Concluded.

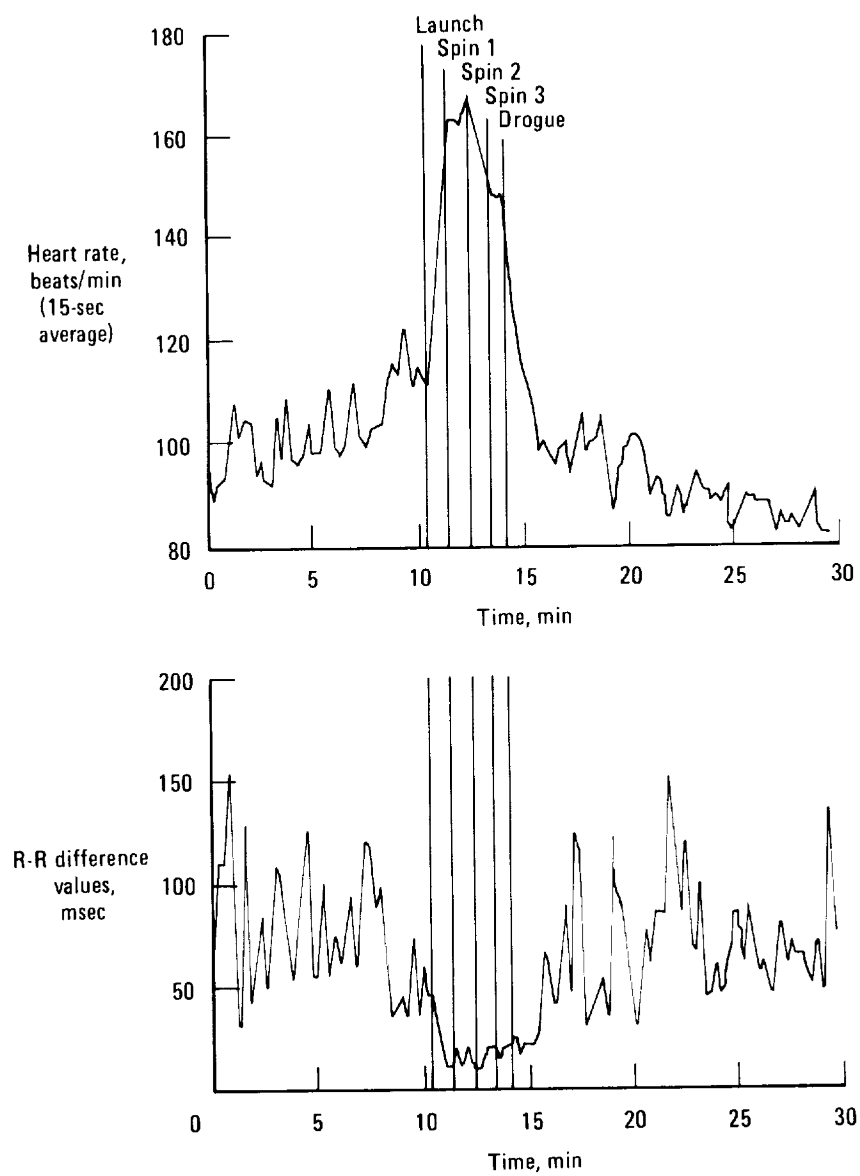

Figure 2.—Electrocardiograph data from the F-15 remotely piloted vehicle.

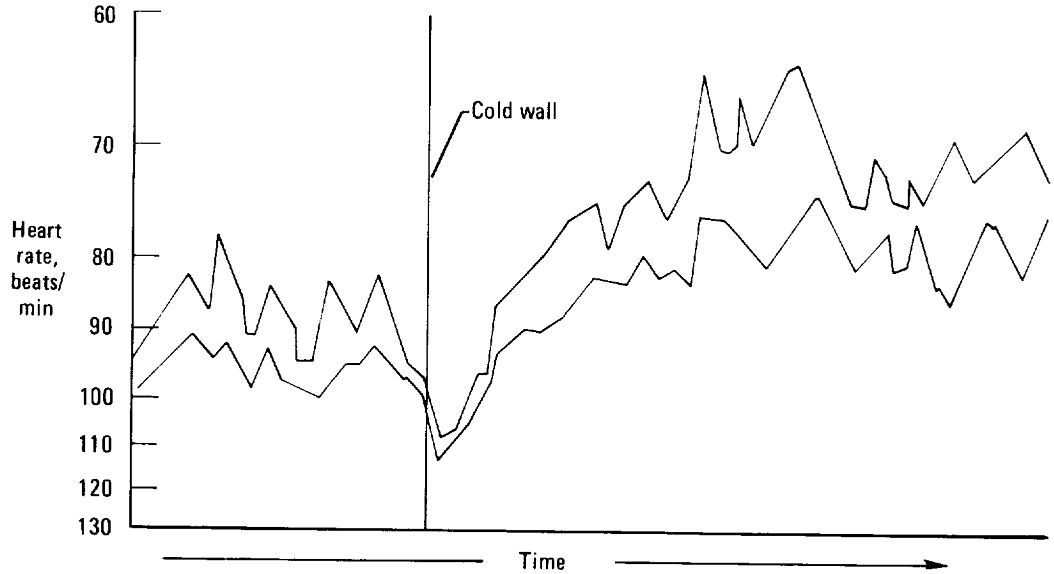

Figure 3.—Sequential histogram showing ECG envelope during cold wall experiment.

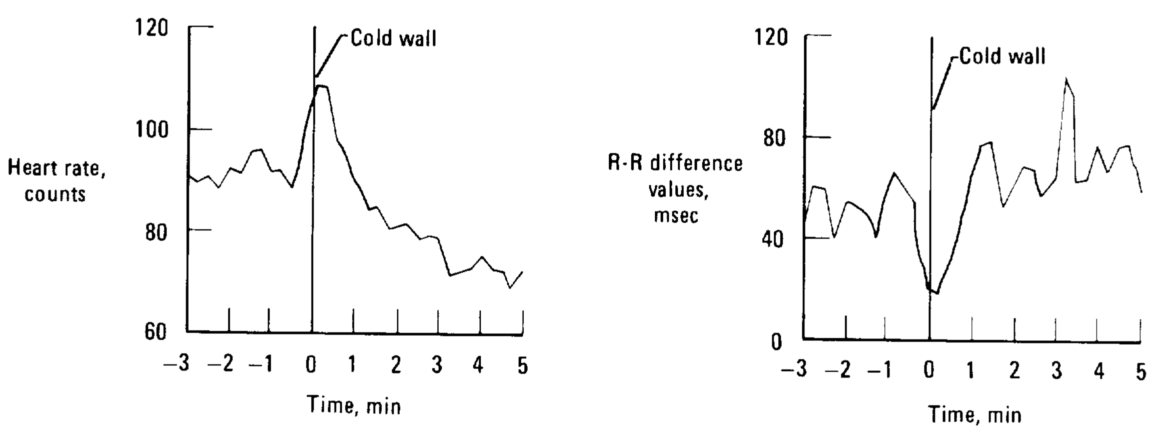

Figure 4.—Electrocardiograph data from YF-12 cold wall experiment showing R-R variation.