Stanley Roscoe
Louis Corl
Jean LaRoche

KEEPING THE PICTURE

The Measurement of Flow Control

HELIO PRESS
Acrobat™ 4.0 Version
July 1999

Intentionally blank

KEEPING THE PICTURE

The Measurement of Flow Control

Canadian Cataloguing in Publication Data

Roscoe, Stanley Nelson, 1920-

Keeping the Picture : The Measurement of Flow Control

Includes bibliographical references and index.

ISBN 2-9803682-5-3

1. Human engineering - Computer programs. 2. WOMBAT-FC (Computer file). 3. Performance - Evaluation - Computer programs. 4. Stress (Psychology) - Evaluation - Computer programs. I. Corl, Louis, 1944- . II. LaRoche, Jean, 1962- . III. Title.

TA166.R672 1999 620.8'2'02855369 C99-941007-5

970 Montée de Liesse, Suite 210
Saint-Laurent, QC H4T 1W7 Canada
Phone 1.514.336.9310
Fax 1.514.336.9400
Email: info@aero.ca

Legal Deposit: July 1999
National Library of Canada
Bibliothèque nationale du Québec
ISBN 2-9803682-5-3

Printed in Canada

Stanley N. Roscoe
Louis Corl
Jean LaRoche

KEEPING THE PICTURE

The Measurement of Flow Control

HELIO PRESS
Acrobat™ 4.0 Version
July 1999

Intentionally blank

Dedicated to
Dick Campbell, Transport Canada (Ret.),
who had a vision of a controller selection test
based on "keeping the picture."

Intentionally blank

Contents

List of Figures

List of Illustrations

Intentionally blank

FOREWORD

The world has many highly effective controllers of complex situations involving multiple moving elements. Beyond basic intelligence and modest motor skill, controller performance depends largely on what is now called **situational awareness**, the overarching ability to:

- attend to multiple information sources,
- evaluate alternatives,
- establish priorities,
- estimate probable outcomes for different courses of action,
- work on whatever has the highest momentary urgency without losing sight of the routine,
- reorder priorities as situations deteriorate or improve,
- act decisively in the face of indecision by others.

The solo WOMBAT-FC™ test is designed to measure **situational awareness**, **stress tolerance** and the **attention management abilities** of individuals keeping track of a continuous flow of information in dynamic situations involving many moving elements.

When two or more controllers are working in teams, their interactions add a social dimension to individual performance that is now addressed by training in **total resource management** (TRM). The DuoWOMBAT-FC™ addresses the abilities of team members to manage their **collective resources**.

About this book

This book is intended to provide the basic operational information about the WOMBAT-FC Situational Awareness and Stress Tolerance Test. The first five chapters present the historical and scientific backgrounds that led to today's WOMBAT-FC system. The subsequent chapters deal with the technical side of WOMBAT-FC, from its installation to operation and score interpretation.

The Adolescence of Engineering Psychology

1

This retrospective account of the emergence of engineering psychologists–in the military, in academia, in the aviation industry, in troubleshooting system problems, in consulting, and in course setting for civil and military agencies–is based largely on the recollections of the senior author and many years of correspondence with others of similar vintage or older.

Context

Engineering psychology is the science of human behavior in the operation of systems. Consequently, engineering psychologists are concerned with anything that affects the performance of system operators–whether hardware, software, or liveware. They are involved both in the study and application of principles of ergonomic design of equipment and operating procedures and in the scientific selection and training of operators. The goal of ergonomics is to optimize machine design for human operation, and the goal of selection and training is to produce people who get the best performance possible within machine design limitations.

Principles of Design

Engineering psychologists are concerned first with the distribution of system functions among people and machines. System functions are identified through the analysis of system operations.

Engineering psychologists typically work backward from the goal or desired output of the system to determine the conditions that must be satisfied if the goal is to be achieved. Next they predict–on the basis of relevant, validated theory or actual experimentation with simulated systems–whether the functions associated with each subgoal can be satisfied more reliably and economically with automation or human participation.

Usually it turns out that the functions assigned to people are best performed with machine assistance in the form of sensing, processing, and displaying information and reducing the order of control. Not only should automation unburden operators of routine calculation and intimate control, but also it should protect them against rash decisions and blunders. The disturbing notion that machines should monitor people, rather than the converse, is based on the common observation that people are poor watchkeepers and, in addition, tend to be forgetful. This once radical notion is now a cornerstone of modern system design.

Selection and Training

The selection and training of system operators enhance performance within the limits inherent in the design of the system. Traditional operator selection criteria have tended to emphasize general intelligence and various basic abilities believed to contribute to good psychomotor performance. Although individuals without reasonable intelligence and skill do not make effective operators, it has become evident that these abilities are not sufficient. To handle emergencies while maintaining routine operations calls for breadth and rapid selectivity of attention and flexibility in reordering priorities.

The more obstinate a system is to operate and the poorer the operator-selection criteria, the greater the burden on training. Modern training technology is dominated by computer-based teaching programs, part-task training devices, and full-mission simulators. Engineering psychologists pioneered the measurement of the transfer of training in synthetic devices to pilot performance in airplanes starting in the late 1940s and demonstrated the effectiveness of these relatively crude machines. More importantly, some general principles were discovered that can guide the design of training programs for systems other than airplanes.

Application

Fortunately, improved human performance in system operations can come from all directions. Ergonomic design can make the greatest and most abrupt differences in performance, but improvements in selection and training can be made more readily by operational management. More immediate, though usually less dramatic, improvements in system effectiveness can be made through the redesign of the operational procedures used with existing systems. A brief history of how all this got started during and immediately following World War II is best told by focusing on the pioneers who made it happen.

The Trail Blazers

Among the earliest experimental studies of the human factors in equipment design were those made during World War II at the Applied Psychology Unit of Cambridge University, England, under the leadership of Sir Frederick Bartlett. In 1939 this group began work on problems in the design of aviation and armored-force equipment (Bartlett, 1943; Craik, 1940). Prominent among the early contributors to engineering psychology at APU were Norman Mackworth, K. J. W. Craik, Margaret Vince, and W. E. Hick. Mackworth explored problems of human vigilance. Craik, Vince, and Hick performed classic studies on the effects of system design variables on manual control performance (e.g., Craik, 1944; Craik & Vince, 1943, 1944; Hick, 1945) and on direction-of-motion relationships between controls and displays (Vince, 1945).

Also in 1939, in the United States of America, the National Research Council Committee on Aviation Psychology was established. This committee, first chaired by Jack Jenkins of the University of Maryland and later by Morris Viteles of the University of Pennsylvania, stimulated a wide range of research in aviation psychology primarily at universities. With support from the NRC, Alexander C. Williams, Jr., working with Jenkins at the University of Maryland, began flight research in 1939 on psychophysiological "tension" as a determinant of performance in flight training. These experiments, involving the first airborne polygraph, also appear to have been the first in which pilot performance was measured and correlated with physiological responses in flight. The report of this research was completed in 1941 but was not released until after the war (Williams, Macmillan, & Jenkins, 1946).

In 1940 John Flanagan was recruited to set up a large aviation psychology program for the US Army. Several dozen leading psychologists were commissioned, starting with Arthur Melton, Frank Geldard, and Paul Horst (Koonce, 1984). With America's entry into the war, Flanagan's original organization, the Applied Psychology Panel of the National Defense Research Committee (NDRC), was greatly expanded and its work was extended into what was later to be known as the US Army Air Forces Aviation Psychology Program (Flanagan, 1947).

The history of the NDRC Applied Psychology Panel was recorded by Charles W. Bray (1948), who served as its chief, succeeding Walter S. Hunter. One of the projects started in 1942 was a study of Army antiaircraft artillery at Tufts College, directed by Leonard Mead and William Biel, which led to the development of a gun-director tracking simulator (Parsons, 1972). Early efforts in the United States to study manual control problems systematically were stimulated by the experiments of Harry Helson and W. H. Howe (1943) on the effects of friction and inertia in controls.

Human Engineering

While most of the psychologists in the British Royal Air Force and the United States Army and Navy were involved hands-on in aviator selection and training, others were occasionally called on to deal directly with the subtle problems aviators were having in operating their newly developed machines. During the war the term "pilot error" started appearing with increasing frequency in training and combat accident reports. It is a reasonably safe guess that the first time anyone intentionally or unknowingly applied a psychological principle to solve a design problem in airplanes occurred during the war, and it is possible that the frequent wheels-up-after-landing mishap in certain airplanes was the first such case (Roscoe, 1992).

It happened this way. In 1943 Lt. Alphonse Chapanis was called on to figure out why pilots and copilots of P–47s, B–17s, and B–25s frequently retracted the wheels instead of the flaps after landing. Chapanis, who was the only psychologist at Wright Field until the end of the war, was not familiar with the ongoing studies of human factors in equipment design. Still, he immediately noticed that the side-by-side wheel and flap controls—in most cases identical toggle switches or nearly identical levers—could easily be confused. He also noted that the corresponding controls on the C–47 were not adjacent and their methods of actuation were quite different; hence C–47 copilots never pulled up the wheels after landing.

Chapanis realized that the so-called pilot errors were really cockpit design errors and that by coding the shapes and modes-of-operation of controls the problem could be solved. As an immediate wartime fix, a small, rubber-tired wheel was attached to the end of the wheel control and a small wedge-shaped end was attached to the flap control on several types of airplanes, and the pilots and copilots of the modified planes stopped retracting their wheels after landing. When the war was over, these mnemonically shape-coded wheel and flap controls were standardized worldwide, as were the tactually discriminable heads of the power control levers found in conventional airplanes today.

Psychoacoustics

In the human engineering area of psychoacoustics, the intelligibility of speech transmitted over the noisy aircraft interphones of World War II presented serious problems for pilots and their crews. At Harvard University's Psycho-Acoustic Laboratory, S. S. Stevens, J. C. R. Licklider, and Karl D. Kryter, with help by a young George A. Miller, later the 77th president of the American Psychological Association, conducted a series of articulation tests of standard and modified interphones at altitudes of 5,000 and 35,000 feet in a B-17 bomber (Licklider & Kryter, 1944). Intelligibility was improved by peak clipping the powerful vowel sounds in human speech and then amplifying the remaining balanced mixture of vowels and consonants (Licklider & Miller, 1951). Incidentally, the psychologists also showed that the B–17 could operate effectively at 35,000 feet, which the Air Force had not yet fully realized.

ENTER THE ENGINEERING PSYCHOLOGISTS

IN THE MILITARY

None of the wartime "human engineers" had received formal training in engineering psychology; indeed, the term hadn't even been coined. Those who became involved in the study of human factors in equipment design and its application came from various branches of psychology and engineering and simply invented the budding science on the job. B. F. Skinner stretched the concept a bit by applying his expertise in animal learning to the design of an air-to-sea guidance system that employed three kamikaze pigeons who learned to recognize enemy ships and voted on which way to steer the bomb they were riding (Skinner, 1960). It worked fine (and still would), but there were moral objections.

After the war, the field of engineering psychology quickly gained momentum. The Applied Psychology Unit in Cambridge, England, was expanded under the leadership of Donald Broadbent, who succeeded Sir Frederick Bartlett as Director. Christopher Poulton's comprehensive work at APU on the dynamics of manual control (integrated in his 1974 book) stands as a major contribution, as does his work in other areas. The psychologists of the Royal Aircraft Establishment at Farnborough conducted a wide range of research under the direction of Air Marshal William Stewart, with John Rolf leading the flight simulation work. Alan Burrows, who learned his trade under Stewart, later headed a human factors unit at Douglas Aircraft in Long Beach, California.

In the summer of 1945, the US Army Air Forces (AAF) Aviation Psychology Program included Colonels John Flanagan, Frank Geldard, J. P. Guilford, and Arthur W. Melton (Flanagan, 1947). By this time the program's personnel had grown to about 200 officers, 750 enlisted men, and 500 civilians (Alluisi, 1994). Their wartime work was documented in 1947 in a series of 19 publications that came to be known as the "blue books." Volume 19, edited by Paul Fitts (1947) and titled *Psychological Research on Equipment Design*, was the first major publication on human factors engineering, or simply "human engineering" as it was referred to in those times.

In August of 1945, with the war about to end, the AAF Aero Medical Laboratory at Wright Field near Dayton, Ohio, established a Psychology Branch. The group, under Lt. Col. Paul Fitts, included 21 officers, 25 enlisted men, and 10 civilians that first year (Fitts, 1947). Prominent psychologists included Majors Judson S. Brown, Launor F. Carter, Albert P. Johnson, and Walter F. Grether; Captains Richard E. Jones and H. Richard Van Saun; First Lieutenants Julien Christensen, John Cowles, Robert Gagne, John L. Milton, Melvin J. Warrick, and Wilse B. Webb; and civilian William O. Jenkins. Fitts was succeeded as Technical Director by Grether in 1949.

Meanwhile, Arthur W. Melton and Charles W. Bray were building the Air Force Personnel and Training Research Center, commonly referred to as "Afpatrick," into a huge research organization with laboratories at Mather, Sted, Williams, Tinker, Goodfellow, Lowry, Tyndall, Randolph, and Lackland Air Force Bases. Prominent psychologists included Edward Kemp at Mather, Robert Gagne at Lackland and later at Lowry, Lloyd Humphreys at Lackland, Jack Adams at Tyndall, and Bob French at Randolph. In 1958 this farflung empire was dismantled by the Air Force. Most of the psychologists returned to academia, while others found civilian research positions in other laboratories.

The Navy was not to be outdone by the Air Force. In late 1945, human engineering in the Navy was centered at the Naval Research Laboratory in Washington, DC, under Franklin V. Taylor. The stature of NRL was greatly enhanced by the originality of Henry Birmingham, an engineer, and the writing skills of Taylor, a psychologist. Their remarkable 1954 work, *A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems*, had an unanticipated benefit; to understand it, psychologists had to learn about the electrical engineering concepts Birmingham had transfused into the psychology of manual control.

Another fortunate development in 1945 was the establishment of the Navy's Special Devices Center at Port Washington on Sands Point, Long Island, with Leonard C. Mead heading its Human Engineering Division. SDC invented and developed many ingenious training devices on site and monitored a vigorous university program for the Office of Naval Research, including the original contract with the University of Illinois Aviation Psychology Laboratory. Task

Order XVI, as it was known, was renewed for 20 consecutive years. Mead went on to head an engineering psychology program at Tufts College and from there to the upper management of the college and eventually of the Smithsonian Institution.

Project Cadillac, the first complex manned-system simulation study was conducted at the Sands Point facility from 1948 until 1955, with experiments actually getting underway in 1951 (Parsons, 1972). The project, initially directed by New York University, grew out of the Navy's early problems with airborne combat information centers (CICs) designed to perform surveillance functions and later interception control. Robert Chapman, Vince Sharkey, and James Regan were prominent contributors. H. M. "Mac" Parsons cut his human engineering teeth on Project Cadillac in 1950 while still a graduate student at Columbia University. He stayed with the project when the NYU Electronic Research Laboratories split off as the Riverside Research Institute in 1952.

In 1946, a Human Engineering Division was formed at the Naval Electronics Laboratory in San Diego under Arnold Small, whose first criterion for hiring, it seemed, was that an applicant could play the violin in the San Diego Symphony. Small, who had majored in music and psychoacoustics and played in the symphony himself, hired several musicians at NEL, including Max Lund, who later moved on to the Office of Naval Research in Washington, and Wesley Woodson, who published his *Human Engineering Guide for Equipment Designers* in 1954. Outstanding contributions were also made by John Stroud (1955), known for his "psychological moment" concept, and Carroll White (1956), who discovered and validated the phenomenal effect of "visual time compression" on noisy radar and sonar displays.

Similar to the pattern after the First World War, some psychologists remained in uniform, but more, including Grether, Melton, Bray, Kemp, Gagne, Humphreys, Adams, French, Taylor, Mead, and Small, stayed on as civil servants for varying tenures, as did Julien Christensen and Melvin Warrick, who had long careers at the Aero Medical Laboratory at Wright Field. Colonel Paul Fitts wore his uniform until 1949, then joined academia and opened his Laboratory of Aviation Psychology at Ohio State University. Many who had not completed their doctorates went back to graduate school on the GI Bill. A few who had earned Ph.D.s before the war joined universities where they could apply their wartime experiences to the training of a new breed of psychologists.

In Academia

On January 1, 1946, Alexander Williams, who had served both as a selection and training psychologist and as a naval aviator, opened his Aviation Psychology Laboratory at the University of Illinois (Roscoe, 1994). The laboratory initially focused on the conceptual foundations for mission analysis and the experimental study of flight display and control design principles (Williams, 1947, 1980). Soon a second major thrust was the pioneering measurement of transfer of pilot training from simulators to airplanes (Williams & Flexman, 1949; Flexman, Roscoe, Williams, & Williges, 1972). And by 1951 experiments were underway on the world's first air traffic control simulator (Johnson, Williams, & Roscoe, 1951).

Also on January 1, 1946, Alphonse Chapanis, who had served as a psychologist but not as a pilot, joined the Psychology Department of Johns Hopkins University. Initially, Chapanis concentrated on writing rather than building up a large research program with many graduate students, as Williams was doing at Illinois. The result was the first textbook in the field, *Applied Experimental Psychology*, a monumental work for its time and still a useful reference (Chapanis, Garner, & Morgan, 1949). With the book's publication and enthusiastic reception, engineering psychology had come of age, and aviation was to be its primary field of application in the years ahead.

Strong support for university research came from the Department of Defense, particularly from the Office of Naval Research and its Special Devices Center and from the Air Force's Wright Air Development Center and its Personnel and Training Research Center. The Civil Aeronautics Administration provided funds for human engineering research via Morris Viteles and his NRC Committee on Aviation Psychology. In 1950 that committee was composed of Viteles as chairman, N. L. Barr, Dean R. Brimhall, Glen Finch, Eric F. Gardner, Frank A. Geldard, Walter F. Grether, W. E. Kellum, and S. Smith Stevens.

The research sponsored by the CAA via the committee was performed mostly by universities and resulted in a series of studies that became known as "the gray cover reports." Number 84, by A. C. Williams, Jr., and S. N. Roscoe (1949), described the first experimental study of instrument displays designed for use with the new VOR/DME radio navigation system. Number 92, by S. N. Roscoe, J. F. Smith, B. E. Johnson, P. E. Dittman, and A. C. Williams, Jr. (1950), reported the first simulator evaluation of a map-type VOR/DME navigation display employing a CRT in the cockpit. Number 122 described the previously mentioned first air traffic control simulator (Johnson, Williams, & Roscoe, 1951).

When Paul Fitts opened his Laboratory of Aviation Psychology at Ohio State in 1949, he attracted a flood of graduate students, many of them veterans, as Alex Williams had been doing since 1946 at Illinois. Charles W. Simon, Oscar Adams, and Bryce Hartman started the flow of Fitts doctorates in 1952. Simon joined the Rand Corporation in Santa Monica and Adams the Lockheed-Georgia Company in Marietta. Hartman embarked on his long career at the Air Force School of Aviation Medicine in San Antonio. By that time the air traffic control studies for Wright Air Development Center were under way, and Conrad Kraft was developing his "broad band blue" lighting system for radar air traffic control centers (Kraft & Fitts, 1954).

Williams stayed at Illinois until 1955 when he joined Hughes Aircraft Company and fashioned a second career, this time as a practicing engineering psychologist (Roscoe, 1980, 1994). He was succeeded at Illinois by Robert C. Houston for two years and then by Jack A. Adams until 1965, when the laboratory was temporarily closed. Fitts remained at Ohio State until 1958 when he rejoined his wartime friend Arthur Melton, who had moved on to the University of Michigan when Afpatrick was being dismantled (Pew, 1994). Fitts was succeeded by another brilliant psychologist, George Briggs (Howell, 1994). Williams, Fitts, and Briggs all died of heart attacks at early ages, Williams and Briggs at 48 and Fitts at 53.

The laboratories of Williams at Illinois, Chapanis at Johns Hopkins, and Fitts at Ohio State were by no means the only ones involved in the engineering psychology field in the 1940s and early '50s, but they were the ones that produced the lion's share of the engineering psychologists during that period. Other universities with outside support for graduate students doing human engineering research in aviation included Harvard, MIT, California at Berkeley, UCLA, Southern California, Tufts, Purdue, Michigan, Columbia, and Maryland. Several prominent engineering psychologists were mentored by Ernest McCormick at Purdue in the late 1950s and early '60s.

In the Aviation Industry

The students of Williams and Fitts invaded the aviation industry in the early 1950s. The boom was on, especially in southwest Los Angeles where one could park along Airport Boulevard at the east end of LAX Runway 25 Left and see new North American and Douglas planes being rolled out and tested every day. Douglas-El Segundo alone had five different production lines running simultaneously in 1952. From a small hill near the airport, one could see the plants of Douglas, North American, Northrop, and Hughes, which were growing to enormous size, and Lockheed was just over the Hollywood Hills in Burbank. Strange planes like the Northrop flying wing flew low over the Fox Hills Golf Course.

Stanley N. Roscoe was Williams' first student at Illinois and received his Ph.D. in 1950 but stayed on at the lab for two years to complete a flight-by-periscope project for the Navy's Special Devices Center (Roscoe, Hasler, & Dougherty, 1952/1966). Then, in 1952, Roscoe was recruited by Hughes Aircraft Company to organize a Cockpit Research Group and went on to become manager of the Display Systems Department. Earlier that year Walter Carel, who had completed all but his dissertation at Columbia University, was hired by General Electric to do research on flight displays, and William B. Knowles joined GE soon thereafter. In 1955 Charles Hopkins and Charles Simon joined Williams and Roscoe at Hughes, and Knowles and Carel soon followed.

Starting in 1953, several of the airplane and aviation electronics companies hired psychologists, but few of these had training in engineering psychology, and fewer yet had specialized in aviation. As the graduates of the universities with aviation programs started to appear, they were snapped up by industry and by military laboratories as it became painfully apparent that not all psychologists were alike. In a few cases groups bearing such identities as cockpit research, human factors, or human factors engineering were established. In other cases the new hires were assigned to the "Interiors Group," traditionally responsible for cockpit layouts, seating, galleys, carpeting, and rest rooms.

In this environment, Neil Warren in the Psychology Department at the University of Southern California and John Lyman in the Engineering Department at UCLA introduced advanced degree programs for many who would distinguish themselves in the aerospace field. Starting in the late 1940s, Warren had used the human centrifuge on the University of Southern California campus (at that time the only one on the west coast) to do display research. It was in Warren's facility where it was first demonstrated that a single "drag" on a cigarette would measurably reduce the number of *G*'s a pilot could withstand before "graying out" in the centrifuge.

Harry Wolbers, a 1955 Warren graduate, was the first engineering psychologist hired by the Douglas Aircraft Company. Wolbers was the human factors leader for Douglas in their prime contract for the Army-Navy Instrumentation Program (ANIP). Another Warren product was Glenn Bryan, who became the first director of the Electronics Personnel Research Group at the University of Southern California in 1952 and went on to head the Psychological Sciences Program at the Office of Naval Research for more than 20 years. Gerald Slocum, who joined Hughes Aircraft in 1953 and later earned his master's degree with Lyman at UCLA, would rise to be a Vice President of the company and eventually of General Motors.

In the east, Jerome Elkind, a student of J. C. R. Licklider at MIT, formed the original human factors engineering group at RCA in the late 1950s. Lennert Nordstrom, a student of Ross McFarland (Ritchie, 1994) at Harvard, organized the human factors program at SAAB in Sweden in the late 1950s. Thomas Payne, Douglass Nicklas, Dora Dougherty, Fred Muckler, and Scott Hasler, all students of Alex Williams, brought aviation psychology to The Martin Company in the mid 1950s. And Charles Fenwick, a student of Ernest McCormick at Purdue, became the guru of display design at Collins Radio in the early 1960s. Managers in industry were gradually recognizing that aviation psychology was more than just common sense.

In Troubleshooting System Problems

In the late 1940s and early '50s, an unanticipated technological problem arose in the military community, one that obviously had critical human components. The new and complex electronics in both ground and airborne weapon systems were not being maintained in dependable operating condition. The weapon systems included radar and infrared guided missiles and airplanes with all-weather flight, navigation, target-detection, and weapon-delivery capabilities. These systems had grown so complex that more often than not they were inoperable and, even worse, unfixable by ordinary technicians. Few could get past the first step—"troubleshooting" the failures. It was becoming evident that something had to be done.

The first alert on the scale of the problem came from the Rand Corporation in the "Carhart report" which documented a host of people problems in the care of electronic equipment (Carhart, 1953). The technicians needed better training, aiding by built-in test circuits, simulation facilities for practicing diagnoses, critical information for problem solving, and objective performance evaluation. To address these problems, the Office of Naval Research in 1952 contracted with the University of Southern California to establish an Electronics Personnel Research Group with the mission of focusing on the people aspects of maintaining the new systems coming on line.

The original EPRG, organized by Glenn Bryan, included Nicholas Bond, Joseph Rigney, Laddie LaPorte, William Grings, L. S. Hoffman, and S. A. Summers. The reports published by this group during the 1950s (e.g., Bryan, Bond, LaPorte, & Hoffman, 1956; and Bryan, Rigney, Bond, LaPorte, Hoffman, & McAllister, 1959; and Grings, Rigney, Bond, & Summers, 1953) had a major impact on the subsequent efforts of the military to cope with the problems of maintaining electronic systems of ever increasing complexity. The lessons learned from this early work were later set forth in Nick Bond's 1970 *Human Factors* article, "Some Persistent Myths about Military Electronics System Maintenance," which won the Jerome H. Ely Award of the Human Factors Society.

In Consulting

In parallel with these developments, several small companies were organized to provide research, design, and consulting services to industry and the government. Among the earliest of these were Jack Dunlap's Dunlap and Associates, Bob Sleight's Applied Psychology Corporation, Harry Older's Institute of Human Relations, and John Flanagan's American Institutes for Research (Alluisi, 1994, p.16). Of these, the American Institutes for Research and Dunlap and Associates expanded into fields other than engineering psychology. Still, Dunlap and Associates warrants extra attention here because of its predominant association with engineering over a long period and the nature of its contributions.

In 1946, Captain Jack Dunlap separated from the US Navy, joined The Psychological Corporation in New York City, and immediately established a bio-mechanics division (Orlansky, 1994). Dunlap's initial recruits were Ralph C. Channell, John D. Coakley, Joseph Gallagher, Jesse Orlansky, and Martin A. Tolcott. Of this group, all but Gallagher, an accountant, left "The Psych Corp" in 1947 to form what would become Dunlap and Associates in 1950. In addition to its main offices and laboratories in Stamford, Connecticut (until 1963), the company had a sizeable branch office in Santa Monica headed by Joseph Wulfeck.

In the 1950s, Jesse Orlansky of "D&A" played a key role in the forward-looking Army-Navy Instrumentation Program (ANIP), working closely with Douglas Aircraft, the prime contractor, and with Walter Carel of General Electric, the originator of the "contact analog" concept (Carel, 1960, 1961). Two of the best minds in the D&A organization were those of Jerome H. Ely and Charles R. Kelley, but in quite different ways. A memorial plaque describes Ely, who died at age 39, as a "scholar, scientist, teacher and gentle man" (Tolcott, 1994). Kelley, on the other hand, saw a perfect continuum between science and mysticism, but his seminal research on predictor displays and his book *Manual and Automatic Control* (1968) were highly creative contributions.

In Course Setting

During the 1950s, "blue ribbon" committees were frequently called on to study specific problem areas for both civilian and military governmental agencies, and aviation psychologists were often included on and sometimes headed such committees. Three of the most influential committee reports, each of which contained major contributions by Alex Williams, included:

Human Engineering for an Effective Air-Navigation and Traffic-Control System. (Fitts et al., 1951).

Human Factors in the Operation and Maintenance of All-Weather Interceptor Systems. (Licklider et al., 1953).

The USAF Human Factor Engineering Mission as Related to the Qualitative Superiority of Future Man-Machine Weapon Systems. (Fitts et al., 1957).

The air-navigation and traffic-control study by the Fitts committee was of particular significance because, in addition to its sound content, it was a beautifully constructed piece that set the standard for such study reports. The group Fitts assembled included Alphonse Chapanis, Fred Frick, Wendell Garner, Jack Gebhard, Walter Grether, Richard Henneman, William Kappauf, Edwin Newman, and Alexander Williams.

The study of all-weather interceptor operation and maintenance by J. C. R. "Lick" Licklider et al. (1953), though not as widely known, marked the recognition by the military and the aviation industry that engineering psychologists in the academic community had expertise applicable to equipment problems not available within the military at that time. Licklider's committee included George Clementson, Joe Doughty, Bill Huggins, Charles Seeger, C. C. Smith, Alex Williams, and Jay Wray.

Not all of the reports of this genre were the products of large committees. Others written in academia, usually under military sponsorship, included:

Handbook of Human Engineering Data, generally referred to as "The Tufts Handbook," produced at Tufts College under a program directed by Leonard Mead for the Navy's Special Devices Center and heavily contributed to by Dunlap and Associates (Tufts College & US Naval Training Devices Center, 1949), followed by:

Vision in Military Aviation by Joseph Wulfeck, Alexander Weisz, and Margaret Raben (1958) for the Wright Air Development Center. Both were widely used in the aerospace industry.

Some Considerations in Deciding about the Complexity of Flight Simulators, by Alex Williams and Marvin Adelson (1954) at the University of Illinois for the USAF Personnel and Training Research Center.

A Program of Human Engineering Research on the Design of Aircraft Instrument Displays and Controls, by Alex Williams, Marvin Adelson, and Malcolm Ritchie (1956) at the University of Illinois for the USAF Wright Air Development Center. (Adelson went on to form the first human factors group in the Ground Systems Division of Hughes Aircraft, and Ritchie formed his own research and consulting company in Dayton, Ohio.)

Perhaps the two most influential articles in the field during the 1950s were:

"Engineering Psychology and Equipment Design," a chapter by Paul Fitts (1951) in the *Handbook of Experimental Psychology* edited by S. S. Stevens, the major source of inspiration for graduate students for years to come, and

"The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information" in the *Psychological Review* by George A. Miller (1956), which encouraged quantification of cognitive activity and shifted the psychological application of information theory into high gear.

Historical Perspective

Taken as a whole, these key reports and articles—and the earlier research on which they were based—addressed not only pilot selection and training deficiencies and perceptual-motor problems encountered by aviators with poorly designed aircraft instrumentation but also flight operations, aircraft maintenance, and air traffic control. All of these problem areas have subsequently received serious experimental attention by engineering psychologists both in the United States and abroad. There are now some established principles for the design, organization, maintenance, and operation of aviation systems that have application beyond the immediate settings of the individual experiments on which they are based.

The early educators in the field—Alex Williams, Al Chapanis, Paul Fitts, Ross McFarland, Len Mead, Lick Licklider, Neil Warren, John Lyman, Jack Adams, George Briggs, and Ernest McCormick—had in common a recognition of the importance of a multidisciplinary approach to aviation problems, and their students were so trained. The early giants, on whose shoulders we walk, could only be delighted by the extent to which all researchers and practitioners now have access to once unimagined information and technology to support creative designs based on sound behavioral engineering principles.

ACKNOWLEDGMENTS

In preparing this historical review, we have drawn on articles by Earl Alluisi (1994), Paul Fitts (1947), and Jefferson Koonce (1984); on the short biographies of George Briggs, Jack Dunlap, Paul Fitts, Ross McFarland, and Jerome Ely, respectively, by Bill Howell, Jesse Orlansky, Dick Pew, Malcolm Ritchie, and Marty Tolcott in the monograph titled *Division 21 Members Who Made Distinguished Contributions to Engineering Psychology*, edited by Henry Taylor and published in 1994 by Division 21 of the American Psychological Association; and on Mac Parsons' 1972 book *Man-Machine System Experiments*. We also received valuable personal communications about "Afpatrick" from Jack Adams (1995) and about the USC Electronics Personnel Research Group and the strange planes flying low over the Fox Hills Golf Course from Nick Bond (1995).

Intentionally blank

Predicting Human Performance

2

Background

While the early engineering psychologists were primarily concerned with the human factors in equipment design, many other psychologists were involved in personnel selection and the transfer of training in simulators to the operation of complex systems—mainly airplanes but also some fairly complex weapon-controlling systems. In aviation, the involvement of psychologists began with pilot selection during the first World War, and during World War II they became deeply involved. The results of the psychomotor and other tests they developed and used were submitted to statistical analysis, and the modest conclusion was that they could account for about 20% of the variance in pass/fail numbers for pilot, navigator, and bombardier trainees.

The Roots of WOMBAT-FC

The Need

For most of this century, psychologists have been developing precise measures of human intelligence and somewhat less precise but nonetheless useful instruments for describing human personality factors. Unfortunately, they have been less successful in assessing human aptitudes for operating nuclear reactors, controlling air and surface traffic, directing civil disaster responses, and providing emergency medical services, to name but a few of the many complex operations humans perform daily. In recent years, with the advent of high-speed computers, the military have invested heavily in the development and validation of selection batteries that now account for more than 25% of the variance in training success but still have no evident correlation with operational performance after training.

The need for valid tests of complex operational aptitude is increasing as the explosion in information technology and associated automation makes more complex operations possible and the cost of placing the wrong person in charge greater than ever. Increasing the information available gives the operator more to attend to, and automation makes it all the more important and difficult to keep track of everything that is going on and decide when some intervention is critical. This is now called **situational awareness** (or situation awareness; Gawron, 1999), and this ability is also centrally involved in **total resource management (TRM)**.

The costs of haphazard personnel selection are not limited to those resulting from bad judgment and mismanagement of critical operations. **It is also costly to invest in the training of individuals who fail to reach criterion performance levels after training or, worse yet, pass all training tests but then are unable to stand up under operational stress**. As so often happens with trainees, the individual may have all of the skills and knowledge normally required but be unable to put them together in the confusion of a complex incident.

The Difficulties

The failure to develop tests of high predictive validity for complex operational aptitude has been caused by several factors, **the first of which is the usual clouding of operational performance criteria against which to validate any such test**. If measures of complex job performance are unreliable, as they typically are, there is no way that the high predictive validity of a test can be shown statistically. The pass-fail criterion would be of value if approximately equal numbers of trainees passed and failed, but when the ratio is four or five to one, as in many training programs, it is almost worthless. Rating scales are no better when almost all trainees are given the same grade.

Aside from the criterion problem, the effectiveness of aptitude tests has been limited by the notion that performance of complex operations depends on a collection of individually simple abilities. Consistent with this idea, batteries have sampled reaction time, manual dexterity, short-term memory, spatial orientation, and the like. The best and most widely used of these was developed for the United States Air Force by LCOL Hector Acosta (USAF, Ret.). The fact that such batteries account for only about 25 percent of the variance in training success results in part because basic abilities measured are highly correlated with one another. **Any one or two of the tests provides almost as much predictive power as the entire battery.**

The Secret

The secret of operational aptitude testing is to recognize the complexity of what we are trying to predict and construct a measuring instrument of similar complexity. The fact that expanding a test battery adds little predictive validity does not mean that a selection test should be short to be cost effective. It is wishful to expect situational awareness and stress tolerance to be revealed reliably in a short test. If a day or even part of two days is required by most candidates to approach a terminal performance level on an aptitude test, its application would still be cost effective if only candidates of high aptitude were selected and the probable failures were rejected before large sums had been invested in their training.

While situational complexity is necessary to test situational awareness, it is not sufficient. To avoid confounding basic aptitude with the effect of prior training in specific tasks, the elements that comprise the test must be unlike any real-world activities such as operating computers or controlling specific vehicles. Furthermore, the individual subtasks must be sufficiently simple to allow their mastery in a short practice period before combining them in the test situation. Sufficient situational complexity can be achieved by the manner in which the individually simple subtasks are combined in an adaptive scenario involving multiple sources of information and multiple response alternatives.

The controller of a complex flow of information must search for, evaluate, and integrate information about all relevant events, conditions, and resources, quickly assess changes in situational priorities, and allocate attention accordingly. To determine an individual's aptitude for meeting these demands requires a complex test in which high scores depend on:

- Finding out what's important now and in the long run and allocating priorities accordingly;
- Perceiving a situation correctly by avoiding preconceived assumptions and subjective biases and being vigilant;
- Discovering rules that are not explicit through induction and deduction;
- Recognizing serendipitous opportunities quickly and seizing them before they pass;
- Ignoring irrelevant distractions and tolerating frustration when things are going badly;
- Coping with the stress of high workload periods and poor performance indications; and finally
- Coping with the boredom of routine tasks and resisting complacency during periods of low workload.

3 WOMBAT-FC

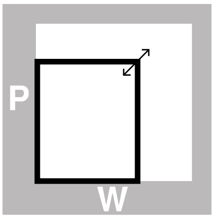
The Scenario

Overview

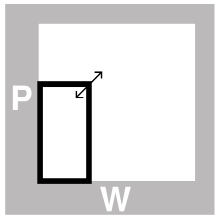
The PC-based **WOMBAT-FC Situational Awareness and Stress Tolerance Test**™ is designed to embody all the demands and constraints described in Chapter 2. The individual tasks involve keeping track of multiple moving targets, spatial orientation, pattern recognition, and short-term (working) memory, and on each a testee can reach his or her asymptotic performance level after a short practice period.

The target tracking involves a number of targets moving on a small grid at a slow and constant speed, making the primary task one of "keeping the picture" rather than placing undue emphasis on motor skills. In a 3-D figure rotation task, two figures have to be rotated manually and/or mentally to reveal whether they are the same, mirror images, or different in some other way. In a quadrant-location task, as each pattern of numbers is learned, it is replaced by a different pattern. A two-back serial digit-canceling task is both tediously boring and frustrating.

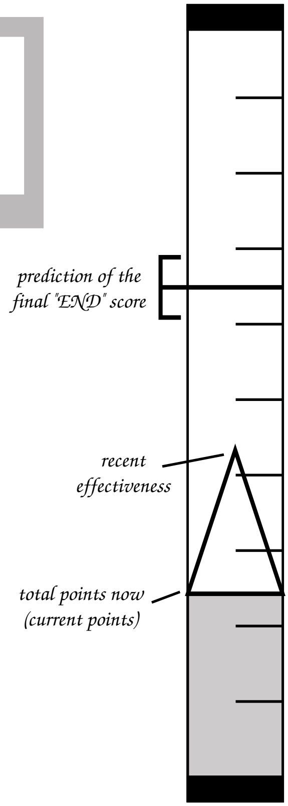
These tasks comprise the menu of scoring alternatives available to the testee on request. Each is relatively culture-free in that it has no real-world counterpart, and each can be learned quickly by the apt testee. **The attention demands of the WOMBAT-FC test are expanded by the ever changing information presented by peripheral indicators**. To score well the testee must monitor the peripheral indicators vigilantly to follow the shifting priorities of the various activities as indicated by their potential scoring worths and current scoring rates and to detect indications of failure modes that may require immediate termination of one activity in favor of another.



Tracking



Bonus



The testee's performance is monitored continuously during the test. Two Worth-Performance indicators are always visible in the top corners of the display and serve as guides to the testee in making the proper choices, establishing the best strategy, and monitoring scoring progress. The left indicator relates to the primary task (target tracking), and the right indicator shows the secondary (Bonus) task worth and performance levels.

The product of the Worth (W) and Performance (P) represents the current scoring rate and is computed into an overall index of recent effectiveness that is also continuously displayed for the testee (shown at the right.)

An indication of total points now (current points) and a prediction of the final score (the "E" symbol), based on current points plus current effectiveness extrapolated for the time remaining, complete the thermometer-shaped total scoring display. While performing the WOMBAT test, the testee receives constant performance feedback and extrapolated outcome based on his/her previous choices. The testee is expected to make good use of these indicators in determining the best course of action.

The Primary Task: Keeping Track of Moving Targets

The Target Tracking Task displays up to 12 targets moving along the lines of a 5-by-5 grid as shown in Figure 1. Each target receives a route assignment from the computer before entering the grid at one of the 12 numbered intersections. The targets follow their individual routes at a constant and rather slow speed before leaving the grid at another of the numbered intersections.

Targets are identified by a single letter of the alphabet. Some targets are said to be active, others to be inactive, waiting for activation. Inactive targets have not yet entered the grid, but each, in turn, will become active when it does. They are displayed in gray at the top of the list of letters representing targets at the left edge of the display.

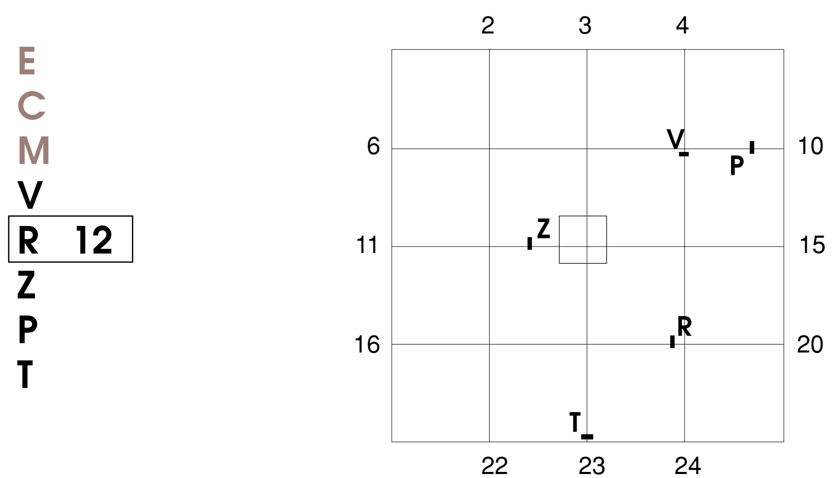


Figure 1. WOMBAT-FC Target Tracking Task. The left-hand task is to position a rectangle over the desired letter representing a target in a list at the left side of the WOMBAT display. This function is called "selecting a target" and is used to find important information about the targets. The first three letters on top of the list represent inactive targets that have not yet entered the grid. Below are colored letters that match active targets moving on the grid. The right-hand task is to position a square cursor on the 5-by-5 grid. This cursor is used to find a missing target that disappeared from the grid and also to designate colliding targets.

Active targets are displayed in one of three bright colors, and the letters are the same colors as their associated targets. Each active target (and its identifying letter) moves along its specific route. Once assigned, the route cannot be changed. When one target leaves the grid, it can be replaced by another target entering at a different intersection.

WOMBAT-FC's Autotrack Mode

The computer normally keeps track of all moving targets as they proceed along their assigned routes on the grid. This automatic tracking mode is called "Autotrack." Autotrack's only function is to track targets on the grid; it has nothing to do with determining their number, their routes, or their separation.

When Autotrack is performing properly, it frees the testee to work on other tasks to earn additional points. However, Autotrack is prone to failure, and it must be monitored continually even though the testee is working on another task. When Autotrack fails, it loses track of one of the moving targets which becomes invisible, resulting in a loss of tracking performance and an indication of a failure on the tracking-performance display.

When Autotrack loses track of a target, the target continues along its assigned route on the grid, even though it is no longer visible. The testee must then reacquire the missing target and return it to Autotrack as quickly as possible.

Reacquiring a Missing Target

When a target goes missing, the testee will have to take manual control and begin the search procedure to find the missing target quickly. Emphasis is placed on quick resolution of Autotrack failures because it demonstrates good mental modeling ("keeping the picture") of the dynamic situation on the grid, even when the grid is not visible because other tasks are temporarily being performed.

Testees interact with WOMBAT-FC by means of a console containing two joysticks and a 13-button keypad. The keypad consists of 10 numeric keys (0-9), left and right arrow keys, and a key labeled "Bonus." The left-hand joystick controls the position of a small rectangle that frames a target-identifying letter in a list displayed along the left edge of the WOMBAT display. The right-hand stick is used to position a cursor on the grid for reacquiring a missing target (or for predicting a collision between two moving targets, to be explained later).

To reacquire a missing target, the testee must first select the target's identifying letter from the list of letters at the left edge of the display using the left control stick.

The right control stick is then used to display and move a square cursor on the grid. When the cursor is positioned over the invisible missing target for a brief time, the missing target reappears inside the cursor. When the testee turns the cursor off with the missing target visible inside, the missing target is reacquired by Autotrack.

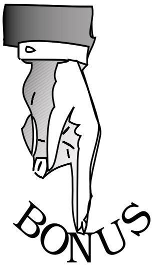
The number of targets on the screen at any given time will depend in part on how well the testee keeps track of targets and how quickly the testee reacquires missing targets. The number of targets entering the grid increases until the testee can handle no more and decreases when missing targets are not found quickly.

Predicting Collisions of Targets

Two targets of the same color (blue, yellow, or purple) will collide if they meet on the grid (even though one of the targets may be "missing" and invisible). When this happens, both targets disappear from the grid and a "starburst" symbol marks for 10 seconds the spot where they collided. The letters representing these targets in the column on the left of the display will also disappear following the collision.

A testee with good situational awareness can use the square cursor to designate one or the other of the targets before the collision. When the prediction is correct, the computer will record a good collision prediction, change the color of the designated target to resolve the conflict situation, and reward the candidate accordingly.

SECONDARY TASKS: THE BONUS POUCH



There is another source of activity for the testee. Known as WOMBAT's "Bonus pouch," it is the source of three side tasks, each of one-minute duration. Performing these tasks can yield rewards and penalties in various forms. A bonus task can be requested whenever the testee elects to move away from the primary task into this "secondary" level of activity.

Keeping the picture of the flow of traffic and quickly reacquiring missing targets is the "primary" task in the sense that it cannot be ignored without serious penalty (the routine must be maintained). The bonus tasks are "secondary" in that the testee may at any time suspend them and return to the primary task without any penalty in the active bonus task when its play is resumed. The secondary tasks provide the problems and opportunities for the testee to demonstrate not only situational awareness and procedural compliance, but also spatial orientation with mental rotation, temporal and spatial pattern recognition, and short-term working memory.

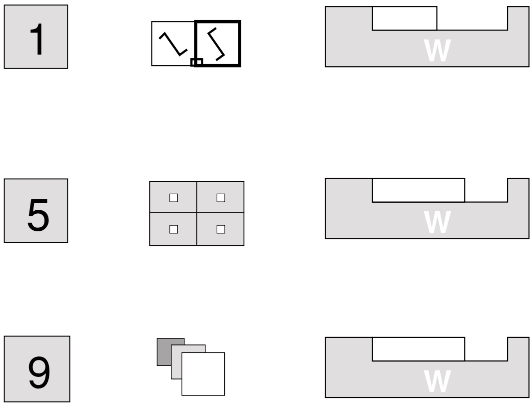


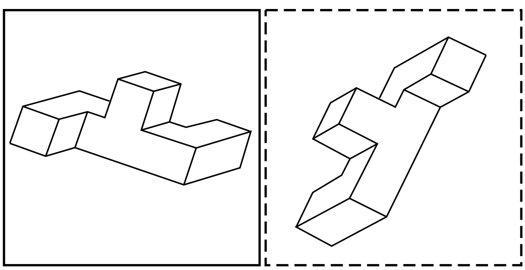
Figure 2. WOMBAT-FC Bonus Menu. This represents the Bonus menu. In WOMBAT-FC each choice of a bonus task is presented to the testee as a pictogram to minimize the use of language and reduce the risk of culture-based biases. The widths of the horizontal bars to the right of the pictograms indicate the momentary worths of the three tasks. Choosing and performing Task 9 will decrease its subsequent worth and increase the worths of Tasks 1 and 5.

The three selectable bonus tasks are:

- a 3-D **Figure Rotation** and matching task adapted from Shepard and Metzler (1971), requiring spatial orientation (labeled "1" in Figure 2),
- a sequential **Quadrant-Location** task involving graphically presented temporal mazes. Each pattern of numbers recurs until learned, and then it is replaced by a different pattern (labeled "5" in Figure 2), and
- a Two-Back **Digit-Canceling** task of short-term memory (labeled "9" in Figure 2).

Vigilance is an important aspect of situational awareness when monitoring a complex traffic situation, and time-sharing of attention is required during performance of any bonus task to catch and respond to Autotrack failures. Let's have a quick look at each bonus task.

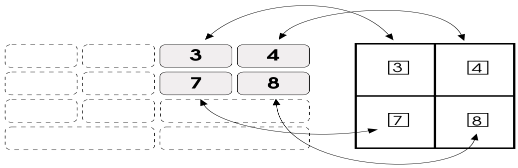
The **Figure-Rotation** task displays two 3-D figures, side-by-side. One figure will be inside a solid square and the other figure will be inside a dashed square. The solid square means that the figure can be rotated using the two sticks on the WOMBAT console; it is the "active" figure.



The testee rotates the active figure until he/she sees all the details of its construction. Then the testee transfers control to the other figure by pressing the right (or left) arrow key on the keypad (as appropriate), and studies it as well. The goal is to find out, as soon as possible, whether the two figures are identical, mirror images, or otherwise different. Maximum points are earned by giving the computer the correct answer as quickly as possible.

When a correct answer is given, the testee is offered the option of beginning another 3-D problem while there still is some time left. The testee should try to solve as many problems as possible during the one-minute trial. When an incorrect answer is given, no other problem is offered for the remainder of the minute.

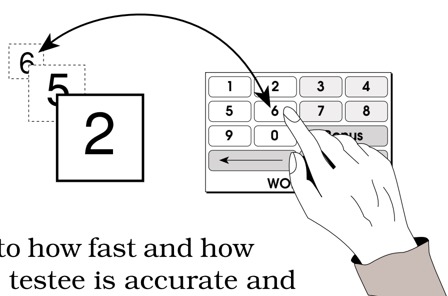
In the **Quadrant-Location** task, the numbers "1" to "32" are placed in groups of eight in the four quadrants on the display. The task is to find each number in ascending order and press the button on the keypad that corresponds to the quadrant in which it lies, thereby canceling the numbers in sequence, 1 through 32.



The quadrant designation number appears in a box in the center of each quadrant. The "3" button corresponds to the upper-left quadrant, the "4" button to the upper-right quadrant, the "7" button to the lower-left quadrant, and the "8" button to the lower-right quadrant. The illustration shows the display quadrants and their respective keypad buttons.

When all 32 numbers have been canceled (with few errors) before the end of a 60-second trial, the testee is offered the option of beginning another Quadrant-Location problem while there still is some time left. The computer records and displays in the scores sheet the number of sequences mastered by the candidate. Until a sequence is canceled efficiently, it reappears each time the Quadrant-Location task is chosen from the Bonus Menu. When a given sequence is mastered, a different one is automatically presented next time the task is invoked.

The Two-Back **Digit-Canceling** task briefly displays a single digit from 1 to 8 inside a square drawn within the center of the display. Starting with the third digit displayed, as soon as each new digit appears, the testee is required to press the key on the keypad matching the digit displayed two back in the sequence. Once the answer has been given, right or wrong, another digit appears, inviting the testee to match the next "two-back" digit in the sequence. Each time an answer is given, the interval between the digits is adjusted according to how fast and how accurate the answers are. If the testee is accurate and fast, the interval will become shorter and the reward will increase. If the testee is wrong or slow, the reward decreases.



The testee must remember the digits as they briefly appear and must also remember the last two digits if an interruption becomes necessary to deal with the primary task. The illustration shows the action of answering digit "6" after "5" and "2" have appeared.

Testing Strategy

All four tasks, one primary and three secondaries, are relatively culture-free in that each is unlike anything called for in controlling any complex real-world system, and each can be learned quickly by the apt testee. As you will see later in this manual, typical group learning curves on the WOMBAT-FC test initially show a slight, gradual improvement in the rate of scoring, with relatively uniform scoring after the third to fifth 10-minute segment.

The attention demands of the test are increased by situational information presented by the peripheral indicators (as seen earlier). To score well, the testee must monitor the indicators vigilantly to follow the shifting priorities of the various tasks as indicated by their potential scoring worths and current scoring rates and to detect indications of failure modes that may require immediate termination of one activity in favor of another.

The task structure places a high premium on a rational attention allocation strategy. The background target tracking task, with its unreliable Autotrack generating a performance score on the appropriate indicator, rewards the allocation of some portion of the testee’s attention to the known signals of variability: the changing worths of the different tasks and the running overall indices of current scoring and predicted final score.

How well the tracking and the bonus tasks are performed is important, but only in proportion to their momentary worths. To maintain high worths for all tasks requires that all be performed on a regular basis. To maintain maximum target tracking worth, bonus tasks must be performed frequently, and conversely, to maintain maximum bonus worth, missing targets must be reacquired quickly. Each time a bonus task is performed, it loses an increment of worth, and the others gain increments. As a result the testee's scoring rate depends more on what task is chosen moment-to-moment than on how well it is performed. The key to a high final score is effective management of the task worths.

Tolerance of frustration is tested in part by intervals in the test when the very best that can be done will produce a low rate of scoring. At other times, for example, when a bonus task has ended with an incorrect answer in less than a minute and Autotrack is performing well, there is nothing to be done but rest and think.

The Origins of WOMBAT Bonus Tasks

We are frequently asked why we chose these particular Bonus tasks. Where did they come from? What do they measure individually?

Figure Rotation

"Mental imagery" tasks have received far more experimental attention than any other type in recent years, largely as a consequence of the brilliant work of Roger Shepard, the Ray Lyman Wilbur Professor of Social Science at Stanford University. Our figure rotation task is an adaptation of the task used by Shepard. In 1995, President Clinton awarded Shepard the National Medal of Science, "For his creative theoretical and experimental work elucidating how the human mind perceives and represents the physical world and why the human mind has evolved to represent objects as it does ... ."

In 1968 the field of cognitive psychology was dominated by theories of artificial intelligence based on the assumption that all thinking involved the manipulation of discrete mental symbols. But Shepard was convinced that some thought processes were nonsymbolic, that they are more like continuous simulations of external events. Then Shepard hit upon a great idea. Using solid block figures, he and two new graduate students, Jacqueline Metzler and Lynn Cooper, embarked on a series of experiments on mental rotation later reported in a bombshell paper in *Science* (Shepard & Metzler, 1971).

They had shown students pairs of pictures of objects in different spatial orientations. Sometimes the objects were the same and sometimes not, and they measured the time the students took to decide. The greater the difference in the orientation of the two objects, the greater the decision time. It became apparent that the students were making comparisons by "mentally rotating" one of the two objects into the same orientation as the other. The time differences even provided an indication of the rate of mental rotation.

But once again, there were wide differences in the mental rotation speeds of individuals, and these differences are believed to be directly related to how people perceive and interpret complex visual scenes in everyday life or in controlling dynamic operations—as well as in the laboratory. Our Figure-Rotation Task differs from Shepard's in that the computer-animated figures can be rotated manually, but there are still wide differences in decision times depending on how much individuals rotate the figures manually versus their facility in mental rotation.

Quadrant Location

This pattern recognition test is an adaptation and extension of a paper-and-pencil test used by Professor Donald Johnson in an experimental psychology class at the University of Illinois in 1946. We have no idea where that test came from and have lost touch with Professor Johnson, but graduate students were mightily impressed with the wide individual differences in the scores of an already highly selected group of psychology students. Our version of the task has some of the characteristics of a temporal as well as spatial maze, and it is inherently motivating—people like it. The premium is on the quick recognition of each new pattern of numbers.

Digit Canceling

Running memory tasks have a long history, with much of the early work done in England. At Cambridge University, Harry Kay (1953) systematically investigated delayed digit canceling, with 1-back, 2-back, 3-back, and 4-back responses called for. Immediate and 1-back responses were almost without error. From 1-back to 4-back, error rates rose rapidly. The 4-back task was impossible for many; a few were able to develop rehearsal strategies that worked occasionally. Later it has been found that a longer interstimulus interval than Kay used allows some to handle the 4-back responses consistently.

For his Ph.D. dissertation at the University of Illinois, Robert North (1977) paired the 1-back task concurrently with each of three other tasks designed to measure immediate memory, classification, and tracking abilities, respectively. He also varied the priorities of the tasks and the difficulty of the tracking task to elevate the attention demands to the individual testee's saturation point. North's dual-task measures and others developed by Diane Damos (1972) proved effective in predicting pilot performance in primary flight training (notably, in experiments by North & Gopher, 1974, and Jacobs, 1976). The WOMBAT-FC test is an extension of their concept.

VALIDATION 4

IN QUEST OF THE IDEAL

The perfect system for selecting dynamic situation controllers would have several qualities. It would be:

- **COMPREHENSIVE**, meaning that the system would not depend on any single attribute of successful controllers but would address cognitive, psychomotor, and medical factors as well as situational awareness.
- **RELIABLE**, meaning that any test used would yield test-retest scores that are highly correlated.
- **DISCRIMINANT**, meaning that any test used would yield a wide range of scores normally distributed over the spectrum of human performance.
- **OBJECTIVE**, meaning that the scores of individual candidates would not be biased by the subjective evaluations of individual examiners.
- **CULTURE FREE**, meaning that performance on any test used would not be biased or otherwise affected by race, gender, or prior training and experience in operating specific complex systems.
- **VALID**, meaning that any test used in the selection system would measure what it is intended to measure, namely, the criterion of future success in dynamic situation controllers.

The history of selection-test validation has been a frustrating exercise yielding, at best, quite modest results. The well-known criterion problem (discussed in the next section) is not the only culprit. Subjective measures such as interviews and peer ratings are notoriously unreliable and hence of low validity. Batteries of individual "basic abilities" tests have some predictive power but still account for only about 25 percent of the variance in training success and have no documented correlation with operational performance. The WOMBAT family has a theoretical basis supported by the experiments done at Illinois in the '70s, and its demonstrated operational validity is strong evidence that this approach brings selection testing to a new level.

During the 1970's at the Institute of Aviation at the University of Illinois, 44 graduate students earned masters degrees and 18 earned doctorates based on experiments done at the Institute's Aviation Research Laboratory. Several of these experiments dealt directly with the predictive validity of various divided-attention tests calling for high degrees of situational awareness. In others, the same or similar tests were used to extract individual differences among trainees through analyses of covariance and thereby increase the power of the experiments.

The tests that showed the highest predictive validities were ones having much in common with the WOMBAT family. These tests involved multiple sources of information with shifting priorities calling for time-sharing and frequent reallocation of attention. In all cases the subtasks to be performed were unlike any subtasks in complex real-world systems. This was done both to guard against the possible transfer of skills acquired through prior training in operating specific devices or systems and to minimize any biasing effects from cultural differences due to race, language, physical activities such as sports, or computer facility.

The other key feature of such tests is that they must extend the testee to full attention-capacity saturation, at which point situational awareness is taxed to the edge of breakdown. To achieve this effect, such tests need to be automatically adaptive, increasing the flow of information and response demands until performance starts to deteriorate, then backing off to keep from crossing the threshold that leads to breakdown. The adaptive logic of the WOMBAT tests has been fine-tuned over the past decade to achieve precisely this necessary balance between stress and overload that keeps the testee working at his or her situational awareness limit.

Validation Requirements

The difficulty of developing tests of high predictive validity for operational aptitude involves several factors, the first of which is the usual clouding of operational performance criteria against which to validate any such test. As discussed in Chapter 2, if measures of operational performance are unreliable, there is no way that the true predictive validity of the test can be shown statistically. The pass-fail criterion is virtually useless when all operational personnel are given whatever amount of simulator refreshment is needed for periodic recertification, and rating scales are no better when almost everyone receives the same grade.

The objective evaluation of a test of situational awareness requires a valid criterion of operational success, one that is unlike any of our traditional validation criteria. In aviation, given the fact that instructors' ratings and pass-fail tests do not discriminate among pilots accurately, where can the investigator turn? Surely measures of performance during training, no matter how objective, are not ideal criteria, because the ultimate purpose of aptitude tests is not to predict immediate success but *distant* future success as a complex system operator or dynamic situation controller.

In an *ideal* validation study, a large number of controller applicants would be tested, all would be trained, and all who completed training, whether certificated or not, would be assigned to control operations and their performances observed and evaluated objectively over an extended period. In addition, all would be retested on the original selection test, and a control group that received no training in the interim would be tested and retested to assess the effect of the first administration on the second. In the *real* world of air traffic control training and other information monitoring functions, none of these conditions, save the first, is practical.

Because a longitudinal study of all the same people over several years is not a feasible approach, an alternative plan is required—one that will still address the predictive validity question in a realistic manner. One such approach is a stratified experimental plan in which independent groups of individuals representing the various stages in the sequence of controller training and increasingly complex operations are tested. The mean test scores for the successive groups are then compared statistically to assess the reliability of the anticipated successive increases in the group means.

In a general way, such studies support the original premise that the individual cognitive tasks and the underlying abilities they are designed to measure are relatively unimportant in the context of situational awareness. Rather, what is important is how these relatively easy and largely culture free tasks are managed to maximize overall performance. To be sure, the ability to perform the individual tasks does have an impact on the rate of scoring, but its contribution is secondary to the management of the relative worths of the four tasks by optimum attention allocation—staying on top of the situation and working on the task with the highest momentary priority.

Introduction to DuoWOMBAT-FC

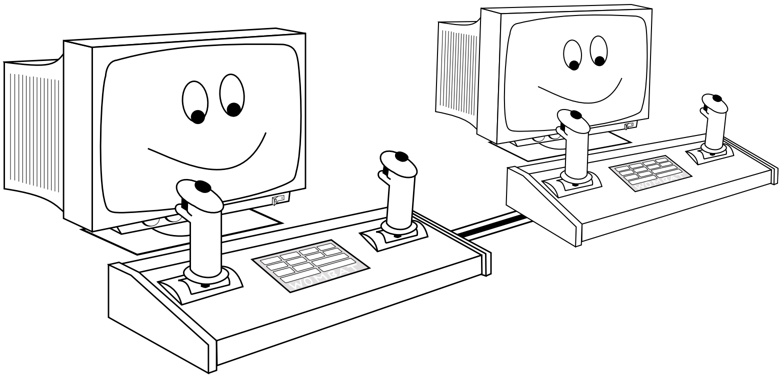
5

Team Testing and Training

The solo WOMBAT-FC Situational Awareness and Stress Tolerance Test is designed to assess the inherent aptitude of individuals to monitor and control complex traffic situations without regard to their interactions with other teammates. Team operations call for additional personal attributes, primarily social in nature but with a cognitive component. Recognition of this fact by air traffic and transit agencies as well as government regulators has led to worldwide formal training in team resource management. By consensus, TRM works.

However, until recently there was no test specifically designed to call for the working exercise of those traits. As the WOMBAT-CS test came into use by airlines and air and surface traffic control agencies, it soon became apparent that the higher-order cognitive demands it imposes on individuals could be extended to encompass the social aspects of team performance. To measure how well team resources are managed, the solo WOMBAT-CS was expanded into the DuoWOMBAT-CS Crew Resource Management Test, and now the same expansion is being applied to the DuoWOMBAT-FC Team Resource Management Test.

DuoWOMBAT-FC



Sitting side-by-side at two linked WOMBAT stations, testees jointly work out their strategies for trading off duties to maximize the team's combined score. The original solo scenario was expanded by the addition of duet versions of the primary tracking task and each of the three secondary bonus tasks. Other modifications were made to generate potential conflict situations between teammates, and the scoring logic and weights were adjusted appropriately. The fully scriptable scenario allows the test administrator to tailor the environment for team assessment or teamwork training. A typical testing scenario would consist of two 30-minute phases of dual performance sandwiched between three 10-minute solo phases (10-30-10-30-10). The three solo phases provide a baseline for each individual against which the combined team performance is evaluated.

Evaluating TRM

Objective evaluations of TRM training and testing are surely not beyond reach, but a new experimental approach may be called for. By consensus, the TRM training programs have an observable effect in the desired direction on team behavior in the workplace. If that is indeed the case, the change in the individual team member's behavior should be reflected by improved performance on the DuoWOMBAT-FC test (beyond that attributable solely to taking and retaking the test). Conversely, continued practice on the DuoWOMBAT-FC would be expected to develop team behavioral attitudes and strategies that would readily transfer to the operational situation.

Installation of WOMBAT-FC 6

The following chapters are intended to assist the supervisor in charge of the WOMBAT-FC implementation. Described are different procedures regarding:

- proper installation techniques,
- how to run the software,
- language issues,
- accessing and understanding the scores,
- running an in–house validation program,
- contacting the ÆRO INNOVATION Technical Support Department for supplementary information.

To benefit fully from the second part of this book, the reader should know the following basic MS-DOS commands and <keyboard keys>:

- **<Enter>**
- **<ESC>**
- **A: B: C:**
- **CD**
- **CD**
- **DIR**
- **DIR/W**
- **DIR/P**
- **DEL**
- **EDIT**
- **COPY**
- **XCOPY**
- **<CTRL BREAK>**
- **AUTOEXEC.BAT**

INSTALLATION

WOMBAT-FC runs on most PC–compatible computers. If you need to install WOMBAT in a computer other than the one supplied by Aero Innovation please refer to the annex at the end of the present manual. You may contact Aero Innovation Technical Support for the proper installation procedures and the latest updates on the WOMBAT-FC program. The phone number is +1.514.336.9310, the email connection is: info@aero.ca, and technical information is also found on the Internet at: http://www.aero.ca.

Each WOMBAT-FC system requires:

- a PC–Compatible computer equipped with a math coprocessor (sometimes called FPU or 387; not required with Pentium-equipped computers) that also contains the following:
  - — an internal hard disk
  - — a high density 3.5" floppy disk drive
  - — a VGA 512kb (min) graphics board
  - — a WOMBAT interface card (supplied) equipped with one DB-25 socket connector
- a keyboard
- a VGA color graphic monitor
- a WOMBAT console (supplied)
- a 2-meter long DB-25 connector cable (supplied)

If you unpack the computer from Aero Innovation, check that it has all the internal components listed above installed. Choose a quiet environment for WOMBAT-FC where the candidate can be left in isolation while he/she takes the test. Arrange the console on a table in a comfortable position. Connect the VGA monitor and the computer keyboard, then plug the DB-25 cable into the connector on the interface board. Push it home firmly taking care not to bend any of the pins. The plugs are polarized, so that each end can have only one position. Connect the other end to the socket on the console. At this point the computer and monitor can be turned on. After the computer has booted in MS-DOS mode, the screen shows:

C:\>

LAUNCHING WOMBAT-FC

We have stored the WOMBAT-FC program and all the necessary files in a subdirectory named after the version of the WOMBAT-FC program current at the time of shipment. For example, Version 1.3 of WOMBAT-FC will be stored on the hard disk of your computer in a directory named **WOMBATFC.V13**.

Subsequent updates of the program can be stored in different directories to insure a smooth transfer from old versions to new ones, prior to deleting superseded versions. To run WOMBAT-FC, switch to the appropriate subdirectory, the one that contains the WOMBAT-FC software, by typing:

CD WOMBATFC.V13 <Enter>

The screen will then show:

C:\WOMBATFC.V13>

If you received your computer from Aero Innovation, or if you installed new WOMBAT-FC software using the supplied installation routine, then the batch command **WOMBAT.BAT** will load in sequence everything that is required for the WOMBAT-FC test. Just type:

WOMBAT <Enter>

and you will have **HARDWARE**, **METAWINDOW** and **WOMBAT-FC** in the appropriate sequential order. You can also invoke a fresh WOMBAT-FC session by typing the same command stored in the **WOMBAT.BAT** batch command:

WOMBATFC <Enter>

The HARDWARE™ Program

The procedures described in this manual will ensure that each time you invoke the WOMBAT-FC program the diagnostics software named **HARDWARE**™, will appear on the monitor. You can also explicitly invoke this program whenever you want by typing the command **HARDWARE <Enter>**. This self-explanatory program is used to verify the integrity of all the cable connections and hardware assembly before you begin the WOMBAT-FC test. While watching the screen:

1. Move systematically each joystick and observe that full deflection causes the associated red cursor to be displayed full scale. Also check that each motion has no effect on any other axis. A noticeable delay between your hand motion and the cursor movement is an indication that there is a problem with the associated joystick and that **the test should not be administered**. Contact Aero Innovation for a replacement joystick.
2. Activate each of the joysticks' thumb and trigger switches and observe the respective red cursors appearing and disappearing on the scales.
3. Press each button on the WOMBAT keypad and observe the change of color of the respective keys on the screen.

HARDWARE™ will verify the integrity of the circuitry of each control you use. You can access its HELP by pressing the F1 key on the computer keyboard. Once the hardware test is satisfactory, press the ESC key on the computer keyboard to exit the program **HARDWARE**™. You are then ready to enter the WOMBAT-FC test.

If one or more connections do not test properly, please make sure that the supplied DB-25 cable is properly connected and fastened to both the console and the WOMBAT computer interface connector located at the back of the computer. The interface card is marked "WOMBAT" just below the DB-25 connector. Be careful not to connect the cable from the console to the parallel printer connection on the back of the computer, which uses the same connector type as the WOMBAT interface card.

If the cable doesn't seem to be the problem, something might have happened during shipping. Please call Aero Innovation's

Technical Support Department. During the warranty period do not open the WOMBAT console without prior authorization from an Aero Innovation Technician.

WOMBAT-FC's Graphics Environment

The WOMBAT-FC program needs a graphics environment named METAWINDOW to run properly. Although the command to load METAWINDOW is normally stored in the batch command **WOMBAT.BAT**, you can manually load it by typing:

METASHEL/I <Enter>

The graphics environment will remain loaded until the command:

METASHEL/K <Enter>

is used to unload it or the power is turned off. These two commands will only work if MS-DOS can find the program METASHEL in the current directory or somewhere that the path environment variable points to. Refer to the MS-DOS User Manual for details on paths.

Intentionally blank

RUNNING WOMBAT-FC 7

REFERENCE STRING AND FILE ANNOTATION

When WOMBAT-FC is launched as a new start, the user is asked to enter a reference string and a file annotation and then to confirm the current date (this is a Year 2000 protection.) The mandatory reference string is used by the program to create a DOS file for the storage of the test results. If the string has a "." then the string as you enter it will be used to name a DOS file. If there is no "." in the string, then **.SCR** (for **SC**o**R**e file) will be added to the end and used to name a DOS file. In both of these cases an error will be generated if the string is not a legal DOS name (such as a maximum of 8 characters.)

It is recommended that you create and maintain a consistent system for determining a unique reference string assigned to each person tested. One example of such a system is using secret codes while keeping track of the relation between these codes and the names of the applicants in a separate, confidential file (the "key").

The file annotation is optional and can be any message up to 30 characters long. It is not processed by MS-DOS. The information is only copied by the program into the final results file for future reference. If you do not wish to enter a file annotation, just press **<Enter>**.

Within a given scores subdirectory, if WOMBAT-FC detects another file with the same name as the one created from the reference string, you will be asked to decide whether to erase it or to go back and reenter a different reference string to avoid duplicating the file name. This check only guarantees that the current scores file will not overwrite any other scores file existing in the **current scores subdirectory**. If you decide to erase an old scores file at this point, the data collected and stored under that filename in the **composite spreadsheet data file** will remain, thus creating two records (or rows of numbers) with the same scores filename (see **Composite Spreadsheet Data File**). It is therefore good practice not to erase an old scores file and to select another reference string if prompted to do so.

It is highly recommended that you periodically archive (or backup) the score files. If you are currently in the WOMBATFC.V13 subdirectory, the simplest DOS command to archive scores files onto a floppy disk in drive A: is:

XCOPY RESULTS.DIR A:\RESULTS.DIR\ <Enter>

This command will send all the files contained in **RESULTS.DIR** to the floppy disk in drive A:. It will even create the directory RESULTS.DIR on the drive if it doesn't already exist. Of course, you can always revert to a MS Windows™ interface and use the mouse to copy whole directories onto floppy disks. To delete the specific scores file **1234.SCR** from the **RESULTS.DIR** directory, the DOS commands are:

CD RESULTS.DIR <Enter>
DEL 1234.SCR <Enter>

or just:

DEL RESULTS.DIR\1234.SCR <Enter>

Refer to MS-DOS manuals for more information on the use of the **COPY** and **DEL** commands as well as the use of wildcards * and **?** in filenames.

The WOMBAT.CFG File

Locate the file named **WOMBAT.CFG** in the **WOMBATFC.V13** subdirectory. This text file lists important parameters controlling a number of functions. The file delivered with Version 1.3 is shown in Figure 3.

WOMBAT.CFG is a user-modifiable text file. You can use your favorite word processing package (but be careful not to add any formatting codes to the text) or a simpler application such as the DOS EDITOR to open and modify the contents of a text file. For example, at the DOS prompt, type:

EDIT WOMBAT.CFG <Enter>

Figure 3 shows what you can expect if you open the **WOMBAT.CFG** file. Each parameter is self-explanatory. If you collect WOMBAT data on more than one WOMBAT station, there is a question about where you will maintain a single archive of all the scores files. **WOMBAT-FC's embedded safety features do not insure that duplicate names will not be on some other WOMBAT station**. Consequently copying all scores together from different stations could result in loss of some scores files with duplicate names. For this reason we suggest that you use a unique scores subdirectory name on each of your WOMBAT stations. Then copying each subdirectory to the single archive will keep the files from different stations from ever mingling.

```
 * See the manual for advice on changing the parameters in this file *
562  The code for a standard VGA display (see README for others)
60  is the whole number of minutes of instruction time (1 or more)
10  minutes for each scoring interval (1 to 90)
9  scoring intervals in the complete test (1 to 90)
2 composite data file order (1 = .SCR file order; 2 = grouped scores)
RESULTS.DIR is the name of the scores subdirectory
COMPOSIT.TXT is the name of the composite spreadsheet data file
RESEARCH.TXT is the name of the research data spreadsheet file
? language directory to be used
```

Figure 3. WOMBAT.CFG file as delivered .

If you decide to change any of the following:

- the interval duration,
- the number of intervals, or
- the composite spreadsheet data file order

through their respective parameters in the **WOMBAT.CFG** file, you should also change the default name of the composite and research spreadsheet data files and/or the name of the default scores subdirectory. Either of these actions will produce new composite and research spreadsheet data files with headers appropriate to the parameters. That way the new data will not be appended to composite and research spreadsheet data files that are only appropriate for the old parameters.

Time Required, Time Allowed

The **WOMBAT.CFG** shown in Figure 3 specifies that the WOMBAT-FC test is to run during 9 consecutive intervals of 10 minutes each for a total of 90 minutes. Moving from one interval to the next is totally transparent and goes unnoticed by the candidate. At the end of each interval, WOMBAT-FC records the scores in the scores file and the temporary composite spreadsheet data file, **WOMBATFC.SPD**.

At the end of the test the overall scores are also recorded in the scores file and the complete set of all scores is added to the end of the composite spreadsheet data file. If the composite spreadsheet data file does not exist at this time, it is created with a first record containing column headers that are appropriate to the kind of data being collected. After writing the composite spreadsheet data file, the temporary file **WOMBATFC.SPD** is erased.

Research organizations may wish to vary the number and duration of intervals. Modifications must be made carefully as they will affect the way scores are labeled inside the two spreadsheet data files. It is highly recommended that you change the scores subdirectory **parameter** each time you change one or more of the parameters mentioned above. This will insure that tests administered according to new parameters will have their scores saved into an independent subdirectory, and consequently in new spreadsheet data files, thereby protecting any previous data. Feel free to discuss this issue with an Aero Innovation representative before altering the number and duration of the intervals.

The following pages assume the default settings found in the **WOMBAT.CFG** file upon delivery, namely, the length and number of intervals and the names of the scores subdirectory and the spreadsheet data files: **RESULTS.DIR**, **COMPOSIT.TXT**, and **RESEARCH.TXT,** respectively. Also assume that the reference number 1234 has been assigned to a candidate and is supplied to WOMBAT-FC Version 1.3 at the startup screen.

Instruction Time Limitation

Candidates go through a succession of instruction pages and accompanying exercises, each being known as an **instructional phase**. These instructional phases constitute the complete **instructions period**. The default time limitation for the instructions period is 60 minutes. You can make the instructions period longer or shorter to suit your requirements by editing a new value into **WOMBAT.CFG**.

Present users of WOMBAT-FC tell us that the use of the **Candidate Manual** before the test shortens the time spent reading the on-screen instructions. We recommend that you distribute the Candidate Manual one week ahead of time to allow the testees to read it carefully.

Test Duration

By default, the test duration is 90 minutes divided in 9 intervals of 10 minutes each. This duration does not include the time spent in the instructions period. The interval duration parameter and the number of intervals parameter are both found in the **WOMBAT.CFG** file (Version 1.3 & up). We recommend that you DO NOT shorten the **total test duration** (determined by the number of intervals times the duration of each) until you have acquired sufficient data to assess the consequences of administering a shorter WOMBAT-FC test. Please contact an Aero Innovation representative to discuss the possibilities of shortening the test.

IMMEDIATE EXIT

Pressing and holding the Control key (CTRL) and then pressing the Break key on the computer keyboard will immediately terminate a WOMBAT-FC session at any time in the instructions period or during the test. The incomplete data collected to that point will be stored in the scores file and the two spreadsheet data files in the current scores subdirectory.

EXITING AND RESTARTING A WOMBAT-FC SESSION

To accommodate the possible need to break the instructions and testing into two separate periods, a mechanism for exiting and then restarting a session has been included (see Figure 4). If the **ESC** key is pressed **anytime before the testing begins**, the program will stop when the testee has spent all the time allowed or when "9" and "0" are pressed. The time spent in each of the instructions phases is then written to the scores file and the program is suspended.

The previously suspended session of candidate **1234** can be re-invoked by typing:

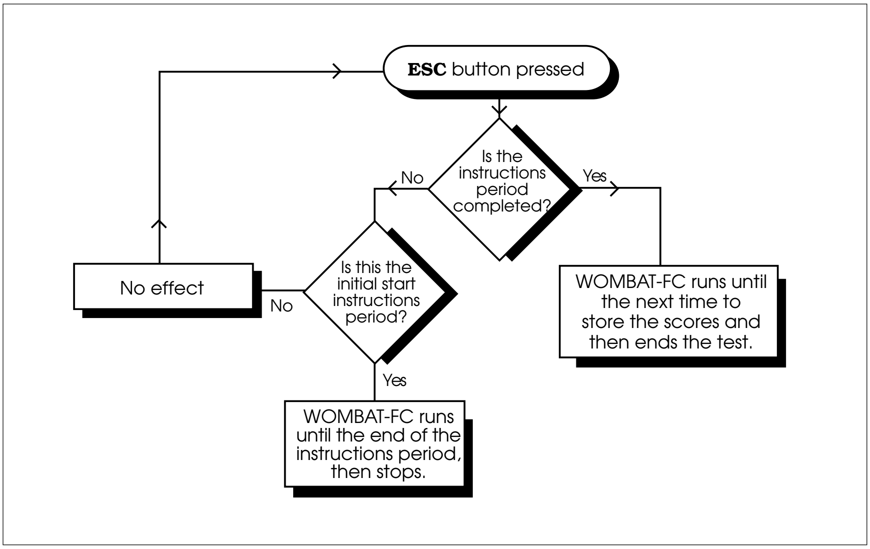


Figure 4. Use of the ESC Button.

WOMBAT 1234 <Enter>

The system will search for the file **1234.SCR** created during a previous session and stored in the current scores subdirectory. If the file is not found, the program stops with an error message. If the file is found, the program will append the new data to the end of the previous data in the file, no matter whether the previous session was a complete test or a shortened test. There is no limitation to the number of sessions that can be appended to a single scores file.

If the file **1234.SCR** is present in the current scores subdirectory, the instructions period is entered near its end and the user is given up to 10 minutes of additional practice. During this practice, the scheduled exit mechanism is disabled. That is, pressing **ESC** will have no effect. After the 10 minutes are up, the test begins and can be stopped by the second escape mechanism to be described below.

Shortening the WOMBAT-FC Testing Time

A second mechanism for scheduling an escape has been included in WOMBAT-FC. **During the test**, it is possible to terminate early by pressing the **ESC** key on the keyboard (see Figure 5). The exit is queued for the next point when scores would be stored on the disk up to the maximum time of the test (i.e. the end of the current interval). A message will appear at the bottom of the screen giving the exit time that has been queued.

Uncontrolled Program Exit

If a power failure, a hardware failure, or a program problem results in an unscheduled exit from the program, the normal exit process for saving the latest data to the scores file or two composite spreadsheet data files will not take place.

If there are any on-screen messages, they should be copied before taking any further actions. The directory of the WOMBAT-FC program should be examined at the earliest opportunity for the temporary file **WOMBATFC.SPD**. If it is found, it contains a record of the last scores saved before the unscheduled exit and should be copied somewhere (possibly with a new name) before the program is run again to collect WOMBAT data; otherwise it will be replaced with the new data.

If you save the temporary file left by an unscheduled exit, there are ways to incorporate the data into the composite spreadsheet data file, which an Aero Innovation representative can help you with if you want.

LANGUAGE ISSUES 8

The text that appears during the instructions and during the test has been stored in text files inside language subdirectories on the computer's hard disk. You can have as many language subdirectories as you wish (see Figure 5), each one containing all the files of one language.

THE BATMSG FILES

As seen in Chapter 7, the instructions period consists of instructional pages and accompanying exercises, both of which are called **phases**. Each instructions page comes from a separate file. There are as many text files as pages of text in the instructions period.

Each file is named **BATMSG.x**, where x is a number between 000 and 999. You <u>must not change the name</u> of any BATMSG file. Accompanying exercise phases have no text associated with them.

The BATMSG files are standard user-modifiable ASCII files. You can use your favorite word processing package, or a simpler application such as the DOS Editor to change the text in any of the files. For example, to edit the text found at Phase 3 of the instructions using the DOS Editor, at the DOS prompt type:

EDIT BATMSG.301 <Enter>

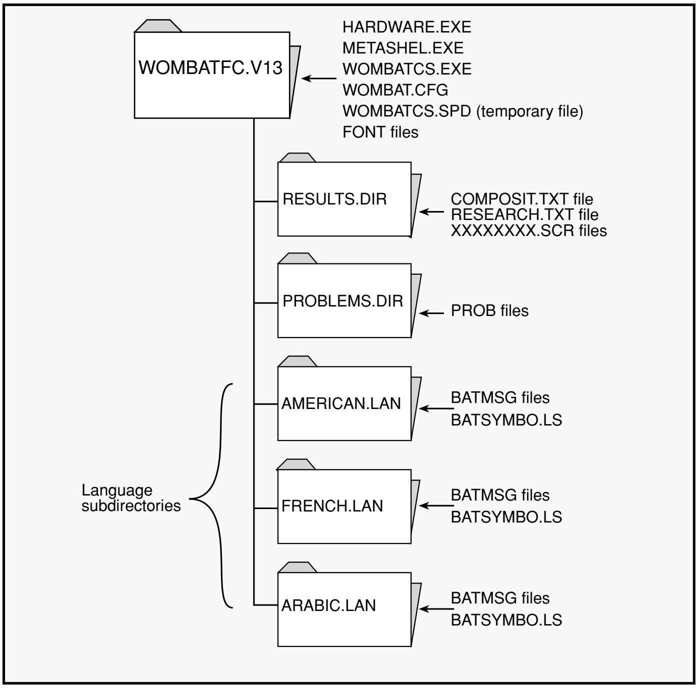


Figure 5. WOMBAT-FC File Structure as Delivered. This figure represents the subdirectory structure and the enclosed files as delivered. RESULTS.DIR, COMPOSIT.TXT, and RESEARCH.TXT are default names for the scores subdirectory and two spreadsheet data files, respectively. There are virtually no limitations on the number of language subdirectories you can create. To be recognized by the program, each language subdirectory must be located inside the WOMBATFC.V13 directory and have its name end with the suffix LAN.

With the **EDIT** command, you have access to the text of any BATMSG file and can make all the editorial changes you want, provided that you **don't add too many lines** of text to the file and that each individual line **ends with a carriage return** (Enter). Keep in mind the fact that WOMBAT-FC uses fixed width characters and uses both the text mode of the terminal and the graphics fonts displayed by METAWINDOW. If new lines are added to the file, some lines from the top of the page could go out of sight once invoked by WOMBAT-FC. In this case, just edit the same file again and reduce the number of lines of text. The limit on the number of lines varies depending on the contents of the screen.

If one or more lines of text in a BATMSG file is longer than 80 characters, the WOMBAT-FC will suffer from a fatal runtime error and crash. In this case, identify which BATMSG file is defective and insert a carriage return at or before the 80th character of that line. A few message screens must be limited to 76 or fewer characters or else there will be conflict between graphics elements of the display and the text.

Remember not to delete any BATMSG files in a given language subdirectory.

THE BATSYMBO.LS FILE

In each language subdirectory, there is a text file named **BATSYMBO.LS**, which contains a list of words or short sentences used by WOMBAT-FC to display commands, menus or comments in the program. Some of the words of BATSYMBO.LS will appear inside colored rectangles during the test, like the words TRIGGER, AUTOTRACK, or BONUS. See a printout of the American English version of BATSYMBO.LS in Appendix 3.

If you take a look at the BATSYMBO.LS file using the **EDIT** command, you will see that some lines begin with one or more key words, followed (either on the same line, or on the next indented line) by a comment to help you understand the meaning or context of the key words. Sometimes, a **Maximum Number of Characters** will be specified to indicate the maximum length of the words, so they fit inside the colored rectangles when displayed in the test. If the comment is the next line then the whole preceding line is read into a string and the program attempts to display the whole string. Sometimes the string is centered in the available area and other times it starts at the left edge of that area.

If you change or translate some or all of the key words contained in the BATSYMBO.LS file, make sure that you change ONLY the key words, nothing else. The number of lines in the file and their order must remain unchanged for WOMBAT-FC to run without problems.

TRANSLATING WOMBAT-FC

WOMBAT-FC is normally delivered in American English and Canadian French languages. WOMBAT-FC can also display several other languages if certain conditions are met. Languages based on the Roman, Cyrillic, or Arabic alphabets can easily be used in the **BATMSG** files. Other languages, such as the ones based on Japanese or Chinese characters may possibly be used with some restrictions. Contact Aero Innovation's Technical Support for more details on translating WOMBAT-FC.

On occasion, it may be more cost-effective to print the translated version of the instructions on paper. The candidate can keep the document next to the console and follow the WOMBAT-FC on-screen instructions step-by-step in his/her native language. If you want to translate the Candidate Manual, Encapsulated PostScript (EPS) vectorial graphics and Adobe PageMaker™ 6.5 templates are available free-of-charge through Aero Innovation to help you publish your own version of the WOMBAT-FC Candidate Manual.

To add a new language to your WOMBAT-FC software (German in the following example), proceed as follows:

- At the **C :\WOMBATFC.V13>** prompt, type:
  XCOPY AMERICAN.LAN GERMAN.LAN\ /V <Enter>
  Then type the following command:
  CD GERMAN.LAN <Enter>.
  This action will create a whole new set of instructions, located inside a new subdirectory labeled **GERMAN.LAN**.

- Translate one by one all the BATMSG files in German using the **EDIT** command. Just overwrite the BATMSG files, don't remove any BATMSG file.

- Translate desired words in the BATSYMBO.LS file, again using the **EDIT** command.

- Once a few files are translated, and once the **BATSYMBO.LS** file is translated and saved, launch **WOMBATFC**. You will see at the top of the monitor the new language menu showing the German addition. Choose it and start the program to check the translation done to that point. You can have **different versions of the same language** to suit different needs, clients, or populations. Just follow the procedure above and give each subdirectory a unique name with the LAN suffix (such names are restricted to 8 characters before the dot).

If you intend to edit the original distribution files in **AMERICAN.LAN**, we suggest that you make a copy of that subdirectory with some other name so you don't lose the originals. They are your reference to the ideas we have thought appropriate to cover in that instruction phase.

Intentionally blank

The WOMBAT-FC Scores

9

Reading the Scores File

You can access the desired scores file, either with a word processor program or, if you are in the current scores subdirectory, by typing:

EDIT 1234.SCR <Enter>

Each scores file is divided into two sections. The first part contains a table of instruction phases and associated elapsed times. WOMBAT-FC keeps track of the total time the candidate devoted to each phase of the instructions, and prints it next to the phase number. At the present time, there is little known benefit to an operator to record or use these times. Such data may be found useful in scientific studies in the future.

The second part is a table of incremental scores collected for each scoring interval during the test and the last interval if the test was terminated before a normal scoring interval. The last line has the summary scores for the whole test. Referring to Figure 6 on a later page, you will find:

- The **Interval** (ms) is the duration of the data collection period in milliseconds. The default interval is 10 minutes or 600,000 milliseconds. The default WOMBAT-FC test consists of 9 scoring intervals for a total of 90 minutes.

- The **Tracking Score** (TS) is calculated as the Tracking Performance multiplied by the Tracking Worth. The TS depends not only on performance but also on the frequent playing of bonus tasks to maintain a high Tracking Worth.

- The **Tracking Performance %** (TP) figure reflects the efficiency of the candidate at keeping track of the moving targets. 100(TS/TP) = **Perfect Tracking Score** (not shown on the scores sheet.)

- The **Collision Detection Score** (CD) reflects the efficiency of the candidate's collision predictions.

- The **Figure Rotation Score** (FRS) is the number of bonus points earned from the Figure-Rotation task (Bonus Performance multiplied by Bonus Worth.)

- The **Quadrant-Location Score** (QLS) is the number of bonus points earned from the Quadrant-Location task (Bonus Performance multiplied by Bonus Worth.)

- The **Sequences Mastered** (SM) is the number of Quadrant Location sequences that were mastered.

- The **Digit-Canceling Score** (DCS) is the number of bonus points earned from the Two-Back Digit-Canceling task (Bonus Performance multiplied by Bonus Worth.)

- The **Total Bonus Score** (TBS) is just the bonus component of the overall score and is the sum of the QLS, FRS, and DCS.

- The **Overall Score** (OS) is the sum of TS, CD, and TBS. It is the final Overall Score (204.0 points in Figure 6) that should be used as the criterion for selection. The other results are shown mainly for research purposes.

- The **Predicted Final Score** (PFS) is the scoring rate for the current interval extrapolated to the end of the test and added to the Overall Score to the present time.

```
Sir Winston Churchill

WOMBAT-FC Version 1.3
Initial Instruction Phase - Date (M/D/Y): 1/16/1944
Using filename: RESULTS.DIR\1234.SCR
And composite spreadsheet data file: RESULTS.DIR\COMPOSIT.TXT
And research-data spreadsheet file: RESULTS.DIR\RESEARCH.TXT


Instruction phase and elapsed time in milliseconds

 1   88523        2   56107        3     46862        4   30222        5   18240
 6   29688        7  154960        8     60591        9   24307       10   31192
11   25652       12   27533       13     11465       14   15549       15   23099
16   30613       17   11317       18     30677       19   11096       20   30409
21    3439       22   30971       23    108772       24   90885       25    4179
26   30521       27    8022       28     45127       29    1592       30   45169
31   46157       32   31530       33     22902       34    8986       35   19164
36   69622       37   48042       38     63922       39   17079       40   61731
41   65268       42   83833       43     41705       44   88063       45  120579
46   54651       47   35666       48     18841       49   40462       50   11011
51  194893       52       0       53         0       54       0
2+...+50  33 minutes and 7 seconds


Test Phase

1. Interval (ms)              2.Tracking Score          3.Tracking Performance (%)
4. Collision-Detection Score  5.Figure-Rotation Score
6. Quadrant-Location Score    7.Sequences Mastered      8.Digit-Canceling Score
9. Total Bonus Score         10.Overall Score          11.Predicted Final Score

        1      2      3      4      5       6      7      8       9      10       11

   600027    6.6   77.5   0.6    0.7     1.4     0     1.7     3.8    11.0     99.1
   599987   10.0   95.2   0.6    0.0     1.3     0     3.8     5.1    15.7    136.9
   600027   12.0   86.2   1.3    0.8     1.3     0     0.0     2.1    15.4    134.3
   599961   13.9   91.7   7.0    0.0     1.9     0     1.7     3.7    24.6    189.7
   599999   17.0   92.4   3.2    0.0     2.9     0     1.9     4.9    25.1    192.0
   600026   18.1   97.6   5.1    2.1     0.0     0     3.9     5.9    29.1    208.2
   599981   16.0   80.9   4.5    1.0     2.6     0     0.0     3.6    24.1    193.1
   599994   18.0   94.0   5.8    1.2     3.3     0     2.1     6.7    30.7    206.3
   600101   20.8   94.6   3.2    0.3     2.0     0     2.1     4.4    28.3    204.0
  5400103  132.2   90.7  31.3    6.0    16.8     0    17.3    40.2   204.0
```

Figure 6. WOMBAT-FC Scores Sheet. Actual scores sheet at the completion of a WOMBAT-FC test. To view a scores sheet, simply use a word processor or text editor program. This scores sheet is named 1234.SCR and is found in the RESULTS.DIR subdirectory on the computer's hard drive in the WOMBATFC.V13 subdirectory.

Spreadsheet Data Files

Whenever the number of tests justifies it, we recommend the use of a commercial spreadsheet or database program to help you analyze the scores of your candidates. Such commercial packages include Excel™, Lotus 1-2-3™ or DBase™ to name just a few. Not only will it be easier to compare one candidate's performance against the group, but you will be able to view the candidate's progression in the test by plotting the score intervals onto a graph, as shown below.

Inside the scores subdirectory (default **RESULTS.DIR**), you will find the **composite spreadsheet data file** (default **COMPOSIT.TXT**) and the **research spreadsheet data file** (default **RESEARCH.TXT**).

The **composite spreadsheet data file** contains the scores from each test conducted from a WOMBAT-FC subdirectory provided WOMBAT.CFG called for adding spreadsheet data to a file with this name. Each time a test ends, a line of data is appended to the composite spreadsheet data file. If you rename or delete this file from the scores subdirectory, WOMBAT-FC will create a new file automatically using whatever name is specified in WOMBAT.CFG (this may be required if you modify some parameters in the WOMBAT.CFG file). While you are archiving all the files containing results, consider renaming the composite spreadsheet data file after a relocation to avoid overwriting it with another file of the same name later.

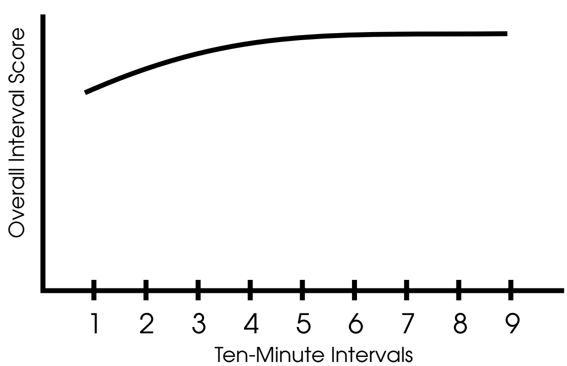


Figure 7. Typical WOMBAT-FC Group Learning Curve.

The **research-data spreadsheet file** is similar in format to the composite file but it contains data to be used by psychometrists during scientific studies. No data contained in this spreadsheet file should be used in isolation for selection purposes. The test administrators usually share the data contained in this research file with Aero Innovation and with other psychometrists for the purpose of improving the test's scoring algorithms.

The data contained in both spreadsheet data files are saved in columns (vertical) delimited by tabs and rows (horizontal) delimited by carriage returns. To open the file, first launch your spreadsheet application, then select OPEN from the FILE menu and locate the desired composite spreadsheet data file. Once the loading is completed, you will observe a number of long rows of data. The first row on top contains all the column labels, and each subsequent row represents one WOMBAT-FC test. The data numbers are exactly the same as in the .SCR files; they were just formatted so as to speed up spreadsheet loading operations and avoid error-inducing retyping.

Scores-File Order vs Grouped-Scores Order

There are two different orders for the data in a row as controlled by the spreadsheet data files order **parameter** in **WOMBAT.CFG**. In the *scores-file order,* the data shown in Figure 6 are written into the spreadsheet from left to right on each line and from the top to the bottom through all the lines.

In the *grouped-scores order,* the instruction phase times are written in the farthest right columns of the spreadsheet and the test data are written from top to bottom of each column of the table and from left to right through the table columns. This results in interval times being in adjacent spreadsheet columns followed by all tracking scores and so on.

How Predictive Is WOMBAT-FC in My Operation?

To answer this question, you will have to conduct a comparison between the WOMBAT-FC scores and ratings based on a recognized criterion such as a standardized performance test administered at a specified level of training or operational experience. The criterion performance requirements for the test must be clearly understood

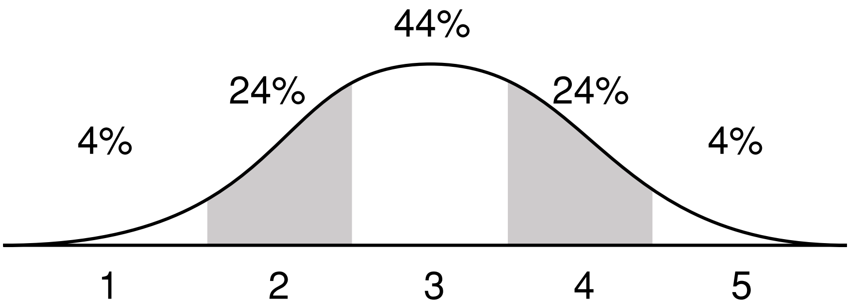


Figure 8. Normal Probability Distribution. This curve represents the normal probability distribution of human performance, and the five scoring categories represent equal intervals based on the variability of scores.

by the candidates and understood and agreed to by the evaluators who will rate the candidates.

To assure maximum objectivity, at least two and preferably three isolated evaluators should make and record simultaneous, independent assessments of the same test performance by each candidate. Once both the performance ratings and the WOMBAT-FC scores are completed, a statistical correlation or regression analysis will show how well WOMBAT-FC predicts the criterion.

For a meaningful validation study, test 40, 50, or more candidates at some specified level of training or operational experience. Have each evaluator independently select the top and bottom 4% of the total number tested and assign them ratings of 5 (best) and 1 (worst), respectively. Then have them assign ratings of 4 (good) to the next best 24% of the group and ratings of 2 (poor) to the next worse 24% of the group. The middle 44% will receive a rating of 3 (average). Then take the means of the ratings of each candidate by the independent evaluators.

Admittedly a 5-point rating scale does not provide as fine a discrimination as one would desire, but evaluators will find it difficult enough to pick the best and worst 4% and the next best and worst 24% according to the distribution shown in Figure 8.

Once the WOMBAT-FC testing and the difficult task of rating candidates are completed, keep the results confidential until the statistical comparison is finished. If you desire you can release all the figures later, normally with the exception of the identities of the candidates.

To do the statistical tests, refer to your spreadsheet Functions manual and calculate the Pearson coefficient of correlation between the two sets of scores, using the WOMBAT-FC scores on one side and the 5, 4, 3, 2, 1 ratings on the other side. Feel free to contact Aero Innovation for support in your data analysis. Make sure that your criterion-based rating procedures closely followed the ones outlined above before you draw any conclusions about the predictive value of WOMBAT-FC, or of any other selection tests you may want to include.

What If I Test a Candidate More Than Once?

According to the present WOMBAT-FC users, it appears that once a candidate has reached his/her asymptotic performance, typically at mid-time during the test, there is little that a testee can do to improve his or her rate of scoring (Figure 9). Data typically show a constant, flat learning curve after the first 30 to 40 minutes of the first test.

A study at George Mason University near Washington, D.C. (Bruce 1994) using WOMBAT-CS, showed, however, a small increase of performance on the order of 6% of the overall score per

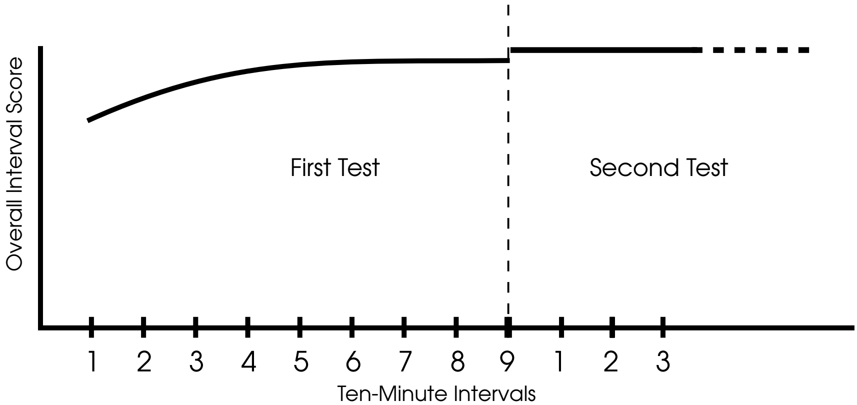


Figure 9. Typical Two-Tests Group Learning Curve. A typical WOMBAT-FC group learning curve during a first administration of the test followed by the first 30 minutes of a retest a day or two later.

interval between the first and second administration. The amount of the performance increase, known as a "reminiscence effect," is generally attributed to the testee's thinking about strategy and mental rehearsal during the time between the two tests. Despite the reminiscence effect, Bruce demonstrated a correlation of 0.88 between the first test's ranking and the second test's ranking. This indicates that if a whole group of candidates were tested twice, they would keep nearly the same relative ranking at the outcome of the second test.

In your own organization, if you are particularly concerned about testees' requests for additional testing sessions, or if you believe some testees have received pretest training using a WOMBAT-FC system other than yours, just have everybody retested for 30 minutes a day or two later, and use those scores to rank the testees. This process can be aided by using two different WOMBAT.CFG files (possibly in two different subdirectories), the first governing a 90-minute test and the second specifying a 30-minute partial retest.

When testing everybody twice, you will short-circuit any attempt by testees to trick the system by training themselves on WOMBAT-FC prior to the screening process. If pretraining is not a potential factor in your organization, you can probably limit the administration frequency to once per year, thereby virtually eliminating any past-exposure advantage.

FREQUENTLY ASKED QUESTIONS

Appendix 1

1. HAS WOMBAT-FC BEEN VALIDATED?

Version 1.3 of the Flow Control (FC) WOMBAT was first released in July 1999, so no large-scale validation study of the new WOMBAT-FC has been published to date, However, "in-house" validations are being conducted by traffic controller training agencies given early access to the software. To avoid possible litigation and for other proprietary reasons, agencies do not normally publish their in-house study results. but those who have used earlier versions freely discuss their unprecedented success in selection and are its strongest boosters.

The original WOMBAT test was based on the results of a series of experiments done at the University of Illinois in the 1970s, and the WOMBAT-FC is a modification and extension of the earlier version based on a successful experiment at New Mexico State University in the 1980s. The studies at Illinois dealt with the ability to attend to multiple sources of information, prioritize courses of action, and allocate attention effectively while under operational stress. The New Mexico experiment measured the specific ability to "keep the picture" of the flow of multiple targets on display screens, as required in controlling traffic.

Few validation studies of any flow-controller selection test have been published and none against operational performance criteria. Administrators typically determine the predictive validity of any selection test and establish their own norms against criteria appropriate to their respective populations, organizational cultures, and specific operational environments. The WOMBAT users have always been strongly encouraged to conduct in-house validation and standardization studies on a continuing basis.

2. Is WOMBAT-FC a Psychomotor Test?

WOMBAT-FC has a minor psychomotor component involving the positioning of symbols with joysticks, but it does not require the degree of psychomotor ability called for by the WOMBAT-CS test. In the primary target-designation task, only simple, single-axis motion by the left hand and slow, imprecise, two-axis motion by the right hand are required. The Figure-Rotation Task calls for imprecise rotation of three-dimensional figures that is intuitive and easy.

WOMBAT-FC does put a premium on timely response but even more on good judgment of what's important in the long run, vigilance, attention allocation, and keeping an accurate mental picture of a dynamic situation.

3. What Is the Minimum Acceptable Score?

There is no magical "Pass/Fail" criterion score with any WOMBAT test. Administrators of the WOMBAT tests set different minimum scores below which candidates are rejected depending on their respective operational situations—the availability of candidates, the number of positions to be filled, and the levels of situational awareness deemed necessary for specific operations.

Different versions of the WOMBAT-FC software may also require minimum score adjustments. Aero Innovation will help test administrators set minimum scores during initial implementation and periodic normalization exercises.

4. How Long Is the Test?

The nominal test duration is 90 minutes, but this may be reduced when enough candidates have been tested to determine at what point the predicted final scores correlate almost perfectly with the actual final scores. The nominal instruction period is 60 minutes, but this can also be shortened with advance distribution of the Candidate Manual. Based on the default numbers above, up to two and a half hours may be required for a complete WOMBAT-FC session.

After verifying the candidate's identification, the administrator gives the candidate a short briefing, and the usual questions are answered. This takes two minutes, at most. The candidate is then left alone in a quiet room with the door closed. No further intervention by the administrator is required. When the candidate finishes the test, the administrator archives the results before beginning a new session.

5. What Is a Typical Pretest Briefing?

This is an example of the verbal briefing given to each candidate. First have the candidate sit in front of the WOMBAT console, ready to begin. The WOMBAT displays the first page of instructions. The verbal briefing should be short and clear, and contain at least the following. The test administrator should question the candidate to make sure everything is understood.

"*The instructions time is limited to XX minutes after you leave the present page. You will have plenty of time to read and practice everything. The time remaining before the test begins is displayed at the bottom of every page. As long as you have some time left, you can go back and review some exercises as you wish. Practice different ways, make mistakes, try different things. You can't make anything go wrong and nothing gets recorded during the instructions.*

If you want to break, ask me a question or use the washrooms, you may do so provided you are back before the test begins. The test will begin at the end of the instruction time, whether you are ready or not. You may begin the test anytime you are ready. The last page of instructions will tell you how to start the test.

WOMBAT will advise you when the test is complete. Then just leave this room with the door open so I can see you're finished. Once the test has started, give your absolute best effort until the end.

Do you have any question?"

6. Do Different Cultures React Differently with WOMBAT?

Thousands of WOMBAT-CS scores from five continents in several languages have failed to show significant differences in mean scores for various races, geographic areas, educational or vocational levels, or training in operating specific systems such as airplanes or computers, which tends to support the claim that WOMBAT tests are culture-free.

7. How Much Training Does the Administrator Need?

WOMBAT is extremely easy to administer. Most of the WOMBAT users have not received any specific training and have simply followed the operating instructions contained in this book. When required, training is always free of charge and supplied directly from Aero Innovation. Test interpretation is equally simple and can be coached by telephone or e-mail using the first few scores gathered after the initial installation.

Most of the test administrators share their test results with Aero Innovation in confidence after having removed candidate identifications. This helps Aero Innovation develop better operating manuals and instructions.

8. Can We Change the Text Displayed on Screen?

Yes. The text printed on the screen during the instructions and throughout the test phase is contained in a series of text files easily accessible though any word processor or text editor. The files are unprotected. Administrators are invited to modify, adapt, or translate the text to suit their operational and language requirements. Refer to Chapter 8 for step-by-step instructions on editing these files.

9. Are There Any Software Updates?

Based on a continuing collection of data provided in confidence by most of the test administrators, Aero Innovation occasionally updates the WOMBAT software. Free updates are guaranteed in writing to all users for a period of two years following acquisition.

10. Are There Any Royalties or User Fees to Pay?

All WOMBAT tests are free of user fees or royalties. Administrators are invited to run as many sessions as desired, since there are no restrictions. This also extends to commercial organizations involved with reselling or leasing testing services based on their WOMBAT systems.

Interface Installation on Another Computer

Appendix 2

This section of the manual is provided in case you have to install the WOMBAT-FC system in another computer.

System Requirements

The minimum configuration for WOMBAT-FC is a PC-compatible system that includes a 386 processor with a clock speed of 25 MHz or faster, a math coprocessor unit (387) and 1 megabyte of RAM. It is not recommended that you install WOMBAT-FC on a system with a clock speed of less than 386/25 MHz, if you want to compare scores between two or more WOMBAT-FC systems. The graphics system must be color VGA. Nowadays, WOMBAT-FC is delivered on Pentium 100 MHz PC-compatibles, but this level of performance is far from being a requirement to run WOMBAT-FC.

Interface Installation

The WOMBAT interface board occupies 5 locations in the Input/Output space from 300 (Hex) to 304 (Hex). Other boards in your system that conflict with this (such as modem, scanner, or sound cards), will need to be removed. If you do have to remove one or more cards from the PC Computer, make sure you also disable their drivers in the **CONFIG.SYS** or **AUTOEXEC.BAT** files.

The base address switches on the WOMBAT board have been set for these values, but you should check them before proceeding to a new installation. The correct dip settings are shown in Figure 10.

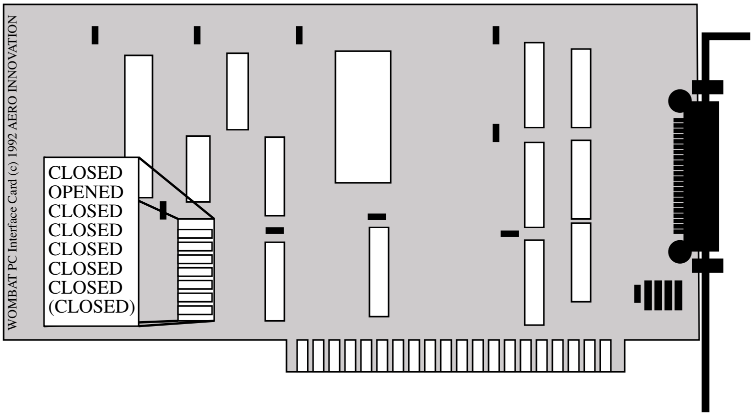


Figure 10. WOMBAT Interface Board. Base address switches on the WOMBAT board have been set for the values shown. You should check them before installation.

To install the WOMBAT interface board, proceed as follows:

1. Turn off the power to the computer and remove the cover.
2. Remove one of the rear cover plates and insert the board firmly into the connector on the mother board.
3. Refit the cover plate screw to hold the board in place.
4. Fit the cable into the socket on the interface board. Push it home firmly taking care not to bend any of the pins. The plug is polarized so that it can have only one position. Connect the other end to the socket on the WOMBAT console.
5. Arrange the control box on a table in a comfortable position and switch on the computer.

The BATSYMBO.LS File

Appendix 3

The following three pages contain the American English text normally found in the BATSYMBO.LS file from Version 1.3 of the WOMBAT-FC software disk. The characters, words, or lines in **bold** may be translated into another language. The text in light characters contains comments not displayed by WOMBAT. The line numbers on the left hand side of the pages are printed here for reference only.

1 **AUTOTRACK** the AutoPilotLabel string
2 **CURSOR** the ManualLabel string
3 **MISSING** the MissingLabel string
4 **COLLISION** the CollisionLabel
5 **Trigger** the TriggerLabel string
6 **P** the PerformanceCh character
7 **W** the WorthCh character
8 **Bonus Pouch Options**
9 the BonusPouchOptionsSt string
10 **Sixty**
11 the SixtySt string
12 **Seconds**
13 the SecondsSt string
14 **Figure**
15 the SolidFigureSt string
16 **Rotation**
17 the RotationTaskSt string
18 **Quadrant**
19 the QuadrantLocationSt string
20 **Location**
21 the TaskSt string
22 **Digit**
23 the TwoBackDigitSt string
24 **Canceling**
25 the CancelingTaskSt string
26 **End of test scheduled at**
27 the ScheduledEndOfTestSt string followed by a time in minutes
28 **minutes**
29 the MinutesSt string for ending ScheduledEndOfTestSt + time
30 **Press the ENTER key when the two control sticks are at spring centers.**
31 the CenterSticksSt string

32 **The offset values of the sticks at these positions are being measured.**

33 the MeasuringOffsetSt string

34 **Channel**

35 the ChannelSt string to be followed by a number 0 to 3 and BeingMeasuredSt

36 **is being measured.**

37 the BeingMeasuredSt string

38 **The scheduled**

39 the ScheduledSt to be followed by game length in minutes

40 **minutes for the full test have expired**

41 the complete test end message FullTestSt

42 **User break at time =**

43 the user break string UserBreakSt to be followed by test length in seconds

44 **seconds since test start**

45 the seconds since test start string ElapsedSecondsSt to follow UserBreakSt

46 **Restarted**

47 the restarted label RestartedSt for the score file followed by InstPhaseSt

48 **Initial**

49 the initial start label InitialSt for the score file followed by InstPhaseSt

50 **Instruction Phase - Date (M/D/Y):**

51 the instruction phase and date string InstPhaseSt followed by date

52 **Using filename:**

53 the scores file name string UsingFileSt following InstPhaseSt and date followed by the scores file name

54 **And composite spreadsheet data file:**

55 the spreadsheet data file string CompDataSt followed by composite file name

56 **And research-data spreadsheet file:**

57 the research data file string ResDataSt followed by research-data file name

58 **Instruction phase and elapsed time in milliseconds**

59 the instruction phase and time table header string PhaseTimeSt

60 **Test Phase**

61 the test scores header string TestPhaseSt

62 **1. Interval (ms) 2. Tracking Score 3. Tracking Performance (%)**

63 the first test scores title string FirstTitleSt

64 **4. Collision-Detection Score 5. Figure-Rotation Score**

65 the second test scores title string SecondTitleSt

66 **6. Quadrant-Location Score 7. Sequences Mastered 8. Digit-Canceling Score**

67 the third test scores title string ThirdTitleSt

68 **9. Total Bonus Score 10. Overall Score 11. Predicted Score**

69 the fourth test scores title string FourthTitleSt

70 **Only one /t or /T is needed or allowed on the command line.**

71 the /T command-line-option error string SlashTErrorSt

72 **Only one restart file can be specified on the command line.**

73 the multiple-restart-file error string RestartErrorSt
74 **The file**
75 the start of the restart file not found message TheFileSt followed by filename
76 **was not found for a restart.**
77 the end of the restart-file not found message NoRestartSt following filename
78 **Same figures**
79 the key 2 legend for the solid-figure-rotation task Key2SolidSt string
80 **Mirror-image figures**
81 the key 5 legend for the solid-figure-rotation task Key5SolidSt string
82 **Other differences**
83 the key 7 legend for the solid-figure-rotation task Key7SolidSt string
84 **Excellent! Your**
85 the ConfidentCorrectStartSt
86 **answer was correct.**
87 the ConfidentCorrectEndSt
88 **Oops! Your**
89 the ConfidentWrongStartSt
90 **answer was wrong.**
91 the ConfidentWrongEndSt
92 **Starts at**
93 theStartsAtSt
94 **Heading**
95 the HeadingSt
96 **Off grid**
97 the OffGridSt
98 **Has collided**
99 the HasCollidedSt
100 **ESWN?**
101 5 characters for directions East, South, West, North, and off grid
102 **New Problem**
103 the NewProblemSt
104 **Map for**
105 the MapForSt
106 **Instruction Phase**
107 the InstructPhaseSt
108 **out of**
109 the OutOfSt
110 **Remaining Time -**
111 the RemainingTimeSt

Intentionally blank

INSTRUCTION PHASES VS BATMSG FILES

Appendix 4

When a testee is going through the instruction phases there will be a display at the bottom of each screen that gives that phase a number out of the total number available for that session of the instructions. For an initial start there will be 55 instruction phases available numbered 1 to 55. For a restart there will be 54 instruction phases available numbered 2 to 54 and 56. The following table gives the numbers seen by the testee for an initial start in the first column, the numbers seen by the testee for a restart in the second column, the numbers used by the computer to put data into the instruction phase times table, and the name of the message file used to display test for that phase (if not a demonstration phase) in the fourth column.

INSTRUCTION PHASES AS SEEN BY SUBJECT

Initial Start 55 Files	Restart 54 Files	Assigned Phase Numbers	Corresponding Text Files
1		1	BATMSG.101
2	1	2	BATMSG.300
3	2	3	BATMSG.301
4	3	4	BATMSG.302
5	4	5	demo
6	5	6	BATMSG.303
7	6	7	demo
8	7	8	BATMSG.304
9	8	9	demo
10	9	10	BATMSG.305
11	10	11	demo
12	11	12	BATMSG.306
13	12	13	demo
14	13	14	BATMSG.307
15	14	15	demo
16	15	16	BATMSG.308

17	16	17	demo
18	17	18	BATMSG.309
19	18	19	demo
20	19	20	BATMSG.310
21	20	21	demo
22	21	22	BATMSG.311
23	22	23	demo
24	23	24	BATMSG.312
25	24	25	demo
26	25	26	BATMSG.313
27	26	27	BATMSG.314
28	27	28	demo
29	28	29	BATMSG.315
30	29	30	BATMSG.316
31	30	31	demo
32	31	32	BATMSG.317
33	32	33	BATMSG.318
34	33	34	demo
35	34	35	BATMSG.102
36	35	36	BATMSG.103
37	36	37	BATMSG.104
38	37	38	BATMSG.800
39	38	39	BATMSG.801
40	39	40	demo
41	40	41	BATMSG.802
42	41	42	demo
43	42	43	BATMSG.803
44	43	44	demo
45	44	45	BATMSG.850
46	45	46	demo
47	46	47	BATMSG.851
48	47	48	BATMSG.900
49	48	49	demo
50	49	50	BATMSG.901
51	50	51	BATMSG.105
52	51	52	BATMSG.106
53	52	53	BATMSG.107
54	53	54	BATMSG.108
55		55	BATMSG.109
	54	56	BATMSG.110
56	55	57	BATMSG.111

REFERENCES

Adams, J. A. (1995). Personal communication.

Alluisi, E. A. (1994). Roots and rooters. In H. L. Taylor (Ed.), *Division 21 members who made distinguished contributions to engineering psychology*. Washington, DC: Division 21 of the American Psychological Association.

Bartlett, F. C. (1943). *Instrument controls and display—Efficient human manipulation* (Report No. 565). London: UK Medical Research Council, Flying Personnel Research Committee.

Birmingham, H. P., & Taylor, F. V. (1954). *A human engineering approach to the design of man-operated continuous control systems* (Report NRL 4333). Washington, DC: Naval Research Laboratory, Engineering Psychology Branch.

Bond, N. A., Jr. (1970). Some persistent myths about military electronics maintenance. *Human Factors*, *12*, 241-252.

Bond, N. A., Jr. (1995). Personal communication.

Bray, C. W. (1948). *Psychology and military proficiency. A history of the Applied Psychology Panel of the National Defense Research Committee*. Princeton, NJ: Princeton University Press.

Bruce, D. (1994). Personal communication.

Bryan, G. L., Bond, N. A., Jr., LaPorte, H. R., Jr., & Hoffman, L. S. (1956). *Electronics troubleshooting: A behavioral analysis*. Los Angeles: University of Southern California, Electronics Personnel Research Group.

Bryan, G. L., Rigney, J. W., Bond, N. A., Jr., LaPorte, H. R., Jr., Hoffman, L. S., & McAllister, B. N. (1959). *The role of humans* *in complex computer maintenance.* Los Angeles: University of Southern California, Electronics Personnel Research Group.

Carel, W. L. (1960). *Human factors in ANIP* [Contract Nonr 1076(00)] Ithaca, NY: General Electric, Advanced Electronics Center.

Carel, W. L. (1961). *Visual factors in the contact analog* [Report R61ElC60, Contract Nonr 1076(00)]. Ithaca, NY: General Electric, Advanced Electronics Center.

Carhart, R. R. (1953). *A survey of the current status of the electronic reliability problem* (RM-1131-PR). Santa Monica, CA: Rand Corporation.

Chapanis, A., Garner, W. R., & Morgan, C. T. (1949). *Applied experimental psychology.* New York: Wiley.

Clinton, (USA) President Bill. (1995). Presentation of the National Medal of Science to Roger Shepard. Washington, DC: Office of the President.

Craik, K. J. W. (1940). *The fatigue apparatus (Cambridge cockpit)* (Report 119). London: British Air Ministry, Flying Personnel Research Committee.

Craik, K. J. W. (1944). *The psychological and physiological aspects of control mechanisms with special reference to tank gunnery* (Unpublished Report). Cambridge, UK: Cambridge University, Applied Psychology Unit.

Craik, K. J. W., & Vince, M. A. (1943). *Psychological and physiological aspects of control mechanisms with special reference to tank gunnery. Part I.* London: UK Medical Research Council, Military Personnel Research Committee.

Craik, K. J. W., & Vince, M. A. (1944). *Psychological and physiological aspects of control mechanisms. Part II.* (Report BPC 44/322). London: UK Medical Research Council, Military Personnel Research Committee.

Damos, D. L. (1972). *Cross-adaptive measurement of residual attention to predict pilot performance* (TR ARL-72-25/AFOSR-72-12). Savoy: University of Illinois at Urbana-Champaign, Aviation Research Laboratory.

Fitts, P. M. (1947). *Psychological research on equipment design* (Research Report 19). Washington, DC: US Army Air Forces Aviation Psychology Program.

Fitts, P. M. (1951). Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1287-1340). New York: Wiley.

Fitts, P. M. (Ed.), Chapanis, A., Frick, F., Garner, W. R., Gebhard, J., Grether, W. F., Henneman, R., Kappauf, W., Newman, E., & Williams, A. C., Jr. (1951). *Human engineering for an effective air-navigation and traffic-control system.* Washington, DC: National Research Council Committee on Aviation Psychology.

Fitts, P. M., Flood, M. M., Garman, R. A., & Williams, A. C., Jr. (1957). *The USAF human factor engineering mission as related to the qualitative superiority of future man-machine weapon systems*. Washington, DC: US Air Force Scientific Advisory Board, Working Group on Human Factor Engineering Social Science Panel.

Flanagan, J. C. (Ed.). (1947). *The aviation psychology program in the Army Air Force* (Research Report 1). Washington, DC: US Army Air Forces Aviation Psychology Program.

Flexman, R. E., Roscoe, S. N., Williams, A. C., Jr., & Williges, B. H. (1972). Studies in pilot training: The anatomy of transfer. *Aviation Research Monograph, 2*(1).

Gawron, V. J. (1999). Situational Awareness: Definitions, measures, technologies, and lessons learned (Handout text for workshop, 10th International Symposium on Aviation Psychology). Columbus, OH: The Ohio State University, Department of Aviation.

Grings, W. W., Rigney, J. W., Bond, N. A., Jr., & Summers, S. A. (1953). *A methodological study of electronics troubleshooting skill: II. Inter-comparisons of the MASTS test, a job sample test, and ten reference tests administered to fleet electronics technicians*. Los Angeles: University of Southern California, Electronics Personnel Research Group.

Helson, H., & Howe, W. H. (1943). *Inertia, friction and diameter in handwheel tracking* (OSRD Report 3454). Foxboro, MA: The Foxboro Company.

Hick, W. E. (1945). *Friction in manual controls* (Report 18). Cambridge: UK Medical Research Council, Applied Psychology Unit.

Howell, W. C. (1994). George Edward Briggs, 1926-1974. In H. L. Taylor (Ed.), *Division 21 members who made distinguished contributions to engineering psychology*. Washington, DC: Division 21 of the American Psychological Association.

Jacobs, R. S. (1976). *Simulator cockpit motion and the transfer of initial flight training* (TR ARL-76-8/AFOSR-76-4). Savoy: University of Illinois at Urbana-Champaign, Aviation Research Laboratory. (Ph.D. dissertation, University of Illinois at Urbana-Champaign).

Johnson, B. E., Williams, A. C., Jr., & Roscoe, S. N. (1951). *A simulator for studying human factors in air traffic control systems* (Report 11). Washington, DC: National Research Council Committee on Aviation Psychology.

Kay, H. (1953). *Experimental studies in adult learning.* Ph.D. dissertation, Cambridge University, UK.

Kelley, C.R. (1968). *Manual and automatic control.* New York: Wiley.

Koonce, J. M. (1984). A brief history of aviation psychology. *Human Factors*, *26*, 499-508.

Kraft, C. L., & Fitts, P. M. (1954). *A broad band blue lighting system for radar air traffic control centers* (Technical Report TR 53-416). Wright-Patterson Air Force Base, OH: Wright Air Development Center.

Licklider, J. C. R., & Kryter, K. D. (1944). *Articulation tests of standard and modified interphones conducted during flight at 5,000 and 35,000 feet* (OSRD Report 1976). Cambridge, MA: Harvard University, Psycho-Acoustics Laboratory.

Licklider, J. C. R., & Miller, G. A. (1951). The perception of speech. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1040-1074). New York: Wiley.

Licklider, J. C. R. (Chair), Clementson, G. C., Doughty, J. M., Huggins, W. H., Seeger, C. M., Smith, C. C., Williams, A. C., Jr., & Wray, J. (1953). *Human factors in the operation and maintenance of all-weather interceptor systems: Conclusions and recommendations of Project Jay Ray, a study group on human factors in all-weather interception* (HFORL Memorandum 41). Bolling Air Force Base, DC: Human Factors Operations Research Laboratories.

Lintern, G. (1994). Personal communication.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

North, R. A. (1977). Task components and demands as factors in dual-task-performance (TR ARL-77-2/AFOSR-77-2). Savoy: University of Illinois, Aviation Research Laboratory. (Ph.D. dissertation, University of Illinois at Urbana-Champaign).

North, R. A., & Gopher, D. (1974). Basic attention measures as predictors of success in flight training. *Proceedings of the Human Factors Society 18th Annual Meeting.* Santa Monica, CA: Human Factors Society.

O'Hare, D. (1996). Personal communication.

O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. In *Proceedings of the Ninth International Symposium on Aviation Psychology*, Columbus: The Ohio State University, Department of Aviation.

O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. *Human Factors, 39*, 540-552.

Orlansky, J. (1994). Jack W. Dunlap, 1902-1977. In H. L. Taylor (Ed.), *Division 21 members who made distinguished contributions to engineering psychology*. Washington, DC: Division 21 of the American Psychological Association.

Parsons, H. M. (1972). *Man-machine systems experiments*. Baltimore & London: The Johns Hopkins Press.

Pew, R. W. (1994). Paul Morris Fitts, 1912-1965. In H. L. Taylor (Ed.), *Division 21 members who made distinguished contributions to engineering psychology*. Washington, DC: Division 21 of the American Psychological Association.

Poulton, E. C. (1974). *Tracking skill and manual control*. New York & London: Academic Press.

Ritchie, M. L. (1994). Ross A. McFarland, 1901-1976, In H. L. Taylor (Ed.), *Division 21 members who made distinguished contributions to engineering psychology*. Washington, DC: Division 21 of the American Psychological Association.

Roscoe, S. N. (1980). *Aviation psychology*. Ames: The Iowa State University Press.

Roscoe, S. N. (1992). From the roots to the branches of cockpit design: Problems, principles, products. *Human Factors Society Bulletin*, *35*(12), 1-2.

Roscoe, S. N. (1993). An aid in the selection process—WOMBAT. *Civil Aviation Training, 4*(2), 48-51.

Roscoe, S. N. (1994). Alexander Coxe Williams, Jr., 1914-1962. In H. L. Taylor (Ed.), *Division 21 members who made distinguished contributions to engineering psychology*. Washington, DC: Division 21 of the American Psychological Association.

Roscoe, S. N., Hasler, S. G., & Dougherty, D. J. (1952/1966). *Flight by periscope: II. Making takeoffs and landings; the influence of image magnification, practice, and various conditions of flight* (Human Engineering Report SDC 71-16-9, Contract B6ori-71, Task Order XVI). Port Washington, NY: Office of Naval Research, Special Devices Center. (Eventually declassified and published in 1966 in *Human Factors, 8*, 13-40.)

Roscoe, S. N., Smith, J. F., Johnson, B. E., Dittman, P. E., & Williams, A. C., Jr. (1950). *Comparative evaluation of pictorial and symbolic VOR navigation displays in a 1-CA-1 Link trainer* (Report No. 92). Washington, DC: Civil Aeronautics Administration, Division of Research.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.

Skinner, B. F. (1960). Pigeons in a pelican. *American Psychologist, 15*, 28-37.

Stroud, J. M. (1955). The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology.* Glencoe, IL: Free Press.

Taylor, H. L. (Ed.). (1994). *Division 21 members who made distinguished contributions to engineering psychology.* Washington, DC: Division 21 of the American Psychological Association.

Thorne, D. R., Genser, S. G., Sing, H. C., & Hegge, F. W. (1985). The Walter Reed performance assessment battery. *Neurobehavioral Toxicology and Teratology, 7*, 415-418.

Tolcott, M. A. (1994). Jerome Hirsch Ely, 1924-1963. In H. L. Taylor (Ed.), *Division 21 members who made distinguished contributions to engineering psychology.* Washington, DC: Division 21 of the American Psychological Association.

Tufts College & US Naval Training Devices Center. (1949). *Handbook of human engineering data.* Medford, MA: Author.

Vince, M. A. (1945). *Direction of movement of machine controls* (Report 233). London: UK Medical Research Council, Royal Naval Personnel Research Committee; also (Report 637). Cambridge: UK Medical Research Council, Flying Personnel Research Committee.

White, C. T. (1956). *Time compression of radar and sonar displays* (Report 671). San Diego: US Naval Electronics Laboratory.

Williams, A. C., Jr. (1947). *Preliminary analysis of information required by pilots for instrument flight* (Contract N6ori-71, Task Order XVI, Interim Report 71-16-1). Port Washington, NY: Office of Naval Research, Special Devices Center.

Williams, A. C., Jr. (1980, posthumously). Discrimination and manipulation in goal-directed instrument flight. In S. N. Roscoe. *Aviation psychology* (pp. 11-30). Ames: The Iowa State University Press.

Williams, A. C., Jr., & Adelson, M. (1954). *Some considerations in deciding about the complexity of flight simulators* (Research Bulletin AFPTRC-TR-54-106). Lackland Air Force Base: Air Force Personnel and Training Research Center.

Williams, A. C., Jr., Adelson, M., & Ritchie, M. L. (1956). *A program of human engineering research on the design of aircraft instrument displays and controls* (WADC Technical Report 56-526). Wright Patterson Air Force Base: Wright Air Development Center.

Williams, A. C., Jr., & Flexman, R. E. (1949). *An evaluation of the Link SNJ operational trainer as an aid in contact flight training* (Contract N6ori-71, Task Order XVI, Technical Report 71-16-5). Port Washington, NY: Office of Naval Research, Special Devices Center.

Williams, A. C., Jr., Macmillan, J. W., & Jenkins, J. G. (1946). *Preliminary experimental investigation of "tension" as a determinant of performance in flight training* (Report 54, Publication Bulletin L 503 25). Washington, DC: Civil Aeronautics Administration, Division of Research.

Williams, A. C., Jr., & Roscoe, S. N. (1949). *Evaluation of aircraft instrument displays for use with the omni-directional radio range* (Report 84). Washington, DC: Civil Aeronautics Administration, Division of Research.

Woodson, W. (1954). *Human engineering guide for equipment designers*. Berkeley: University of California Press.

Wulfeck, J. W., Weisz, A., & Raben, M. (1958). *Vision in military aviation* (TR-WADC 58-399). Wright-Patterson Air Force Base, OH: Wright Air Development Center.

Intentionally blank

INDEX

A

B

C

D

E

F

P

Q

R

S

T

U

V

Intentionally blank

Inside Back Cover

Beyond their basic intelligence and perceptual-motor skills, successful managers of dynamic situations depend largely on what is now called **situational awareness**. In the aviation community where it was first studied, it is often referred to as "airmanship" and in the engineering psychology community as "residual attention." By whatever name, situational awareness is the overarching ability to glean information from many concurrent sources, evaluate alternatives, establish priorities, estimate probable outcomes of alternative courses of action, and work on whatever has the highest momentary urgency without losing control of routine operating demands. In the world of dynamic environments, there has always been a need for a reliable means of selecting controllers and coordinators who possess this ability to a high degree.

KEEPING THE PICTURE, The Measurement of Flow Control traces the involvement of psychologists during and following World War II in the scientific study of human factors in system design and in the selection and training of personnel to operate complex machines such as airplanes, submarines, and various air-traffic and weapon control systems. This book also addresses the difficulties in developing and validating tests that measure situational awareness as a predictor of future operational performance and identifies the rules to be followed if such tests are to be free of cultural biases due to race, gender, or training and experience in operating specific systems. Finally, *KEEPING THE PICTURE* is a concise yet detailed guide for implementing the assessment of situational awareness in programs to select and train controllers of complex dynamic situations.

ISBN 2-98036-825-3

9 782980 368257

1.3