Archives of Clinical Neuropsychology 22S (2007) S15–S37

ANAM® Genogram: Historical perspectives, description, and current endeavors☆

Dennis L. Reeves[a,*], Kathryn P. Winter[b], Joseph Bleiberg[c], Robert L. Kane[d]

[a] Clinvest, Springfield, MO, United States
[b] Spawar Systems Center, NAS, Pensacola, FL, United States
[c] National Rehabilitation Hospital, Washington, DC, United States
[d] VA Maryland Healthcare System, University of Maryland School of Medicine, United States

Abstract

This paper presents a historical overview and current perspective of the Automated Neuropsychological Assessment Metrics (ANAM®) test system. We discuss the history of its development along with a synopsis of the evolution of computerized testing that has occurred and led to ANAM® over the past 30 years within the Department of Defense (DoD). We include a description of our current system and test library. Finally, we present an overview of advanced development projects that are presently underway.

We have intentionally avoided addressing issues of reliability, stability, clinical sensitivity, and construct validity in this paper. These issues are presented in other reports in this special issue.

Keywords: Neuropsychological assessment; Performance assessment; Computerized assessments; Computerized neuropsychological assessment

1. Introduction

The Automated Neuropsychological Assessment Metrics (ANAM®) is a library of computerized tests and test batteries designed for a broad spectrum of clinical and research applications. The current ANAM® system is the result of 30 years of computerized psychological test development sponsored primarily by the U.S. Military. It is the direct decedent of the joint services Unified Tri-Service Cognitive Performance Assessment Battery (UTC-PAB). ANAM® has been constructed to meet the need for serial testing and precision measurement of cognitive processing in a variety of contexts that include neuropsychology, fitness for duty, neurotoxicology, pharmacology, military operational medicine, human factors engineering, aerospace and undersea medicine, and most recently, sports medicine.

ANAM® operates in the Microsoft Windows™ environment on IBM-compatible notebook and desktop systems and is compatible with Windows™ 2000, and XP. It also has Internet "Web-enabled" and handheld Palm OS modules. Throughout ANAM® development, we have maintained the integrity of the basic test units. Hence, data collected with MSDOS versions such as ANAM 3.11 are comparable to data from the present Windows™ based version of ANAM®. Tests in the ANAM® system have been designed to assess attention and concentration, working memory, mental

☆ Individual programmers that have significantly contributed to the creation and advanced evolution of ANAM include, S. Lacour, K. Raynsford, K. Culligan, M. Flannigan, and P. Muldoon, SPAWSAR, Pensacola, FL.

* Corresponding author.
E-mail address: dreeves@clinvest.com (D.L. Reeves).

0887-6177/$ – see front matter
doi:10.1016/j.acn.2006.10.013

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**
4. TITLE AND SUBTITLE **ANAM? Genogram: Historical perspectives, description, and current endeavors**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Clinvest,Springfield,MO**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**23**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

flexibility, spatial processing, cognitive processing efficiency, memory recall, mood, fatigue level and psychomotor performance.

2. History of ANAM®

In the early 1980s the Department of Defense (DoD) was acutely aware of the use of chemical weapons by countries such as Iraq. In 1984, the commanding general of the Army Medical Research and Development Command, MG G. Rapmund, directed the Division of Neuropsychiatry at the Walter Reed Army Institute of Research (WRAIR) to establish a program to assess the impact of chemical defense medications and antidotes on normal performance. Subsequently, the director of the Division of Neuropsychiatry, COL F. Tyner, established the Joint Working Group on Drug Dependent Degradation on Military Performance, with F.W. Hegge (the senior research psychologist in the Army) as the director. The mission of this tri-service Joint Working Group was to develop a standardized performance assessment system that included cognitive, emotional, perceptual, psychomotor, and physiology measures. The overall objective was to assess the extent to which a pre-treatment drug or antidote would produce an appreciable and possibly problematic performance decrement, potentially endangering lives and missions during military operations. The core of the Joint Working Group charter included development of a psychological performance testing system that would enable the working group, as Dr. Hegge often stated, to "write the warning label on the antidote and pre-treatment bottle".

Under Dr. Hegge's direction and inspiration, the Joint Working Group designed a multiple-level drug evaluation program that provided for transition from basic laboratory science through three levels of clinical trials, to simulator and field validation, and finally to computer modeling of military performance and drug interactions (Hegge, 1983). The program included task analysis studies which translated military jobs, such as flying a Black Hawk helicopter or operating an M-1 Abrams tank, into behavioral taxonomies of critical task elements linked to behavioral constructs measured by cognitive and psychomotor tests (Cooper et al., 1987). The final level of the program was the development of a personal computer (PC)-based network simulation technology named MicroSaint. This breakthrough technology enabled behavioral scientists to construct and run computer simulations of medical and field operations on their PCs (Laughery, Archer, Plott, & Dahn, 2000; Laughery, LeBiere, & Archer, 2004).

During this era, a number of research and clinical procedures were requiring participants to be evaluated repeatedly over various time intervals. These included: (1) monitoring effects of medication on mental functioning; (2) assessing the effects of medical interventions, such as serial spinal taps or ventricular shunts; (3) tracking disease progression; (4) monitoring the acute effects of brain injury and concussion; (5) assessing the effects of behavioral and environmental treatment interventions in brain injured patients. It was recognized that traditional psychological and neuropsychological test batteries were not designed for repeated testing paradigms. For example, a number of the most sensitive neuropsychological measures, such as the Halstead-Reitan Neuropsychological Test Battery, were vulnerable to the effects of practice (Dodrill & Troupin, 1975; McCaffrey et al., 1992). As such, traditional measures were not designed to be used in extended baseline studies required by the Joint Working Group mission.

In contrast, the DoD had been sponsoring computerized test development with emphasis on repeated measures paradigms, generation of alternate forms and the ability to assess changes in performance over time. These computerized test systems became the obvious choice to meet the research and clinical objectives and led to the development of ANAM®.

Tests selected for inclusion in this system were those that had been supported by prior research and which had evidence for construct validity, reliability and sensitivity. Specifications for both the original UTC-PAB (Englund et al., 1987) and later ANAM® were derived from well established computer test systems from all three uniformed services. These include: the Walter Reed Performance Assessment Battery (WRPAB) (Thorne et al., 1985); the Air Force Criterion Task Set (CTS) (Shingledecker, 1984); Navy Performance Evaluation Tests for Environmental Research (PETER) (Bittner, Carter, Kennedy, Harbeson, & Krause, 1986); Naval Medical Research Institute Performance Assessment Battery (NMRI-PAB) (Thomas & Schrot, 1988); the North Atlantic Treaty Organization (NATO) Standardized Tests for Research with Environmental Stressors battery developed by the NATO Aerospace Medical Panel Working Group 12 led by C. Bates, G. Santucci, and J.A. Winship (Reeves et al., 1991; Santucci et al., 1989). The original UTC-PAB system development was subsequently sponsored by the Army Medical Research and Development Command, F.T. Detrick, MD. The depth and breadth of these historical roots for are significant and establish the validity of the ANAM® system. Key characteristics are described in the following sections.

2.1. The Walter Reed Performance Assessment Battery

The WRPAB was designed to assess performance changes in sustained operations and sleep deprivation paradigms, and to test the efficacy of countermeasures of performance degradation due to fatigue and lack of sleep (Thorne et al., 1985). According to the test's authors, tasks were selected: (1) to represent samples of basic skills presumed to underlie real-world tasks; (2) for brevity and repeatability; (3) for their ability to be implemented on the computer; (4) for their demonstrated or presumed sensitivity to physiological, psychological, or environmental variables. The original WRPAB contained 21 tests and operated on Apple II computers. These tests provided measures of attention, working memory, spatial processing, logical reasoning, overall cognitive performance efficiency, and mood. An important aspect of the WRPAB system was Dr. D. Thorne's creation of a cognitive efficiency measure that he labeled "throughput". This metric translates literally as the number of correct responses per minute. As such, it is a measure of the rate of producing correct responses on a given test and is derived from a formula that in simple terms is response time divided by accuracy. The throughput measure has been adopted by the ANAM® system and has proven to be one of the most sensitive metrics for detecting change in performance.

2.2. Performance Evaluation Test for Environmental Research

The development of the PETER battery began in 1977 (Bittner et al., 1986). This program focused on identifying traditional cognitive measures that were suitable for repeated administration. Over 100 paper and pencil tests were reviewed and, from these, a subset of tests was selected which had been determined to stabilize quickly and were considered suitable for repeated administrations. Subsequently, these tests were implemented for automated administration on Apple II computers (Irons & Rose, 1985). In a similar effort, psychologists J. Thomas and J. Schrot developed the Naval Medical Research Institute Performance Assessment Battery (Thomas & Schrot, 1988). This battery was one of the first systems to operate on IBM-compatible computers and included original tests as well as components of the PETER battery and the WRPAB. The purpose of the battery was to provide standardized assessment for use in operational military research in heat and cold-stress environments.

2.3. The Criterion Task Set

The CTS was developed in a Human Factors Engineering laboratory at the USAF Harry G. Armstrong Aerospace Medical Research Laboratory, Wright Patterson AFB, OH (Shingledecker, 1984). This test system operated on Commodore 64 computers. It was the first computerized test battery to be firmly based on theoretical models of perceptual-motor and cognitive behavioral theories such as Wickens' resource allocation model (Wickens, 1984). This model hypothesized three primary stages of processing and associated resources dedicated to perceptual input, central processing, and motor output or response activities. Additionally, the CTS was accompanied with a large normative database that had been collected during validation studies (Schlegel & Gilliland, 1990).

CTS development was followed closely by specification of the NATO STRES battery, which was designed for aerospace and environmental medicine applications and operation on IBM-PC systems (Reeves et al., 1991). With the advent of special software timers that permitted near millisecond timing accuracy without additional hardware, the STRES battery enhanced the portability of DoD computerized test systems. Since timing cards and special input devices were no longer required, it was possible to administer this battery on a desktop or laptop computer. Two other test programs related to aerospace medicine that also supported ANAM® development were sponsored by NASA in the 1990s. The first was directed by Dr. Sam Schiflett, at the School of Aerospace Medicine, Brooks Air Force Base, and San Antonio, TX. This project produced the Performance Assessment Work Station (PAWS) (Schlegel, Shehab, Schiflett, French, & Eddy, 1995). The PAWS and was flown on two space shuttle missions in 1994 and 1996 as a behavioral science research instrument (Schiflett, Eddy, Schlegel, & Shehab, 1998). The other project created a neurocognitive subset of ANAM® for use on the International Space Station known as Spaceflight Cognitive Assessment Tool (SCAT) or the later windows version WinSCAT (Kane, Short, Sipes, & Flynn, in press).

The Joint Working Group evolved into the Office of Military Performance Assessment Technology (OMPAT), and continued under the direction of Dr. Fred Hegge. The OMPAT recognized the potential of the UTC-PAB technology for use in clinical medicine as well as for environmental research and sponsored formal development of the ANAM® system in 1990. As the direct decedent of the UTC-PAB, ANAM® development included objectives that tests were repeatable, offered millisecond precision in response time measurement, assessed a broad range of cognitive, perceptual, and psychomotor domains, and could detect alterations in central nervous system (CNS) function. Additionally, ANAM® development focused on clinical applications in field and medical settings.

A variety of normative data collection efforts have been conducted over the years. These include U.S. Marine Corps Recruits (Gastaldo, Reeves, Levinson, & Winger, 2001) and Navy divers (Lowe & Reeves, 2002). However, one program that has made ANAM truly clinically operational in brain injury is being conducted by the Defense and Veterans Brain Injury Center under the direction of Dr. D. Warden and in collaboration with Dr. J. Bleiberg and the National Rehabilitation Hospital, Washington, DC (Warden et al., 2001). This program has focused on the diagnosis and treatment of traumatic brain injury. The foundation for this work began in the 1990s with a series of basic psychometric studies to determine ANAM's suitability for mild brain injury research and clinical practice (Bleiberg, Garmoe, Halpern, Reeves, & Nadler, 1997; Bleiberg, Kane, Reeves, Garmoe, & Halpern, 2000; Daniel et al., 1999), culminating in a recently published large-scale sports concussion study (Bleiberg et al., 2004).

Early versions of ANAM® employed technology from the NATO STRES battery (Reeves et al., 1991). As ANAM® evolved, more tests were developed to meet specific needs in neuropsychological assessments. Throughout production, ANAM® retained the design philosophy that has been part of DoD-sponsored computerized test systems. Namely, ANAM® remains a library of test modules that can be configured into batteries designed for different clinical and research applications. Additionally, it also includes pre-configured batteries designed for a range of clinical populations.

Consistent with the philosophy underlying the UTC-PAB, ANAM® includes a flexible test system built around a library of test modules. Each module has modifiable parameters such as number of stimuli and rate of stimulus presentation. These options permitted clinicians and researchers to adjust the number of items, inter- and intra-test intervals, and the appearance of some of the stimuli. A test-menu system allows the user to create batteries using different test combinations for different purposes. Additionally, pseudo-randomization procedures allowed the creation of multiple alternative forms from item sets to permit tests to be used for performance monitoring and in repeated measures designs. Separate companion utilities were developed that allowed convenient data extraction and summary. The early versions of ANAM® were developed for computers using the MS DOS operating system. The latest version of ANAM®, to be presented here, runs in Microsoft Windows™ versions 2000 and XP.

ANAM® software development continues to be sponsored by the U.S. Army Medical Research and Material Command (AMRMC). COL Karl Freidl was the initial Program Director and as a dynamic driving force for the program. He has recently been succeeded by COL Brian Lukey who is furthering development of the ANAM system. Other important benefactors from AMRMC are Dr. Stephen Grate who has provided program management and direction that has led to the successful granting of a U.S. Patent. The ANAM® program, software has been written by a team at SPAWAR System Center Charleston, Pensacola Detachment, NAS, Pensacola, FL. Kathy Winter continues as Principal Investigator and Programming Team Leader. Dennis Reeves has served as a project scientist and system design coordinator along with Robert Kane, Joseph Bleiberg, Alan Lewandowski, and Jack Spector. At present ANAM is being transitioned for commercial production at C-Shop, University of Oklahoma under the auspices of Kirby Gilliland and Robert Schelegel.

3. ANAM® system description

System requirements. The ANAM® system has been designed for use on IBM-compatible computer systems. Minimum requirements include a Pentium 90 MHz microprocessor, 32 MB RAM; 4 MB free disk space. Response devices include Microsoft or Logitech compatible mice. Windows 2000 or XP is also required.

Alternate forms. ANAM® tests produce multiple alternate test forms. The test session number is placed into a formula that uses pseudo-randomization to select items for inclusion during the administration of each test. Pseudo-randomization means there are restrictions on item selection to conform to certain pre-set limits. For example, items are selected in the Mathematical Processing Test so that there is a balance in the number of items whose answers are greater than or less than 5. The use of the session number as a seed number in the pseudo-randomization procedure ensures that a subject taking the test for the first time, second time, third time, etc., gets the same test items as any other subject at each sequence. However, the items differ for each test session. So items for session number one are always the same but they differ from item sequences used in subsequent test administrations.

Parameter modification. Tests in ANAM® can be configured in a variety of ways. Test parameters are controlled by "Command line switches" that specify items such as the number of stimuli presented, the size, color and type of stimuli used, and the control of inter- and intra-stimulus intervals. The menu list file is an ASCII text file that is read by the "master" program, A2_Menu.exe, which runs the test battery. This design allows the construction of customized batteries using a text editor such as Microsoft® Notepad.

Participant interface. The computer keyboard, mouse, and most recently, voice recognition may be used for response devices. The conventional device has been the computer Microsoft® or Logitech serial mouse. On most tests, participants respond by pushing either the left or right mouse button. The mouse interface simplifies the computer for the participant and keyboard familiarity becomes irrelevant for most tests. The person taking the test rests their hand on the mouse in the usual fashion and pushes the appropriate button. They are able to focus their attention on the stimuli being presented on the computer screen and not on the response device.

Data output. All tests within the ANAM® system produce a consistent series of scores. These scores include: (1) percent correct, (2) mean response time for all responses, (3) mean response time for correct responses, (4) standard deviation of the mean response time for all responses, (5) standard deviation of the mean response time for correct responses, (6) median response time for all responses, (7) median response time for correct responses, (8) throughput (calculated number of correct responses per minute), (9) response omissions (lapses), and (10) premature responses (i.e., "bads" defined as responses of less than 130 ms). Some scores can be displayed immediately following the administration of a test battery. Immediate feedback is an option that may be selected when starting the battery. ANAM® also has three companion modules that allow the examiner to view all test scores in selected formats. The original Statview and ANAM view programs were updated to work in the Windows environment. However, they do not accommodate all of the most recent additions to the ANAM library of tests. A new program called ANAM Xmlviewer allows data to be retrieved in a format that is compatible with standard statistical packages. This utility retrieves data so that all data for a subject are displayed in a one-subject per-row format.

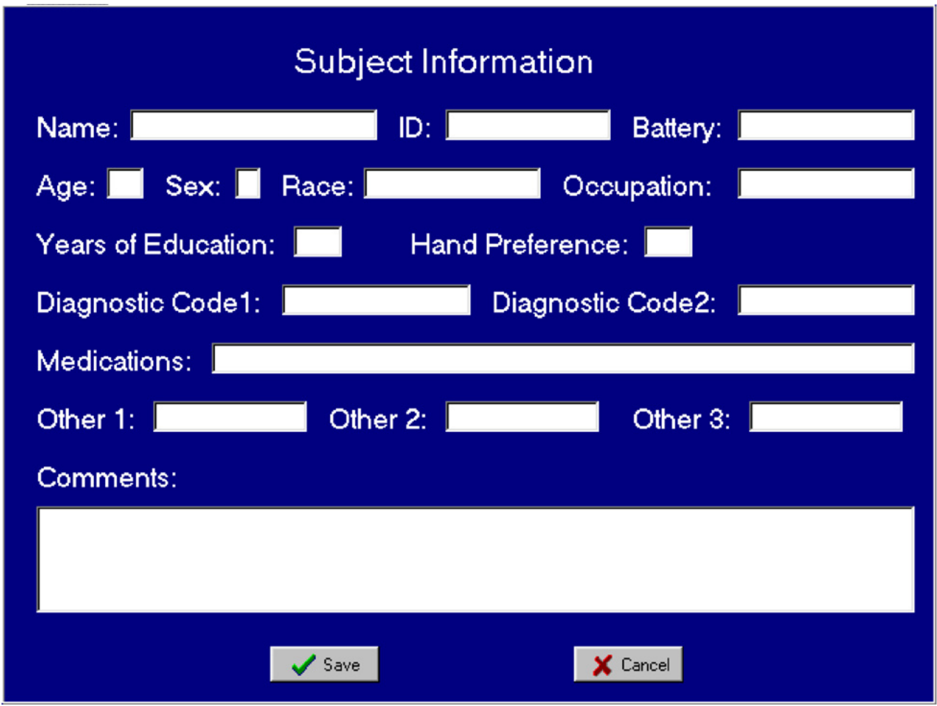
Two third-party ANAM® data management utilities have been developed through the National Rehabilitation Hospital. These have been and continue to be distributed free of charge. The ANAM® Access Database (AADB) automates large-scale importation of ANAM raw and summary data. Simple interface screens assist the user to navigate through pre-existing tables for each ANAM subtest, and an automatic query feature allows the user to rapidly and accurately organize and view data from any ANAM® subtest variable. File and error log tables assist in identifying questionable raw and summary data files. Secondly, ANAM® ClinicView© is a software utility for creating graphical and tabular displays of a subject's ANAM performance across multiple sessions. It thus is useful for research and clinical monitoring of an individual subject's status over time. These may be obtained at no-cost on the Internet at: http://www.nrhresearch.org/ccn.cfm?id=700).

4. ANAM® test modules

ANAM® presently includes 31 test modules and several companion utilities designed for recording demographic information and extracting summary data for research purposes. Additionally, there are several projects being sponsored by AMRMC that are specifically designed to develop third-party utilities that provide for extraction and computer assisted interpretation of clinical results. The following is a brief description of the current tests and support modules in the ANAM® system. In general, each test is preceded by practice items so that the test taker will understand what is required by each task.

4.1. Subject and patient information

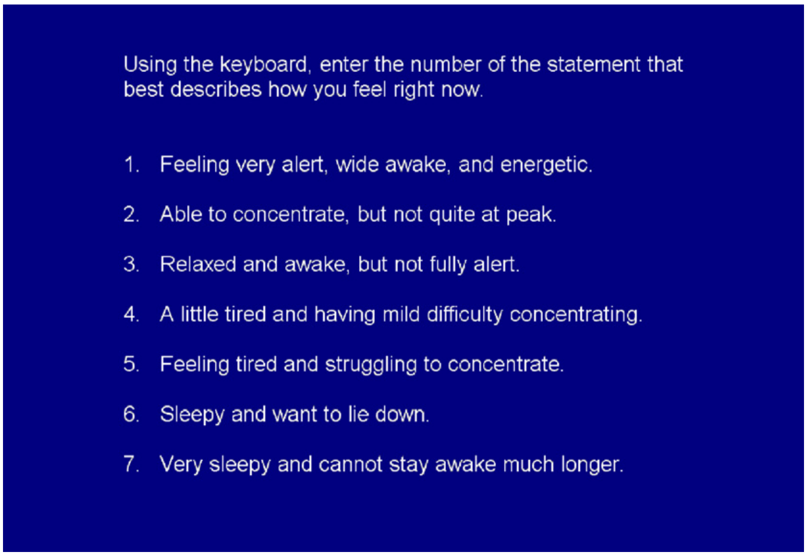
1. *Demographics Module*: This demographics module is designed to allow the recording of relevant historical and medical information. Participant's coded identifications (i.e., IDs) and the name of the battery being administered are automatically registered. Fixed and modifiable fields allow for the recording of pertinent background information such as age, gender, and education.



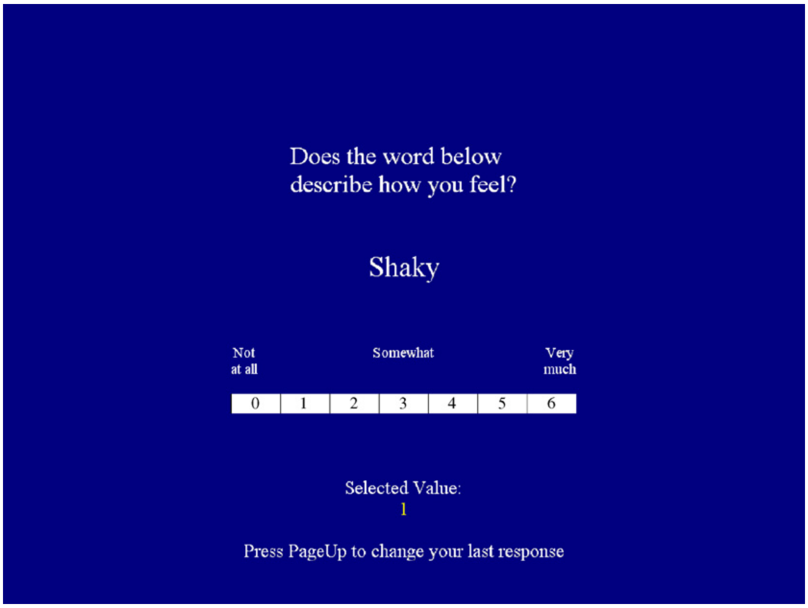
2. *Information Module*: This is a module that permits the addition of information screens between tests that may be used to provide additional help to the person taking the test. It can incorporate both text files created with Windows Notepad and graphical files that are formatted as bitmap (bmp) files.

4.2. Emotional state and alertness

1. *Stanford Sleepiness Scale-Revised*: This module consists of a listing of seven statements that describe states of alertness. The original version was created and validated by Hoddes et al. (1973). The version used in ANAM® was derived from the WRPAB (Thorne et al., 1985). In the ANAM® version, the descriptive statements have been simplified to minimize the effects of education and culture. The statements implemented in the ANAM® version of the Sleepiness Scale are presented below.

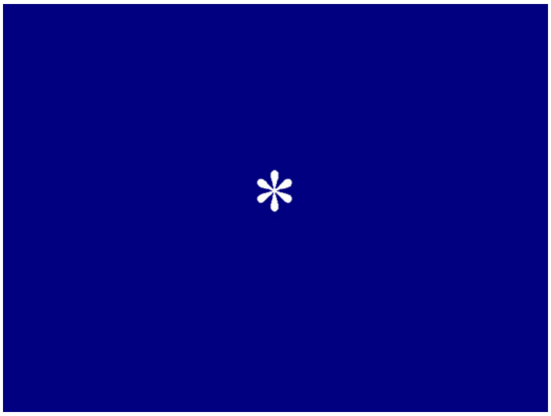


1. *Mood Affect Scale*: This test was derived from a paper-and-pencil adjective checklist constructed and validated by Ryman, Biersner, and LaRocco (1974). The ANAM® version was derived from the WRPAB Moodscale-2 (Thorne et al., 1985). The ANAM® Mood Affect Scale presently consists of seven subscales that include six adjectives each. The subscales include Vigor, Fatigue, Happiness, Depression, Restless, Anxiety, and Anger. Participants respond by moving a mouse cursor across a seven point variable analog scale displayed under adjectives as they are presented one at a time on the computer screen. This module used an Xml file that includes and categorizes adjectives for the various subscales. As such, it has been designed to allow expansion or contraction of mood or affect subscales and modification of adjectives to accommodate cultural and language translations (this includes characters/symbols such as occur in Chinese, Japanese calligraphy).

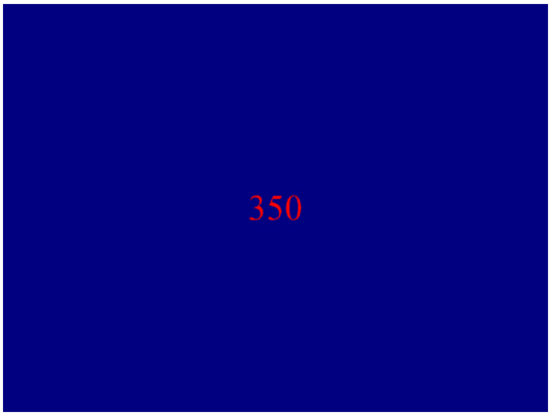


4.3. Reaction time and general cognitive processing

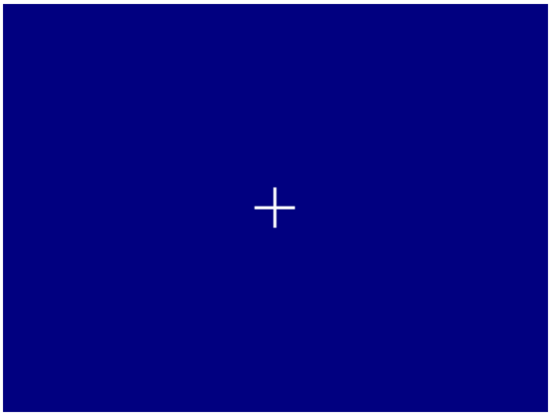
2. *Simple Reaction Time* (*SRT*): The ANAM® version of Simple Reaction Time serves two purposes. The first is to provide a measure of pure reaction time. The importance of such a basic measure is frequently overlooked in neurocognitive assessment. The second purpose is to provide a means to partial out the effects of motor response speed from actual cognitive processing time. This test presents a simple stimulus on the screen (e.g. an asterisk: *). The participant is instructed to press a specified response key each time the stimulus is presented.



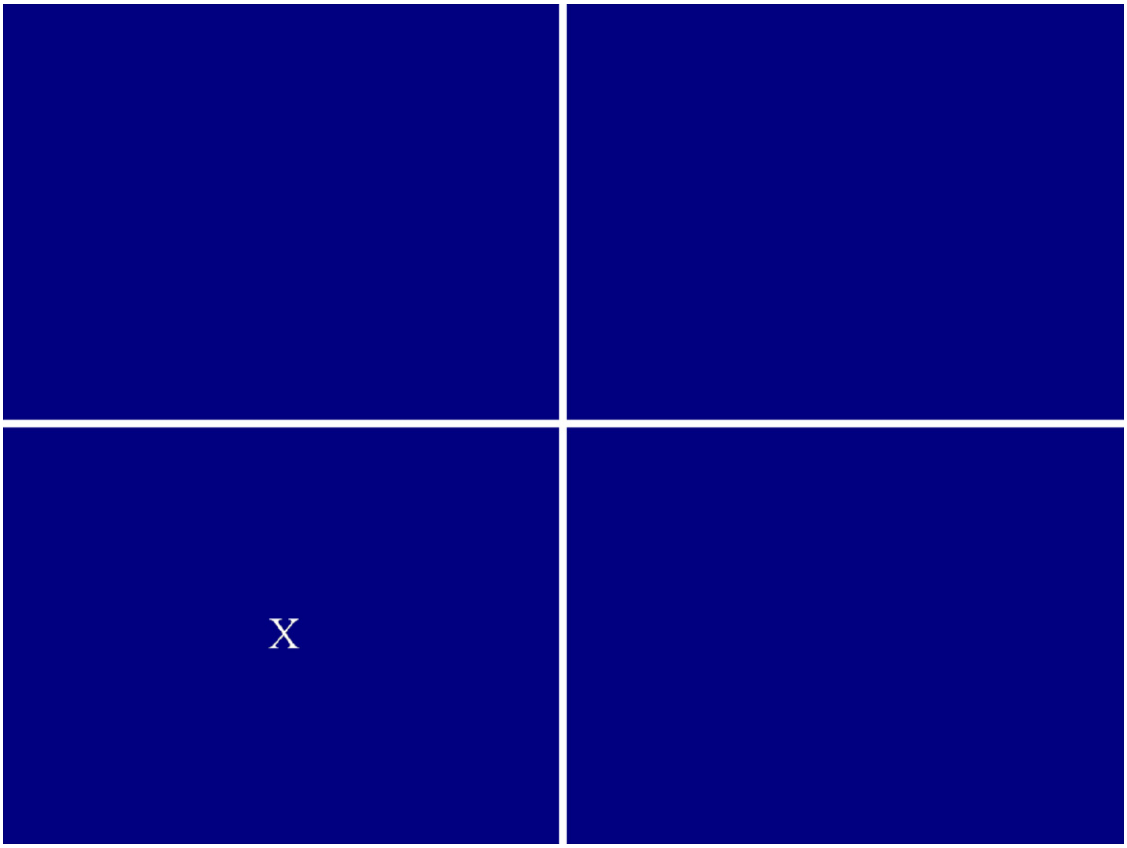
3. *Digital Reaction Time*: An alternate form or the SRT is the Digit Reaction Time task. In this version, the ANAM® traditional "asterisk" is replaced with a digital count down display that records subject response in milliseconds. This alternate form was modeled after the Performance Vigilance Task, (Dinges et al., 1997) that is frequently used in fatigue and sleep deprivation research.



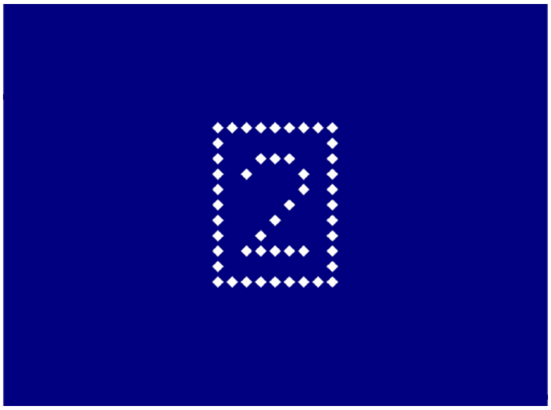
4. *Two-Choice Reaction Time*: Two-Choice Reaction Time was included in the ANAM® system to measure the ability to shift mental set in elderly individuals undergoing evaluations for dementia. The test presents one of two stimuli on the screen (e.g. + or *). The participant presses a specified response button corresponding to the presented stimulus.



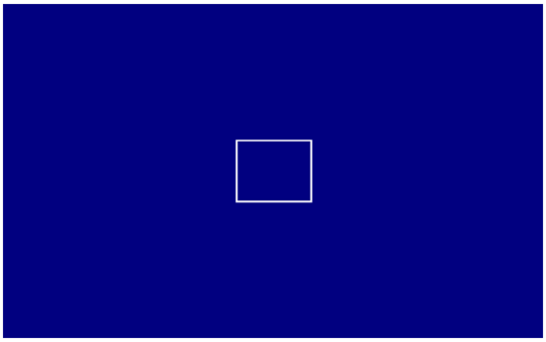
5. *Four-Choice Reaction Time*: This task is modeled after the Wilkinson's four-choice reaction time test (Wilkinson & Houghton, 1975). It consists of presentation of a small square in one of four quadrants on the computer screen. The object of this task to to move the mouse cursor and touch the square as soon as it appears. An alternate method is to use the computer keyboard as a response device. In this case the 1, 2, 3, and 4 numeric keys are used and the object is to press the key that corresponds to the quadrant that has the small square.



6. *Procedural Reaction Time-Standard*: This is a choice reaction time measure that requires the participant to differentiate between two sets of characters. The test presents a stimulus on the screen, a 2, 3, 4, or 5. The participant is required to press the left mouse key if a 2 or 3 is presented or the right mouse key if a 4 or 5 is displayed. The test is preceded by a practice session.



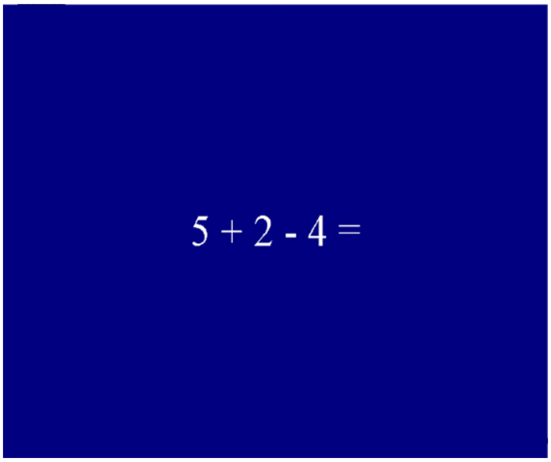
7. *Procedural Reaction Time-Symbolic*: This test is identical to the Procedural Reaction Time-Standard with the exception that the numbers are replaced with a circle, triangle, square and cross.



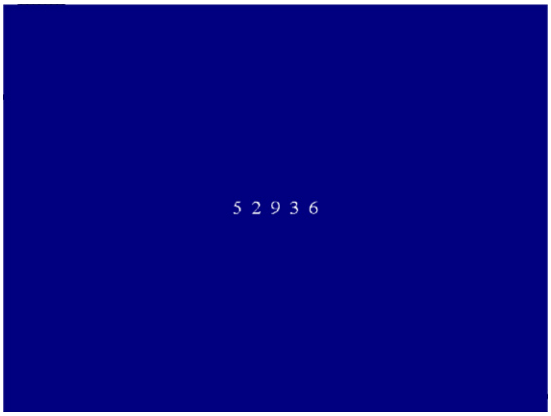
4.4. Memory tasks

8. *Memory Search* (*Sternberg Task*): This ANAM® adaptation of the Sternberg serial reaction time paradigm (Sternberg, 1966) requires participants to memorize a string of six letters. Once participants indicate they have memorized the string, it disappears from view and individual letters are presented one at a time. The object of the test is to decide whether the letter presented belongs or does not belong to the string. The test can be configured to use shorter strings of two and four letters if this is required for special populations.

9. *Mathematical Processing*: This task was designed to assess working memory by requiring the subject to perform basic arithmetic operations. All problems require an addition and subtraction sequence in the form of "$x + y - z =$". The participant indicates if the solution to the problem is greater than or less than 5. The operators and operandi are selected at random with the following restrictions: only the digits 1 through 9 are used; the correct answer may be any number from 1 to 9 except 5; test problems whose answers are greater than or less than 5 are equally probable; cumulative intermediate totals have a positive value; working left to right the same digit cannot appear twice in the same problem unless it is preceded by the same operator on each occasion (e.g. +3 and +3 are acceptable, while +3 and −3 are not); the sum of the absolute value of the digits in a problem must be greater than 5. A version of this test is also available that uses a four number equation (e.g. $w - x + y - z$).

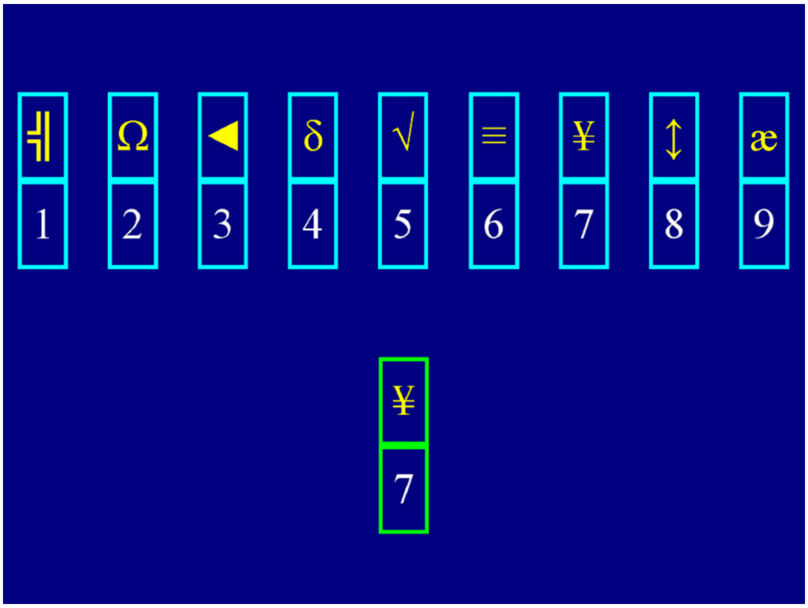


10. *Digit Set Comparison*: This test is an analogue to the classic digit span-forward test. A string of digits ranging in length from 2 to 10 numbers is presented in the center of the screen. After a specified period, this first string disappears and a second string of digits is presented. The objective is to compare the two strings and decide if they are the same digits in the same order.

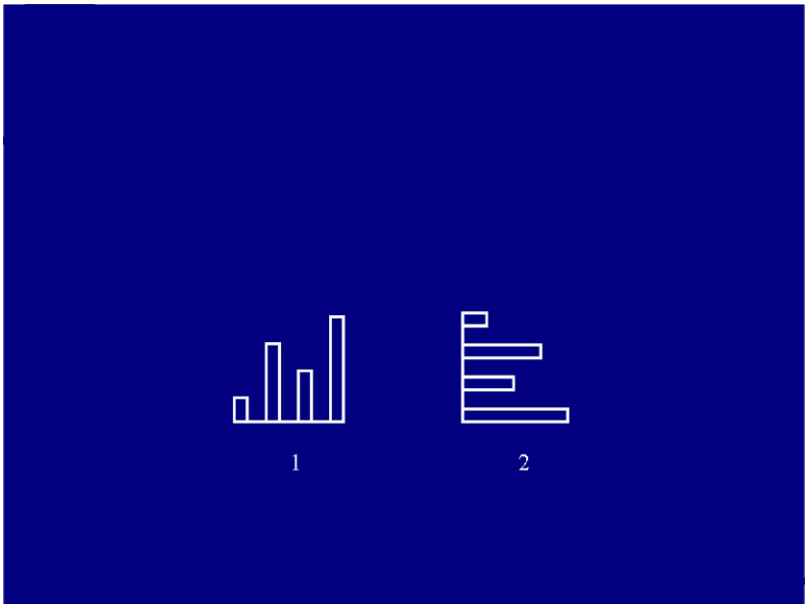


11. *Logical Reasoning-Symbolic*: This test is an adaptation of a task developed by Baddeley (1968). It differs from the original paper and pencil version in that stimulus pairs are presented one at a time and are screen-centered rather than left-justified to reduce differences in visual search times. On each trial, the symbol pair "# &" or "& #" is displayed along with a statement that correctly or incorrectly describes the order of the symbols as depicted in the example below:

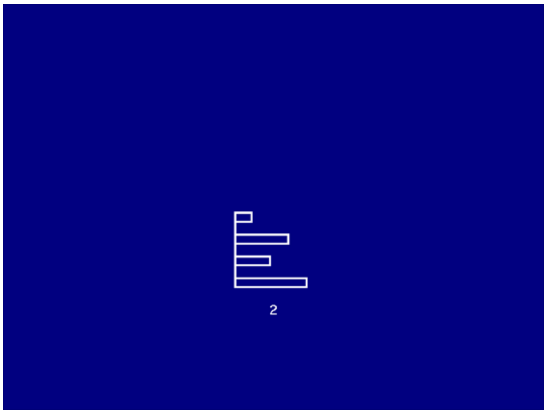
12. *Code Substitution Learning*, *Immediate and Delayed Memory Tests*: This test is based on the symbol digit-coding paradigm. The ANAM® rendition emphasizes scanning and paired associate learning of the symbol–number pairs. The test also includes a recall component. A string of up to nine symbols and nine digits are paired in a "key" across the upper portion of the screen. During the test, there is a box at the bottom of the screen that contains a single symbol–number pair. Pairings at the bottom change and at times match the pairing in the key at the top. Other times the pairing is incorrect. During the learning phase, the participant indicates whether or not the pairings at the bottom match the key and receives feedback for incorrect responses. An immediate and/or a delayed recall trial can also be included. During the recall phase, there is no key at the top and the participant must indicate if the pairings appearing at the bottom are correct or incorrect from memory.



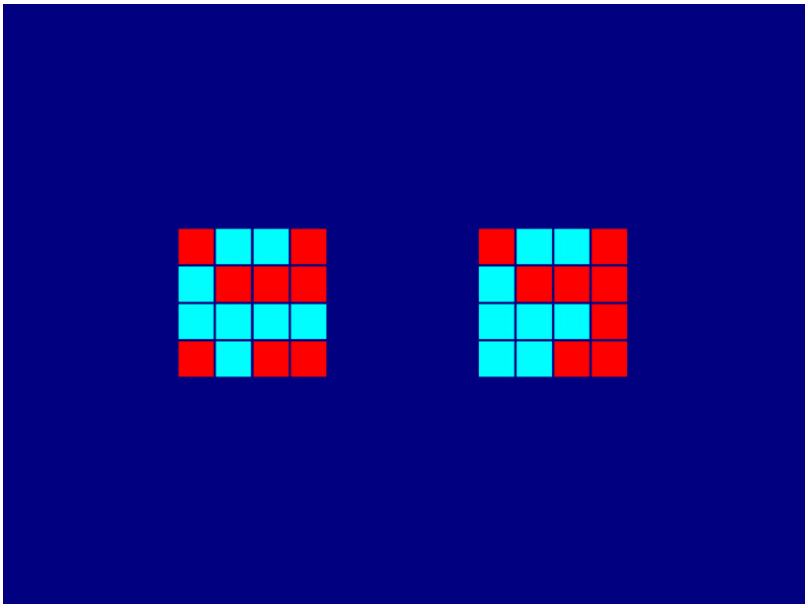
13. *Spatial Processing Simultaneous*: During this test, pairs of four-bar histograms are presented simultaneously on the monitor and the participant is requested to determine whether they are identical. One histogram is always rotated either 90° or 270° with respect to the other histogram.



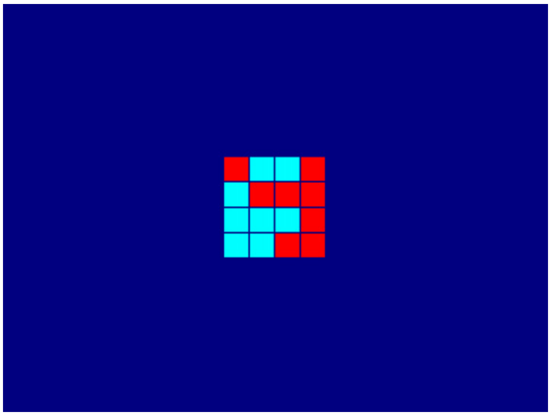
14. *Spatial Processing Successive*: This task is similar to the Spatial Processing Simultaneous task except that the two histograms are presented successively. The participant first sees one and then sees the other. The time interval between the presentations of the two histograms can be set by the examiner.



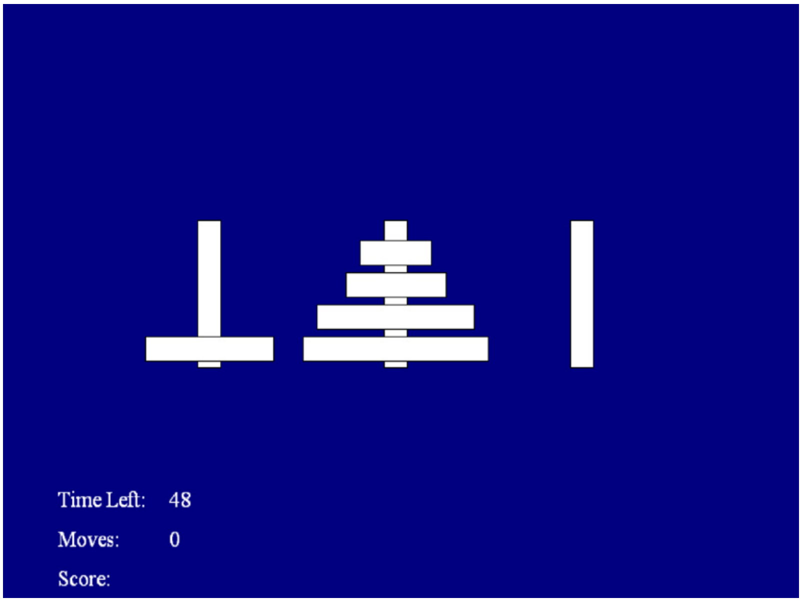
15. *Matching Grids and Matrix Rotation*: These tasks require the participant to match two 4×4 matrix (checkerboard) patterns. In the Matching Grids Task, the patterns are presented side-by-side and in the same orientation. In the Matrix Rotation task, the patterns are presented side-by-side; however, one 4×4 pattern is rotated. In both tasks, participants must indicate if the patterns are the same or different.



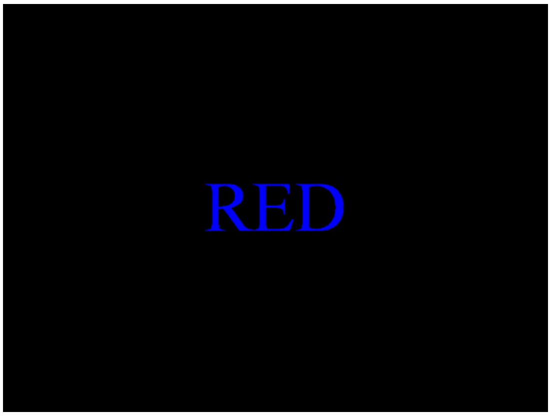
16. *Delayed Matching to Sample*: A 4×4 checkerboard matrix is presented for approximately two seconds and then disappears. After an interval, two matrices are presented side-by-side. The participant indicates which of these two matrices matches the first exactly. The examiner can determine the delay interval between the initial stimulus and the response options. The default value is 5 s.



17. *Tower Puzzle* (*Tower of Hanoi*): In this implementation of the Tower Puzzle, the participant moves pieces between three poles using either the number keys on the keyboard or the mouse. This test leads participants through instructions and practice trials, and then presents nine test puzzles. Scores are based on time and efficiency. Efficiency is scored by assessing a participants performance against the minimum number of moves needed to solve each puzzle.



18. *Stroop Color Word Test*: The ANAM® version of the Stroop task is modeled after Dr. Charles Golden's adaptation of this test. This task consists of an introductory exercise that requires an individual to press one of three computer keys to identify each of three colors, (i.e., red, green, or blue). The keys on the computer keyboard should be labeled with colored dots such that: (1) red, (2) green, (3) blue. The test begins with an initial practice trial that consists of 60 colored Xs. This is followed by a trial in which the words, Red, Green, and Blue are displayed individually in black and white. The correct response entails pressing the key that corresponds to the name of the color. The second task consists of the presentation of four Xs (XXXX) in three colors: Red, Green and Blue. The correct response for this task is to press the key that corresponds to the color of the Xs. The third task consists of presenting the original words RED, GREEN, BLUE in colors that do not match the meaning of the word. This is the interference task and requires pressing the key that corresponds to the color of the letters rather than the color indicated by the word.



4.5. Continuous performance and sustained attention

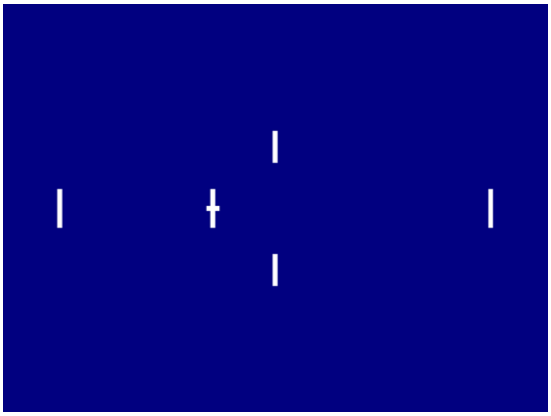
19. *Running Memory*: This task is a continuous reaction time test using a "one back" paradigm to assess working memory along with sustained attention. Specifically, this subtest requires subjects to recall the last character that appeared on the screen and decide if the current character displayed on the screen is the same or different. The task is forced-paced and subjects are given only a limited amount of time to respond to each character. This task can be configured to run for both shorter and longer time intervals. The default duration is 5 min. Examiners have the option of using numbers or letters as the test stimuli.

20. *Continuous Performance Test-Traditional*: This is a traditional continuous performance task that requires the participant to monitor the computer screen for the occurrence of an infrequent event. The test can use number, letters, or symbols as stimuli. The examiner can set the frequency with which the target stimulus appears during the test.

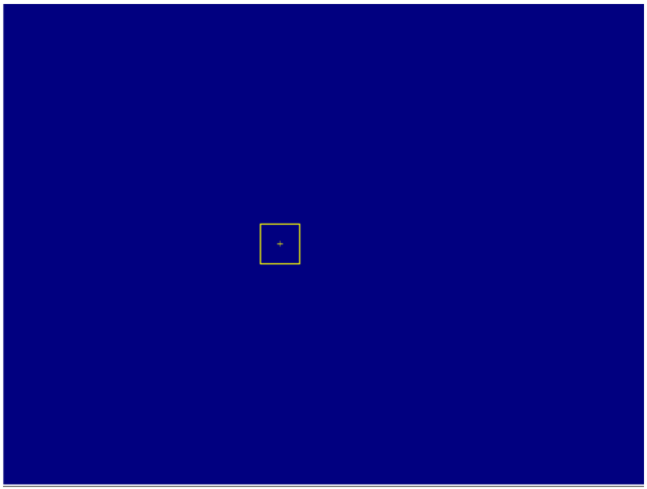
4.6. Psychomotor speed and coordination

21. *Tapping Test* (*Left and Right index finger*): The ANAM® tapping test was modeled after the Finger Tapping Test (FTT) (Reitan & Wolfson, 1993) from the Halstead-Reitan Neuropsychological Test Battery. The keyboard space bar replaces the finger-tapping device. Five 10-s trials, with 15-s rest periods for trials 1, 2, 4, 5, 7 and 8, are performed using each hand. On trials 3, 6, and 9, a 90-s rest period is imposed. The score for each index finger is the average of five trials in which the scores were within a range of 5. If this criterion is not met, the score is the average of 10 trials.

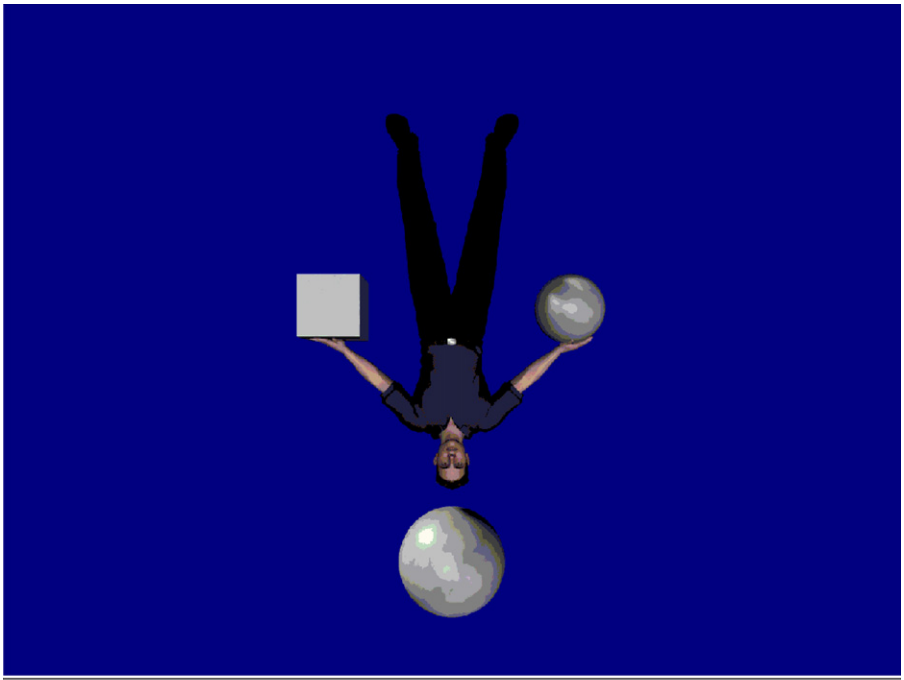
22. *Unstable Tracking*: The participant is instructed to keep the cursor centered in the middle of the screen by moving the mouse to control the movements of the cursor. The cursor moves horizontally and resists being centered. The cursor is also sensitive to the efforts of the participant to keep it centered. The faster the participant moves the mouse the more the cursor resists being centered. Hence, the task incorporates velocity as well as position sense.
23. *Adaptive Unstable Tracking*: Adaptive tracking is a special implementation of the Unstable Tracking task. With practice, participants improve their performance of tracking tasks and it becomes easier for them to keep the cursor near the center of the screen. Adaptive tracking adapts to participants as they get better at this task. As the participant improves, the task becomes more difficult by increasing the resistance of the cursor to being centered.



24. *Pursuit Tracking*: This is a very flexible module that is designed to assess the ability to track the mouse cursor while manipulating the mouse with the forearm and wrist. Options include smooth or unstable horizontal, vertical sine and square waves. The original design for this task was created by T. Elsmore and D. Reeves.

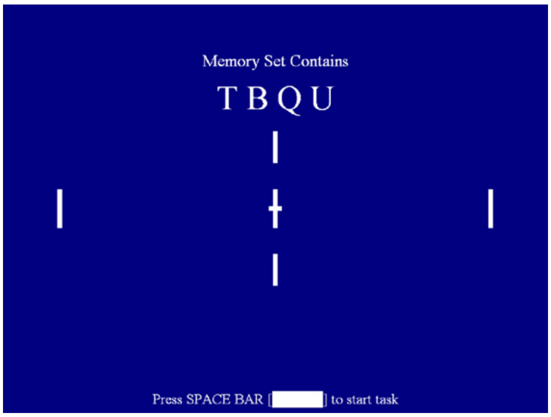


25. *Manikin*: In this task a human figure is presented holding a object in one hand. The figure appears in various orientations: standing, upside down, facing the test taker, facing away from the test taker. The task is to judge which hand is holding the object.

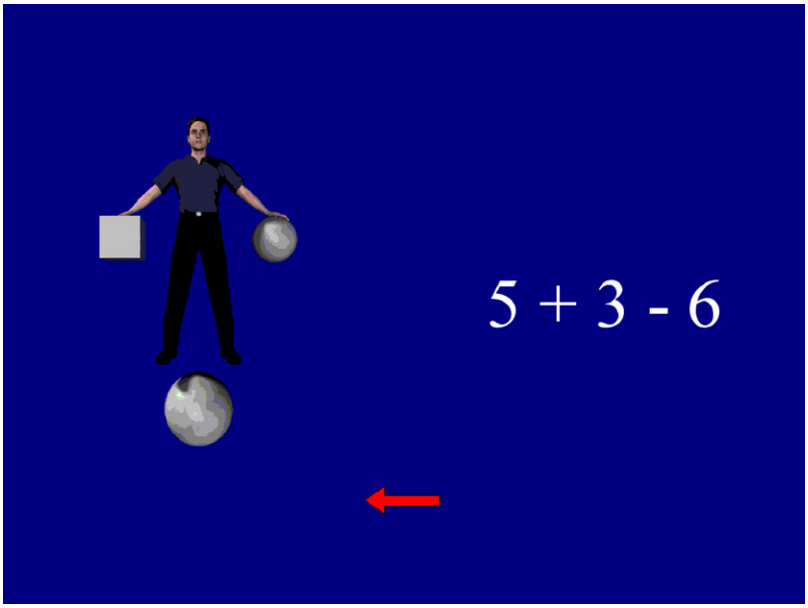


4.7. Divided attention and dual tasks

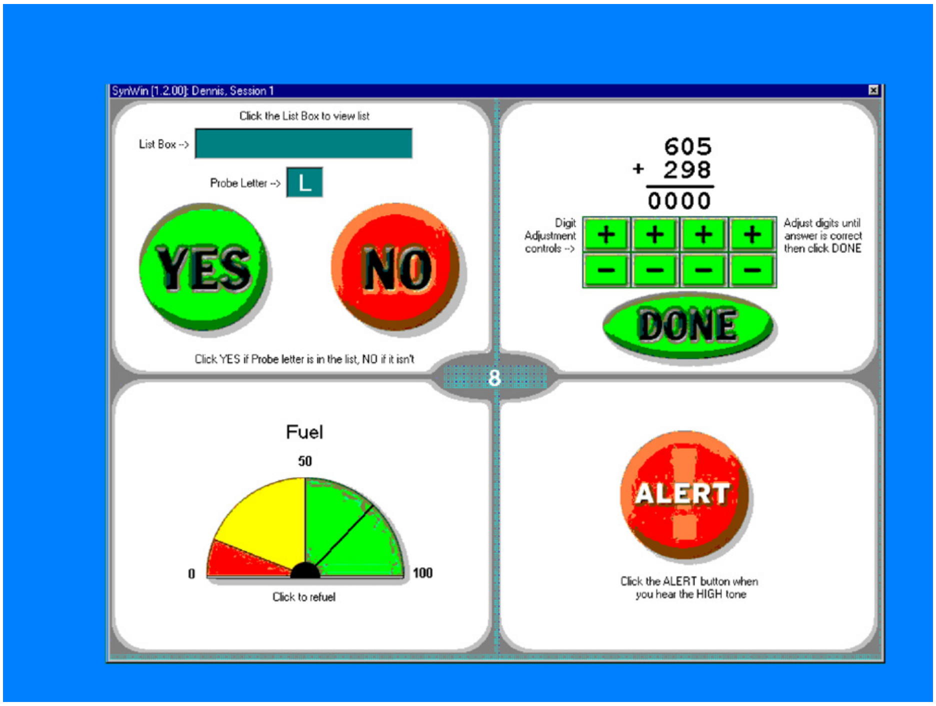
26. *Dual Task* (*Memory Search and Tracking*): Dual task combines the Unstable Tracking and Sternberg memory tasks. These two tasks are presented simultaneously. Participants work at centering the cursor while holding in mind a memory set and indicating if letters appearing on the screen are members of this set.



27. *Switching Task*: The Switching Task is a combination of the Manikin and Mathematical Processing Task. During the administration of this test, the Manikin and Math test may be displayed simultaneously or in a variable interval successive fashion. The participant is required to respond to the indicated task and then when indicated by an arrow on the screen switch to responding to the other task in the simultaneous mode. In the successive mode, one task is displayed for a number of trials and then the test switches to the second task, and goes back and forth from one task to the other on a variable interval schedule. This is designed to be an executive function task that requires the ability for mental flexibility and shifting set. It was derived from the PAWS battery (Schiflett et al., 1998).

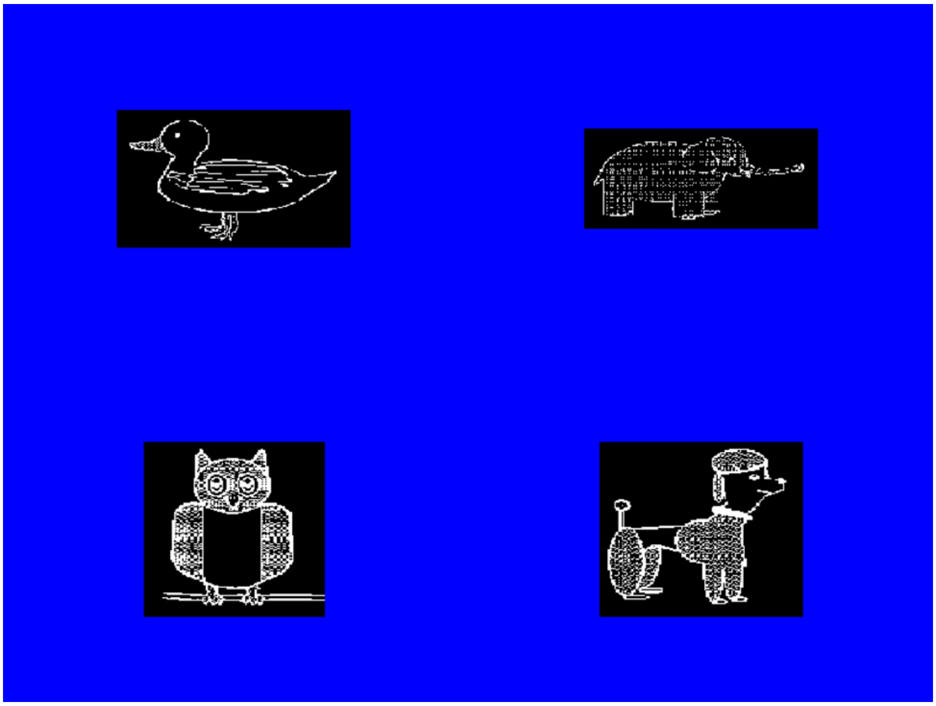


28. *Synthetic Work* (*SynWork*): SynWork is an independently developed test module that can run within the ANAM® system. It simulates a work environment by presenting up to four tasks on the screen simultaneously. These tasks include versions of the Sternberg memory task, mathematical calculation, gauge monitoring, and auditory vigilance. Tasks can be presented individually or in various combinations. Each task is scored individually. In addition, SynWork combines the participant's performance on all administered tasks into a single proficiency score.



4.8. Language module

29. *Automated Aphasia and Memory Scale*: This task was developed initially for special neuroimaging and neuromedical procedures, such as WADA testing. It combines object naming and recognition memory.



4.9. Malingering module

30. *Malingering Task*: This task was based on the Delayed Matching to Sample paradigm described above (16). It was modified to function as a test of effort. Stimuli vary with respect to degree of difficulty and there is no extended delay interval. The task yields different scores reflecting performance relating to different types of stimulus presentation.

5. ANAM® batteries

As currently released, ANAM® includes two standard batteries: (1) Neurocognitive Screening Test Battery (NeuroCog); (2) Moderate Cognitive Screening Test Battery (MODcog).

NeuroCog has served as ANAM®'s standard or "default" assessment battery. It was constructed for general use in clinical and laboratory settings. Tests included in this module are listed in the below section.

5.1. NeuroCog screening battery

1. Demographics
2. Sleepiness Scale-R
3. Mood Scale-R
4. Simple Reaction Time
5. Code Substitution 9 (Learning)
6. Matching Grids
7. Delayed Matching to Sample
8. Mathematical Processing
9. Logical Relations-Symbolic
10. Running Memory Continuous Performance Test (160)
11. Code Substitution 9 (36) (Delayed Recall Test)
12. Memory Search (Stern 6)
13. Simple Reaction Time-2nd Trial

MODcog was designed for use in screening more impaired indviduals. It includes essential components of NeuroCog with some modifications. In MODcog, we incorporated 2-Choice Reaction Time and Code Substitution 9 (Immediate Recall Test) and removed Mathematical Processing, Simple Reaction Time-2nd Trial, and the Running Memory Continuous Performance Test, since Running Memory is likely to be too demanding for the more impaired patient. Tests included in MODcog are listed the below section.

5.2. MODcog screening battery

1. Demographics
2. Sleepiness Scale-R
3. Mood Scale-R
4. Simple Reaction Time
5. 2-Choice Reaction Time
6. Code Substitution 9 (Learning)
7. Code Substitution 9 (Immediate Recall Test)
8. Matching Grids
9. Delayed Matching to Sample
10. Logical Relations
11. Code Substitution 9 (Delayed Recall)
12. Memory Search (Stern 6)

Other pre-configured batteries are or will be available. The ANAM® Sports Medicine Battery (ASMB), noted previously, is a specialized subset designed for baseline assessment of athletes and recovery from concussion. Tests in the ASMB were selected to assess sustained attention, mental flexibility, cognitive-processing efficiency, arousal/fatigue level, learning, recall, and working memory. The battery was also designed for repeated-measures testing. The ASMB includes both the full test module and a brief demonstration module. ASMB tests include:

1. Demographics Module
2. Modified Stanford Sleepiness Scale
3. Simple Reaction Time
4. Code Substitution Learning
5. Running Memory Continuous Performance Task
6. Mathematical Processing
7. Delayed Matching to Sample
8. Code Substitution Delayed Memory Test

WinSCAT is a battery developed for NASA that runs in its own shell and which contains a built in data display program. It contains the following ANAM measures:

1. Code Substitution Learning
2. Running Memory Continuous Performance Task
3. Mathematical Processing
4. Delayed Matching to Sample
5. Code Substitution Delayed Memory Test

ANAMgp is a general-purpose battery which will contain a larger selection of tests. These tests are presently being normed on a broad sample of subjects and will enhance the use of the battery for single as well as repeated measures assessment.

6. Continued development and future directions

6.1. ANAM® website

Recently, a website was established to facilitate the distribution of the battery, communication among users, and archiving of data from different studies. This website is sponsored by the United States Army Medical Research and Materials Command and is located at the U.S. Navy SPAWAR command in Pensacola, FL. The WEB address is https://www.anam.army.mil.

6.2. ANAM® Internet oracle database

An Oracle database is being developed which reads the electronic data files generated for each ANAM® test. This database is an especially powerful tool. It will facilitate the development of norms as well as studies of cognitive changes resulting from different clinical and environmental conditions.

6.3. ANAM® audio–visual animation for instructions

ANAM® tests traditionally include written instructions. A newer option is now available using Microsoft Agents. Agents are cartoon characters that speak to the participants and show them how to take the tests. Agents have advantages over streaming video clips as a method of instruction. Streaming video clips have to be remade if instructions need to be changed or translated into different languages. Agents verbalize typed text allowing for greater flexibility and ease of modification and translation. In addition, a cartoon character, such as a parrot, is culturally neutral.

6.4. Enhanced data output and interpretation

Current developments are underway to facilitate data retrieval, enhance data presentation, and to expand the normative base for both single and repeated measures administrations. Interpretation of change scores will be augmented with the implementation metrics such as a reliable change index or regression techniques.

6.5. Remote assessment and tele-health

Providing interventions to locations not served adequately by sub-specialists and experts is becoming an important aspect of healthcare. In addition, automated test batteries like ANAM® have a potential role to play in the time- and cost-efficient screening of individuals at risk for neurocognitive impairment. The major challenge to providing automated neurocognitive assessment to remote locations has been the fact that computerized tests require precise near millisecond timing to produce the most clinically useful results. The technical impediments to providing time sensitive neurocognitive assessment over the Internet are considerable and are discussed in a technical paper that is part of this special issue. A recent development pioneered at SPAWAR, Pensacola, FL, provides a means of implementing remote neurocognitive assessment while maintaining timing accuracy. This approach permits the downloading of software to a remote site, password enabled usage, the retraction of key program elements following conclusion of a test session, and encryption of data for secure transmission back to a remote site. This approach combines both store and forward and real-time techniques.

6.6. Handheld ANAM

Cognitive status assessments are frequently desired in situations where it is inconvenient or impossible to use desktop or laptop computer systems. To meet these needs, many of the ANAM tests have been implemented in the Palm Operating System for handheld computers. The first such implementation was Migraine Early Warning Tool (MEWT), a program designed for at-home monitoring of cognitive status in migraine patients (Farmer, Cady, & Reeves, 2003). This program incorporates several ANAM tests into a single Palm OS program, and provides individualized post-test feedback on performance to assist the patient and physician with decisions regarding treatment options. The ARES system (Elsmore, Reeves, & Reeves, this volume) is a general-purpose Palm OS based test system with the flexibility to configure test batteries containing many of the ANAM tests. Development of this system has been supported by the U.S. Army Medical Research and Materiel Command, and has been used successfully in a variety of military field operations.

6.7. Clinical interpretation

Many neuropsychological tests originally designed for the laboratory have evolved into clinical instruments. There is no APA or FDA set of regulations or standards to describe the process by which a laboratory neuropsychological instrument transitions into a clinical assessment instrument. Even a cursory review of contemporary neuropsychological assessment batteries reveals enormous variability in the psychometric underpinnings of different commonly employed clinical neuropsychological tests. Further, clinical instrument development often is an iterative–reiterative process, where an instrument is released after a minimally acceptable degree of psychometric development, and then undergoes multiple rounds of further development after it has been in clinical use and been incorporated into research projects. Indeed, often more is learned about an instrument after it is released than is known about it at the time of release.

ANAM, having been shown to be useful as a research instrument, now is being developed as a clinical instrument. The primary research issue of showing a statistically significant difference between groups is being replaced with the clinical task of demonstrating accurate classification and discrimination of diseased from healthy subjects. Clinical classification introduces concepts such as sensitivity and specificity, but also introduces human factors issues related to "packaging" ANAM data in ways which facilitate its use by clinicians to promote accurate diagnosis. Ongoing work in this area indicates that clinicians, even when they are psychologists, have difficulty with the enormous amount of data generated by ANAM. There thus is the need to develop methods for data reduction, simplification, and display. This includes developing ANAM summary and other aggregate scores, much like the common impairment indices used in traditional neuropsychological batteries, but also learning through clinical experience how ANAM data fit into the larger clinical diagnostic endeavor, one where multiple sources of data are available in addition to ANAM, and one where "clinical judgment" is the statistic of choice (Reeves et al., 2006).

References

Baddeley, A. D. (1968). A 3-min reasoning test based on grammatical transformations. *Psychonomic Science*, *10*, 341–342.

Bleiberg, J., Cernich, A., Cameron, K., Sun, W., Peck, K., Ecklund, J., et al. (2004). Duration of cognitive impairment after sports concussion. *Neurosurgery*, *54*(5), 1–6.

Bleiberg, J., Garmoe, W., Halpern, E., Reeves, D., & Nadler, J. (1997). Consistency of within-day and across-day performance after mild brain injury. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *10*(4), 247–253.

Bleiberg, J., Kane, R. L., Reeves, D. L., Garmoe, W. S., & Halpern, E. (2000). Factor analysis of computerized and traditional tests used in mild brain injury research. *The Clinical Neuropsychologist*, *14*(3), 287–294.

Bittner, A. C., Carter, R. C., Kennedy, R. S., Harbeson, M. M., & Krause, M. (1986). Performance evaluation tests for environmental research (PETER): Evaluation of 114 measures. *Perceptual and Motor Skills*, *63*, 683–708.

Cooper, M., Schemmer, M., Fleishman, E., Yarkin-Levin, K., Harding, F., & McNelis, J. (1987). *Task analysis of navy and marine corps occupations: A taxonomic basis for evaluating CW antidote/pretreatment drugs* (Final Report FR-3-87-1). Bethesda, MD: Advanced Research Resources Organization.

Daniel, J., Olesniewicz, M., Reeves, D., Tam, D., Bleiberg, J., Thatcher, R., et al. (1999). Repeated measures of cognitive processing efficiency in adolescent athletes: Implications for monitoring recovery from concussion. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *12*(3), 167–169.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., et al. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 h per night. *Sleep, 20*, 267–277.

Dodrill, C. B., & Troupin, A. S. (1975). Effects of repeated administrations of a comprehensive neuropsychological battery among chronic epileptics. *Journal of Nervous and Mental Disorders*, *161*(3), 185–190.

Englund, C. E., Reeves, D., Shingledecker, C., Thorne, D., Wilson, K., & Hegge, F. W. (1987). *The Unified Tri-Service Cognitive Performance Assessment Battery (UTC-PAB). I. Design and specification of the battery* (Report No. 87-10). San Diego: Naval Health Research Center.

Farmer, K., Cady, R., & Reeves, D. (2003). The effects of prodrome on cognitive efficiency. *Headache*, *43*, 518.

Gastaldo, E., Reeves, D., Levinson, D., & Winger, B. (2001). ANAM USMC normative data: Effects of heat stress on neurocognitive performance. *Archives of Clinical Neuropsychology*, *16*(8), 825.

Hegge, F. (1983). *The effect of chemical warfare treatment drugs on military performance: A Tri-Service Drug Screening Program*. Fort Detrick, MD: U.S. Army Research and Development Command.

Hoddes, E., Zarcone, V., et al. (1973). Quantification of sleepiness: A new approach. *Psychophysiology*, *10*(4), 431–436.

Irons, R., & Rose, P. (1985). Naval biodynamics laboratory computerized cognitive testing. *Neurobehavioral Toxicology and Teratology*, *7*, 395–397.

Kane, R. L., Short, P., Sipes, W., & Flynn, C. F. (in press). Development and validation of the Space-flight Cognitive Assessment Tool for Windows (WinSCAT). *Aerospace and Environmental Medicine*.

Laughery, K. R., Archer, S., Plott, B., & Dahn, D. (2000). Task network modeling and the MicroSaint family of tools. In *The Proceedings international ergonomics society meeting*

Laughery, K. R., LeBiere, C., & Archer, S. G. (2004). Modeling human performance in complex systems. In G. Salvendy (Ed.), *The handbook of human factors* (Chapter).

Lowe, M., & Reeves, D. (2002). *Automated Nuropsychological Assessment Metrics: Norms for U.S. Navy divers*. (Report No. NEDU TR 02-05), Panama City, FL: Navy Experimental Diving Unit.

McCaffrey, R. J., Ortega, A., et al. (1992). Practice effects in repeated neuropsychological assessments. *Clinical Neuropsychologist*, *6*(1), 32–42.

Reeves, D. L., Bleiberg, J., Roebuck-Spencer, T., Cernich, A. N., Schwab, K., Ivans, B., et al. (2006). Reference values for performance on the Automated Neuropsychological Assessment Metrics V3.0 in an active duty military sample. *Military Medicine*, *171*(10), 982–994.

Reeves, D. L., Winter, K. P., et al. (1991). *The UTC-PAB/AGARD STRES battery: User's manual and system documentation*. Pensacola: Naval Aerospace Medical Research Laboratory.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation* (2nd ed.). Tucson, AZ: Neuropsychology Press.

Ryman, D., Biersner, R., & LaRocco, J. (1974). Reliabilities and validities of the mood questionnaire. *Psychological Reports*, *35*, 479–484.

Santucci, G., Farmer, E., Grissett, J., Wehterell, A., Boer, L., Gotters, K., et al. (1989). AGARDograph #308, Human Performance Assessment Methods (ISBM-92-835-0510-7). North Alant6ic Treaty Organization Advisory Group on Aerospace Research and Development, Working Group 12, Seine, France.

Schiflett, S. G., Eddy, D. R., Schlegel, R. E., & Shehab, R. L. (1998). Cognitive performance aboard the Life and Microgravity Spacelab. *Acta Astronautica*, *43*, 193–210.

Schlegel, R. E., & Gilliland, K. (1990). *Evaluation of the Criterion Task Set. Part I. Appendices A and B: Univariate summaries* (Technical Report AAMRL-TR_90-008). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Schlegel, R. E., Shehab, R. L., Schiflett, S. G., French, J. R., & Eddy, D. R. (1995). The NASA Performance Assessment Workstation: Astronauts vs. a ground-based reference group. In *Proceedings of the 11th IAA Man-in-Space Symposium*.

Shingledecker, C. A. (1984). *A task battery for applied human performance assessment research* (Technical Report AFAMRL-TR-84-071). Wright-Patterson Air Force Base: Air Force Aerospace Medical Research Laboratory.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(736), 652–654.

Thomas, J. R., & Schrot, J. (1988). *Naval Medical Research Institute Performance Assessment Battery (NMRI PAB) documentation (NMRI 88-7)*. Bethesda, MD: Naval Medical Research Institute.

Thorne, D., Genser, S., et al. (1985). The Walter Reed Performance Assessment Battery. *Neurobehavioral Toxicology and Teratology*, *7*, 415–418.

Warden, D. L., Bleiberg, J., Cameron, K., Ecklund, J., Walter, J. D., Sparling, M., et al. (2001). Persistent prolongation of simple reaction time in sports concussion. *Neurology*, *57*(3), 524–526.

Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles E. Merrill Publishing Company.

Wilkinson, R., & Houghton, D. (1975). Portable four-choice reaction time test with magnetic tape memory. *Behavioral Research Methods and Instrumentation*, *7*, 441–446.