

SPECIAL PRESENTATION

Assessment of Aids-Related Cognitive Changes:
Recommendations of the NIMH Workshop on
Neuropsychological Assessment Approaches*

Nelson Butters, Igor Grant, James Haxby, Lewis L. Judd, Alex Martin,
Jay McClelland, Willo Pequegnat, Daniel Schacter, and Ellen Stover

ABSTRACT

This article presents an extended (7-9 hours) and a brief (1-2 hours) battery designed to evaluate early cognitive changes associated with seropositive, asymptomatic persons. The battery was recommended by an NIMH Workgroup which was guided by 10 principles in its development. The domains assessed by the battery are: (1) Indicators of Premorbid Intelligence; (2) Attention; (3) Speed of Processing; (4) Memory; (5) Abstraction; (6) Language; (7) Visuoception; (8) Constructional Abilities; (9) Motor Abilities; and (10) Psychiatric Assessment. Although the battery assesses a wide range of psychological functioning, specific emphasis has been placed on divided and sustained attention as well as speed of processing and retrieval from working and long-term memory. Descriptions of both the traditional clinical tests and tasks used in cognitive psychology are provided. Although the Workgroup strongly recommends the use of the extended battery in order to

* A two day Workgroup on Neuropsychological Assessment Approaches was convened by the Office of AIDS Programs, Office of the Director, National Institute of Mental Health (NIMH), Alcohol, Drug, and Mental Health Administration, Public Health Service, Department of Health and Human Services. The Workgroup was Co-Chaired by Nelson Butters, Ph.D. and Lewis L. Judd, M.D. and the following were participants: Hans Bergman, Ph.D., Karolenska Sjukhuset; Eric Caine, M.D., Rochester, NY; Jeffrey Cummings, M.D., West Los Angeles VA Medical Center; Igor Grant, M.D., University of California at San Diego; James Haxby, Ph.D., National Institute of Health; Albert Heyman, M.D., Duke University Medical Center; Alex Martin, Ph.D., National Institute of Mental Health; Jay McClelland, Ph.D., Carnegie Mellon University; Willo Pequegnat, Ph.D., National Institute of Mental Health; Daniel Schacter, Ph.D., University of Arizona; Ellen Stover, Ph.D., National Institute of Mental Health.

Requests for reprints should be addressed to Dr. Willo Pequegnat, Office of AIDS Programs, NIMH, Parklawn Building, Room 17C-06, 5600 Fishers Lane, Rockville, Maryland 20857, USA.

Accepted for publication: June 8, 1990.

ensure the most sensitivity, it recognizes that there may be situations in which this is not possible. In order to increase the likelihood that neuropsychological tests will identify neurologically affected CDC Stage II and III seropositive individuals, the Workshop recommends that each patient's protocol be rated by two trained neuropsychologists using the same clinical criteria. The Workgroup also recommends that a concerted effort be made to incorporate data from the extended and the brief batteries in some central data bank.

BACKGROUND AND GOALS

On April 10-11, 1989, the NIMH-sponsored "AIDS Workshop: Neuropsychological Assessment Approaches" met in Rockville, Maryland. The major goal of the Workshop was to recommend a neuropsychological test battery which would be suitable for investigators who are focusing upon the early detection of the AIDS Dementia Complex (ADC) and other neurologic complications in HIV-positive asymptomatic individuals (Grant et al., 1987; Janssen et al., 1989; McArthur et al., 1989). Since some preliminary data had suggested that dementia due to the neurotropic effects of HIV may precede other signs of AIDS (Grant et al., 1987), early detection was deemed critical for therapeutic interventions. That is, antiviral agents that are in the process of being developed are most likely to be successful if administered to patients in the earliest stages of the disease. Neuropsychological tests highly sensitive to the subtle cognitive changes associated with the early stages of the ADC and with various opportunistic infections of the brain (e.g., encephalitis, meningitis) would seem a most valuable instrument in facilitating this early diagnosis. Also, neuropsychological evaluations may play an important role in judging the efficacy of the forthcoming experimental treatments. Agents which are successful in destroying or inhibiting HIV might lead to improved cognitive status as measured by neuropsychological instruments.

In determining the composition and clinical application of a neuropsychological battery for the early detection of ADC, the Workshop participants arrived at six conclusions.

- (1) To ensure the overall sensitivity of the battery (i.e., to reduce the likelihood of false negatives) an extensive array of tests requiring 7-9 hours to administer should be used.
- (2) If an extended battery is not feasible in a given clinical setting, at least a brief (1-2 hours) battery comprised of selected, highly sensitive tests should be utilized. The Workshop participants recommended that this brief battery might serve as part of a national register to which all of the NIMH AIDS Research and Clinical Centers might contribute a specified number of cases on a yearly basis.
- (3) Both well-known, standardized tests and relatively new experimental paradigms borrowed from cognitive psychology should be represented in the battery. The need for these newer, albeit unstandardized, cognitive tasks emanates from their greater complexity and potentially their increased sensitivity to subtle cognitive changes. However, the Workshop participants cautioned that the choice

of these experimental tests should be guided by some preliminary evidence of their diagnostic value with either individuals infected with HIV or with other neurologic disorders characterized by progressive dementia.

(4) The selected neuropsychological battery should emphasize the assessment of those functions most likely to be compromised. Since both the current neuropathological and neuroradiological literature strongly suggest that HIV has its initial and perhaps most deleterious effects on subcortical white matter and nuclei, tests focusing upon the attentional and speed of processing deficits shown to be associated with various "subcortical" dementias (e.g., Huntington's disease, progressive supranuclear palsy) should receive special emphasis in any neuropsychological assessment.

(5) The Workgroup felt that the test performances of patients (i.e., HIV-positive, asymptomatic individuals) meeting the criteria for the Centers for Disease Control (CDC) Stages II and III are likely to be highly variable. Therefore the recommendation was made that clinical ratings should be used in studies to evaluate the degree of impairment of each individual's cognitive profile; otherwise this variability may be obscured. If only 25-40 % of Stage II and III patients have impairments in specific areas of functioning, group means and comparisons with HIV-negative controls might mask the true incidence of neurologic involvement.

(6) Because the incidence of reactive depression and anxiety among HIV positive individuals is likely high, some scales should be included as part of the total neuropsychological battery. Such an assessment should address not only quantitative measures of anxiety and depression, and their effects on neuropsychological tests, but also the identification of distinct DSM-III syndromes.

Given these six considerations, the Workshop participants recommended the extended and brief batteries of neuropsychological and psychiatric tests described in the following sections of this report. Although the proposed batteries assess a wide range of psychological functions, special emphasis has been placed on divided and sustained attention as well as on speed of processing and retrieval from working and long-term memory. Descriptions of those tests that are well-known to clinicians have been kept brief, whereas more detail has been provided for those tasks that have their origins and primary employment in cognitive psychology.

RECOMMENDED EXTENDED (7-9 HOURS) NEUROPSYCHOLOGICAL TEST BATTERY

A. Indicators of premorbid intelligence

Vocabulary (Wechsler, 1981). The *Vocabulary* subtest of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) correlates highly with verbal IQ and is considered a measure of "crystallized" intelligence and an indicator of premorbid intellectual abilities. The subjects are required to define 35 words, the frequency

of which range from very common to moderately uncommon. The subject's definitions are scored according to the criteria described by Wechsler (1981). Administration time: 20 min.

Table. 1. Domains of the NIMH core neuropsychological battery

-
- A. Indication of Premorbid Intelligence
 - 1. *Vocabulary (WAIS-R)*
 - 2. National Adult Reading Test (NART)
 - B. Attention
 - 1. Digit Span (WMS-R)
 - 2. *Visual Span (WMS-R)*
 - C. Speed of Processing
 - 1. Sternberg Search Task
 - 2. Simple and Choice Reaction Times
 - 3. *Paced Auditory Serial Addition Test (PASAT)*
 - D. Memory
 - 1. *California Verbal Learning Test (CVLT)*
 - 2. Working Memory Test
 - 3. Modified Visual Reproduction Test (WMS)
 - E. Abstraction
 - 1. Category Test
 - 2. Trails Making Test, Parts A and B
 - F. Language
 - 1. Boston Naming Test
 - 2. Letter and Category Fluency Test
 - G. Visuospatial
 - 1. Embedded Figures Test
 - 2. Money's Standardized Road-Map Test of Direction Sense
 - 3. Digit Symbol Substitution
 - H. Construction Abilities
 - 1. Block Design Test
 - 2. Tactual Performance Test
 - I. Motor Abilities
 - 1. Grooved Pegboard
 - 2. Finger Tapping Test
 - 3. Grip Strength
 - J. Psychiatric Assessment
 - 1. Diagnostic Interview Schedule (DIS)
 - 2. *Hamilton Depression Scale*
 - 3. *State-Trait Anxiety Scale*
 - 4. Mini-Mental State Examination
-

Italic indicates instruments in abbreviated version of the NIMH neuropsychological battery.

National Adult Reading Test (NART) (Nelson & O'Connell, 1978). The American version of this test (Grober, Sliwinski, & Buschke, 1990) provides an estimate of premorbid intelligence level by measuring an individual's previous familiarity with (i.e., ability to correctly pronounce) relatively uncommon words. The patients must read aloud 50 words which are "irregular" with respect to the common rules of pronunciation (e.g., naive, detente). Thus, the words can only be correctly read if the individual knows and recognizes them in their written form. The number of correct and incorrect pronunciations are recorded. Since oral reading is relatively preserved until the later stages of most dementias, it can be used as a valid index of premorbid intelligence. Both Nelson and O'Connell (1978) and Grober et al. (1990) have developed valid formulae for estimating an individual's premorbid verbal IQ with this measure of oral reading. Administration time: 10 min.

B. Attention

Digit Span (Wechsler, 1987). This task, one of the subtests of the Wechsler Memory Scale-Revised (WMS-R), requires the subject to repeat a sequence of single-digit numbers read aloud by the examiner. In the first condition, the subject must repeat the digits in the same order (i.e., digits forward); in the second condition, the digits must be repeated in the reverse order (i.e., digits backward). In both the forward and backward conditions, the lengths of the sequences increase progressively from as few as three digits to a maximum of nine (two to eight digits in the backward condition). Two trials are presented with each sequence length. Administration time: 10-15 min.

Visual Span (Wechsler, 1987). On this nonverbal span task from the WMS-R, subjects watch the examiner touch a series of colored squares distributed on a white card and then immediately attempt to repeat the sequence in either the same (forward) or reverse (backward) order. The forward condition is always administered before the backward condition. The lengths of the sequences increase progressively from two to eight items (two to seven items in the backward condition). Two trials are presented with each sequence length. Administration time: 10-15 min.

Visual Search Test (VST) (Rennick, 1979). This visuoperceptual task requires the subject to find an exact match for a relatively simple stimulus located in the middle of a 41-stimulus matrix. On each trial, the subject is shown a matrix of 41 white blocks printed in a checkerboard pattern (alternating black and white blocks) on a 12 X 8 inch page. The subject is asked to look at the white block in the middle of the matrix (i.e., the *standard stimulus*) and to note that it contains two little black squares. It is then pointed out to the subject that each of the remaining 40 white blocks (*comparison stimuli*) surrounding the middle block also contains two little black squares. The subject must then identify the one (of the 40) that has little black squares in the identical spatial locations as the middle (i.e., *standard*) block.

After two practice items have been successfully completed, eight test trials

are administered. On each trial the subject is presented with a different matrix of black and white blocks and asked to locate among the 40 comparison stimuli the one which is identical to the standard white block in the middle of the matrix. The location of the little black squares within the white standard and the 40 comparison blocks is different on every trial.

The examiner records the time required by the subject to locate the exact match. If a subject makes an error (i.e., chooses the wrong white block), the subject is told that the selection is not an exact match and to continue with the search. The subject continues searching for the matching comparison stimulus until he commits three errors or fails to locate the correct stimulus within 90 s. The subject's scores on this visual search task are the total number of errors committed on the eight trials and the total number of seconds required to complete the eight matches. Administration time: 10-15 min.

C. Speed of processing

Sternberg Search Task (Sternberg, 1966). The Sternberg test examines the speed with which subjects can search information being held in short-term or working memory.

The subject is shown a small number of digits called a memory set, ranging in size from one to four digits which are displayed singly at a fixed locus for 1.2 s each. The memory set is followed by a signal (XXX) for 1.2 s indicating that the memory set is complete, and then followed immediately by a probe digit which is terminated by the subject's response or is displayed a maximum of 5 s. The task for the subject is to decide whether or not the probe digit *was* or *was not* in the memory set by pressing a key labeled "YES" or a key labeled "NO." A warning signal "READY" is presented for 1.2 s before each memory set.

This task has 96 trials. For 48 trials, the probe digit is one of the digits in the memory set (Positive Set); on the remaining 48 trials it is not (Negative Set). For both the 48 positive trials and 48 negative trials, 12 trials are memory set size one, 12 trials are memory set size two, 12 trials are memory set size three, and 12 trials are memory set size four. For the positive trials, the probe appears an equal number of times for each serial position in the memory set (e.g., for memory set size three, the probe appears four times in the first position, four times in the second position, and four times in the third position). The ordering of the 96 trials is randomized for each subject.

Reaction time in ms is measured from the onset of the probe digit to the key press measuring the response. The time to make the decision that the probe *is* or *is not* in the memory set is typically a linear function of the number of items in the memory set. The slope of the function is viewed as a measure of the speed of accessing one item in memory.

The critical quantitative measures are reaction times for positive and negative trials as well as the number of errors. Patients having difficulty initiating a systematic search of information in short-term (working) memory would be ex-

pected to have longer reaction times, especially for probes from larger memory sets (i.e., three or four digits). Administration time: 30 min.

Simple and Choice Reaction Times (Martin, Robertson, & Edelstein, 1989; Dubois, Pillon, Legault, Agid, & Lhermitte, 1988). Comparisons of simple and choice reaction times (RTs) have often been used as an indicator of simple speed of responding and of decision time. For simple RTs, subjects are told to press a button as soon as a specified stimulus (e.g., red light) appears on a screen before them. The subject receives a warning stimulus (e.g., brief tone) 1-3 s prior to the onset of the target stimulus. Simple RT is the amount of time (measured in ms) between the onset of the target stimulus and the subject's response. Subjects are given three practice (warm-up) trials, followed immediately by 12 experimental trials. The length of the temporal interval between the warning and target stimuli is randomized over trials.

For choice RTs, subjects are instructed to respond (i.e., *go condition*) to one target stimulus (e.g., red light) and *not* to respond (i.e., *no go condition*) to another target stimulus (e.g., blue light). Again, a warning stimulus (e.g., brief tone) precedes the presentation of the target stimuli by 1-3 s. Two blocks of 24 trials each are administered to each subject. On 12 of the trials of each block, the go-stimulus is presented; on the other 12 trials, the no-go-stimulus appears on the screen before the subject. The only difference between the two blocks of trials is a switch in the value of the stimuli. For example, a red light which is the go-stimulus during the first block becomes the no-go-stimulus during the second block, while a blue light changes from the no-go to the go-stimulus between the two blocks of trials. The interval between the onset of the warning stimulus and one of the two target stimuli is randomized over the 48 test trials. The critical measures from the choice RT procedure are *total correct hits*, *median RT* (interval between onset of go-stimulus and the subject's response) *for correct hits*, and *total false positives* (responses to a no-go-stimulus). The difference between choice and simple RTs is also a critical measure because it represents a measure of decision time (i.e., the amount of time needed to determine whether the presented light is a go- or no-go-stimulus). Administration time: 20 min.

Paced Auditory Serial Addition Test (PASAT) (Gronwall, 1977). This is a test of attention and speed of information processing. In this test, four sets of randomized digits between 1 and 9 are serially presented via tape recording. Fifty numbers are presented in each set at a constant presentation rate. Subjects are asked to add the current number to the number which preceded it and to respond with the total. Thus, after each new digit is presented, a new total is achieved. The test becomes more difficult with each set as the rate at which the digits are presented increases. Administration time: 15 min.

D. Memory

California Verbal Learning Test (CVLT) (Delis, Kramer, Kaplan, & Ober, 1987). The CVLT is a list learning task which assesses multiple cognitive parameters associated with learning and memory, and thereby provides an evaluation of the

learning process as well as a measure of how much information is acquired and retained. Measures are provided for the total number of words recalled (or recognized) on each trial, serial position effects, learning rates across trials, vulnerability to proactive and retroactive interference, semantic and serial learning strategies (e.g., clustering) and perseverations and intrusions in recall.

On each of five trials, 16 words are presented orally at the rate of one word per second and immediate free recall of the words is elicited. The 16 words consist of four words from each of four semantic categories (e.g., tools, herbs and spices, fruits and clothing). Immediately following the fifth presentation/recall trial with the original list, a single presentation/recall trial is presented using a new list of 16 words. The subject is then asked to again recall the original 16 words, and then to recall the same words when provided with semantic cues (the four semantic categories represented in the word list). Following a 20-min delay filled with unrelated neuropsychological testing, free recall and cued recall of the original 16-word list is again elicited. Finally, a yes/no recognition test is administered consisting of the original 16 words and 28 randomly interspersed distractor words. Administration time: 25 min.

Working Memory Test (Baddeley, Logie, Bressi, Della Sala, & Spinnler, 1986). The working memory task that has been selected involves both a *primary task* and a *secondary task*. The primary task is a pursuit tracking task that involves computerized presentation of a moving 2X2 cm white square on a color monitor. The subject's task is to use a light sensitive pen to follow the movement of the square. When the pen moves off the square, an error is indicated by a change in the color of the square. For each subject, a pretesting phase is completed initially in which the square moves slowly around the screen, and the speed is increased until the subject stays on target approximately 60% of the time. The speed of movement should remain constant for about 20 s before it is increased. After establishing a speed at which the subject is on target for 60% of the time, a brief rest period should be allowed, and then a further series of trials is administered to ensure that performance is stable within a range of 40%-60% time on target.

The secondary task is a digit span task, with the exact length of the span determined individually for each subject in a pretest session. Starting with presentation of a single digit, span length should be increased until subject is unable to recall correctly two of three sequences of a given length. The previous list length (i.e., the last length at which they recall three lists correctly) is taken as that subject's digit span.

After both pretests are completed, the subject should perform the tracking task for 2 separate periods of 2 min each. During one 2-minute span, the subject should perform the tracking task alone (single-task condition). During the other, the subject should perform it while at the same time being required to perform the digit span task (dual-task condition). The subject should be instructed to do his best to stay on target while repeating the digit sequences being read to him. The exact number of spans administered during the 2-minute dual task condition will vary from subject-to-subject depending on the length of the span (i.e., in the

pected to have longer reaction times, especially for probes from larger memory sets (i.e., three or four digits). Administration time: 30 min.

Simple and Choice Reaction Times (Martin, Robertson, & Edelstein, 1989; Dubois, Pillon, Legault, Agid, & Lhermitte, 1988). Comparisons of simple and choice reaction times (RTs) have often been used as an indicator of simple speed of responding and of decision time. For simple RTs, subjects are told to press a button as soon as a specified stimulus (e.g., red light) appears on a screen before them. The subject receives a warning stimulus (e.g., brief tone) 1-3 s prior to the onset of the target stimulus. Simple RT is the amount of time (measured in ms) between the onset of the target stimulus and the subject's response. Subjects are given three practice (warm-up) trials, followed immediately by 12 experimental trials. The length of the temporal interval between the warning and target stimuli is randomized over trials.

For choice RTs, subjects are instructed to respond (i.e., *go condition*) to one target stimulus (e.g., red light) and *not* to respond (i.e., *no go condition*) to another target stimulus (e.g., blue light). Again, a warning stimulus (e.g., brief tone) precedes the presentation of the target stimuli by 1-3 s. Two blocks of 24 trials each are administered to each subject. On 12 of the trials of each block, the go-stimulus is presented; on the other 12 trials, the no-go-stimulus appears on the screen before the subject. The only difference between the two blocks of trials is a switch in the value of the stimuli. For example, a red light which is the go-stimulus during the first block becomes the no-go-stimulus during the second block, while a blue light changes from the no-go to the go-stimulus between the two blocks of trials. The interval between the onset of the warning stimulus and one of the two target stimuli is randomized over the 48 test trials. The critical measures from the choice RT procedure are *total correct hits*, *median RT* (interval between onset of go-stimulus and the subject's response) *for correct hits*, and *total false positives* (responses to a no-go-stimulus). The difference between choice and simple RTs is also a critical measure because it represents a measure of decision time (i.e., the amount of time needed to determine whether the presented light is a go- or no-go-stimulus). Administration time: 20 min.

Paced Auditory Serial Addition Test (PASAT) (Gronwall, 1977). This is a test of attention and speed of information processing. In this test, four sets of randomized digits between 1 and 9 are serially presented via tape recording. Fifty numbers are presented in each set at a constant presentation rate. Subjects are asked to add the current number to the number which preceded it and to respond with the total. Thus, after each new digit is presented, a new total is achieved. The test becomes more difficult with each set as the rate at which the digits are presented increases. Administration time: 15 min.

D. Memory

California Verbal Learning Test (CVLT) (Delis, Kramer, Kaplan, & Ober, 1987). The CVLT is a list learning task which assesses multiple cognitive parameters associated with learning and memory, and thereby provides an evaluation of the

Baddeley et al. [1986] study it ranged from 11 to 15 sequences). Half the subjects participate in the single-task condition first and dual-task condition second; the other half are given the tasks in the reverse order. The critical measure on this working memory test is the accuracy of the subjects' recall of the digit sequences during the dual-task condition. It is anticipated that subjects who encounter difficulty with divided attention will repeat fewer digit sequences under the dual-task condition. It is also possible that subjects with such attentional difficulties will exhibit reduced accuracy on the tracking task while they are attempting to recall digit sequences. Administration time: 30-40 min.

Modified Visual Reproduction Test (VRT) (Wechsler, 1945; Russell, 1975). This modified version of the VRT from the original Wechsler Memory Scale provides measures of immediate and delayed retention of geometric forms. On each of three trials, the subject must reproduce a complex geometric figure from memory immediately following a 10-s study period. Three increasingly complex stimuli containing from 4-10 components are presented on successive trials. As a measure of long-term retention, the subject is asked after 30 min of unrelated testing to again reproduce the figures from memory. Finally, the subject is asked to simply copy the stimulus figures in order to assess any visuoperceptual dysfunction that may be contaminating visual memory performance. The subject's reproductions are scored for the number of components correctly represented from the original stimulus drawings. Administration time: 10 min.

E. Abstraction

Category Test (Halstead, 1947; Reitan & Davison, 1974). This is a measure of complex reasoning and conceptual skills. This task involves a sequential presentation of 208 stimulus slides on a projector screen. Each stimulus can be associated with a number between 1 and 4, and the subjects are asked to indicate (by pushing on a numbered button) which number each target suggests. Immediate and automatic feedback is given by the test apparatus as to whether each answer is right or wrong. The test is divided into seven subtests, and in all but the last of these a single principle or concept is in force throughout. The seventh (memory) subtest contains slides from the previous six subtests. Administration time: 30-40 min.

Trails A and B (Reitan & Davison, 1974). This is a measure of psychomotor speed, attention and cognitive sequencing. In Part A of this test, subjects are asked to quickly connect, in ascending order, a series of randomly arranged circles each of which contains a different number from 1 to 25. Part B requires subjects to quickly connect, in ascending order, a series of randomly arranged circles, alternating between sequential numbers and letters (from 1-13, A-L). The task requires the same abilities as in Part A, but additionally requires the shifting of cognitive sets (numbers to letters, etc.). Administration time: 10 min.

F. Language

Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983): An abbreviated

version of this test requires the subject to name 30 outline drawings of objects (either the odd- or even-numbered items from the full 60-item Boston Naming Test). The drawings are graded in difficulty, with the easiest drawings presented first. If a subject encounters difficulty in naming an object, a stimulus or phonemic cue is provided. The number of spontaneous and cued correct responses, perceptual errors, circumlocutions, paraphasias, and perseverations are used to evaluate the subject's performance. Administration time: 15 min.

Letter and Category Fluency Tests (Borkowski, Benton, & Spreen, 1967; Butters, Granholm, Salmon, Grant, & Wolfe, 1987). On the Letter Fluency test the subject is asked to generate orally as many words as possible that begin with the letters "F," "A," and "S," excluding proper names and different forms of the same word. For each letter, the subject is allowed 1 min to generate words. Performance is measured by the total number of correct words produced for the three letters. Perseverations (i.e., repetitions of a correct word) and intrusions (i.e., words not beginning with the designated letter) are also recorded.

On the Category Fluency test the subject is asked to generate orally as many different kinds of animals, fruits, and vegetables as possible within a given time limit. For each category, the subject is allowed 1 min to generate items. The subject's score is the total number of items correctly named in each category during the 1-min time period. Perseveration and intrusion errors are also noted.

Although fluency tests are sensitive to language dysfunctions, they also reflect an individual's capacity to retrieve information from semantic memory. If other language abilities, such as confrontation naming and comprehension are intact, impaired fluency may indicate an inability to initiate systematic retrieval of information in semantic memory (Butters et al., 1987). Administration time: 10 min.

G. Visuoperception

Embedded Figures Test (Witkin, Ohman, Raskin, & Karp, 1971). This visuoperceptual task requires the subject to locate a simple figure within the contours of a complex figure. On each of 24 trials, the subject is shown a simple geometric figure for 15 s and then asked to locate the simple stimulus within the boundaries of a complex colored geometric form. The examiner records the time required for the subject to find the embedded simple figure. If the subject makes an error, he is told of his mistake and asked to continue looking for the simple figure. On each trial, testing is continued until the subject locates the simple figure embedded within the complex pattern or fails to find the simple stimulus within 180 s. The subject's score on the Embedded Figures Test is the *average* time to find a hidden simple figure. Administration time: 20-30 min.

Money's Standardized Road-Map Test of Direction Sense (Money, Alexander, & Walker, 1965; Money, 1976). The subject is shown a street map of a small town. On this map are drawn two routes taken by two hypothetical travelers when they visited the town. One traveler encountered four street intersections and made four turns; the second encountered 32 intersections and made 32 turns.

The subject is instructed to assume that he is taking the same routes as the travelers and is asked to indicate whether he will be taking a right or left turn at each intersection along the routes. The street map always remains in a fixed position in front of the subject, and the subject is not allowed to turn his body or to alter his position to facilitate right-left distinctions. The examiner traces each route with his fingers and at each intersection asks the subject, "Now, at this corner would you be turning right or left?" For both routes, half the turns are rights, with the sequence of turns randomized.

The short route is administered first and serves as a practice item to ensure that the subject understands the nature of the task. The long route (32 turns) is presented immediately after completion of the short route. The number of correctly identified turns and the amount of time needed to complete the routes are recorded for each subject.

To perform well on this task, the subject has to imagine himself traveling along the specified route and to spatially rotate himself (in imagery) to ascertain whether a right or left turn is demanded at various intersections. Since these spatial rotations and distinctions (right vs left) are made with reference to the subject's own body, this Road-Map task is considered a test of personal (i.e., egocentric) space. Administration time: 15 min.

Digit Symbol Substitution (Wechsler, 1981). This visuoperceptual and motor task from the WAIS-R requires the subject to associate single digit numbers with unfamiliar symbols. A stimulus set of nine printed digit-symbol pairs (i.e., the digit-symbol code) is presented above rows of numbers printed without the appropriate symbols. The subject is instructed to draw the correct symbol below each of the numbers using the digit-symbol code presented above. After four practice items, the subject completes as many substitutions as possible in 90 seconds. Administration time: 4 min.

H. Constructional abilities

Block Design Test (Wechsler, 1981). On this constructional task from the WAIS-R the subject is presented with four or nine red and white blocks and asked to construct replicas of nine designs. The blocks are red on two sides, white on two sides, and half white-half red on two sides. The four-block designs have 1-minute time limits; the nine-block designs 2-minute limits. The subject's score depends both upon accuracy and speed. For the first block design, the subject copies the examiner's block construction; for the remaining eight designs, he or she copies two-dimensional pictures of the designs. Administration time: 15-20 min.

Tactual Performance Test (Halstead, 1947; Reitan & Davison, 1974). This is a complex measure of sensory-motor integration, psychomotor speed and nonverbal problem solving. The task requires subjects to place 10 blocks, one at a time, in their respective cut out spaces on a board that faces them. Since the subjects are blindfolded during the entire task, they must rely on tactile and proprioceptive feedback. There are three trials: one with the dominant hand only, a second in

which only the nondominant hand is used, and a third trial in which both hands are used. The total time required to place all blocks is recorded for each trial. A maximum of 10 min per trial is allowed, regardless of whether all blocks have been placed or not. After the board and blocks have been removed from the subjects' view, they are asked to draw a picture of the board they were just working with. They are asked to draw as many blocks as they can remember and place them in the correct location relative to each other on the board. Three scores are derived from the TPT: total time, memory (number of shapes recalled), and location (number of shapes recalled in the correct locations). Administration time: 20-30 min.

I. Motor abilities

Grooved Pegboard (Kløve, 1963). This is a test of fine motor coordination and speed. In this test, subjects are required to place 25 small metal pegs into 25 holes on a 3" X 3" metal board. All the pegs are alike and have a ridge along one side which corresponds to a randomly positioned slot in each hole on the board. Each peg must be rotated to match the slot on the hole before it can be inserted. The subjects are asked to place the pegs in the holes as fast as they can, first with their dominant hand, then with their nondominant hand. The total time for each hand is recorded as the score. Administration time: 15-20 min.

Finger Tapping Test (Halstead, 1947; Reitan & Davison, 1974). This is a measure of a simple motor speed with the upper extremities. Subjects are required to tap on a key counter using the index finger of their dominant and their nondominant hands for 10 s. This is repeated until five consecutive trials are obtained that are within five taps of each other, or until a total of 14 trials is administered. The final score is the average of the five trials or the average, after discarding the two highest and two lowest scores, of 10 trials. The examiner must make allowances for frequent breaks between trials. Administration time: 15 min.

Grip Strength (Reitan & Davison, 1974). A hand dynamometer is used first with dominant and then nondominant hands to assess bilateral grip strength. Two trials are administered to each hand, alternately, unless the readings (in kilograms) for that hand are not within 5 kg of each other. In the latter case, two additional trials are administered to each hand after a short break. The mean grip strength for dominant and nondominant hands is recorded as the score. Administration time: 5 min.

J. Psychiatric assessment

Diagnostic Interview Schedule (DIS) (Robins, Helzer, Croughan, & Ratcliff, 1981). This interview was designed for use in a set of large-scale epidemiological studies. The DIS can make a diagnosis using three systems (DSM-III, Feighner Criteria, and Research Diagnostic Criteria) based on a computer algorithm. This is a very structured interview designed to be used by lay interviewers who do not compose questions or make diagnostic judgments. A probe Flow Chart provides for

nondirective follow-up to positive answers. The diagnosis is made on a lifetime basis first and, then, questions are asked to determine how recently the symptoms have been experienced so that a current diagnosis can be made. The latter can be defined for four time periods: (1) the last 2 weeks, (2) the last month, (3) the last 6 months, and (4) the last year. In addition to diagnostic results, the DIS provides a total symptom count across diagnoses for each of the three systems and a count of the number of criteria met for each diagnosis. Administration time: 45-75 min.

Hamilton Depression Scale (HDS) (Hamilton, 1960; Williams, 1988). This scale was devised to quantify the results of an interview with a patient already diagnosed as suffering from an affective disorder. It has, however, emerged as a widely used scale for patient selection and follow-up in research studies for treatment of depression. This scale provides a simple way of assessing the severity of depression and showing changes in that condition. The content includes cognitive, behavioral and psychological symptoms associated with depression and, in addition, some less usual symptoms (e.g., paranoid and obsessional thinking) are included. A guide is provided for rating each variable, and some special directions are provided for rating women. Overall reliability of this scale has only been fair. Williams (1988) has concluded that this lack of reliability is due to the variability in the way the information is *obtained* and has prepared a structured interview guide to standardize the administration of the scale. Administration time: 15-20 min.

State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, & Lushene, 1970). The State-Trait Anxiety Inventory was originally developed to study the relationship between anxiety and learning and has been used in a variety of research contexts to evaluate changes in the level of anxiety as both a predictor and outcome. The inventory is composed of separate self-report psychometric scales for two distinct but related concepts of anxiety. State anxiety is conceptualized as a transitory emotional state or condition that involves consciously perceived feelings of tension and apprehension and a heightened autonomic nervous system. Trait anxiety, on the other hand, is a relatively stable personality feature which reflects differences in individuals' tendencies to respond anxiously to situations perceived as threatening. Each scale contains 20 items which indicate the presence or absence of a specific anxiety symptom. Scores can range from 20 to 80 on each scale. Administration time: 5-10 min.

Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975). This is an instrument designed to evaluate quickly the cognitive component of a standard psychiatric mental status exam. It is divided into two sections: the first half requires oral responses and assesses orientation, short-term and immediate memory, and attention. The second half requires both oral and written responses and evaluates both receptive and expressive functions. There is no measure of abstraction. The entire exam is comprised of 30 questions and tasks. The maximum score is 30. Administration time: 5-10 min.

RECOMMENDED BRIEF (1-2 HOURS) NEUROPSYCHOLOGICAL TEST BATTERY

If the full neuropsychological battery cannot be utilized in a particular clinical setting, a concerted effort should be made to administer at least the following tests: Vocabulary (WAIS-R), CVLT, PASAT, VST, HDS, and STAI. The Workshop group urged that these indices be incorporated into any emerging AIDS national register. In view of the brevity of this battery all AIDS research and clinical centers should not be unduly burdened by the collection and transmission of this neuropsychological data to a central data bank. It seemed likely that the large number of entries collected by such a national registry would greatly facilitate the identification of any neuropsychological features associated with the various stages of HIV infection.

THE UTILITY OF CLINICAL RATINGS

While direct statistical comparisons of group means has proven of great value in studies of patients with marked brain damage, this procedure has substantial limitations when trying to detect subtle, but real, deficits on a neuropsychological test battery. If the capacities affected are inconsistent across patients within a group (e.g., Stage II and III HIV-positive patients), the changes may not be apparent with the group's mean score. That is, on any particular test the "normal" performances of the unaffected individuals may tend to mask the impaired scores of the affected individuals.

To increase the likelihood that neuropsychological tests will identify neurologically affected Stage II and III HIV-positive individuals, the Workshop group recommended that each patient's scored protocol be rated by two trained neuropsychologists using the same clinical criteria. These ratings should involve an estimate of global cognitive functioning as well as judgements of specific areas of neuropsychological ability (e.g., language, memory, visuoperception). It is important to note that previous research has demonstrated the reliability and validity of clinical ratings in neuropsychology, both for neurodiagnostic purposes and for the detection of subtle changes in intellectual functioning (Grant, Heaton, McSweeney, Adams, & Timms, 1982; Heaton, Grant, Anthony, & Lehman, 1981; Heaton, McSweeney, Grant, Adams, & Petty, 1983). Thus, while investigators should continue to examine group means, they should also employ clinical ratings to insure an accurate estimate of the incidence of neuropsychological impairment in the population of HIV-positive individuals. These clinical ratings are obviously more applicable to the well-standardized than to the experimental tests recommended in this report. Due to a lack of age- and education-corrected normative data, it may not be possible at this time to specify the exact clinical criteria for judging the presence and severity of deficits on many of the experimental tasks. However, this limitation should not deter either the development of these experi-

mental tasks for neurodiagnostic purposes or the statistical comparisons of overall group performances on these tests.

REFERENCES

- Baddeley, A., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology*, 38A, 603-618.
- Borkowski, J.G., Benton, A.L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia*, 5, 135-140.
- Butters, N., Granholm, E., Salmon, D.P., Grant, I., & Wolfe, J. (1987). Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology*, 9, 479-497.
- Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *The California Verbal Learning Test*. New York: The Psychological Corporation.
- Dubois, B., Pillon, B., Legault, F., Agid, Y., & Lhermitte, F. (1988). Slowing of cognitive processing in progressive supranuclear palsy: A comparison with Parkinson's Disease. *Archives of Neurology*, 45, 1194-1199.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Grant, I., Atkinson, J.H., Hesselink, J.R., Kennedy, C.J., Richman, D.D., Spector, S.A., & McCutchan, J.A. (1987). Evidence for early central nervous system involvement in the acquired immunodeficiency syndrome (AIDS) and other human immunodeficiency virus (HIV) infections. *Annals of Internal Medicine*, 107, 828-836.
- Grant, I., Heaton, R.K., McSweeney, A.J., Adams, K.M., & Timms, R.M. (1982). Neuropsychologic findings in hypoxic chronic obstructive pulmonary disease. *Archives of Internal Medicine*, 142, 1470-1476.
- Grober, E., Sliwinski, M., & Buschke, H. (1990). Premorbid intelligence in the elderly. (Abstract). *Journal of Clinical and Experimental Neuropsychology*, 12, 30.
- Gronwall, D.M.A. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367-373.
- Halstead, W.C. (1947). *Brain and intelligence*. Chicago: University of Chicago Press.
- Hamilton, M.A. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, 23, 56-62.
- Heaton, R.K., Grant, I., Anthony, W.Z., & Lehman, R.A.W. (1981). A comparison of clinical and automated interpretation of the Halstead-Reitan Battery. *Journal of Clinical Neuropsychology*, 3, 121-141.
- Heaton, R.K., Grant, I., McSweeney, A.J., Adams, K.M., & Petty, T. L. (1983). Psychologic effects of continuous and nocturnal oxygen therapy in hypoxic chronic obstructive pulmonary disease. *Archives of Internal Medicine*, 143, 1941-1947.
- Janssen, R.S., Saykin, A.J., Cannon, L., Campbell, J., Pinsky, P.F., Hessol, N.A., O'Malley, P.M., Lifson, A.R., Doll, L.S., Rutherford, G.W., & Kaplan, J.E. (1989). Neurological and neuropsychological manifestations of HIV-1 infection: Association with AIDS-related complex but not asymptomatic HIV-1 infection. *Annals of Neurology*, 26, 592-600.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983) *Boston Naming Test*. Philadelphia: Lea & Febiger.
- Kløve, H. (1963). Clinical neuropsychology. *Medical Clinics of North America*, 46, 1647-1658.
- Martin, E.M., Robertson, L.C., & Edelstein, H.E. (1989). Decision-making speed is impaired in early-stage HIV infection. (Abstract). *Journal of Clinical and Experimental Neuropsychology*, 11, 78.

- McArthur, J.C., Cohen, B.A., Selnes, O.A., Kumar, A.J., Cooper, K., McArthur, J.H., Soucy, G., Comblath, D.R., Chmiel, J.S., Wang, M., Starkey, D.L., Ginzburg, H., Ostrow, D.G., Johnson, R.T., Phair, J.P., Polk B.F. (1989). Low prevalence of neurological and neuropsychological abnormalities in otherwise healthy HIV-1-infected individuals: Results from the Multicenter AIDS Cohort Study. *Annals of Neurology*, 26, 601-611.
- Money, J. (1976). *A Standardized Road-Map Test of Directional Sense*. San Rafael, CA: Academic Therapy Publications.
- Money, J., Alexander, D., & Walker, H.T. (1965). *A Standardized Road-Map Test of Direction Sense*. Baltimore: Johns Hopkins Press.
- Nelson, H.E., & O'Connell, A. (1978). Dementia: The estimation of premorbid intelligence levels using the new adult reading test. *Cortex*, 14, 234-244.
- Reitan, R.M., & Davison, L.A. (1974). *Clinical neuropsychology: Current status and applications*. New York: Winston/Wiley.
- Rennick, P.M. (1979). *Color-naming and visual search tests for repeatable cognitive-perceptual motor battery*. Grosse Point Park, MI: Axon Publishing Company.
- Robins, L.N., Helzer, J.E., Croughan, J., & Ratcliff, K.S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry*, 38, 381-389.
- Russell, E.W. (1975). A multiple scoring method for the assessment of complex memory functions. *Journal of Consulting and Clinical Psychology*, 43, 800-809.
- Spielberger, C.D., Gorsuch, R.C., & Lushene, R.E. (1970). *Manual for the State-Trait Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153, 652-654.
- Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology*, 19, 87-95.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised Manual*. New York: The Psychological Corporation.
- Wechsler, D. (1987). *Wechsler Memory Scale-Revised*. New York: Psychological Corporation.
- Williams, J.B.W. (1988). A structured interview guide for the Hamilton Depression rating scale. *Archives of General Psychiatry*, 45, 742-746.
- Witkin, H., Ohman, P., Raskin, E., & Karp, S. (1971). *Embedded Figures Test*. Palo Alto, CA: Consulting Psychologists.