Neurotoxicology and Teratology, Vol. 18, No. 4, pp. 365–370, 1996

0892-0362/96 $15.00 + .00

ELSEVIER

PII S0892-0362(96)00041-4

Development of a Computer-Based Battery Designed to Screen Adults for Neuropsychological Impairment

RICHARD LETZ,* ROBERT C. GREEN† AND JOHN L. WOODARD†

**Department of Behavioral Sciences and Health Education, Rollins School of Public Health, Emory University, Atlanta, Georgia 30322, USA*
†Neurobehavioral Program, Department of Neurology, Emory University School of Medicine, Atlanta, Georgia 30322, USA

LETZ, R., R. C. GREEN AND J. L. WOODARD. *Development of a computer-based battery designed to screen adults for neuropsychological impairment.* NEUROTOXICOL TERATOL **18**(4)365–370, 1996.—The initial step in the development of a new computer-based neuropsychological testing system is described. A new hardware–software system and a screening battery consisting of an orientation task and five cognitive tests has been implemented. This novel screening battery is designed for use in identifying individuals among exposed groups who may require more extensive follow-up neuropsychological evaluation. Sophisticated, but currently available, and relatively inexpensive hardware and software technology are employed, introducing an improvement over existing computer-based batteries. Use of a digitized speech production device for producing instructions facilitates testing of illiterate subjects, use of auditory stimulus materials, and relatively easy translation of the testing instructions to other languages. Use of a pen-based notebook computer enables responding by the subject in a manner that is both natural and analogous to that used in existing paper-and-pencil testing instruments. Five neuropsychological tests were developed to cover a range of cognitive domains: 1) verbal list learning and recognition memory, 2) visual memory span, 3) conceptual and motor tracking, 4) psychomotor speed and accuracy, and 5) delayed verbal recognition. The screening instrument has been pilot-tested for feasibility of use among outpatients at an occupational medicine clinic and among community-dwelling older adults. This instrument is intended to provide a standardized, efficient, cost-effective method for widespread use in occupational medicine and public health to detect and track subclinical neurotoxic effects and to prevent additional harmful exposures. Additional aspects of development of the computer-based testing system are discussed.

Neurobehavioral testing   Neuropsychological tests   Computer-based testing

THE effects of exposure to neurotoxic agents are often detected by functional testing using cognitive and behavioral instruments. Because of the complexity of the nervous system and the range of functions it performs, batteries, or sets, of neuropsychological tests are commonly used (10). From the mid-1960s through 1989, more than 250 neurobehavioral tests had been used in 185 occupational epidemiologic studies reviewed by Anger (1). In epidemiologic studies neurobehavioral tests are sensitive indicators of effects of exposures on central nervous system function (1). Existing computer-based systems provide standardized, efficient data collection methods for epidemiologic studies (3,4,6,8,11,20).

The Neurobehavioral Evaluation System (NES), a computer-based testing system, was developed specifically for use in epidemiologic investigations of the potential neurotoxic effects of agents in the work or general environment (3). More than 15 tests of reaction time, hand–eye coordination, and simple cognitive skills are included within NES. The system is inexpensive, requires minimal hardware, has software that is easy to use by professionals having some personal computer experience, and has flexibility for optimizing test parameters to different exposure situations. NES has been used extensively in the occupational and environmental health field (12). Over 80 laboratories have obtained the current version of the

Requests for reprints should be addressed to Richard Letz, Ph.D., Department of Behavioral Sciences and Health Education, The Rollins School of Public Health of Emory University, 1518 Clifton Road, Room 520, Atlanta, GA 30322. Tel: (404) 727-8742; Fax: (404) 727-1369; E-mail: letz@sph.emory.edu

software, NES2 (13), and the instructions have been translated into nine foreign languages. It has been used in a large number of epidemiologic and laboratory studies, more than 6000 subjects have been tested with it, and more than 40 articles concerning its use have been published in peer-reviewed journals. NES2 has been described as the computerized system likely to be most used in the 1990s in worksite research (1,17).

Despite the success of NES2, a number of limitations have been acknowledged: 1) all of the instructions for NES2 tests are presented as text on a video screen; 2) NES2 also uses either a relatively crude IBM-PC joystick or the numeric keys on the keyboard for response input; 3) the NES2 test–administrator interface requires some knowledge of how to use the computer, and substantial computer knowledge is required to manage the NES2 data and perform data analysis. It is clear that auditory presentation of instructions and some stimuli would be more natural, more like traditional neuropsychological tests, and would allow testing of illiterate subjects. In addition, with auditory instructions, translation to other languages, including nonalphabetic ones, does not present as much of a technical challenge as it does with the printed text instructions of NES2.

For some time we have been planning a major revision of the NES software (NES3). Enhanced computer hardware and software capabilities are presently available to address some of the limitations of NES2, including use of high-quality, inexpensive digitized speech output hardware and integration of audio stimulus presentation and pen input (pointing at pictures). Currently available software development tools can also provide a consistent, robust, and efficient interface between the tests and both the test administrator and the individual taking the tests. A summary of the major differences between NES2 and NES3 is given in Table 1.

As the first step in developing new hardware and software for NES3, it seemed advisable to focus on implementing a limited set of tests that, at the same time, meet an important need not met by currently available computer-based testing systems (i.e., screening of *individuals* exposed to neurotoxicants). There is increasing awareness of potential neurotoxic effects from many other chemicals in common use and at exposure levels formerly assumed to be safe (17). This awareness of potential neurotoxic effects highlights the need for tools that can screen persons who have experienced exposures. Currently available methods are not well suited for this purpose. The standard neurologic examination requires a physician, is not sensitive to subtle (particularly cognitive) impairments, and is performed inconsistently by different physicians. Despite the sensitivity of conventional neuropsychological evaluation, not all potentially affected individuals can undergo conventional clinical neuropsychological testing for diagnostic purposes, as this requires on average 6.5 h of testing by a trained neuropsychologist at an average cost of $700 (18). Brief instruments (e.g., the Mini-Mental State Examination) are often used today for screening, but their standardization in practice, limited sensitivity and reliability among mildly impaired patients, and cultural and educational biases limit their utility [e.g., (2)].

Only one computer-based neurobehavioral testing system was designed specifically for *clinical use* in occupational neurotoxicology: the Cognitive Scanner developed by Laursen (11). However, the time required for administration is almost as long as a manually administered clinical neuropsychological examination, the norms provided are derived from a rather homogenous Danish population, and the system is expensive to purchase. There is no computer-based neurobehavioral testing system available that has been demonstrated to be valid for *screening* workers exposed to potential neurotoxic agents.

We describe in this article the hardware–software basis for NES3 and the development of a battery of computer-based tests designed specifically for screening of individuals, which is the first set of tests implemented in the new NES3 hardware–software environment. The computer-based system described was designed to assist the clinician in deciding whether individual workers exposed to potential neurotoxicants should be referred for more expensive full-scale clinical neuropsychological assessment. It is intended to provide a standardized, efficient, cost-effective method for widespread use in occupational medicine and public health to detect subclinical neurotoxic effects and to prevent additional harmful exposures.

METHOD

Apparatus

The design criteria for the data acquisition computer were that it be lightweight, portable, rugged, as inexpensive as possible, and readily available for an extended period into the future. The software will run on any PC-compatible computer capable of running Microsoft Windows 3.11 for Pen Computing and having 16-bit audio capability. However, for standard test administration, an *IBM THINKPAD 360PE notebook computer* is recommended. This computer is equipped with a 50Mhz 80486 DX microprocessor, VGA-compatible video,

TABLE 1
COMPARISON OF NES2 AND NES3

System Attribute	NES2	NES3 Screen Battery
Input	Joystick, keyboard	Stylus (pen-based)
Instructions	Text on video screen	Digitized audio
Task emphasis	Predominantly simple psychomotor speed tests	Graphical and auditory verbal stimuli emphasized
Computer's role in testing	Computer assisted	Computer assisted
Operating environment	DOS	Windows
Data output format	ASCII file	MS Access
Intended Purpose	Epidemiologic tool	Screening tool (initially, as an epidemiologic and clinical tool later)

color liquid crystal display, pen input, and built-in 16-bit audio capabilities. The audio output device (headphones or self-powered speakers) is plugged directly into an audio jack on the computer. Thus, all instructions are auditory, and all responses are made by simply touching the pen-like stylus to the screen.

Software System

The software system utilized by NES3 was developed under, and runs in, the Microsoft Windows for Pen Computing environment. Commercially available software for relatively easy programming of tests and recording graphics images and sounds are available. Specifically, *Microsoft Visual Basic* was used to program the neuropsychological tests. Files were created and maintained in standard Windows digital formats (i.e., BMP format for video and WAV format for audio), such that they can be manipulated by software available from numerous vendors.

Neurobehavioral Tests in the Screening Battery

For the computer-based battery to be used effectively for screening, brevity was considered an important criterion, and, although coverage of a wide range of functions might be theoretically desirable, it was not considered essential that the tests chosen for inclusion isolate particular nervous system functions. In fact, multifactoral tests were considered potentially the most efficient, as disruption of any of the component functions may result in observable decrements in the outcome. The five tests selected for this battery were chosen to sample the cognitive domains of memory (Word List Learning with Delayed Recognition), attention (Visual Span), conceptual motor tracking (Sequences), and psychomotor speed (Digit–Symbol). Each of these tasks is based on established neuropsychological tests with demonstrated sensitivity to global cognitive dysfunction [see (15)]. More specific measures of discrete cognitive abilities (e.g., language and visuospatial reasoning) were not selected for inclusion because of their potential insensitivity to generalized cognitive impairment. Thus, these tests were selected with the goal of maximizing the ability to detect general neuropsychological deficits.

In addition to having programs for the orientation task and five tests chosen for inclusion in the screening battery, the software includes an executive program that contains a menu to select which tests are administered. The executive program also launches each of the individual neuropsychological test programs, allows entry of covariate information such as age and education, and retrieves data from the database. The six tasks in the battery are listed in Fig. 1 as they appear on the computer screen in the executive program.

The *Orientation Task* is aimed at determining whether the subject can hear verbal instructions, can follow simple verbal requests for manual responses, and is physically capable of manipulating the stylus to respond to the tests. If the subject is confused or physically unable to respond, then it is pointless to administer additional tests. The subject is asked to make a simple response (i.e., use the stylus to touch a lighted area of the screen that contains a geometric figure). Other information describing general features of appropriate responding in later tasks (e.g., touching with the pen perpendicular to the screen) is also presented.

The *Word List learning task* determines whether the subject can perceive auditory verbal information, can discriminate simple visual objects, is capable of immediate memory, and displays learning (i.e., improves performance with repeated presentations). A list of 10 unrelated names of concrete objects is presented verbally, followed by presentation on the screen of 20 simple pictures, two at a time. Ten of the pictures represent the words presented and 10 are distracter items. The process continues for three trials. The number of correct items is recorded on each trial. There is a single trial with a second set of 10 words and 20 pictures. In the *Delayed Recognition task*, 20 pictures, half of which are from the first list and half from the second, are presented without cueing some minutes later, after the other tests described below are completed. An example of the response alternatives presented in both the acquisition and delayed recognition tasks is shown in Fig. 2.

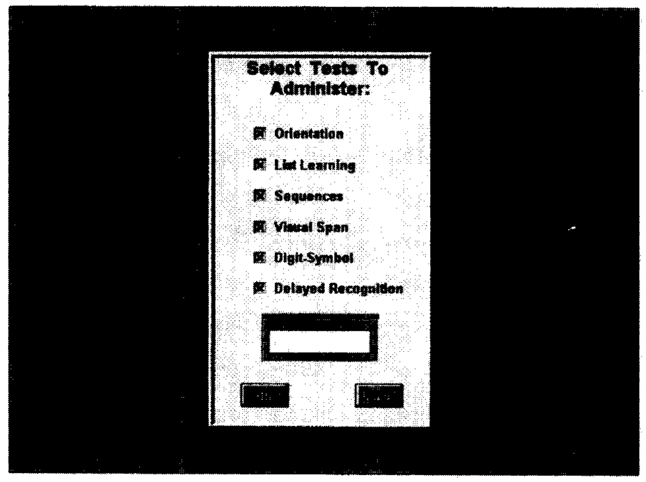

FIG. 1. Tests in the screening battery as shown on the computer screen in the executive program of NES3.

In the *Visual Span task* the subject demonstrates that (s)he can perceive a visual–spatial–temporal pattern and execute a spatial–temporal reproduction of it. The present implementation is similar to the Corsi blocks neuropsychological task [see (15)] and is similar in principle to the SimonSez task in the Microtox System (6). It is a visual–spatial analogue of the Digit Span test. Small square areas of the screen are lighted in

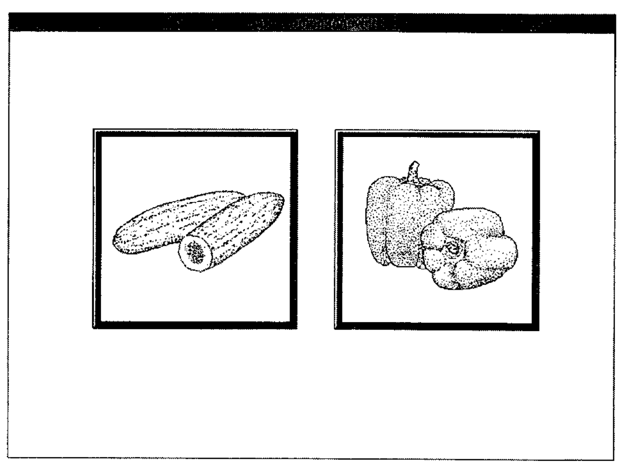

FIG. 2. Example of a response screen shown during the Word List Learning Task after a list of 10 items containing the word "cucumber" has been presented through the computer's speakers.

sequence, and when the sequence is completed, the subject must touch the same areas in the same sequence with the stylus. After a practice trial with a three-item sequence, the test begins with a four-item sequence, and the length of the "span" keeps increasing on subsequent trials, as long as the subject reproduces the subject answers correctly. If the subject answers incorrectly, a second trial is given at that span length. When two incorrect trials occur at one span length, the number of items in the sequence is decreased until a trial is answered correctly. The procedure is continued until there are five reversals of direction of the change in length of the span or successful completion of a nine-item sequence.

The *Sequences task* is similar to the paper-and-pencil Trail–Making task, which is usually administered as part of the Halstead–Reitan battery and is highly sensitive to the effects of brain damage [see (15)]. This test of visual conceptual and visuomotor tracking involves motor speed and attention functions. It is given in two parts. In the first part, 16 small squares with digits inside are presented, and the subject must touch the square with the digit "1" inside and continue by touching the other numbered squares in order as fast as (s)he can. The latency to complete the task and the number of errors are recorded. In the second part of the test, squares with both digits and letters are presented. The subject's task is to alternate between the two sequences, touching "1", "A", "2", "B," etc., up to the letter "H" as quickly as possible. Practice trials having only six squares on the screen are given for both parts. The appearance of the computer screen for the practice trial of the second part of this test is shown in Fig. 3.

The *Digit–Symbol task* involves visual scanning, mental tracking, and graphic responding. It is analogous to the Digit–Symbol task of the Wechsler Adult Intelligence Scale (WAIS) (19). The Digit–Symbol task is the WAIS subtest that is most sensitive to diffuse brain damage (15). It, and the similar NES Symbol–Digit task, have been found to be affected more often than other tests in occupational epidemiologic studies (1). In this implementation a row of digits is matched with a row of symbols at the top of the screen, and a row of the same symbols in scrambled order is presented alone at the bottom of the screen. Digits are presented one at a time in the middle of the screen, and the subject must touch the symbol in the row at the bottom of the screen that is matched with that same digit at the top of the screen. After completing practice by matching five digits with correct symbols, the latency required to complete 18 digits with symbols, as well as the number of errors committed, is recorded. The test determines the efficiency with which the subject can discriminate visual symbols and execute a simple coding task.

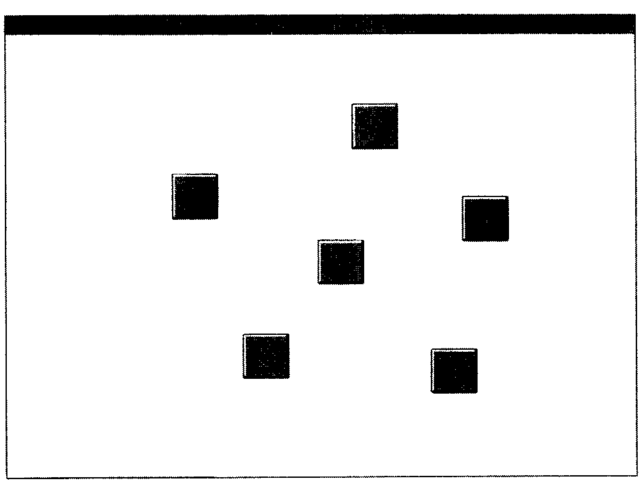

FIG. 3. Appearance of the response screen during the practice trial of Sequence B, in which the subject must alternate between touching letters and numbers in ascending order.

The raw data from each of the five tests are written directly into a common Microsoft Access data base. This procedure allows flexibility for individual users to develop tools for manipulating the data in a manner best suited for each testing situation. The data base management functions most likely to be developed include calculating summary data for a particular subject's session and reporting them to the video screen or printer, creating back-up files of the data, and transferring a set of data for input to a statistical package.

PILOT TESTING

Two separate pilot administrations of the NES3 were carried out. In the first, a total of 29 outpatients at the Emory Occupational Medicine Program at the Emory Clinic in Atlanta, GA, were recruited as subjects to evaluate the functionality of the screening tests' software and hardware. They ranged in age from 24 to 81 years; 12 were males and 17 were females; 16 were Caucasian, 12 were African American, and one was Asian; two were native Spanish speakers and the rest were native English speakers; they ranged in educational attainment from 8 to 19 years. A structured interview combining multiple choice and open-ended questions was employed to evaluate the patient's responses to the screening test procedures. The questions addressed the patient's general perceptions and experiences of the tests: for example, the smoothness of the person–machine interface, the clarity of verbal instructions, the timing of interchanges, the level of anxiety induced, and the level of performance displayed. In addition, there were open-ended questions soliciting comments about each test.

Review of the questionnaires from these subjects identified a few minor modifications to the tests, which were implemented. In general, all subjects found the testing system understandable and interesting. Two changes were made to the tests on the basis of the performance of these subjects. In these generally healthy adults, the level of difficulty of the List Learning task seemed too low, and the level of difficulty for the Sequences task seemed too high. The order of presentation of the word list, which had been constant across trials, was changed to vary from trial to trial. In the Sequences task the number of stimulus squares on the screen was reduced from 20 to 16, which gave the computer screen a less cluttered look and removed potential confusion between the number "1" and the letter "I" in the second part of that task.

In the second pilot study a sample of convenience of 38 older, community-dwelling men was recruited. The mean age of these men was 71.4 years (SD = 6.3), and all were Caucasian. Although none of the subjects carried diagnoses of dementia, administration of standard neuropsychological screening and staging instruments such as the Mattis Dementia Rating Scale (DRS) (16) and the California Verbal Learning Test (5) revealed that this group of subjects exhibited considerable variation in cognitive function. For example, the range of DRS total scores was from 123 to 144, the maximum possible score on the DRS. A cut-off score of 122 is recommended (16), but 133 has recently been demonstrated to be a sensitive cut-off score for the detection of early dementia (9). Subjects were seated in front of the computer, the audio level was adjusted to assure that the subject could hear the instructions, the response stylus was handed to the subject, and the subject was asked to follow the instructions given by the computer.

All 38 of the older adult male subjects were able to complete the computer-based tasks with minimal assistance from the interviewer. Minor difficulties were experienced by subjects in adjusting for the ambient light levels, the angle of the screen, and the force with which the subject tapped the pen to the screen. Occasionally, subjects encountered difficulty when a superficial pen touch did not register during a series of sequential responses, even though auditory feedback indicating an incorrect responses was provided. This difficulty led to the inclusion of some additional verbal instructions in the Orientation task to point out the meaning of the feedback sounds.

DISCUSSION

Computer-based testing may have a number of advantages for screening or use in conjunction with conventional neuropsychological assessment. Well-designed, computer-based tests can reduce error variance in test administration and test scoring, require minimal administration skill, present complex stimuli creatively and accurately, and reduce examiner error or bias by accurately, precisely, and objectively recording a variety of patient responses, including both verbal or motor responses. However, most currently available computer-based tests provide assessment of a rather narrow range of behavioral functions, overemphasize visual presentation of heavily language-based stimuli (including reading, which creates bias against educationally limited subjects), and severely limit response formats. Since the results of a computer-based evaluation is typically a test score, it may be argued that reliance upon test scores neglects nonquantitative aspects of the patient's behavior that may be critical in interpreting the test score or in making a diagnostic or treatment planning decision. Expressive and organizational aspects of responding that may be particularly important in ascertaining the nature and extent of deficits are usually not measured (12). A major concern of health professionals has been the fear that "mechanizing these tests will mechanize the psychologist and that he would become increasingly distant from the patient" (7). These concerns may be alleviated when computer-based testing is used mainly in a screening context to identify at-risk individuals who then may be directed to more complete evaluation (7). In the future, a well-designed computer-based assessment system for individual subjects may be particularly well suited for screening in a physician's office by a nurse or technician. An immediate listing of results could be provided to the physician when the patient was seen, permitting further exploration of deficits or referral for further evaluation. Test results from different visits could be compared to allow quantitative tracking of the patient's mental status to determine if changes had occurred.

With this in mind, the screening battery described represents the initial development of a new pen-based neuropsychological testing system, NES3. This first set of tests implemented in the NES3 hardware–software environment is intended to be used for detecting cognitive impairments in individuals exposed to neurotoxic agents in the workplace. Available, and relatively inexpensive, hardware and software technology was employed. Use of a pen-based notebook computer with a speech production device is an improvement over existing computer-based batteries because they allow both auditory presentation of tests and recording of responses in a manner that is more familiar to subjects in general, and more analogous to response modes required in conventional paper-and-pencil neuropsychological tests.

Pilot testing of an early version of the screening battery software has provided information with which to refine its programming and presentation features. Initial testing of the revised pen-based computerized neuropsychological screening test demonstrated that adults of modest educational level are capable of completing it with minimal supervision and find it interesting. Software development of the screening battery is complete and the system is ready to be employed in validation studies and field trials.

Initial development of the screening battery represents the first set of tests implemented in the new NES3 hardware–software environment. Much additional development work remains to be performed. First, a validation study assessing the performance of the screening battery among workers known to be cognitively impaired and other workers known not to be impaired is required to estimate the sensitivity and specificity of the battery for screening workers for cognitive impairment. A normative study with the screening battery is also required to estimate the effects of covariates of performance such as age and education (14). We plan also to explore expansion of application of the screening battery to other populations, particularly older adults, children, and various ethnic/cultural groups. Some of the most useful NES2 tests will be adapted to the new NES3 environment to facilitate future epidemiologic applications. Finally, some additional neuropsychological tests will be implemented in the NES3 system to provide assistance to clinicians in performing in neuropsychological evaluations. We welcome collaboration in performing this additional development work.

ACKNOWLEDGEMENTS

This work was supported in part by the US National Institute for Occupational Safety and Health (R01 OH03078-01). The hard work, creativity, and attention to detail of Ms. Katie Moore in helping to design the screening battery tasks is gratefully acknowledged. Dr. Roberta White of Boston University and Dr. Edward Baker of the Centers for Disease Control and Prevention provided much constructive input and support.

REFERENCES

1. Anger, W. K. Worksite behavioral research: Results, sensitive methods, test batteries, and the transition from laboratory data to human health. Neurotoxicology 4:627–718; 1990.
2. Anthony, J.; LeResche, L.; Unaiza, N.; Von Koroff, M.; Folstein, M. Limits of the "mini-mental state" as a screening test for dementia among hospital patients. Psychol. Med. 12:397–408; 1982.
3. Baker, E. L.; Letz, R.; Fidler, A. A computer-administered neurobehavioral evaluation system for occupational and environmental epidemiology. Rationale, methodology, and pilot study results. J. Occup. Med. 27:206–212; 1985.
4. Cassitto, M. G.; Gilioli, R.; Camerino, D. Experiences with the Milan Automated Neurobehavioral System (MANS) in occupational neurotoxic exposure. Neurotoxicol. Teratol. 11:571–574; 1989.
5. Delis, D. C.; Kramer, J. H.; Kaplan, E.; Ober, B. A. California verbal learning test adult version, research edition manual, version 1. San Antonio, TX: The Psychological Corporation; 1987.
6. Eckerman, D. A.; Carrol, J. B.; Foree, D.; Gullion, C. M.; Lansman, M.; Long, E. R.; Waller, M. B.; Wallsten, T. S. An approach to brief field testing for neurotoxicity. Neurobehav. Toxicol. Teratol. 7:387–383; 1985.

7. Elithorn, A.; Mornington, S.; Stavrou, A. Automated psychological testing: Some principles and practice. Int. J. Man Machine Stud. 17:247–263; 1982.
8. Gamberale, F.; Iregren, A.; Kjellberg, A. SPES: The computerized Swedish performance evaluation system: Background, critical issues, empirical data, and a user's manual. Solna, Sweden: National Institute of Occupational Health; 1989.
9. Green, R. C.; Woodard, J. L.; Green, J. Validity of the Mattis Dementia Rating Scale for Detection of cognitive impairment in the elderly. J Neuropsychol. Clin. Neurosci. 7:357–360; 1995.
10. Johnson, B.L., ed. Prevention of neurotoxic illness in working populations. New York: John Wiley & Sons; 1987.
11. Laursen, P. A computer-aided technique for testing cognitive functions validated on a sample of Danes 30 to 60 years of age. Acta Neurol. Scand. Suppl. 131:1–108; 1990.
12. Letz, R. The Neurobehavioral Evaluation System (NES): An international effort. In: Johnson, B. L.; Anger, W. K.; Durao, A.; Xintaras, C., eds. Advances in neurobehavioral toxicology: Applications in environmental and occupational health. Chelsea, Michigan: Lewis Publishers; 1990.
13. Letz, R. Neurobehavioral Evaluation System 2 (NES2) user's manual (version 4.4). Winchester, MA: Neurobehavioral Systems Inc.; 1991.
14. Letz, R. Covariates of computerized neurobehavioral test performance in epidemiologic investigations. Environ. Res. 61(1):124–132; 1993.
15. Lezak, M. D. Neuropsychological assessment (3rd ed.). New York: Oxford University Press; 1995.
16. Mattis, S. Dementia rating scale professional manual. Odessa, FL: Psychological Assessment Resources, Inc.; 1973.
17. National Research Council. Environmental neurotoxicology. Washington, DC: National Academy Press; 1991.
18. Putnam, S.; DeLuca, J. The TCN Professional Practice Survey: Part 1: General practices of neuropsychologists in primary employment and private practice settings. Clin. Neuropsychol. 4:199–244; 1990.
19. Wechsler, D. The measurement of adult intelligence (4th ed.). Baltimore, MD: Williams & Wilkins; 1958.
20. Williamson, A. M. The development of a neurobehavioral test battery for use in hazard evaluations in occupational settings. Neurotoxicol. Teratol. 12(5):509–514; 1990.