Self-administration and Reliability of Computerized Neurobehavioral Tests among Egyptian Pesticide Workers.

Fayssal M. Farahat

ABSTRACT: Background: Behavioral Assessment and Research System (BARS) is a computer-based test system designed to assess neurobehavioral function and neurotoxicity in humans. This system is available in 5 languages including Arabic. Objectives: To assess the potential use of the Arabic version of computerized neurobehavioral system (BARS) and to compare performance of Egyptian pesticide workers to non-exposed populations using the same computerized battery. Participants and Methods: This study involved the administration of eight neurobehavioral tests from the Arabic computerized neurobehavioral test battery (BARS) to Egyptian workers occupationally exposed to pesticides (n= 25) as compared to non-exposed Egyptian workers (n= 25). One-week test-retest reliability of the computerized battery was measured among non-exposed participants. Results: Performance of pesticide non-exposed Egyptian workers did not show any significant differences between test and re-test (i.e., after one week). Performance of pesticide exposed workers was significantly lower in most of the administered computerized tests as compared to non-exposed Egyptian and US populations. Conclusions: The current findings demonstrate the potential utility of the Arabic computerized BARS in occupational epidemiological research especially in the short-term intervals. Stability of the administered Arabic BARS tests over the short-term interval makes it broadly applicable in assessing exposures at different workplaces and with different cultural and educational levels. Key Words: Arabic, neurobehavioral, computerized battery

INTRODUCTION

Since the late 1960’s neurobehavioral performance tests have been used to assess the effects of occupational exposure in adult workers\(^1\text{-}\text{3}\). The use of neurobehavioral tests to assess workplace exposure has continued to increase and neurobehavioral tests have become the most efficient methods (in terms of cost and time) to screen for adverse effects of neurotoxic exposures in adult workers\(^4\text{-}\text{6}\). Behavioral Assessment and Research System (BARS) is a computer-based test system designed to assess...
neurobehavioral function and neurotoxicity in humans\textsuperscript{(7,8)}. It has been effectively applied among poorly educated populations and young children, with minimal support from a human examiner\textsuperscript{(9-15)}. Instructions were written in very direct and simple words. Each simply-stated step was presented on the screen, followed by practice on that step or concept. Feedback was provided for correct (smiling face) and incorrect (frowning face) performance at each step of the instructions/practice\textsuperscript{(9)}. BARS tests (originally administered in English) have been translated into Spanish\textsuperscript{(13)}, Portuguese\textsuperscript{(15)}, Korean\textsuperscript{(14)}, and Arabic (translation was done by the researcher during his training at Center for Research on Occupational and Environmental Toxicology, CROET, Oregon, USA)\textsuperscript{(9)}. Fig. 1 demonstrates an instruction screen from the Serial Digit Learning test in Korean, Portuguese, Arabic, and Spanish.

To achieve the goal of testing non-educated participants, spoken instructions were implemented. Thus, BARS could be used with participants who had low (or no) reading skills and the instructions could be presented in any language by using the appropriate graphics (pictures) of the instructions and associated sound files.

One of the biggest challenges in using computers for testing in the workplace is that many workers with low-level of education are not at all familiar with computers and they are reluctant to even touch the keyboard. This led to the development of the 9BUTTON (formerly named DataSled) unit\textsuperscript{(9,16)} that was placed over and replaced the keyboard (on laptops). With the 9BUTTON, participants pressed larger buttons rather than the relatively smaller computer keys. With the 9BUTTON unit in place over the keyboard, reluctance to touch the keyboard disappeared as an issue, and the
participants’ attention became focused on the tests.

Although reliability of the original English BARS has been tested\(^{16,17}\) and was applied to identify neurobehavioral deficits in several studies\(^{10,12,13,18,19}\), reliability of the Arabic BARS has not yet assessed. The test-retest procedure is the most widely used paradigm to assess reliability \(^{20, 21, 22}\). Meanwhile, neurobehavioral test batteries are often repeatedly administered to the same individuals to study changes over time, the progress of pathologies, or the effects of exposure and clinical interventions\(^{20}\). Test-retest reliability involves administering the same test to a group of participants on two different occasions\(^{21,23}\). The interval of time may be as short as same day or it can be as long as several years\(^{24,25}\). Studies of neurotoxic exposure tend to rely on shorter intervals\(^{26}\). This study was conducted to assess the potential use of the Arabic version of computerized neurobehavioral system (BARS) and to compare the performance of a group of Egyptian workers occupationally exposed to pesticides to that of non-exposed populations using the same computerized battery.

**MATERIAL and METHODS**

**Participants**

This study involved administration of the Arabic computerized neurobehavioral tests (BARS) to Egyptian workers occupationally exposed to pesticides (n= 25) as well as to a group of non-exposed population (n= 25). Egyptian pesticide workers were recruited among agricultural engineers (university education) and technicians (secondary school education) who are working at the Pesticide Control Department, while non-exposed participants were recruited among those working at the administrative departments of the Menoufia General Directorate of Agriculture. This study was conducted during the period from July – August 2006. Participants were excluded if
they are diabetics, taking medication for any neurological, psychological or hepatic condition, or reporting a history of previous head injury.

**Measures**

A battery of 8 neurobehavioral tests from the Arabic BARS neurobehavioral test battery was administered as follow\(^9,10\):

1. **Symbol Digit (SDT; complex function):** This is described as a coding test in which digits are paired with symbols in a 2 squares by 9 squares matrix. A similar matrix at the bottom of the screen contains the symbols but not the digits. The subject is instructed to type the correct numbers (i.e., that correspond with the respective symbols) in the empty matrix spaces.

2. **Simple Reaction Time (SRT; response speed):** The subject is instructed to respond by pressing a 9BUTTON as quickly as possible after they see a stimulus presented on the screen or when a 9BUTTON response button becomes backlit.

3. **Continuous Performance Test (CPT; sustained attention):** A series of stimuli are presented one at a time for several minutes, typically 5-10. Subjects are instructed to press a key when a target is presented.

4. **Digit Span Forward (DSF; attention and memory):** A series of numbers is presented sequentially on the computer screen, and the subject is instructed to reproduce the sequence of numbers by pressing the numbered 9BUTTON buttons in the same order (forward). The number of digits increases until a failure criterion is met.

5. **Digit Span Backward (DSB; attention and memory):** Same as DSF except that subjects press the numbered 9BUTTON in the reverse order.

6. **Match-to-Sample (MTS; visual memory):** A 10×10 matrix of blocks is followed by three choices, among which one is the same as the sample stimulus.

7. **Selective Attention (SAT; sustained attention):** A small dot is briefly presented
inside or outside one of two squares, one on the left and one on the right half of the screen. The subject is instructed to press one button when a dot appears in the square on the left, a different button when the dot appears on the right, and to not press a button when the dot appears outside of either square.

8. **Finger Tapping (TAP; response speed coordination):** The subject is instructed to press (tap) button(s) as rapidly as possible using the index finger of one or both hands on one or two buttons. Responses incrementally increase the height of a dark bar to suggest progress to the subject.

Each test in BARS allows the investigator to set test parameters such as instruction language (i.e., English, Arabic, Spanish, Portuguese, Korean), test duration, and stimulus set. This allowed for the use of alternate forms for the Symbol Digit, Digit Span, and Match-to-Sample tests(9).

**Procedure:** Test-retest reliability of the Arabic computerized BARS was measured among the non-exposed group. They were tested at their place of work on two separate occasions, one-week apart, keeping conditions constant (e.g., examiner, procedures, time of day, and environment). On the other hand, pesticides exposed workers were tested once at the agricultural units. Figure (2) shows two of the pesticide workers during test administration. At the time of the first test session, informed consent was obtained and demographic information was collected from all the participants. Assessment of all groups was done with one examiner present during testing. Participants were tested during regular working hours (between 8:00 AM - 4:00 PM). The test session lasted for one hour on average. BARS tests were administered on Apple PowerBook computers and participants responded by pressing buttons.
(numbered 1-9) on a 9BUTTON external response unit (Figure 2).

**Statistical Analysis**

The data were statistically analyzed using SPSS version 13. Performance of the participants was assessed by computing means and standard deviations for the BARS tests and compared by Student’s $t$ test, paired $t$ test, Analysis of Variance (ANOVA) for quantitative normally distributed data; chi square test for qualitative data, and Mann-Whitney test for quantitative abnormally distributed data. Kolmogorov-Smirnov (K-S) was used to determine normal distribution. Level of significance was determined at $p < 0.05$.

**RESULTS**

Demographic data of all participants are presented in Table (1). Exposed and non-exposed participants were males with mean (SD) of age of 34.2 (11.1) and 37.8 (9.5), respectively. Forty percent of exposed workers have more than 12 years of education compared to 32% among non-exposed. Fifty-two percent of the exposed participants were current smokers compared to 40% among non-exposed workers. No statistically significant differences were detected between exposed and non-exposed participants (Table 1).

Data of the test-retest reliability of the Arabic BARS among non-exposed participants was compared to that of the original English BARS among US population (Table 2). Data of US population was obtained from the test-retest reliability study by Farahat et al., (2003)\(^{(16)}\). Means and standard deviations of the performance for the non-exposed participants and US population at Time 1 and Time 2 are summarized in Table 2. The mean difference for each measure (Time 2 – Time 1) is also presented. Paired $t$-test was used to evaluate performance differences on the measures from Time 1 to Time 2. Performance of the non-exposed participants in all the administered
tests did not show statistically significant differences from Time 1 to Time 2 (i.e., after one week). However, although difference scores of the US population on all measures were fairly small, significant differences were detected on Symbol Digit and Finger Tapping (preferred hand) (Table 2). Mean difference scores between T1 and T2 of both non-exposed Egyptian and US populations were compared using Mann-Whitney test, where no significant differences were reported in any of the administered tests (Table 2).

Pairwise comparisons were done to compare performance of the Egyptian pesticide workers to that of the non-exposed Egyptian and US populations at the first session (Time 1) (Table 3). Egyptian non-exposed participants showed better performance than pesticide workers. Differences were statistically significant in 6 out of 8 administered tests (Digit span backward and Continuous performance were not statistically significant). On the other hand, performance of the US population was significantly better than Egyptian pesticide workers in all the administered tests. Non-exposed US population showed statistically significant better performance than non-exposed Egyptian workers in all tests except for tapping and reaction time tests (Table 3).

**DISCUSSION**

The Behavioral Assessment and Research System (BARS), a computerized neurobehavioral test system, was initially targeted for use with a broad range of working populations with different educational levels and cultural backgrounds\(^7,8\). Previous research has found the English version of BARS tests to be reliable across a one-week interval in a normative sample. The correlation coefficients ranged from 0.44 to 0.92 between session 1 and session 2 on seven standard neurobehavioral tests\(^{16,17}\).

This study assessed the reliability of the Arabic computerized neurobehavioral
test battery among Egyptian population (pesticide non-exposed workers) over one-week interval. Furthermore, performance of a group of Egyptian pesticide exposed workers was compared to both Egyptian and US non-exposed populations. The Arabic computerized neurobehavioral tests were reliable as shown by absence of significant differences in any of the administered tests between time 1 and time 2 among Egyptian non-exposed participants. Additionally, mean differences of performance of the Egyptian non-exposed workers from time 1 to time 2 did not show any statistically significant differences as compared to the US population. These data strongly support the use of the Arabic computerized neurobehavioral tests as reliable screening measures where short-term evaluation may be required (e.g., within one-week) for the assessment of exposures to different neurotoxicants at the workplace (e.g., pesticides). Furthermore, it supports the reliability of the Arabic computerized neurobehavioral tests among different populations with different backgrounds and educational levels.

The significant differences in test-retest of finger tapping and symbol digit among US population have been attributed by the authors to the effect of practice\textsuperscript{16}. Meanwhile, the magnitude of practice effects among current Egyptian non-exposed workers was fairly small in all measures as reported by absence of statistical significance of paired t-tests. Relatively, absence of practice effects among Egyptian non-exposed workers can be attributed to the use of alternate forms in the Arabic BARS. Computerized neurobehavioral tests allow the use of alternative forms that can be modified before test administration\textsuperscript{(10,13)}. The use of alternate forms may reduce the amount of practice effect, a finding consistent with the literature\textsuperscript{27}. However, presence of practice effects should not be surprising as re-
administration of the same measures\textsuperscript{(28)} or even alternate forms of the measures tends to result in improved performance especially at short time intervals\textsuperscript{(20)}. Dikmen \textit{et al.}, (1999)\textsuperscript{(20)} reported significant practice effects in more than half of the measures of the Wechsler Adult Intelligence Scale (WAIS) including Symbol Digit, Object Assembly, Block Design and Picture Arrangement, and Picture Completion. Practice effects, can also occur even on purely motor measures (e.g., Finger Tapping)\textsuperscript{(29)}.

Lower performance of the Egyptian pesticide workers as compared to non-exposed Egyptian and US populations may be attributed to pesticide exposure effects. Although demographic data (known neurobehavioral confounders)\textsuperscript{(6)} of exposed and non-exposed Egyptian participants were not significantly different, performance of exposed workers was significantly lower in 6 out of 8 administered computerized tests. Previous studies of non-computerized neurobehavioral tests among pesticide workers reported significant impact of pesticide exposure on the performance of workers\textsuperscript{(30)}. Higher educational levels among US population (84% of US population have more than 12 years of education) beside other socio-cultural factors including more familiarity with computers may contribute to the significantly better performance among US population compared to Egyptian non-exposed workers and add more to the difference as compared to Egyptian exposed workers.

In the current study, although evidence can be inferred regarding the effect of pesticides on neurobehavioral performance of exposed workers, other measures of exposure monitoring (e.g., serum acetylcholinesterase AChE) should be assessed and correlated with neurobehavioral performance in the future studies.
In conclusion, the current findings demonstrate the potential utility of the Arabic computerized BARS in occupational epidemiological research especially in the short-term intervals. Stability of the administered Arabic BARS tests over the short-term interval makes it broadly applicable in different types of research where short-term intervals are required. However, these findings about the reliability of this computerized test battery should be extended to address longer-time intervals (e.g., months), such as those due to chronic occupational exposures, in clinical research, and with different cultural groups.

ACKNOWLEDGEMENT

I would like to acknowledge Dr WK Anger and Dr DS Rohlman (Experimental Psychologists) from Center for Research on Occupational and Environmental Toxicology (CROET), Oregon, USA for their great help and technical support during my training at OHSU and during the development of the computerized Arabic BARS.

**Table 1: Demographic characteristics of the participants**

<table>
<thead>
<tr>
<th>Studied Variable</th>
<th>Egyptian Pesticide Workers (n = 25)</th>
<th>Non-exposed workers (n = 25)</th>
<th>Test of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years (mean ± SD)</td>
<td>34.32 (± 11.12)</td>
<td>37.82 (9.53)</td>
<td>Student’s t = 1.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P = 0.24</td>
</tr>
<tr>
<td>Years of education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;12 y {n (%)}</td>
<td>2 (8.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>12 y {n (%)}</td>
<td>13 (52.0)</td>
<td>17 (68.0)</td>
<td>X² = 2.76</td>
</tr>
<tr>
<td>&gt;12 y {n (%)}</td>
<td>10 (40.0)</td>
<td>8 (32.0)</td>
<td>P = 0.25</td>
</tr>
<tr>
<td>Smoking: Current smoker {n (%)}</td>
<td>13 (52.0)</td>
<td>10 (40.0)</td>
<td>X² = 0.72</td>
</tr>
<tr>
<td>Non-smoker {n (%)}</td>
<td>12 (48.0)</td>
<td>15 (60.0)</td>
<td>P = 0.39</td>
</tr>
</tbody>
</table>
Table 2: Mean and Standard Deviation of Test, Retest, and Test-retest Mean Differences of the non-exposed Egyptian and US populations.

<table>
<thead>
<tr>
<th>Studied Tests</th>
<th>Egyptian non-exposed</th>
<th>US population*</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>(Time 1)</td>
<td>(Time 2)</td>
<td>Difference (1)</td>
<td>(T2 – T1)</td>
<td>Difference (2)</td>
</tr>
<tr>
<td>Tapping (preferred hand)</td>
<td>103.8 (4.89)</td>
<td>104.5 (7.07)</td>
<td>0.64 (6.47)</td>
<td>101.7 (13.3)</td>
<td>105.1 (13.8)</td>
</tr>
<tr>
<td>Symbol digit (latency)</td>
<td>2332.4 (296.5)</td>
<td>2234.6 (18.2)</td>
<td>18.2 (68.3)</td>
<td>1810.6 (310.4)</td>
<td>1653.6 (377.3)</td>
</tr>
<tr>
<td>Reaction time (latency)</td>
<td>327.2 (243.4)</td>
<td>345.4 (4.2)</td>
<td>0.24 (0.72)</td>
<td>332.5 (332.5)</td>
<td>340.1 (332.5)</td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>4.8 (0.76)</td>
<td>5.3 (0.63)</td>
<td>-0.96 (6.36)</td>
<td>7.3 (1.3)</td>
<td>6.4 (1.6)</td>
</tr>
<tr>
<td>Digit Span Reverse</td>
<td>4.2 (0.88)</td>
<td>4.5 (0.51)</td>
<td>-3.24 (6.02)</td>
<td>6.0 (1.7)</td>
<td>480.9 (480.9)</td>
</tr>
<tr>
<td>Selection Attention (number)</td>
<td>398.4 (38.9)</td>
<td>394.0 (59.6)</td>
<td>0.4 (1.4)</td>
<td>65.2 (24.2)</td>
<td>60.9 (25.5)</td>
</tr>
<tr>
<td>Match-to-Sample (%) correct</td>
<td>60.5 (9.0)</td>
<td>59.6 (26.5)</td>
<td>1.6 (1.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>91.3 (10.4)</td>
<td>88.0 (12.7)</td>
<td>3.4 (10.9)**</td>
<td>0.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Performance (% hits)</td>
<td>0.0 (12.7)</td>
<td>0.64 (6.47)</td>
<td>3.4 (10.9)**</td>
<td>0.001</td>
<td>0.02</td>
</tr>
</tbody>
</table>

* data obtained from Farahat et al., 2003 (16).
** p < 0.05 using paired t test.
*** no statistical significance between mean difference (1) and mean difference (2) using Mann-Whitney test.

Table 3: Analysis of Variance (ANOVA) to compare performance on neurobehavioral tests among exposed and non-exposed populations.

<table>
<thead>
<tr>
<th>Studied Tests</th>
<th>(1) Egyptian Pesticide Workers</th>
<th>(2) Egyptian Non-exposed workers</th>
<th>(3) US non-exposed population</th>
<th>Post-Hoc test (LSD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>P1 (1) Vs (2)</td>
</tr>
<tr>
<td>Tapping (preferred hand)</td>
<td>91.9 (15.4)</td>
<td>103.8 (4.89)</td>
<td>101.7 (13.3)</td>
<td>0.001</td>
</tr>
<tr>
<td>Symbol digit (latency)</td>
<td>3170 (419)</td>
<td>2332.4 (395.3)</td>
<td>1810.6 (310.4)</td>
<td>0.001</td>
</tr>
<tr>
<td>Reaction time (latency)</td>
<td>512.5 (138)</td>
<td>327.2 (69.2)</td>
<td>332.5 (53.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>4.2 (0.8)</td>
<td>4.8 (0.76)</td>
<td>7.3 (1.3)</td>
<td>0.01</td>
</tr>
<tr>
<td>Digit Span Reverse</td>
<td>3.8 (0.8)</td>
<td>4.2 (0.88)</td>
<td>6.0 (1.7)</td>
<td>0.09</td>
</tr>
<tr>
<td>Selection Attention (#)</td>
<td>356.4 (71.2)</td>
<td>398.4 (53.3)</td>
<td>480.9 (52.9)</td>
<td>0.02</td>
</tr>
<tr>
<td>Match-to-Sample (%) correct</td>
<td>42.7 (10.4)</td>
<td>60.5 (9.0)</td>
<td>32.6 (5.0)</td>
<td>0.001</td>
</tr>
<tr>
<td>Continuous Performance (% hits)</td>
<td>85.3 (8.7)</td>
<td>91.3 (12.7)</td>
<td>65.2 (24.2)</td>
<td>0.06</td>
</tr>
</tbody>
</table>

* p value.

Bull High Inst Public Health Vol.37 No.2 [2007]
Fig. 1. Instruction screen from the Serial Digit Learning test presented in Korean, Portuguese, Arabic, and Spanish (Rohlman et al., 2003).

Figure 2. Two of the participants during testing.
REFERENCES