

Reliability and Validity of Bedside Version of Persian WAB (P-WAB-1)

Reza Nilipour^{1*}, Abbas Pourshahbaz², Zahra Sadat Ghoreyshi¹

1. Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

2. Department of Clinical Psychology, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

Article info:

Received: 22 January 2014

First Revision: 25 March 2014

Accepted: 03 September 2014

Key Words:

Aphasia, Persian,
CVA, Aphasia quotient,
P-WAB-1

ABSTRACT

Introduction: In this study, we reported the reliability and validity of Bedside version of Persian WAB (P-WAB-1) adapted from Western Aphasia Battery (WAB-R)^{1,2}. P-WAB-1 is a clinical linguistic measuring tool to determine severity and type of aphasia in brain damaged patients based on Aphasia Quotient (AQ) as a functional measure. For the purposes of a quick clinical screening of aphasia in Persian, we adapted the bedside version of WAB-R to assess the performance of Persian aphasic patients.

Methods: The data we reported on adaptation, validity and reliability of P-WAB-1 are based on faithful translation and criterion validity ratio (CVR) taken from the expert panel and the performance of 60 consecutive brain damaged patients referred to different university clinics for rehabilitation and 30 healthy subjects as norms and 40 age-matched epileptic patients as the control group.

Results: Based on the results of this study, P-WAB-1 has internal consistency ($\alpha=0.71$) and test-retest reliability ($r=.65$ $P<0.001$) and the subtests are sensitive enough to contribute to Aphasia Quotient (AQ) as a functional measure of severity of aphasia in Iranian brain damaged patients. Based on AQ results, our aphasic patients were classified into four distinct groups of severity.

Discussion: P-WAB-1 is the first clinical linguistic test to determine severity of aphasia based on an operational index and can be considered as a valid baseline for screening and diagnosis of aphasia among Persian speaking brain damaged patients. This study is the initial step on adaptation of different versions of WAB-R to measure the severity of aphasia using AQ, LQ and CQ as operational measures and to classify Persian speaking aphasic patients into different types.

1. Introduction

Aphasia is known to be one of the most fascinating and complex neurolinguistic problems that many clinicians encounter in clinical settings. At the same time, its diagnosis and classification into different types have been controversial. Consequently, different methods of assessment have been proposed for aphasia interpretation and classification by leading researchers in this field (Goodglass & Caplan, 1972; Kertesz, 1982; Paradis et al., 1987).

The history of assessment of aphasia in Iran goes back to 1980s when Persian version of the Bilingual Aphasia Test (BAT) (Paradis & Libben, 1987) as the first multi-lingual clinical and experimental aphasia test was adapted for Persian (Paradis et al., 1987). Later Azari (Paradis et al., 1987) and Kurdish (Paradis et al., 1987) versions of the BAT were developed for 2 major languages spoken in different parts of Iran. Following adaptation of the BAT, a new Persian Aphasia Battery (PAB) (Nilipour, 2011) was developed based on the guidelines of Boston Diagnostic Aphasia Examination (BDAE) (Goodglass

* Corresponding Author:

Reza Nilipour, PhD

Address: Department of Speech Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

Tel/ Fax: +98 (21) 22180043

E-mail: rnilipour@gmail.com

& Caplan, 1972) and the scoring format of the BAT (Paradis & Libben, 1987). A Persian Aphasia Naming Test (Nilipour, 2011) was also developed based on Armstrong Naming Test (Armstrong, 1995) for clinical and therapeutic applications. Based on the available clinical linguistic assessment tools, some case studies have been reported at international conferences or published in international English or Iranian journals by SLP clinicians and neurolinguistic researchers (Nilipour et al., 2012), for more information see Persian Clinical Linguistic Database (PCLD.USWR.AC.IR)*.

Although the available aphasia batteries have been adapted for clinical applications of Iranian aphasic patients, the format of these tests does not provide an operational index for measuring severity of aphasia. The major feature of the present aphasia tests is that these tests can provide a general profile of the language impairments at different levels and residual linguistic skills of the aphasic patients but the statistical transformation of the profile into severity and type has to be formulated mainly by the clinician.

In clinical management of aphasia measurement of severity and type of aphasia are two important factors (Kertesz & Pool, 1974). In this context, the relationship between severity and type of aphasia and lesion site could be very helpful for diagnosis and rehabilitation management. Presently, there is few documented literature on numerical percentile index of severity of aphasia, especially on Persian speaking aphasic population.

Among the present international aphasia batteries Western Aphasia Battery (WAB-R) (Kertesz, 2007) as a clinical tool has been used routinely and widely to evaluate adult language deficits in English and some other languages (Kyoung et al., 2010) and has been reported to have high internal consistency, test-retest reliability, and validity (Kertesz, 1982). The WAB has also been widely used to determine the presence, type, and severity of aphasia based on the 3 quantitative measures (Cynthia et al., 1980; Kertesz & Poole, 1974). The Aphasia Quotient (AQ) is the essential summary value of the individual's aphasic deficit, and is proportional to the severity of aphasia regardless of the type and etiology. The Language Quotient (LQ) combines oral and written language scores to emphasize communicative importance and the relationship between the two modalities. The Cortical Quotient (CQ) includes optional nonverbal tests, apraxia, and written language in addition to the AQ to provide a balanced summary of focal cortical function (Kertesz & Poole, 1974; Kertesz, 2006: 83). Since WAB-R is a criterion-referenced test based on the AQ, LQ and

CQ, it brings some degree of quantification for measuring severity of aphasia for clinical usage and research purposes. It is reported that the prognostic value of the AQ in stroke is considerable and in degenerative conditions can be interpreted to stage Alzheimer's disease and primary progressive aphasia (Kertesz, 2006: 83).

It has also been reported that the Bedside WAB-R scores on content and fluency of spontaneous speech, subtests of comprehension, command, repetition, and naming can be used as baseline measures of the severity of aphasia or ability of the patient prior to rehabilitation or surgery.

The purpose of this study was to report adaptation of bedside version of P-WAB (P-WAB-1) as a clinical measuring tool for quick screening to determine the severity of aphasia among Persian speaking CVA and other brain-damaged patients based on AQ.

2. Methods

In adaptation of Persian version of the WAB-R, the first stage was to obtain a faithful translation of the English version of the clinical version of the WAB into Persian considering the linguistic and cultural differences between English and Persian. At this stage, English version of the WAB was translated into Persian in group discussion sessions using back translation method in collaboration with two linguists, an experienced SLP and a clinical linguist. In this study, we reported the results of standardization procedure of the clinical version (P-WAB-1). The first draft of the P-WAB-1 as a record form for clinical use was made available based on the faithful translation of the subtests of the WAB-R (Kertesz, 2006) into Persian.

In order to measure the content validity of the P-WAB-1, an expert panel group consisted of 9 experienced SLPs were invited and asked to measure the content validity ratio (CVR) of the P-WAB-1. Based on the results obtained from the panelists, we were able to get CVR (Lawshe, 1975) of different sections of the P-WAB-1. The first draft of P-WAB-1 was administered on 30 healthy normal adult native speakers of Persian. They were asked to answer the questions in each sub-test of P-WAB-1 and describe a culturally adapted version of the Nest Story cartoon (see PCLD.USWR.AC.IR) of the BAT (Paradis & Libben, 1987) to obtain a unified spontaneous speech database for the norms to evaluate the content and fluency of spontaneous speech of each patient. Based on the results of the performance of the normal subjects, the second draft of P-WAB-1 as a clinical trial record form was made available. The spontaneous speech samples of the 30 healthy subjects on the

Nest Story were analyzed and considered as the source of evaluation and scoring of the content and fluency of spontaneous speech of each patient (See for details of the guidelines for scoring connected speech in the record form of P-WAB-1).

2.1. Structure of P-WAB-1

The P-WAB-1 in this study is consisted of six linguistic subtests. Each subtest of P-WAB-1 obtains a raw score of 10. As suggested in the manual of WAB-R based on the raw score, a percentile Aphasia Quotient (AQ) can be formulated in order to determine the severity of aphasia (Kertesz, 2006). The record form of P-WAB-1 consists of the following 6 sections and proper space is provided to register the patient's responses in front of each item. The subtests and raw scores of each section are as follows (See PCLD.USWR.AC.IR for record form):

- Spontaneous speech content (10 points):
3 conversational questions (5 points) and content of spontaneous speech (5 points)
- Fluency of spontaneous speech (10 points)
- Auditory comprehension: 10 Yes/No questions (10 points)
- Sequential commands: 5 commands of different complexities (10 points)
- Repetition: 6 words and sentences of different lengths (10 points)
- Naming: 20 different naming categories (10 points)

The language subtests of P-WAB-1 are chosen to represent equally important functions of spoken language in order to arrive at a numerical percentile index of severity (AQ) as proposed by Kertesz (Kertesz & Poole, 1974). The numerical percentile AQ as an operational index of severity does not require statistical transformation by the clinician.

2.2. Subjects

The target population of P-WAB-1 is Persian-speaking adults with acquired neurological disorders such as CVA and head injury. 60 brain damaged patients who were referred by neurologists for therapy to university clinics and 40 patients with epilepsy as the control group were participated in this study for adaptation, reliability and validity of P-WAB-1. The inclusion criteria for aphasic patients were left hemisphere damage, first stroke with no prior history of psychiatry. The demographic characteristics of both groups are given in Table 1. The aphasic patients included in this study were all adult left brain damaged speakers of Persian with mean education of

10.13 (± 4.8). The etiology of the majority of the aphasic patients was CVA except 10% with etiology of trauma.

Table 1. Demographic characteristics of the subjects.

Group	N	Age	Male	Female
Aphasic	60	51.95 (± 8.59)	49	21
Epileptic	40	45.73 (± 8.48)	24	16
Total	100		73	37

NEURSCIENCE

With respect to post onset time, 83.3% of the patients were chronic and 16.6% were during sub-acute stage. The mean of post onset time was 36.05 (± 34.02) months. The neurologic control group subjects who participated in the study constitute 40 age-matched epileptic patients out of 120 patients who were under investigation by the Comprehensive Epilepsy Program Group, Isfahan University of Medical Sciences.

3. Results

3.1. Evidence of Content Validity Ratio (CVR)

The results of CVR obtained from the panelists indicated only two questions: one in connected speech and one in naming received a CVR of 0.67 which was below standard score. Based on the CVR obtained from the expert panelists, required modifications were made to improve the content validity of the P-WAB-1. The new revised version of P-WAB-1 was made available with the expected CVR of 0.78.

3.2. Evidence of Reliability

In order to check the reliability of the P-WAB-1, we evaluated the internal consistency and stability of the scores. The measurement of internal consistency is whether various parts of the P-WAB-1 contribute in a consistent manner to the total score. For the purpose of internal consistency, the scores of 20 aphasic patients who participated in the study were used. The Kuder Richardson coefficient for AQ was 0.71, and for the subtests, it was between 0.71 and 0.91 indicating the high internal consistency of the P-WAB-1.

For the purpose of stability, we evaluated the scores of the same 20 aphasic patients who participated in the study in one week interval. Based on the results of test re-test, the correlation coefficient for AQ was 0.65 ($P < 0.001$), and for sub-tests it was between 0.42 and 0.98 indicating stability of P-WAB-1 (See Table 2).

Table 2. Internal consistency and test t-test reliability of subtests of P-WAB1.

Subtest	N	Test1	Test2	Correlation Coefficient	α
Content	20	3.05	2.85	0.84	0.78
Fluency	20	1.7	1.6	0.98	
Auditory Comprehension	20	4.3	4.7	0.58	0.79
Command Comprehension	20	3.25	3.2	0.58	0.71
Naming	20	3.05	3	0.42	0.91
Repetition	20	1.45	1.3	0.47	0.86
AQ	20	46.7	46.6	0.65	0.79

NEUR SCIENCE

3.3. Evidence of Validity of P-WAB-1 Subtests

The P-WAB-1 subtests correlation matrix in Table 3 is based on the performance of 60 aphasic patients who

participated in this study. As the figures indicate, all subtests have high correlation with AQ (between 0.76 and 0.85). The figures in the correlation matrix also indicate the validity of the test.

Table 3. Correlation matrix for P-WAB-1 subtests scores.

Section	Content	Fluency	Comprehension	Naming	Repetition	AQ
Content	1.00					
Fluency	0.69	1.00				
Comprehension	0.47	0.45	1.00			
Naming	0.69	0.50	0.63	1.00		
Repetition	0.68	0.61	0.72	0.64	1.00	
AQ	0.79	0.72	0.85	0.81	0.90	1.00

*Correlation is significant at the level 0.01.

NEUR SCIENCE

3.4. Evidence of Criterion Validity of P-WAB-1

As mentioned before, 60 CVA patients and 40 patients with epilepsy as control group participated in this study. The comparison of the means and standard deviations of scores of 60 subjects and 40 control group are presented in Table 4. The results of P-WAB-1 represent equal levels of difficulty subtests and in sufficient number to detect patients with aphasia from the control group. The subtest scores indicate that P-WAB-1 items and tasks represent distinct language functions of equal importance.

The results of t-test in Table 4 indicate that the difference between means of sub-tests and AQ of the aphasics and epileptic patients are significant (P -value<0.001). As the results indicate, the control subjects would easily achieve a score of above 90 percent (Table 4).

The comparison of mean scores of aphasics and the control group on P-WAB-1 are also summarized in Figure 1. As the subtests and AQ scores in Figure 1 indicate, the scores can reliably be differentiated between aphasics and the control group as non-aphasic subjects.

Table 4. Number of patients, Mean scores, SDs of subtests and AQ of aphasics and epileptic patients.

Group	No	Content	Fluency	Comprehension**	Repetition	Naming	AQ
Aphasic	60	4.39 (+2.76)	3.17 (+2.26)	6.28 (+2.69)	4.9 (+3.28)	4.52(+3.44)	49.77(+23.49)
Epileptic	40	8.8 (+1.45)	7.98 (+1.37)	9.58 (+0.72)	9.88 (+0.38)	9.88(+0.38)	92.98(+6.76)
T*	100	9.27	12.04	7.56	9.56	10.01	11.31

* P -value< 0.001 **Mean scores of comprehension are means of auditory comprehension & sequential command sub-tests. Degree of freedom was 98.

NEUR SCIENCE

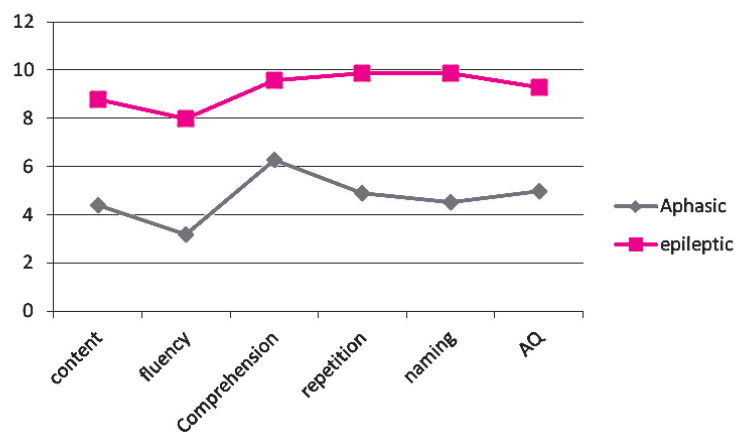


Figure 1. Comparison of performance of aphasics and epileptic patients on AQ & sub-tests

With respect to quantification of the severity of aphasia as suggested by Kertesz (Kertesz et al., 1974), the AQ results in our study were sensitive enough to classify the aphasic patients into 4 different clusters of severity (Table 5). Based on the severity framework proposed by Kertesz (2006), 9 patients in our study were classified as very severe, 24 as severe, 15 as moderate, and 12 as mild aphasics (Table 5).

4. Discussion

The validated P-WAB-1 reported in this study was designed based on the spoken subtests of the bedside version of original WAB format (Kertesz, 2006). P-WAB-1 is meant for research and quick clinical utility to assess language impairments in Iranian brain damage patients and is an initial step in developing P-WAB battery for Persian speaking patients. The P-WAB-1 can be administered in about 15 minutes. The adaptation of subtests of P-WAB-1 for Persian based on the original WAB and the evidence of validity and validity results indicating that P-WAB-1 has clinical utility and can be used as an assessment tool to obtain an operational index for diagnosis and screening Iranian aphasic patients.

As the data indicated, scores of aphasic patients were significantly different from the control group (Table 4). Based on the collected data on 60 aphasic patients,

we were able to cluster the patients into four different groups of severity (Table 5).

As Kertesz has reported, the subtests which correlated best with AQ was information (content), while in our study, the repetition and comprehension were the best correlated subtests with AQ (Table 3).

With respect to difficulty levels of aphasia, the data in our study were quite comparable with the results reported by kertesz (2006), as verified by the similarity of means and standard deviation obtained from 40 control subjects (Table 5).

One of the main findings of this study was that P-WAB-1 can be used as a valid measuring clinical tool to assess language impairments in patients with brain damage and quantify the severity of their impairment based on an operational AQ index proposed by Kertesz (1982).

As suggested by Kertesz (1974; 2006) the Aphasia Quotient (AQ) obtained from WAB is the essential summary value of the individual's aphasic deficit and is proportional to the severity of aphasia regardless of the type or etiology. In other words using an operational definition, AQ of 0-25 is very severe; an AQ of 26-50 is severe; an AQ of 56-75 is moderate, and an AQ of 75 and above is mild. Based on the data obtained from the assessment of

Table 5. Clustering of aphasic patients by severity.

Clusters of Severity	Number	AQ Range
Very Severe	9	0-25
Severe	24	26-50
Moderate	15	51-75
Mild	12	76-92
Total	60	

60 aphasic patients and 40 control group, 9 of the patients were within the range of very severe, 24 patients severe, 15 moderate and 12 patients were mild aphasics.

Finally, the adaptation of P-WAB-1 made it possible to quantify the severity of aphasia among Persian aphasic patients with different lesion sites. It is also possible to classify the patients into different clinical sub-types based on AQ, and define the relationship of the type of aphasia with the lesion site. The taxonomic approach to classify Persian aphasics into different types and developing other versions of P-WAB to measure CQ, LQ and CQ will be the next step in our study. We hope measures of severity of aphasia introduced in this study can help clinicians for screening different patterns of language impairments in degenerative diseases, and also can help SLPs to diagnose and rehabilitate Iranian brain damaged patients.

This study is an on-going project on clinical applications of P-WAB on Iranian brain-damaged patients to improve its clinical utility. The complete sections of P-WAB-1 mentioned in this paper will be available on-line at Persian Clinical Linguistic Database (peld.uswr.ac.ir). The present available sections of P-WAB-1 for clinical applications are: 1- About P-WAB-1, 2- Guidelines to administer and scoring P-WAB-1, 3- P-WAB-1 Record Form 4-Patient's Clinical Questionnaire, 5- Bird Nest Story, 6- Guidelines to evaluate connected speech and scoring content and fluency of connected speech samples.

Acknowledgement

This research was partly supported by a grant to the first author by Iranian National Science Foundation (INSF). The authors would like to thank Dr. M. Barkatain from Division of Neuropsychiatry and Neuropsychology, Comprehensive Epilepsy Program, Kashani Hospital, Isfahan University of Medical Sciences who provided the data on the first draft of P-WAB-1 on patients with epilepsy as the control group in this study. We specially thank Dr. Mahgol Tavakoli who administered P-WAB-1 on patients with epilepsy and analyzed the data. We also would like to thank all our collaborators who collected the data for this study in different SLP clinics. Finally our special thanks go to the anonymous reviewers of the journal for their constructive comments on the earlier version of the paper.

References

- Armstrong, L. (1995). *Armstrong Naming Test*. London: Wur.
- Cynthia, M., Shewan, & Kertesz, A. (1980). Reliability and validity characteristics of the western aphasia battery (WAB). *Journal of Speech and Hearing Disorders*, 45(3), 308-324.
- Goodglass, H., Caplan, E. (1972). *The Boston diagnostic aphasia examination*. Philadelphia: LEA.
- Kertesz, A. (1982). *Aphasia and associated disorders: taxonomy, localization, and taxonomy and recovery*. N.Y.: Grune & Stratton.
- Kertesz, A. (1982). *Western aphasia battery test manual*. New York: Grune & Stratone.
- Kertesz, A. (2006). *Western Aphasia Battery- Revised (WAB-R): Examiner's Manual*. PsychCorp. Harcourt Assessment Incorporation.
- Kertesz, A., & Poole, E. (1974). The aphasia quotient: The taxonomic approach to measurement of aphasic disability. *The Canadian Journal of Neurological Sciences*, 1(1), 7-16.
- Kyoung, E. K., Hae, M. S., Han M. K., Kim, W., Han, T. R., & Paik, N. J. (2010). Severity of post-stroke aphasia according to aphasia type and lesion location in Koreans. *Journal of Korean Medical Sciences*, 25(1), 123-7.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.
- Nilipour, R. (2011). *[Aphasia naming test (Revised version) (Persian)]*. University of Welfare & Rehabilitation Sciences Press, Tehran, Iran.
- Nilipour, R. (2011). *[Persian aphasia battery (Revised version) (Persian)]*. University of Welfare & Rehabilitation Sciences Press, Tehran, Iran.
- Nilipour, R., Rezai, H., Meysami, N., Abasiyan Bidgoli, S. (2012). Neuropsychological Double Dissociation between Linguistic Levels: Clinical Linguistic Evidence from Iranian Aphasia Patients. *Journal of Basic & Clinical Neuroscience*, 4(1), 14-24.
- Paradis, M., & G. Libben. (1987). *The Assessment of Bilingual Aphasia*. LEA. Hillsdale, New Jersey.
- Paradis, M., Nilipour, R., & Dehghan, Y. (1987). *Bilingual Aphasia Test (Azari Version)*. Hillsdale, NJ, Lawrence Earlbaum Associates. USA.
- Paradis, M., Nilipour, R., & Maniei. (1987). *Bilingual aphasia test (Kurdish Version)*. Hillsdale, NJ: Lawrence Earlbaum Associates. US Febiger.
- Paradis, M., Paribakhat, T., & Nilipour, R. (1987). *Bilingual aphasia test (Farsi Version)*. Hillsdale, NJ: Lawrence Earlbaum Associates. USA.
- Zimmerman, Donald, W. (1972). Test reliability and the Kuder-Richardson formulas: derivation from probability. *Educational and Psychological Measurement*, 32(4), 939-54.