

WAB汉化版失语症诊断量表的信度与效度分析

李嘉洋*, 梁韵瑾, 欧阳嘉璐, 林柳彤, 門久扎拉

上海交通大学外国语学院, 上海

收稿日期: 2022年4月6日; 录用日期: 2022年5月5日; 发布日期: 2022年5月13日

摘要

目的: 研究西部失语症评定量表(Western Aphasia Battery, WAB)的汉化版在评测中国失语症患者中的信度与效度。方法: 应用汉化版WAB量表评测47例于上海市阳光康复中心接受治疗的失语症患者以及29例正常人, 使用SPSS 24.0统计学软件, 引用克隆巴赫-阿尔法信度系数, 评估量表的内部一致性信度, 采用结构效度、聚合效度和区分效度评估量表的效度。结果: 汉化版WAB量表信度系数在0.8以上, 整体信度系数高达0.959, 大多数小题区分度较高。结论: 汉化版WAB量表具有较好的信度、效度与结构设计, 可以用于评估中国患者失语症严重程度。但其大题间相关度过高, 部分小题区分度过低, 自发言语部分信度低, 需改进。

关键词

失语症, 西部失语症评定量表, 汉化版WAB, 信度, 效度

The Analysis of Validity and Reliability of the Chinese Version of the Western Aphasia Battery in Chinese Aphasia Patients

Jiayang Li*, Yunjin Liang, Jialu Ouyang, Liutong Lin, Zhala Menjiu

School of Foreign Languages, Shanghai Jiao Tong University, Shanghai

Received: Apr. 6th, 2022; accepted: May 5th, 2022; published: May 13th, 2022

Abstract

Objective: To investigate the reliability and validity of the Chinese version of the Western Aphasia Battery (WAB) among Chinese patients with aphasia. **Methods:** 47 patients with aphasia at Shanghai Rehabilitation Center and 29 normal people were measured by Chinese version of

WAB. To assess internal consistency reliability, Cronbach's alpha coefficient was measured. Construct-related validation, convergent validity and discriminant validity were analyzed for overall validity of Chinese version WAB. Results: The reliability of Chinese version WAB exceeds 0.8, with an overall reliability coefficient as high as 0.959 and an excellent level of discrimination for most of the subtests. Conclusions: Chinese version WAB with good reliability, validity and structural design could effectively evaluate the severity of patients with aphasia. But the correlation between the major components is too high while the distinction between some sub-tests and the reliability of the spontaneous speech section are low and need improving.

Keywords

Aphasia, Western Aphasia Battery, Chinese Version WAB, Reliability, Validity

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

失语症常见于脑卒中、脑内出血、脑部梗塞等脑部损伤后[1], 大约 15%至 42%的脑卒中患者会患上失语症[2] [3], 汉语失语症患者的言语和空间工作记忆均明显受损[4], 造成患者的语言能力, 包括听觉理解能力、口语表达能力、阅读能力、书写能力的丧失或损伤。

国内外有许多常用的失语症量表。国外的有 1962 年提出 1978 年成型的 Token 测验, 由难度逐渐增大的试题组成, 可以检查患者的听理解能力、抽象能力和记忆能力, 但无法区别患者的失语症类型, 且效度较低。1972 年编制的波士顿诊断性失语症检查(Boston Diagnostic Aphasia Examination, BDAE)则能够进行十分全面的检查, 但是花费时间长且信度较低。于是西方推出了(Western Aphasia Battery, WAB [5])作为 BDAE 的缩短版[6]。国内常用的量表还有中国康复研究中心汉语标准失语症检查量表(Chinese Rehabilitation Research Center Standard Aphasia Examination, CRRCAE), 尽管其在评价患者的训练效果方面具有较好的敏感度, 但每次评价需要间隔 3~4 周进行[7]。2003 年有国内学者设计了汉语语法检查量表(Chinese Agrammatism Battery, CAB), 其在西方国家语法量表的理论和框架下选用了依据汉语习惯和规则的语句测验失语症患者[8]。2018 年有国外学者编制了(Quick Aphasia Battery, QAB)量表, 旨在 15 分钟内对失语症患者的失语程度进行评估[9]。由于 WAB 量表不仅能对失语症患者进行十分全面的检查, 且测试时间被控制在了 1 小时以内[7], 是较为理想的失语症量表, 其汉化版被广泛使用, 但其标准化程度不高, 且未经过诊断验证[10]。因此, 本研究旨在通过分析汉化版 WAB 量表的信度与效度, 改进量表, 使得量表得以更准确地反映患者的失语程度, 更好地适用于中国失语症患者。

2. 对象和方法

2.1. 研究对象

研究对象为 47 例失语症患者和 29 例正常人。47 例失语症患者均为在上海市阳光康复中心临床诊断明确的患者, 平均年龄 50 岁, 最大年龄 79 岁, 最小年龄 20 岁。其中 29 例为男性, 18 例为女性, Broca 失语症类型 8 例, 经皮质感觉性失语症类型 6 例, 经皮质混合性失语症类型 5 例, 经皮质运动性失语症类型 10 例, 命名性失语症类型 10 例, 完全性失语症类型 6 例, 未记录失语症类型 2 例。选取 29 例正常人, 其中 12 例为男性, 17 例为女性, 平均年龄 47 岁, 最大年龄 83 岁, 最小年龄 20 岁。

2.2. 研究方法

为确保实验数据的准确, 由上海阳光康复中心的经过系统培训的专业医师作为检查者, 使用汉化版 WAB 量表对 47 例失语症患者进行评测计分, 由已学习过专业医师使用汉化版 WAB 量表的学生对 29 位正常人进行评测计分, 操作流程均按 WAB 量表使用手册进行。整个过程在被试允许的情况下进行录音, 以备数据复核。

2.3. 统计学方法

使用 SPSS 24.0 统计学软件对数据进行处理, 进行了信度检验、因子分析、相关分析等。

3. 结果

首先分析 47 例患者的量表数据, 研究量表的信度、效度, 并根据 29 例正常人的答题情况找出难度过大的题目。

3.1. 信度

本研究引用克隆巴赫-阿尔法信度系数评估量表的内部一致性, 并以此作为量表的信度标准。克隆巴赫系数 > 0.9 时表示信度极好, $0.8\sim 0.9$ 较好, $0.7\sim 0.8$ 可接受, 0.7 以下不太可信。

量表整体和各分项目的信度系数都在 0.8 以上, 整体信度系数高达 0.959 , 说明本量表具有较高的信度。在四个分项目中, 由于自发言语部分没有小题分数记录, 因此无法对本部分进行信度分析, 剩余三部分信度系数都高达 0.85 以上, 可见量表的各分项目内部一致性也较高。检验总量表及各部分内部一致性(Cronbach α 系数)见表 1。

Table 1. Cronbach's alpha coefficient for the test of intrinsic reliability of the scale

表 1. 量表内在信度检验 Cronbach α 系数

维度	条目数	克隆巴赫 Alpha 系数
总分及大题	5	0.838
整体	137	0.959
自发言语	-	-
听力理解	91	0.931
复述能力	15	0.885
命名能力	31	0.942

3.2. 效度

本研究采用结构效度、聚合效度和区分效度来评定总体效度。

3.2.1. 结构效度

本研究采用因子分析法, 对量表的结构进行效度分析。首先对量表进行 Bartlett 球形检验, 得到 KMO 值为 0.74 , 在 0.7 以上, 说明可以做因子分析。测试发现, 量表的四大组成部分初始特征值总计 > 1 , 即只可提取出一个有效成分, 且累积方差贡献率高达 70% 以上(见表 2)。同时, 该公因子在的四大组成部分的方差都大于 0.4 , 表明每部分 40% 以上的方差都可以用公共因子解释(见表 3), 说明量表有较好的结构设计。

Table 2. Total variance explanation**表 2.** 总方差解释

成分	初始特征值			提取载荷平方和		
	总计	方差百分比	累积%	总计	方差百分比	累积%
1	2.926	73.142	73.142	2.926	73.142	73.142
2	0.638	15.944	89.086			
3	0.313	7.833	96.919			
4	0.123	3.081	100.000			

Table 3. Common factor variance**表 3.** 公因子方差

部分	初始	提取
第一部分: 自发言语	1.000	0.749
第二部分: 听力理解	1.000	0.774
第三部分: 复述能力	1.000	0.490
第四部分: 命名能力	1.000	0.912

3.2.2. 聚集效度和区分效度

本研究通过比较量表各大题得分和总得分之间的相关性、各小题得分和所属大题得分间的相关性,评价量表的聚合效度和区分效度。相关系数大于 0.7 为强相关, 0.3~0.7 为中度相关, 小于 0.3 为弱相关。

每大题和总分间的相关系数都>0.7, 为强相关, 第四大题和总分的相关性系数高达 0.94, 说明四部分大题设置合理, 量表的聚集程度较好。但是不同大题之间的相关性系数都大于 0.4, 为中度相关; 第四大题和第二大题、第四大题和第一大题的相关性系数大于 0.7, 为强相关, 说明量表的区分效度较弱, 相关性显著(见表 4)。

Table 4. Pearson's correlation coefficient between major parts and the total score (two-tailed)**表 4.** 各大题与总分之间皮尔逊相关系数(双尾)

	总分	第一大题	第二大题	第三大题	第四大题
总分	1	0.897**	0.843**	0.711**	0.940**
第一大题	0.897**	1	0.686**	0.427**	0.799**
第二大题	0.843**	0.686**	1	0.473**	0.820**
第三大题	0.711**	0.427**	0.473**	1	0.608**
第四大题	0.940**	0.799**	0.820**	0.608**	1**

每小节的区分度取值在-1 到 1 之间, 大于 0.4 表示该小节区分度好, 0.3~0.39 较好, 0.2~0.29 不太好需修改, 0.19 表示不好应淘汰(见表 5 及图 1)。137 道小题中, 大部分小题的区分度好, 有 5 道小题需要修改(见表 6), 5 道小题应被淘汰(见表 7), 后续本文在讨论部分针对这些题目提出了改进建议。

3.2.3. 难度系数

本量表的难度系数及其频率分布如图 2 所示, 各小题难度系数基本服从正态分布, 因此在难度方面, 本量表设计较为合理, 效度较高。

Table 5. Distinctiveness of the questions**表 5.** 各小题区分度情况

区分度范围	频数
>0.4	111
0.3~0.39	16
0.2~0.29	5
<0.19	5

Table 6. Questions requiring revision**表 6.** 需要修改的题目

题目编号	题目	区分度
2A(5)	你住在(患者所住地址)吗?	0.289
2A(6)	你住在郑州吗?	0.217
2A(13)	这是医院吗?	0.250
2A(15)	纸能在火中燃烧吗?	0.294
2A(18)	7 月份下雪吗?	0.259

Table 7. Questions to be eliminated**表 7.** 需要淘汰的题目

题目编号	题目	区分度
2A(1)	你叫张明华吗?	0.183
2A(2)	你叫李飞翔吗?	0
2A(4)	你住在乌鲁木齐吗?	0.004
2A(19)	马比狗大吗?	0.074
2A(20)	你用斧子割草吗?	0.125

3.3. 正常人数据印证

对 29 例正常人数据着重进行难度系数的分析, 从而侧面反映汉化版 WAB 量表的区分度(见表 8)。

本量表 94% 以上的题目对正常人来说都十分简单(难度系数大于 0.9), 但仍有 8 道难度系数低于 0.9 的题目(见表 9), 对于正常人来说亦难度稍大, 不适合用于测试失语症患者的康复程度, 所以本文在讨论部分针对这些题目提出了改进建议。

4. 讨论

本研究结果发现, 汉化版 WAB 除自发言语部分外的三部分信度系数高达 0.85, 内部一致性较高, 信度较高; 有较好的结构设计, 聚集程度较好, 但区分度较弱, 有 8 道过难的题目有待修改。且由于汉化版 WAB 量表的自发言语部分未设有小题, 且自发命名大题中评分规则模糊, 其包含 7 个小题, 需要评测者根据患者回答这 7 个独立小题的信息量评 0~10 分, 再根据流畅度、语法能力和错语评 0~10 分, 但是信息量、流畅度、语法能力和错语达到多少程度算好并未给出标准给, 导致了评测者间的给分差异, 建议在这两个部分均增设可量化的分数等级供评分者参考。此外, 量表的部分题目的区分度过低, 难度过大, 需要改变题目的内容和形式以提高效率。

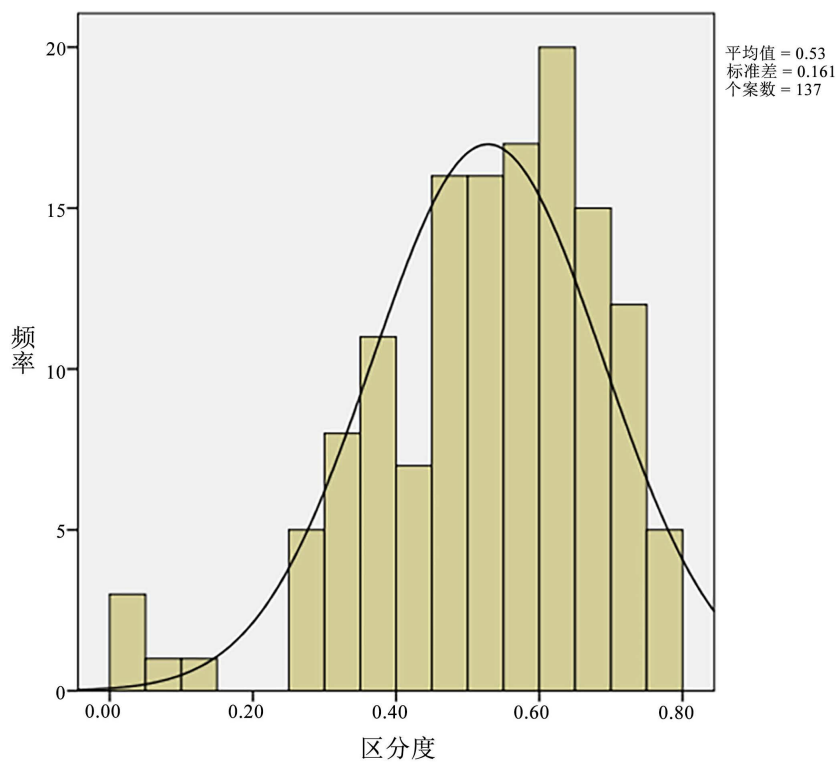


Figure 1. Frequency distribution of differentiation
图 1. 区分度频率分布图

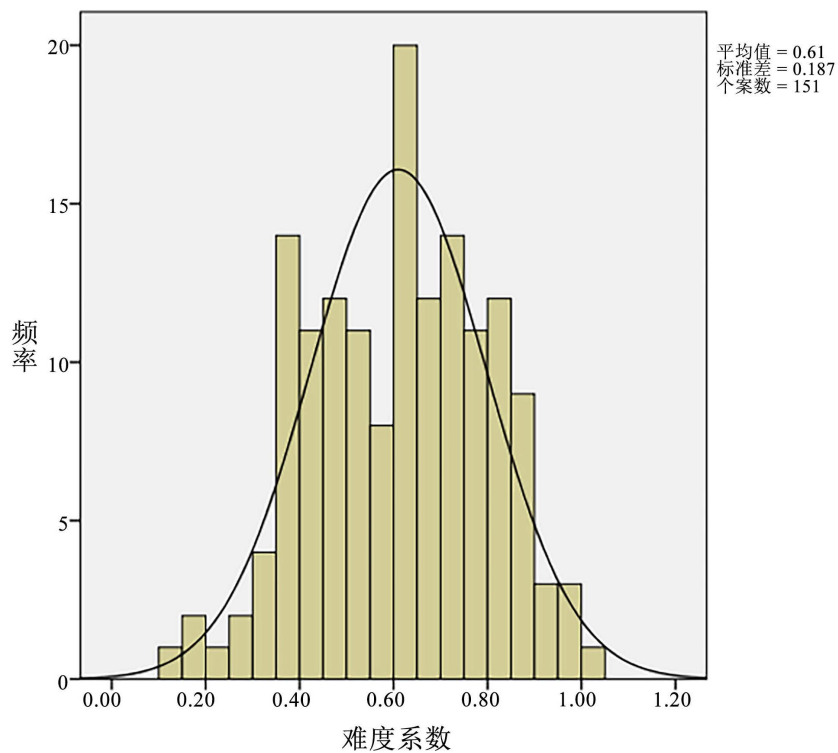


Figure 2. Difficulty coefficients for each question and their frequencies
图 2. 各小题难度系数及其频率

Table 8. Difficulty coefficients (normal)**表 8.** 难度系数(正常人)

难度系数范围	频数
>0.9	129
0.8~0.9	7
<0.8	1

Table 9. Questions with a difficulty factor below 0.9**表 9.** 难度系数低于 0.9 的题目

题目编号	题目
2A(16)	3 月比 6 月先来吗?
2B(10)	身体左右部: 右肩、左膝、左踝、右腕
4A(5)	别针
4A(13)	纸夹子
4A(18)	透明胶纸卷
4C(2)	糖是……的
4C(3)	玫瑰是红的, 向日葵是……的
4C(5)	腊八是在农历……月

本文选择具有代表性的待修改题目提出改进建议。如 2A(2): 你叫李飞翔吗? 其中“李飞翔”不是患者真名, 所以正确答案为“否”。这道题目中所有患者都答对了, 区分度为 0。因为此题与前一小题 2A(1) 和小题几乎一样, 量表应该避免形式和答案都相似性极高的题目短小时内重复出现, 从而使得患者真正经过思考理解再输出语言的过程, 反映出患者语言的真实能力。如 2A(19): 马比狗大吗? 按照常识, 马一般来说都比狗要大, 所以正确答案为“是”。本题区分度为 0.074, 低。首先, 从句法角度分析, “马比狗大吗?”并不符合汉语的语言习惯, 具有翻译的生涩感, 直接承接了英文的语序和逻辑连接词。如 2A(20): 你用斧子割草吗? 按照常识, 斧子不常用来割草, 所以正确答案为“否”。本题区分度为 0.125, 低。由于 WAB 量表的编制年份以及汉化年份都在 20 世纪, 且东西方文化差异有所不同, 所以和国内的日常生活有所脱节, 应该将“斧子”、“割草”等名词替换为更现代更日常且贴合中国文化的物品和劳动。

本研究存在一定的不足之处, 首先在于患者样本以及正常人样本量有限较小; 其次, 未进行失语症治疗前后的效果评定。未来研究可以通过扩大样本量, 同时收集治疗前后数据, 从而形成更大的数据集进行分析比对, 进一步分析量表的信度和效度问题, 为汉化版 WAB 量表在中国失语症患者的诊断和治疗评估中提供有效的临床依据。

尽管汉化版 WAB 量表存在区分效度不够高, 尤其体现在命名部分与自发言语、命名部分与听力理解上。对此, 本文提出了一系列改进量表效度的措施, 包括但不限于: 更新老旧词汇, 用语日常化, 减少翻译生涩感, 提高问题的逻辑, 以及避免相似题目短小时内高频出现等。

5. 结语

综上所述, 汉化版 WAB 量表具有较高的信度, 量表的聚集程度较好, 在同一特质的两种或多种测定方法间有较高的相关性[11], 难度基本符合正态分布, 因此, 汉化版 WAB 量表基本可以用于评估中国

患者的失语症。

参考文献

- [1] Tremblay, P. and Dick, A.S. (2016) Broca and Wernicke Are Dead, or Moving Past the Classic Model of Language Neurobiology. *Brain and Language*, **162**, 60-71. <https://doi.org/10.1016/j.bandl.2016.08.004>
- [2] Kadojic, D., Bijelic, B.R., Radanovic, R., *et al.* (2012) Aphasia in Patients with Ischemic Stroke. *Acta Clinica Croatica*, **51**, 221-225.
- [3] Ryglewicz, D., Hier, D.B., Wiszniewska, M., *et al.* (2000) Ischemic Strokes Are More Severe in Poland Than in the United States. *Neurology*, **54**, 513-515. <https://doi.org/10.1212/WNL.54.2.513>
- [4] 范舒涵, 牛星阳, 古锦敏, 等. 脑卒中后汉语失语症患者的言语和空间工作记忆障碍[J]. 中山大学学报(社会科学版), 2021, 42(5): 738-745.
- [5] Kertesz, A. (2007) Western Aphasia Battery—Revised. <https://doi.org/10.1037/t15168-000>
- [6] 曹京波, 赵纯, 金旻, 等. 失语症的常用评价方法[J]. 中国组织工程研究, 2006, 10(18): 139-141.
- [7] 张庆苏, 纪树荣, 李胜利, 等. 中国康复研究中心汉语标准失语症检查量表的信度与效度分析[J]. 中国康复理论与实践, 2005, 11(9): 703-705.
- [8] 赵丽丽, 李承晏, 毛善平, 等. 汉语语法量表及其信度和效度研究[J]. 卒中与神经疾病杂志, 2003, 10(3): 152-154.
- [9] Wilson, S.M., Eriksson D.K., Schneck S.M., *et al.* (2018) Correction: A Quick Aphasia Battery for Efficient, Reliable, and Multidimensional Assessment of Language Function. *PLoS ONE*, **13**, e0192773. <https://doi.org/10.1371/journal.pone.0199469>
- [10] Rohde, A., Worrall, L., Massey, M., *et al.* (2018) Diagnosis of Aphasia in Stroke Populations: A Systematic Review of Language Tests. *PLoS ONE*, **13**, e0194143. <https://doi.org/10.1371/journal.pone.0194143>
- [11] 蒋小花, 沈卓之, 张楠楠, 等. 问卷的信度和效度分析[J]. 现代预防医学, 2010, 37(3): 429-431.