

运用Delphi法构建基于计算机辅助语言评估系统一级条目的可行性分析

关汉添^{1*}, 周钰^{2#}, 陈卓铭³, 黄祥敏⁴, 黎莹², 曾庆⁵, 陈嘉露¹, 黄咏榆¹, 刘韵¹, 谭赐生¹, 李佩宜¹

¹广州医科大学, 广东 广州

²广州医科大学附属第二医院康复科, 广东 广州

³暨南大学附属第一医院, 广东 广州

⁴深圳市宝安中医院, 广东 深圳

⁵南方医科大学珠江医院, 广东 广州

Email: [#]1061302248@qq.com, 845989506@qq.com

收稿日期: 2020年10月6日; 录用日期: 2020年10月22日; 发布日期: 2020年10月29日

摘要

目的: 运用Delphi法构建基于计算机辅助语言评估系统一级条目, 发挥计算机的辅助作用, 通过构建计算机辅助下的语言评估检测指标, 为计算机语言智能系统量表构建提供理论基础。方法: 文献查阅借鉴国内外语言量表应用情况, 筛选一级指标, 采用Delphi法, 选取25位专家并对指标进行问卷调查和意见咨询。结果: 国内外排名前三量表汉语失语检查法(ABC)、BDAE、WAB、结合(36项目)Token (筛查), 筛选出一级条目池。参与评卷专家25名(平均年龄38.41岁), 62.5%以上专家为硕士及以上学历。专家积极系数88%, 意见提出率40%, Kendall's W系数0.361, 差异有统计学意义($P < 0.001$), 根据专家意见集中程度筛选语言一级筛查条目是: 听觉理解、自发性语言、口语表达、命名、阅读、复述。结论: 专家意见集中, 协调程度一致, 筛选出合适的语言评估系统一级条目, 为计算机构建语言评估智能系统量表提供可靠的理论基础。

关键词

德尔菲法, 语言评估, 计算机化

The Delphi Method Is Used to Construct the Feasibility Analysis of the First Level Items in the Computer Aided Language Assessment System

*第一作者。

#通讯作者。

文章引用: 关汉添, 周钰, 陈卓铭, 黄祥敏, 黎莹, 曾庆, 陈嘉露, 黄咏榆, 刘韵, 谭赐生, 李佩宜(2020). 运用 Delphi 法构建基于计算机辅助语言评估系统一级条目的可行性分析. *心理学进展*, 10(10), 1519-1528.

DOI: 10.12677/ap.2020.1010178

Hantian Guan^{1*}, Yu Zhou^{2#}, Zhuoming Chen³, Xiangmin Huang⁴, Ying Li², Qing Zeng⁵,
Jialu Chen¹, Yongyu Huang¹, Yun Liu¹, Cisheng Tan¹, Peiyi Li¹

¹Guangzhou Medical University, Guangzhou Guangdong

²Department of Rehabilitation, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou Guangdong

³The First Affiliated Hospital of Jinan University, Guangzhou Guangdong

⁴Shenzhen Bao'an Traditional Chinese Medicine Hospital, Shenzhen Guangdong

⁵Zhujiang Hospital of Southern Medical University, Guangzhou Guangdong

Email: #1061302248@qq.com, 845989506@qq.com

Received: Oct. 6th, 2020; accepted: Oct. 22nd, 2020; published: Oct. 29th, 2020

Abstract

Objective: To construct a level 1 item of the computer-aided language assessment system (CATS) based on Delphi method. **Methods:** The first level indicators were screened by referring to the application of cognitive scale at home and abroad, and 25 experts were selected by Delphi method to conduct questionnaire survey and opinion consultation on the indicators. **Results:** Chinese aphasia test (ABC), BDAE, WAB, and combination (36 items) Token (screening) of the top three scales at home and abroad were selected to screen the first-level item pool. Twenty-five experts (average age 38.41 years old) participated in the marking, and 62.5% or more of them had master degree or above. The expert positive coefficient was 88%, the opinion proposal rate was 40%, Kendall's W coefficient was 0.361, the difference was statistically significant ($P < 0.001$). The language level 1 screening items were auditory comprehension, spontaneous language, oral expression, naming, reading, and rehearsal according to the degree of expert opinion concentration. **Conclusions:** The expert opinion is concentrated, the degree of coordination is consistent, and the appropriate level 1 items of the cognitive assessment system are screened out, which provides a reliable theoretical basis for the computer to construct the intelligent language assessment system scale.

Keywords

Delphi Method, Cognitive Assessment, Computerized

Copyright © 2020 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

据世界卫生组织的调查数据显示, 2015 年人口预期寿命达 75 岁(World Health Organization, 2017), 而我国 60 岁以上人口在 2050 年将达到 31.1% (张文娟, 魏蒙, 2016)。随着社会进入老龄化的趋势不断发展, 痴呆及脑卒中发病率增高, 常伴随着语言功能障碍, 表现在语言的理解、表达、复述、命名、阅读、书写等某一方面或几方面的功能障碍, 引起语言交流受限, 导致个人生活能力的严重退步, 因此在 DSM-5 诊断标准中已将语言障碍列为痴呆的六项主症之一(刘增玲, 2013)。

临床心理学家通过心理测验的方法研究大脑和行为的的关系, 酌情使用不同的量表。然而神经心理学测验因需要专门培训, 花费时间长, 且不同评测者信度不一, 在临床应用中受到限制(刘增玲, 2013)。随着医疗科技的发展, 计算机辅助技术和虚拟语言康复逐渐应用于临床。计算机辅助下语言评估检测指标

客观、检测过程严谨明了、检测操作方便、检测评估量化客观,同时,全智能化诊断并提供诊断符合率,使其能真正做到非专业人员也可使用,对缓解我国治疗师不足有重要的现实意义(谢波,康德智,郑建民等,2010)。本研究根据使用的汉语文化及特点,在既往研究基础上,利用标准量表编制流程,筛选出合适的语言评估系统一级条目,为计算机语言智能系统量表构建提供可靠的理论基础。

2. 研究对象

遴选专家小组 25 人,专家小组成员由采用课题小组组长推荐及所选专家推荐的方式为主,推荐对象需在此研究领域具有一定权威和经验,突出专业性和敬业性。

纳入标准(敖漫,2015):1) 工作领域:临床医学、神经病学、语言学、神经心理学、流行病学和医学统计学。2) 医学专家必须来自三级甲等综合性医院,副高级职称以上。3) 本领域工作 10 年以上。4) 硕士及以上学历。5) 对此研究感兴趣,态度积极,能完成整个实验评估过程。

3. 研究方法

3.1. 语言评估系统一级指标的筛选

分析国内外使用可靠、诊断准确、应用广泛的语言评估及筛查量表,分析其构架和特点,并加以改进,确定基于计算机化语言评估筛查量表的一级指标。

3.2. 采用德尔菲法对基于计算机语言智能评估系统的指标进行评价

本课题研究以德尔菲法(曾光,1994)对指标进行评价,采取匿名方式进行函询获取专家意见,议题小组对意见汇总、修改。当查阅文献专家数量接近 25 人时进一步增加人数对预测精度影响很小(孙振球,2006)。

3.2.1. 成立课题小组

课题小组共 3 人,成员分工:1 名组员负责查阅文献,初步筛选出一级指标,初步编制专家函询问卷,统计问卷结果。1 名组员担任联络员,负责与专家沟通,向议题小组报告专家组意见。1 名组长负责统筹及协调、遴选专家,召开小组讨论后决议。

3.2.2. 设计“基于计算机语言智能评估系统”指标的函询问卷

该函询问卷包括四部分:1) 调查说明书,阐述本课题的研究背景、目的、意义。2) 专家的基本情况,包括年龄、性别、职称、学历、工作年限。3) 正文即权威程度表:包括专家对指标判断依据,认为重要程度等级。内容:包括基于计算机语言智能评估系统的一级指标,指标的重要性程度按 Likert 分级评分法划分为相应级别,按照专家意见重要程度性分 7 个级别,并赋值。1 = 很不重要、2 = 不重要、3 = 不太重要、4 = 一般重要、5 = 比较重要、6 = 很重要、7 = 特别重要。4) 专家可以根据自己的临床实践经验提出补充和修改意见,完善量表内容。

3.2.3. 发放函询问卷并收集反馈意见

在会场中以现场评价方式由联络员向专家组发放调查问卷进行评价,回收并整理意见,统计数据,议题小组召开小组会议讨论后做出决议,完善一级指标。

3.2.4. 专家一般情况

对专家的年龄、性别、文化程度、职务、工作年限、职称等情况进行描述性分析,说明专家的专业水平及经验丰富的程度。

3.2.5. 专家积极系数

专家积极系数指的是专家对本课题关心的程度,一般以专家函询问卷的有效回收率和意见提出率表示。回收率和意见数越多,说明专家越积极,对课题的合作程度越大。如果 10 天之内反馈,回收率在 70%以上,表示专家积极系数较高。

回收率=回收函询问卷数/发放函询问卷数。

意见提出率=提出意见专家数/发放专家数。

3.2.6. 专家意见的一致性

专家意见的一致性也称为集中程度,用专家问卷中重要性得分的均值和满分比(K_j)来表示,是维度、条目保留或删除的依据。

意见集中程度以专家对量表指标重要性评分的算数均数(M_j)来表示, $M_j \geq 3.50$ 为较好,说明指标越重要。

$$M_j = \frac{1}{mj} \sum_{i=1}^m C_{ij}$$

m_j 表示参加 j 指标的专家数, C_{ij} 表示 i 专家对 j 指标的分值。 M_j 越大说明 j 指标越重要。

3.2.7. 专家协调程度协调系数(Kendall's W)

协调程度指的是所有专家对全部指标给出意见的分歧程度,用和谐系数(Kendall's W)表示(德威利斯,魏勇刚,龙长权,宋武,2004)。W 取值范围为 0~1,一般在 0.3~0.5 之间,和谐系数需要经过显著性检验以保证其可靠和可信,W 越大说明专家的意见一致性越好,结果越可信。

显著性检验 - 卡方检验:在 SPSS 中采用多个相关样本的非参数检验,可计算出 χ^2 和 p 值,若 $p < 0.05$,则认为和谐系数具有显著性,

肯德尔和谐系数公式:

$$W = \frac{12}{m^2 - (n^3 - n)} \sum_{j=1}^n d_j^2$$

m 表示专家总数, n 表示指标数。

3.3. 统计学分析

利用 SPSS22 完成数据统计分析。专家一般情况、积极系数、意见集中程度采用描述性统计分析方法,协调程度用肯德尔协调系数(Kendall's W)表示,显著性检验采用卡方检验,检验水准 $\alpha = 0.05$ 。

4. 结果

4.1. 语言障碍量表文献检索

4.1.1. CNKI 知识网络服务平台的查询

① 检索方式: 关键词: (评估 or 量表 or 筛查) and (语言障碍 or 言语障碍 or 失语 or 构音障碍 or 语言发育迟缓 or 言语发育迟缓 or 言语失用 or 发声障碍 or 阅读障碍 or 理解障碍), 采用 CNKI 自带计量可视化分析结果显示: 文献总数 5190 篇, 表 1 所示文献数量呈缓慢上升趋势, 逐渐成为研究热点。

Table 1. CNKI language impairment scale annual publication volume

表 1. CNKI 语言障碍量表年度发表文献量

年度	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
文献量(篇)	444	493	535	510	521	558	551	506	521	542

在此检索结果中, 改变检索方式为: (评估 or 量表 or 筛查) and (语言障碍 or 言语障碍 or 失语 or 构音障碍 or 语言发育迟缓 or 言语发育迟缓 or 言语失用 or 发声障碍 or 阅读障碍 or 理解障碍) and (计算机), 搜索文献总数为 35 篇。

4.1.2. PubMed 的查询

① 检索方式: (screening [MeSH Terms]) OR (assessments [MeSH Terms]) OR (scales [MeSH Terms]) AND (comprehension disorders [MeSH Terms]) OR (dyslexia [MeSH Terms]) OR (vocalization disorders [MeSH Terms]) OR (speech apraxia [MeSH Terms]) OR (speech retardation [MeSH Terms]) OR (aphasia [MeSH Terms]) OR (speech disorders [MeSH Terms]) OR (Language disorders [MeSH Terms]) Filters: published in the last 10 years; Humans (Leydesdorff, Rotolo, & Rafols, 2012)。检索到 13,190 篇文献。表 2 所示文献数量呈缓慢上升趋势, 近几年成为研究热点。

Table 2. PubMed language impairment scale annual publication volume

表 2. PubMed 语言障碍量表年度发表文献量

年度	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
文献量(篇)	993	1545	1453	1587	1464	1404	1282	1403	1431	616

② 改变检索方式为(screening [MeSH Terms]) OR (assessments [MeSH Terms]) OR (scales [MeSH Terms]) AND (comprehension disorders [MeSH Terms]) OR (dyslexia [MeSH Terms]) OR (vocalization disorders [MeSH Terms]) OR (speech apraxia [MeSH Terms]) OR (speech retardation [MeSH Terms]) OR (aphasia [MeSH Terms]) OR (speech disorders [MeSH Terms]) OR (Language disorders [MeSH Terms]) AND (computers [MeSH Terms]) Filters: published in the last 10 years; Humans。检索到 65 篇文献。

国内外研究文献结果表明: 一、语言障碍相关评估筛查量表逐渐成为研究热点领域。但计算机结合语言障碍量表等领域研究较少。因此, 基于计算机化的语言障碍量表的研究或可成为新的研究方向。

4.2. 语言智能评估系统一级指标筛选

4.2.1. 国内语言障碍筛查量表应用情况

搜索公认的语言障碍评估筛查量表, CNKI 以全文词进行检索, 检索条件: 时间为 2010 年 1 月 1 日至 2019 年 12 月 31 日这 10 年的文献。检索方式为高级检索。数据库选择为跨库(期刊、特色期刊、博士、硕士、国内会议、国际会议、报纸), 文献分类选择: 所有。词频: 模糊。中文文献。PubMed 以 All Fields 搜索公认的语言障碍评估筛查量表。检索限制条件: 时间: 2010 年至 2019 年。种类: 人类。检索方式为高级检索。

中国知网文献查询系统中语言障碍筛查评估量表数量最多的前三位为汉语失语成套测验(ABC 测验)、汉语标准失语症检查(CRRCAE)、Token 测验; 维普中文科技期刊数据库中文献数量最多的前三位: 汉语标准失语症检查(CRRCAE)、西方失语症成套测验(WAB 测验)、汉语失语成套测验(ABC 测验); PubMed 检索系统文献数量最多的前三位: BDAE、WAB、SLTA。三大文献检索系统中文献数量均大于 1000 条的量表有: BDAE、汉语失语成套测验。均小于 100 条的有: 语言障碍诊治仪、汉语波士顿失语症检查法、Token 测验。详见表 3、表 4。

4.2.2. 语言量表排名前三位量表的一级指标框架

把排名前三位的语言评估量表, 与中国汉语失语检查法(aphasia battery of Chinese, ABC) (汉语失语症康复治疗专家共识组, 2019)、简易语言筛查量表是: 36 项目 Token 相结合, 把语言量表的一级指标(维度)列出, 定为第一轮专家评估内容。详见表 5。

Table 3. Research literature on the language impairment screening scale in China**表 3.** 国内语言障碍筛查量表研究文献

量表	CNKI 相关文献数量	维普相关文献数量
汉语失语成套测验(ABC 测验)	12,213	48
汉语标准失语症检查(CRRCAE)	836	106
Token 测验	734	0
波士顿诊断失语症检查(BDAE)	654	4
西方失语症成套测验(WAB 测验)	502	59
日本标准失语症检查(SLTA)	414	1
汉语波士顿失语症检查法	195	13
语言障碍诊治仪	150	7
Paradis 的双语失语症检测(BAT 测试)	2	0

Table 4. Foreign research literature on the language impairment screening scale**表 4.** 国外语言障碍筛查量表研究文献

量表	PubMed 相关文献数量
BDAE (Boston diagnostic aphasia examination)	11,841
WAB (The western aphasia battery)	134
SLTA (standard language test of aphasia)	84
Token examination (Token battery)	59
BAT (the Bilingual aphasia test)	28
CRRCAE (China rehabilitation research center aphasia examination)	2
Chinese aphasia test battery (ABC test)	13
Chinese Boston aphasia test	2
Language disorder diagnosis apparatus ZM2.1	0

Table 5. Construction of the top three language scales and simple Token Level 1 indicators at home and abroad**表 5.** 国内外排名前三位语言量表+简易 Token 一级指标构建

量表	一级指标	汉语失语检查法(ABC)	BDAE	WAB	(36 项目) Token (筛查)
1		谈话	对话和自发言语	自发言语	听理解
2		理解	听觉理解	理解检查	
3		复述	言语表达	复述	
4		命名	书面语理解	命名	
5		阅读	书写	阅读	
6		书写		书写	
7		结构与视空间		运用	
8		运用		结构能力	
9		计算		视空能力	
				计算能力	

4.3. 采用德尔菲法对语言智能评估系统指标予评价

4.3.1. 专家基本情况

按照专家纳入标准,本研究纳入专家人数 25 人。第一轮发放函询问卷 25 份,回收 22 份,无效量表 3 份。25 位专家中来自三甲综合医院 19 人,高等综合大学 8 人。教育背景均为本科以上,其中博士 11 人。职称有 9 个初级,5 个中级,11 个高级。专业背景:临床医学 16 人(9 人从事神经康复专业、1 人从事儿童康复医学、6 个语言学专业)、神经心理学 1 人,特教 2 人,市场部 1 人,治疗师 5 人。平均年龄为 38.41 岁。对研究领域均为熟悉或很熟悉(表 6)。

Table 6. Profile of experts (n = 25)

表 6. 专家基本情况(n = 25)

基本情况		人数	百分比(%)
年龄	20~29	7	29.17
	30~39	6	25
	40~49	8	33.33
	50~	3	12.5
职务	院长/副院长	3	12.5
	科室主任/副主任	4	16.7
	其他	17	70.8
职称	副主任医师	1	4.1
	主任医师	1	4.1
	副教授	3	12.5
	教授	6	25
	其他	13	54.2
文化程度	本科	9	37.5
	硕士	4	16.7
	博士	11	45.8
从业年限(年)	1~10	12	50
	10~19	2	8.2
	20~29	8	33.33
	30~	2	8.2
研究生导师	硕士生导师	10	62.5
	博士生导师	2	18.7
工作单位	综合三甲医院	13	81.2
	高等综合大学	3	18.7
从业领域	临床医学	12	75
	流行病学和医学统计学	1	6.2
	神经语言心理学	3	18.7

4.3.2. 专家积极系数及一级指标修改建议

意见回收：第一轮专家积极系数为 $22/25 = 88\%$ ，有 10 位专家提出意见，意见提出率 $10/25 = 40\%$ (表 7)。专家选择的一级条目中所提意见有以下几点：自发性语言改为自发表达；听觉理解改为听理解；书面语言理解改为文字理解；听觉理解改为听；口语表达改为说；书面语言理解改为读；书写改为写。

Table 7. Positive coefficient of experts

表 7. 专家积极系数

轮次	发放	回收	回收率(%)	意见率(%)
第一轮	25	19	76	40

4.3.3. 语言专家意见集中程度，详见表 8

语言指标的专家意见集中程度为：15 项指标的平均值为：2.4。故选用听觉理解、自发性语言、口语表达、命名、阅读、复述、书面语理解。因自发性语言表达与口语表达(陈卓铭, 2004)均属于语言表达内容，归为一类；按照语言神经心理(陈卓铭, 2015)，将书面语言理解与阅读归为一类，在阅读的 90 分数档的题目中增加书面语言理解内容；其中听觉理解分为听辨认、听是否，因听指令(周钰, 陈卓铭, 2019)在语言题目中定向题、注意均涵盖，故听觉理解暂列出听辨认、听是否(陈卓铭, 2004)题目。按照专家意见，语言一级指标为：听理解、自发性表达、命名、阅读、复述。本研究依据专家意见对条目进行删改、调整，最终形成了语言一级条目。

Table 8. Concentration of language experts' opinions

表 8. 语言专家意见集中程度

指标	均数
1 听觉理解	6.11 ± 0.33
2 自发性语言	4.65 ± 0.71
3 口语表达	3.06 ± 0.52
4 命名	3.00 ± 0.33
5 阅读	2.72 ± 0.44
6 复述	2.56 ± 0.39
7 书面语言理解	2.56 ± 0.50
8 指令	2.11 ± 0.32
9 书写	1.94 ± 0.30
10 会话	1.72 ± 0.34
11 运用	1.44 ± 0.18
12 计算	1.33 ± 0.24
13 画图	1.00 ± 0.00
14 视空能力	1.00 ± 0.00
15 结构能力	1.00 ± 0.00

4.3.4. 语言一级指标专家意见协调程度，见表 9

专家意见协调程度对单个一级指标用变异系数表示，所有专家对全部一级指标的协调程度用

Kendall's W (宋铁, 代吉亚, 吴发好等, 2014)表示。第一轮语言一级指标的统计结果的协调系数 Kendall's W 为 0.361, 协调系数经检验后具有显著性, $p = 0.000$, 说明专家协调程度好。

Table 9. Level of coordination of the views of language experts
表 9. 语言专家意见协调程度

轮次	指标	Kendall's W	χ^2	P 值
第一轮	语言	0.361	95.967	0.000

5. 讨论

语言交流是一种复杂的心理活动, 涉及人对语言信息的获取、转化、贮存、衍生及表达。在语言交流过程中任何一个环节的损害均可导致语言功能的缺失, 出现语言障碍。所以尽早对痴呆、脑功能障碍的患者在这种情况下, 进行积极、相对准确的语言功能障碍的评估, 制定合理的治疗方案(陈卓铭, 2015), 尽早改善患者的语言交流能力, 促进参与日常生活和社会实践活动, 减轻家庭及社区负担。

分析国内外语言障碍量表文献检索情况, 可得出以下两点: 1) 结合中国文化及语言特色, 国内应用最多的语言障碍评估量表是汉语失语成套测验(ABC 测验)、汉语标准失语症检查(CRRCAE); 国外应用最多的语言障碍评估量表 BDAE、WAB, 虽然 BDAE 已在国内汉化, 但因为语言文化在国内不融合, 在国内尚未普及应用。2) 构建语言障碍评估指标仍需借鉴国内外应用较广的语言评估量表, 在既往研究基础上, 根据中国汉语文化及语言特点进行构建。目前传统的语言障碍评定主要依赖于治疗师使用检测表的主观判断, 无法进行定量分析。随着计算机科学的发展, 特别是多媒体技术、模糊数学、语音识别等技术的应用(陈卓铭, 2004), 目前语言障碍的临床诊断评定进入了一个新的时代。

国外有 Parrot software (Kambeyanda, Singer, & Cronk, 1997)系列产品, 其主要基于英语语言康复训练。国内暨大陈卓铭团队自主研发并转化临床使用的计算机辅助汉语失语症评估软件包括: 语言障碍诊治仪 ZM2.1、失语症计算机评测系统、语言语言训练评估系统等。语言障碍 ZM2.1 诊断设计是基于语言链中每一个环节检测及计算机智能运算的基础之上, 通过优选各种失语症检查方法的敏感指标, 结合汉语和计算机应用的特点设计, 实现病历管理 - 检测评估 - 各残存功能显示 - 康复建议 - 康复实施等程序化管理。该检查可实现自动分析量、语速等语音参数, 并根据汉语特点设计了利手检测。通过对听、视、语音、口语表达四部分共 65 题检查, 囊括表达、理解、复述、命名、阅读等失语症检测的各项指标。进行了信校标研究(陈卓铭, 李巧薇, 唐桂华等, 2006; 陈卓铭, 林玉萍等, 2005), 王红等(李涛, 王红等, 2003)进行了相关性研究, 其检测结果可用于失语症患者语言理解功能的评估, 并为其语言康复治疗提供理论依据。

Delphi 法, 又称专家调查法, 通过反馈匿名函询, 就一定的问题向相关领域的专家进行咨询, 然后对专家的答复意见进行系统的总结归纳, 最终形成意见趋于集中、可靠性较强的意见。本研究首先通过国内外文献查询, 把排名前三位的语言评估量表, 与中国汉语失语检查法(aphasia battery of Chinese, ABC) (汉语失语症康复治疗专家共识组, 2019)、简易语言筛查量表是: 36 项目 Token 相结合, 把语言量表的一级指标(维度)列出, 作为 Delphi 专家咨询法第一轮专家打评内容, 运用统计学方法对专家积极系数意见集中程度及协调程度进行统计分析。本研究根据专家建议修改意见及神经心理通路(Vallila-Rohter & Kiran, 2013), 选用听理解、自发性表达、命名、阅读、复述为一级指标, 相应的题目在视觉上呈现最终形成了语言一级条目, 能较全面反映语言各维度变化, 研究方法科学可靠。

本研究在文献分析、专家咨询及基于计算机的基础上, 结合选题规则, 遵循科学的量表编制原则和统计学分析方法, 完成语言智能评估系统的一级指标的初步构建。相对于翻译修订国外的量表, 本量表

植根于中国本土文化，更具有中国特色，被试更容易理解，同时更符合中国的国情和实际情况。

参考文献

- 敖漫(2015). *ICU 后综合征评估量表的编制研究*. 硕士论文, 荆州: 长江大学.
- 曾光(1994). *现代流行病学方法与应用*. 北京: 北京医科大学、中国协和医科大学联合出版社.
- 陈卓铭(2004). 言语治疗. 见: 黄东锋(编). *临床康复医学(下册)* (页 908-1021). 广东: 汕头大学出版社.
- 陈卓铭(2015). 汉语语言心理加工与失语症评估. *中国康复医学杂志*, 30(11), 1091-1094.
- 陈卓铭, 李巧薇, 唐桂华, 等(2006). 语言障碍诊治系统 ZM2.1 诊断亚项的正常范围研究. *中华物理医学与康复学杂志*, 28(3), 194-196.
- 陈卓铭, 林玉萍, 等(2005). 语言障碍诊治仪 ZM2.1 对失语症患者语言功能评定的效度和灵敏度. *暨南大学学报(医学版)*, 26(4), 552-555.
- 德威利斯, 魏勇刚, 龙长权, 宋武(2004). *量表编制: 理论与应用*. 重庆: 重庆大学出版社.
- 汉语失语症康复治疗专家共识组(2019). 汉语失语症康复治疗专家共识. *中华物理医学与康复杂志*, 41(3), 161-169.
- 李涛, 王红, 尹义臣, 等(2003). 计算机测定失语症语速的相关分析. *中国康复*, 18(6), 341-343.
- 刘增玲(2013). *无痴呆型血管性认知障碍患者认知功能量表与外周血炎症性标志物的研究*. 硕士论文, 苏州: 苏州大学.
- 宋铁, 代吉亚, 吴发好, 等(2014). 基于德尔菲法和层次分析法的疾控机构应急能力评估. *华南预防医学*, (1), 1-6.
- 孙振球(2006). *医学综合评价方法及其应用*. 北京: 化学工业出版社.
- 谢波, 康德智, 郑建民, 等(2010). 计算机辅助语言评定系统的编制及其灵敏度和特异度检测. *中国临床神经外科杂志*, 15(8), 471-473.
- 张文娟, 魏蒙(2016). 中国人口的死亡水平及预期寿命评估——基于第六次人口普查数据的分析. *人口学刊*, 38(3), 18-28.
- 周钰, 陈卓铭(2019). 精准璟云平台语言训练治疗语言障碍合并失语症病例报告. *康复学报*, 29(4), 54-60.
- Kambeyanda, D., Singer, L., & Cronk, S. (1997). Potential Problems Associated with Use of Speech Recognition Products. *Assistive Technology*, 9, 95-101. <https://doi.org/10.1080/10400435.1997.10132301>
- Leydesdorff, L., Rotolo, D., & Rafols, I. (2012). Bibliometric Perspectives on Medical Innovation Using the Medical Subject Headings (MeSH) of PubMed. *Journal of the American Society for Information Science & Technology*, 63, 2239-2253. <https://doi.org/10.1002/asi.22715>
- Vallila-Rohter, S., & Kiran, S. (2013). Non-Linguistic Learning in Aphasia: Effects of Training Method and Stimulus Characteristic. *American Journal of Speech-Language Pathology*, 22, S426-SS427. [https://doi.org/10.1044/1058-0360\(2013/12-0087\)](https://doi.org/10.1044/1058-0360(2013/12-0087))
- World Health Organization (2017). *Global Health Observatory (GHO) Data*.