

· 心理卫生评估 ·

中文版韦氏成人智力量表第四版的常模制订^{*}

崔界峰¹ 王健² 范宏振¹ 姚晶¹ 陈楠¹ 段京辉¹ 邹义壮¹

(¹北京回龙观医院,北京 100096 ²中国中医科学院广安门医院,北京 100053 通信作者:王健 wjmd@263.net; 邹义壮 yzouy@263.net)

【摘要】目的:通过分析韦氏成人智力量表第四版(WAIS-IV)常模数据来证实常模制订过程和方法的科学性和有效性。方法:详细介绍WAIS-IV具体修订过程、计算机辅助系统开发和常模取样计划,分析1757例实际常模数据的分布情况。结果:在全国区域分布方面,与计划取样相比,华北和东北地区取样偏多($\chi^2=78.02, P<0.0001$);在9个年龄段分布方面,除16~17岁和65岁以上年龄段的高学历样本以及30~34岁年龄段的低学历样本取样偏少外,而其余年龄段均达到计划取样要求;在性别分布方面女性稍多于男性,但差异无统计学意义($\chi^2=2.28, P=0.131$);在教育程度分布方面,基本符合计划取样比例($\chi^2=2.74, P=0.603$);在职业、当地居住年限和户籍分布方面,基本符合预定目标。结论:中文版WAIS-IV实际常模数据基本符合计划样本分布,采用的修订和统计学方法适当,为全国常模数据的代表性和可靠性提供了保证。

【关键词】智力;韦氏智力量表;成人智力测验;测验修订;常模

中图分类号: B841.7 文献标识码: A 文章编号: 1000-6729(2017)008-00635-07

doi: 10.3969/j.issn.1000-6729.2017.08.010

(中国心理卫生杂志,2017,31(8):635-641.)

Norm development of the Chinese edition of Wechsler Adult Intelligence Scale-Fourth Edition

CUI Jie-Feng¹, WANG Jian², FAN Hong-Zhen¹, YAO Jing¹,
CHEN Nan¹, DUAN Jing-Hui¹, ZOU Yi-Zhuang¹

¹Beijing Huilongguan Hospital, Beijing 100096, China ²Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing 100053, China

Corresponding authors: WANG Jian, wjmd@263.net; ZOU Yi-Zhuang, yzouy@263.net

【Abstract】 Objective: By analyzing the norm data results of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) in China, to prove the validity of the procedure and methods during the norming development. **Methods:** The whole process of the revision of WAIS-IV, the development of computer-assisted system and norm sampling plan, were introduced in more detail, and the distribution of actual norm data of 1757 cases was analyzed. **Results:** For area distribution, compared with planned sampling number, the number from North and Northeast China was statistically significant different ($\chi^2=78.02, P<0.01$). For age stages distribution, most of cases conformed to the requirements of sampling, except that some cases including high-level educational cases aged 16-17 years and above 65 years, and low-level educational cases aged 30-34 years were less than the planned sampling number. For gender distribution, male subjects were more, but there was no statistically significant difference between male and female subjects ($\chi^2=2.28, P=0.131$). For educational degree distribution, the sampling conformed to the requirements of sampling plan ($\chi^2=2.74, P=0.603$). For occupation, resident years and registered permanent residence, and the sample was basically representative. **Conclusion:** The process of the revision of the Chinese version of WAIS-IV is appropriate, and actual norm sampling basically conforms to planned sample distribution, providing the sufficient representativeness and reliability for national norm data of WAIS-IV.

【Key words】 intelligence quotient; Wechsler Intelligence Scale; Adult Intelligence Test; test revision; norm

* 北京科委2006重大项目 D0906001000091; 首发重点攻关基金项目(2007-1019); 首都医学发展科研基金项目(2009-3149)

(Chin Ment Health J, 2017, 31(8): 635-641.)

韦氏成人智力量表(Wechsler Adult Intelligence Scale-Fourth Edition, WAIS-IV)是以16岁以上的成年人为对象,进行的认知能力评估和个别施测的测评工具。WAIS-IV是当今国际心理学界公认的已被广泛运用的个别智力测验,它的临床效度高、信息丰富、相关的结果解释及科研文献极为丰富,不但广泛应用于医学、教育、残疾评定、职业评估、能力测定等领域,而且心理学工作者常把它作为金标准对其他有关智力的测验进行效度检验。WAIS测验最早在1955年由美国心理公司(The Psychological Corporation)(现在并入Pearson公司)出版,1981年有修订版(WAIS-R)^[1],1997年推出第三版(WAIS-III)^[2],2009年出版第四版(WAIS-IV)^[3]。

在我国,目前所使用的韦氏成人智力量表是湖南医科大学龚耀先等于1981年以WAIS(1955年版本)为蓝本进行修订,结合我国当时的文化背景对某些条目进行了改编,称“中国修订韦氏成人智力量表(WAIS-RC)^[4]”。三十多年过去了,受弗林效应(Flynn effect)的影响,中国修订韦氏成人智力量表已不适合目前国内发展现状,测验条目内容过时,常模数据作为参照的准确性相应下降。为了弥补这一学术空缺,2007年由北京回龙观医院与美国Pearson公司合作,对WAIS最新版(2008年确定为WAIS-IV)进行正版引进、翻译和修订,以及测验的本土化、标准化和常模建立等工作,并采用计算机信息技术,将测验进行计算机软件程序化。

中国人口众多,地区差异很大,制订一个具有代表性的全国常模是一项巨大的工程,会涉及很多技术和现实的难题,如样本量问题、城镇和农村常模问题、常模的统计学方法问题等等。WAIS-IV的常模制订过程引进了一些新的理念和技术,做了一些有益的尝试。本文旨在通过分析WAIS-IV常模数据来证实常模制订过程和科学性,为解决上述技术和现实难题提供一些思路。

1 对象与方法

1.1 对象

研究对象来自于1757例WAIS-IV全国常模取样人群。入组标准:①被试年龄在16岁以上,中文是第一语言;②被试须有正常的精细和粗大运动

能力;③被试须具备用手摆放木块以及用铅笔画图的能力;④被试须能理解分测验的指导语并完全投入测试。排除标准:①有影响测试的未矫正的视力或听力障碍;②患有神经心理障碍或精神类疾病或精神发育迟滞;③有过由脑外伤或其他原因引起的20分钟以上的意识丧失史,或目前或最近2月内有认知功能性障碍;④目前或在过去的两个月内接受化疗或放疗;⑤目前诊断为物质滥用或依赖,或在过去1年内有物质滥用或依赖史;⑥正接受可能影响测试成绩的药物治疗(如抗癫痫药,抗精神病药、苯二氮草类药物、精神兴奋剂、阿片类、三环类抗抑郁药);⑦被试是主试的亲密朋友或家人;⑧过去的6个月里接受过任何韦氏或其他智力测试,或者受过心理专业水平训练的或熟悉韦氏智力量表施测的人;⑨某个被试的双胞胎兄弟姐妹;⑩其他各种不能进行有效评估的不依从测试情形。

1.2 工具

1.2.1 测验结构

根据当今智力理论、认知过程理论及神经心理学研究成果,WAIS-IV对测验结构进行了较大调整和修正。降低了操作测验中速度的比重,拓展分数区间(40~160分)和年龄范围(16~90岁),为低能力端和高能力端设计了许多新条目,在常模样本中加入精神发育迟滞、超常智力等特殊被试,提高常模样本的代表性。并且增加了评估流体智力、工作记忆和加工速度的分测验,使量表整体更符合当代认知理论的模式。

WAIS-IV包括积木(block design, BD)、类同(similarities, SI)、背数(digital span, DS)、算术(arithmetic, AR)、矩阵推理(matrix reasoning, MR)、词汇(vocabulary, VC)、符号检索(symbol search, SS)、拼图(visual puzzles, VP)、译码(coding, CD)、常识(information, IN)10个核心分测验,涵盖了言语理解、知觉推理、工作记忆和加工速度4个认知指数。言语理解指数(verbal comprehension index, VCI)相当于WAIS-RC的言语智商(VIQ),知觉推理指数(perceptual reasoning index, PRI)相当于WAIS-RC的操作智商(PIQ),VCI与PRI合成算出一般能力指数(general ability index, GAI)。同时又增

www.cmhj.cn

加了工作记忆 (working memory index, WMI) 和加工速度 (processing speed index, PSI) 两个指数, WMI 和 PSI 合成算出认知效率指数 (cogni-

tive proficiency index, CPI), 4 个指数合成算出总智商 (FIQ), 详见图 1。

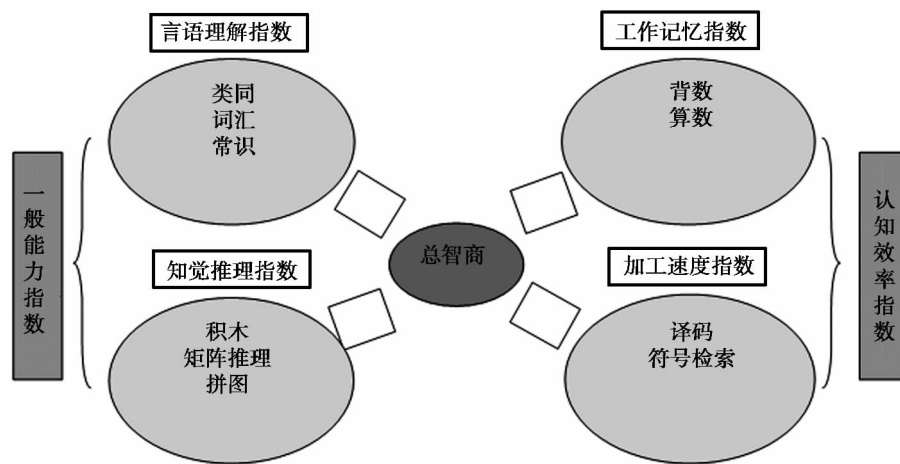


图1 WAIS-IV 中文版的测验结构

1.2.2 测验修订

①制定方案: 修订的准备工作从查阅相关文献, 明确修订目标, 制定基本方法步骤与实施方案开始。

②翻译和回译: 在测验内容方面, 因语言和文化的差异, 重点针对 WAIS-IV 三个言语分测验, 常识、类同、词汇分测验, 先进行翻译 (初译、一校、二校和三校) 和回译, 其中二校、三校和回译者曾经和现正在英语国家工作生活的精通中英文的心理学专家。

③确定测验条目: 在确定测验条目时, 尽量保留 WAIS-IV 原版的条目, 参考 WAIS-RC、WAIS-III 台湾中文版及韦氏儿童智力量表 (WISC) 第四版, 来选取和拟定用于预测的言语测验条目, 在与国际接轨的同时, 兼顾语言文化上的差异。在条目数量上比原题多准备了 20%, 以筛选更优的条目。另外分别让心理测量学专家与语言学专家提出意见, 并进行修订, 修改了部分不符合语言文法的部分参考答案和调整了部分与言语能力不相符的答案评分, 保证了文化兼容性。

④预测试: 根据人口统计学资料, 按各年龄段、教育程度分布取样, 性别比率 1:1。共计划在北京取 130 例的小样本进行预测试, 包括 30 例临床样本 (北京回龙观医院住院患者)。实际完成正常人 109 例, 临床病例 30 例。

⑤分测验参考答案整理及修订: 将各测验被试

原始答案录入电脑, 整理分类, 并参考标准答案进行评分。各分测验所保留题项, 其标准答案与参考答案也予以保留, 个别文化不符的则予以修改。常识分测验增题的标准答案来源于百科全书及其他工具书的正确答案。类同分测验增题的标准答案为借鉴以前各韦氏版本的参考答案或归纳所得, 词汇分测验增题的标准答案来源于新华词典及辞海。将初次评分的各原始答案进行整理, 制定言语测验参考答案初版。将言语测验参考答案初版分发给项目小组成员, 各成员提出修改意见并进行小组讨论修订, 形成言语测验参考答案版本初次修订版。将修订版分别让心理测量学专家与语言学专家提出意见, 并进行修订, 修改了部分不符合语言文法的部分参考答案和调整了部分与言语能力不相符的答案评分, 形成言语测验参考答案二次修订版。

⑥再评分及条目分析: 参考言语测验参考答案二次修订版对预测试被试答案进行重新评分。计算总分并进行条目分析。将正常组按例数基本均等原则, 按年龄排序分为三组, 共计算三组的难度、区分度及测验信度, 以及总的难度、区分度、测验信度; 临床组与正常组分开计算难度、区分度、测验信度。另外请统计学家做条目反应分析, 计算了分测验中各题项所反映总测验的信息量及冗余程度。删除各测验信息量指标 $infit$ 与 $outfit < 0.7$ 及 > 1.2 的条目, 再结合难度梯度、区分度 0.2 以上 (部分难度较小或较大的题项区分度 0.1 也予以保

留), 保证每个难度梯度条目数均衡, 同一难度梯度选取区分度高的条目, 部分难度梯度, 集中在难度较低和较高的难度梯度, 拟增加 1~2 条目。

⑦增补题的预测试: 选取临床样本 30 例, 对增加条目进行预测试, 选取难度与区分度合适的条目, 并增加备选条目 1 条备用。

⑧标准化版本测验条目的确定: 类同分测验共 22 条目, 比最终版要多 4 条目。如将叉子-勺子 (Fork-Spoon) 更换为“筷子-勺子”。词汇分测验共 39 条目, 比最终版要多 9 条目。删去了不合适的词汇, 增加了“绰号”等词汇。常识分测验: 共 30 题, 比最终版要多 4 条目。删去了不适合于中国文化的条目, 增加了“清朝最后一位皇帝是谁?”等条目。用于标准化取样的中文版中, 非言语测验基本保留原版式样, 只是在符号检索分测验中, 将“NO”改为“无”, 其余均未做调整。所有分测验的指导语都严格按照原版翻译并润色。

⑨测验的计算机辅助化: 在测验系统中, 有些分测验完全采用计算机条目呈现和记录, 如背数、算术、矩阵推理、拼图和积木; 有些分测验采用计算机条目呈现和记时, 以及纸质答题卡原始记录相结合的方式, 如类同、词汇、常识; 有些分测验仍完全保留纸质答题卡的形式, 计算机只对测验进行记时和记分, 如符号检索、译码等分测验。为保持测验的生态性, 指导语要求主试来说, 条目呈现采用计算机标准语音播放, 全部测验过程采用标准化操作。由心理学家和软件工程师共同编制计算机辅助测验系统, 对测验系统进行检测, 发现故障并修改。对计算机测验系统(硬件和软件)的记分系统、相关数据进行人工核对和调试。

1.3 常模取样方案

抽样总体为全国 16 岁以上人口。依据北京大学人口研究所^[5-6]提供的 2005 年 1% 抽样城镇人口统计学资料(不包括农村), 并根据人口预测原理进行适当调整(中国社科院人口与劳动经济研究所王广州研究员指导)^[7-8]。取样变量包括年龄、性别、教育程度 3 个主要变量。地区、职业、居住年限和户籍, 作为次要变量。

①年龄段: 16~17、18~19、20~24、25~29、30~34、35~44、45~54、55~64、65 岁以上 9 个年龄段。按照全国人口年龄、性别、受教育程度比例, 每个年龄段拟取 200 例, 理论取样

1800 例, 实际有效取样 1757 例。

②性别: 男和女, 基本上按照 1:1 比率取样。但对于 65 岁以上年龄组, 鉴于人数比例, 可适当增加女性被试的选取。

③教育程度: 分 5 种教育程度: 小学及以下、初中、高中、大专、本科及以上。其中中专和职中视为高中水平。

④地区: 六大行政区, 华北, 东北, 华东, 中南, 西南, 西北。每个行政区取 2~4 个取样点, 全国共 16 个取样点。

⑤职业: 包括国家公务员、专业技术人员、职员、企事业管理人员、工人、农民、学生、现役军人、自由职业者、个体经营者、无业人员、退休人员和其他, 共 13 种职业。为避免因被试职业倚倚而缺乏广泛代表性, 控制一种职业被试的最大取样量 < 总取样量的 10%。

⑥当地居住年限: 规定在城镇居住年限 < 0.5 年的农村户籍人群要小于 10%。

⑦户籍: 取样对象是居住在城镇的人群, 因人口流动现状, 户籍问题并不做严格限制, 但整体上应以城镇居民为主。

1.4 统计方法

对全国收集来的测验数据进行整理, 计算原始分、标准分、合成分, 过程分、标准误和可信区间, 采用连续常模拟合技术 [由美国 Pearson 公司朱建军博士 (Dr. Zhu J.) 指导]^[9] 制定全国常模, 计算出测验的信度(内部一致性、评分者间一致性信度、重测信度、分半信度)和效度(内容效度、探索性因素分析、验证性因素分析得出的结构效度、与 WAIS-RC 的关联效度、特殊样本的临床效度)。WAIS-IV 中文版的上述心理测量学具体结果参见相关论文^[10-16]。

2 结果

2.1 样本行政区域、性别和受教育程度分布

在六大行政区中, 华北地区、东北地区、华东地区、中南地区、西南地区和西北地区的计划取样例数与实际取样例数差异有统计学意义。华北和东北地区取样偏多, 华东和中南地区取样偏少, 导致了这种差异。性别分布差异无统计学意义。受教育程度分布按 2005 年人口学抽样数据计算, 小学及以下为 606 例、初中 747 例、高中 321 例、大专 81

例、本科及以上学历45例，其中低教育程度所占比例过高，明显不符合取样时2012年的实际情况，影响未来常模的使用。参照由生命表、升学率和死亡率等参数所建立的人口预测模型，对各年龄段例群的受教育程度比例进行了调整，计划取样中小学及以下为323例、初中649例、高中412例、大专218例、本科及以上学历198例。在实际取样例数中，小学及以下为340例、初中624例、高中395例、大专226例、本科及以上学历172例，差异无统计学意义（表1）。

2.2 样本年龄段分布

年龄段分布：按照取样方案，每个年龄段200例，9个年龄段计划取样1800例，男女各半，受教育程度参照全国人口学比例来抽取。实际有效常模取样1757例，各年龄段人数见表2。因为缺少16~17岁和65岁以上年龄段的高学历样本、以及30~34岁年龄段的低学历样本，所以导致这3个年龄段取样偏少（180例左右），而其余年龄段均达到计划取样要求（200例左右），整体上基本符

合取样计划。

表1 全国常模计划取样与实际取样在行政区域、性别和受教育程度变量上的数据分布 [例数 (%)]

变量	计划取样 (n=1800)	实际取样 (n=1757)	χ^2 值	P 值
行政区域			78.02	<0.001
华北地区	213(11.8)	326(18.6)		
东北地区	151(8.4)	225(12.8)		
华东地区	524(29.1)	438(24.9)		
中南地区	506(28.1)	344(19.6)		
西南地区	273(15.2)	269(15.3)		
西北地区	133(7.4)	155(8.8)		
性别			2.28	0.131
男性	900(50)	834(47.5)		
女性	900(50)	923(52.5)		
受教育程度			2.74	0.603
小学及以下	323(17.9)	340(19.4)		
初中	649(36.1)	624(35.5)		
高中	412(22.9)	395(22.5)		
大专	218(12.1)	226(12.9)		
本科及以上学历	198(11.0)	172(9.8)		

表2 全国常模样本中9个年龄段按性别和受教育程度的数据分布 (例数)

年龄段 /岁	合计 /例	小学及以下			初中			高中/中专			大本			本科及以上学历		
		男	女	总计	男	女	总计	男	女	总计	男	女	总计	男	女	总计
16~17	182	15	19	34	27	38	65	25	19	44	12	11	23	9	7	16
18~19	190	9	17	26	36	24	60	20	16	36	19	20	39	16	13	29
20~24	205	7	1	8	28	30	58	28	31	59	19	19	38	17	25	42
25~29	197	2	7	9	35	40	75	29	28	57	13	16	29	13	14	27
30~34	180	8	6	14	32	38	70	24	26	50	10	16	26	12	8	20
35~44	205	12	13	25	42	51	93	20	25	45	14	13	27	3	12	15
45~54	201	15	23	38	43	47	90	18	29	47	6	9	15	4	7	11
55~64	216	39	44	83	38	36	74	18	18	36	7	6	13	4	6	10
65~	181	51	52	103	16	23	39	8	13	21	10	6	16	1	1	2
总计	1757	158	182	340	297	327	624	190	205	395	110	116	226	79	93	172

2.3 样本职业、居住年限和户籍分布

在13种职业分布方面，国家公务员、专业技术人员、职员、企事业管理人员、工人、农民、学生、现役军人、自由职业者、个体经营者、无业人员、离退休人员和其他职业的人数分别为19例、218例、200例、31例、301例、98例、307例、30例、70例、68例、93例、212例和110例，涵盖职业范围广泛，样本具有一定代表性。

在取样当地居住年限方面，半年及以下的、半年~2年的、2~4年的、5~9年的、10年及以上的例数分别为61例、90例、110例、132例、

1323例。其中半年及以下本地居住年限的只占不到4%，2年以下居住年限的人群只占不到10%，而2年以上居住年限的人群占90%以上，基本代表了当地常住人口人群的情况。

在户籍分布方面，城镇人群为1237例，农村人群为479例。

2.4 样本智商基本数据

将全国1757例常模样本中所有被试的原始分进行标准化处理，转换成得到量表分和合成分数。WAIS-IV中文版在各分测验（包括积木等10个核心分测验）水平上转换后的量表分属于一个均数

为 10, 标准差为 3 的曲线分布。指数 (包括言语理解指数、知觉推理指数、工作记忆指数、加工速度指数、一般能力指数和认知效率指数等 6 个合成指数) 和总智商都是基于不同分测验量表分相加后, 经过转化得出的合成分数, 是一个均为 100, 标准差为 15 的曲线分布 (表 3)。

表 3 全国常模样本 WAIS-IV 得分基本数据 ($n=1757$)

各分测验和指数	全距	中位数	均数 \pm 标准差 ($\bar{x} \pm s$)
积木测验	1 ~ 19	10	10 \pm 3
类同测验	1 ~ 19	10	10 \pm 3
背数测验	1 ~ -19	10	10 \pm 3
算术测验	2 ~ 18	10	10 \pm 3
矩阵推理测验	1 ~ 19	10	10 \pm 3
词汇测验	2 ~ 19	10	10 \pm 3
符号检索测验	1 ~ 19	10	10 \pm 3
拼图测验	1 ~ 19	10	10 \pm 3
译码测验	1 ~ 19	10	10 \pm 3
常识测验	2 ~ 18	10	10 \pm 3
言语理解指数	55 ~ 142	100	100 \pm 15
知觉推理指数	51 ~ -155	100	100 \pm 15
工作记忆指数	49 ~ 148	100	100 \pm 15
加工速度指数	45 ~ 145	100	100 \pm 15
一般能力指数	52 ~ 151	100	100 \pm 15
认知效率指数	49 ~ 147	100	100 \pm 15
总智商	53 ~ 149	100	100 \pm 15

3 讨论

作为一个国际公认的个体智力测验, 韦氏智力量表从 1981 年引进并修订至今, 已经在中国得到广泛使用和认可。随着 30 年来中国社会状况和人口结构发生的变化, 以及心理测量学的发展, WAIS-IV 修订和常模制订过程中面临新的理论和实际问题, 主要包括以下几个方面。

首先, 分层取样所参照的人口学数据问题。在 1981 年 WAIS-RC 修订时, 当时无法获得全国人口学数据, 是按照 1980 年长沙市和效区的人口组成情况来计算分层比率^[4]。而 WAIS-IV 常模取样启动之初, 正值 2010 年全国第六次人口普查, 而只有 2000 年全国第五次人口普查数据可用, 由于中国人口基数大, 经济和教育发展迅猛变化极大, 按 2000 年的人口学数据明显会失真。鉴于此, 我们采用了北京大学人口研究所提供的 2005 年全国 1% 抽样数据作为计算分层取样的依据, 为了更接

近常模公布后得到广泛使用时 (大约 2016 年) 的实际情况, 我们请教了人口预测学专家, 对 2005 年数据进行了调整, 主要针对各年龄段的受教育程度比例进行调整。从 2010 年第六次人口普查公布的数据看, 与 2000 年相比, 10 年间小学及以下人口比例明显降低 (文盲率下降了 2.64 个百分点), 每 10 万人中具有大学文化程度的由 3611 例上升为 8930 例, 增加了两倍多^[17]。所以, 也证实了这种调整具有前瞻性, 符合人口变化的大趋势, 同时保证了取样工作的顺利进行。

其次, 城镇和乡村版本问题。在修订 WAIS-RC 时, 城市取样 2029 例, 农村取样 992 例, 分为城市和农村两套常模^[4], 符合当时城乡二元化结构这一现实。30 后的今天, 中国是否还需要两套常模, 这成为一个有争议的问题。WAIS-RC 研究显示, 农村青年修订韦氏成人智力量表 (农村式) 总智商分显著高于城市青年 (城市式) ($P < 0.01$), 将农村青年测试粗分换算为城市式后显著低于城市青年 ($P < 0.01$), 提示对文化程度较高的农村青年进行智商测试应选用修订韦氏成人智力量表城市式进行测试^[18]。对于 WAIS-RC 城市样本依老标准的平均智商为 102, 较新标准膨胀 2 个智商, 农村样本依老标准的平均智商为 104, 较新标准膨胀 4 个智商, 差异更为显著^[19]。另外, 中国的现状, 城市、乡镇和农村的格局发生巨变, 人口流动频繁而数量巨大, 通信及交通的迅猛发展, 城乡二元式结构逐渐被打破, 不论在时间和空间方面, 还是在信息共享方面, 都提示我们城乡隔绝的问题已经在渐渐地远去^[17]。另外, 韦氏成人智力量表不但应用于医学、心理学领域, 而且用于司法鉴定和智力残疾评定方面, 所以会涉及法律层面的公平性问题, 也要求采用一套标准来评价, 更符合社会的公平性要求。在常模取样的可操作性方面, 农村取样是极其困难的, 尤其是对于某些样本, 如较高教育程度的和年轻男性, 他们大都是家庭收入的主要承担者, 而且经常远离家乡, 已经成为城镇居民的一部分。本研究显示在户籍分布方面, 城镇人群为 1237 例 (72.0%), 农村人群为 479 例 (28.0%)。中国城市和农村户籍的二元结构是一个历史问题, 随着城乡一体化的加剧, 虽然户籍在一定程度上反映了人群在居住地和知识结构等多方面的差异, 但这种差异已经被打破而使情况变得复

杂起来。在城市居住的农村户籍人群变得越来越多,也成为城镇人群中的一部分,而我们的常模取样也在一定程度上反映了这种现实。所以,无论从科学性、公平性和可行性角度,两套常模并不符合当前的社会结构对智力结果的影响,而以城镇样本为基础制定的 WAIS-IV 全国常模具有合理性和现实性。当然,对于个别偏僻农村居民的智力测定是否更为准确客观,有待于以后的研究来验证。

最后,常模制定的统计学方法问题。本研究常模制定中所采用连续常模法(Continuous Norming)^[20-21]或称为推断常模法(Inferential Norming)^[9-22],与传统的标准方法相比,所制定的常模更为准确,同时样本量的数量更小,但是要求每个取样质量较高,节省了取样成本和取样周期。连续常模法采用曲线拟合技术,根据样本发展的趋势和样本统计数据导出理论总体分布。然后用理论总体分布导出常模,主要步骤包括:先以多变量回归分析的方法估计每个分量表中各年龄组的理论总体分布曲线,并以此为基础计算和修订出各年龄组的正态化标准分数以及量表分数,最后对分布曲线进行修匀完成常模。在 Wilkins, Rolhus, Weiss 和 Zhu 等^[9-23]的研究中,采用来自 WISC-IV 标准化数据库中的小样本来对推断常模法的作用进行研究。采用多正态曲线拟合度,计算出相同分数的病例百分比,平均绝对分数差异,分数分布,分数相关性,区分天才或智力缺陷的敏感度等多个指数,来对常模的质量进行评估。结果显示,采用推断常模法(Inferential Norming),每个年龄组样本量 50 和 75 时的常模,与采用更大样本量所获得的常模相当。所以,当大样本量因实际限制而无法获得时,每个年龄组 50 例可认为是获得比较理想常模的低限。中国人口总数大约是美国人口的 4 倍,依此计算,中国每个年龄组可以定为 50 例的 4 倍,即 200 例,9 个年龄组,共 1800 例,考虑保证科学性的基础上,具有较好的可行性。

综上所述,WAIS-IV 中文版的修订和常模取样工作工程巨大,设计需要考虑的因素复杂。在常模取样时,对智力成绩可能有影响的因素进行控制和考虑后,常模取样人群的性别、年龄、受教育程度、职业、地区、居住年限、户籍等人口学变量分析结果显示,基本符合计划的样本分布,为取得真实可靠有效的全国常模数据提供了保证,为

WAIS-IV 的良好心理测量学性能打下了基础。

参考文献

- [1] Sattler JM. Age effects on Wechsler Adult Intelligence Scale-Revised tests [J]. J Consult Clin Psychology, 1982, 50(5): 785 - 786.
- [2] Tulsy DS, Zhu J. Could test length or order affect scores on letter number sequencing of the WAIS-III and WMS-III? Ruling out effects of fatigue [J]. Clin Neuropsychol, 2000, 14(4): 474 - 478.
- [3] Hartman DE. Wechsler Adult Intelligence Scale IV (WAIS IV): return of the gold standard [J]. Appl Neuropsychol, 2009, 16(1): 85 - 87.
- [4] 修订韦氏成人智力量表全国协作组. 韦氏成人智力量表的修订 [J]. 心理学报, 1983, 15(3): 362 - 370.
- [5] 郑晓瑛, 陈功, 庞丽华, 等. 中国人口、人力资本变化趋势 [J]. 市场与人口分析, 2007, 13(1): 1 - 11 + 72.
- [6] 陈功, 曹桂英, 刘玉博, 等. 北京市未来人口发展趋势预测——利用多状态模型对未来人口、人力资本和城市化水平的预测分析 [J]. 市场与人口分析, 2006, 12(4): 29 - 41.
- [7] 王广州. 人口年龄结构间接估计方法与应用研究 [J]. 中国人口科学, 2001, (5): 19 - 23.
- [8] 王广州. 空间人口数据间接估计方法与应用研究 [J]. 中国人口科学, 2000, (3): 31 - 34.
- [9] Zhu J, Chen H Y. Utility of Inferential Norming With Smaller Sample Sizes [J]. J Psychoeduc Assess, 2011, 29(6): 570 - 80.
- [10] 王健, 邹义壮, 崔界峰, 等. 韦克斯勒记忆量表第四版中文版(成人版)的修订 [J]. 中国心理卫生杂志, 2015, 29(1): 53 - 59.
- [11] 范宏振, 王健, 崔界峰, 等. 加工速度、工作记忆以及推理能力的年龄组和受教育程度水平差别 [J]. 中国心理卫生杂志, 2015, 29(1): 60 - 67.
- [12] 陈饶, 李建明, 王健, 等. 精神发育迟滞患者韦氏记忆测验结果及其相关因素 [J]. 中国康复理论与实践, 2014, 20(6): 558 - 561.
- [13] 何雪玲, 王健, 崔界峰, 等. 住院精神分裂症患者韦氏成人智力测验第四版中文版测评的智力特征 [J]. 中国心理卫生杂志, 2013, 27(11): 858 - 863.
- [14] 王健, 邹义壮, 崔界峰, 等. 韦氏成人智力量表第四版中文版的信度和结构效度 [J]. 中国心理卫生杂志, 2013, 27(9): 692 - 697.
- [15] 晏丽娟, 王健, 李建明, 等. 精神发育迟滞患者韦氏智力测试结果及其相关因素 [J]. 中国康复理论与实践, 2013, 19(7): 691 - 693.
- [16] 何雪玲, 王健, 崔界峰, 等. 精神分裂症智力与记忆、注意和执行功能的关系 [J]. 中国康复理论与实践, 2013, 19(9): 805 - 808.
- [17] 中华人民共和国统计局网站, 中国 2010 年人口普查资料 [OL]. (2010) [2016-6]. <http://www.stats.gov.cn/tjsj/pcsj/rkpc/6/rp/indexch.htm>.
- [18] 魏春燕, 林志筠. 城乡青年修订韦氏成人智力测验对比分析 [J]. 临床心身疾病杂志, 2014, 20(1): 74 - 75.
- [19] 蒋莉, 姚树桥, 陈欢, 等. WAIS-RC 简式测验结果二十年变化的分析 [J]. 中国临床心理学杂志, 2006, 14(2): 121 - 122, 117.
- [20] Zachary RA, Gorsuch RL. Continuous norming: implications for the WAIS-R [J]. J Clin Psychol, 1985, 41(1): 86 - 94.
- [21] Taylor R. Continuous norming: improved equations for the WAIS-R [J]. Br J Clin Psychol, 1998, 37 (Pt 4): 451 - 6.
- [22] Wilkins C, Rolhus E. A simulation study of the efficacy of inferential norming compared to traditional norming (assessment report) [M]. San Antonio, TX: Harcourt, 2004.
- [23] Wilkins C, Rolhus E, Weiss L, et al. A new method for calibrating translated tests with small sample sizes [C]. The 2005 annual meeting of the American Educational Research Association. Montreal, Quebec, Canada, 2005.

编辑: 张卫华

2016-07-06 收稿

www.cmhj.cn