

龚氏非文字智力测验中老年常模的初步编制

任孝鹏*, 龚耀先

(中南大学湘雅二医院医学心理中心 湖南 长沙 410011)

【摘要】 目的:建立龚氏非文字智力测验(GNIT)中老年人群的初步常模,检验它的测量学特征。方法:在华东、中南和西南地区采样,共收集427例样本,按照45岁以上每10岁一组,75岁以上共一组,共分4组。结果:常模包括分测验的粗、量表分、总标准分和百分位数。信度和效度均达到了测量学要求,量表结构模型比较理想。结论:GNIT可以用中老年人。

【关键词】 非文字智力测验;常模;信度;效度

中图分类号:G449.4 文献标识码:A 文章编号:1005-3611(2002)03-0169-04

Establishment of the Adults and Elderly Norms for Gong's Non-verbal Intelligence Test (GNIT)

REN Xiao-peng, GONG Yao-xian

2nd Xiangya Hospital of Central South University, Changsha 410011, China

【Abstract】 Objective: The purpose of this study was to develop a set of age norms of GNIT on subjects from 45 to 98 years old and to examine its psychometric characteristics. **Methods:** 427 subjects (244 males, 183 females) ranging from 45 to 98 years were sampled from East China, Center South China and Southwest China. Four kinds of norms: raw scores, subtest scaled scores, standardized total scores and percentile ranks were obtained. **Results:** Test-retest reliability coefficients were .85 for standardized total score and .43~.88 for subtest scaled scores. Criterion-related validity studies indicated that the standard score of GNIT correlated .56, .58 and .62 with IQs of WAIS-RC respectively. **Conclusion:** GNIT has been established as a standardized instrument to measure the intelligence of people over 45 years.

【Key words】 Non-verbal intelligence test; Norm; Reliability; Validity

龚氏非文字智力测验(GNIT)是龚耀先为我国蒙古族、回族、土家族、苗族和白族等少数民族儿童所编制的非文字智力量表^[1-3]。在编制时按照智力结构说和编制跨文化测验的原则,并且为了弥补非文字智力测验主要测量知觉组织因素的缺陷,增加了测量思维活动和言语流畅性的分测验,设计了六个分测验,用图案而不用文字或图画做测验材料^[2,3]。因为五个民族在所有分测验中均无明显差异,所以建立一个共同常模,成为一套可用于多个民族的跨文化智力测验。后又建立了汉族儿童常模,并且把其中的几个测验用于聋人智力量表,均有满意的信度和效度^[4]。由于该测验不用语言或文字做测验材料,所以可以用于特殊群体。所谓特殊群体,是指智力、感知或者运动损害,不懂国语的移民或本国文盲等^[1]。它用于特殊群体时可以起到传统标准测验所起不了的作用,但是不能够取代传统智力量表,而是作为它的一种补充。

到目前为止,已经有人用GNIT来对某些特殊老年人群进行测量,发现GNIT能够反应他们的智

力特点^[5,6]。但是由于没有相应年龄阶段的常模,使GNIT的广泛应用受到了限制,所以建立本测验的中老年常模就显得尤其重要。本文是对建立GNIT的中老年常模的总结。

1 测验和取样

1.1 测验内容和实施方法

测验内容和实施方法与在少数民族和汉族儿童中所用的完全相同。测验内容包括颜色匹配和编码、认数辨色、分类、填图、填数和接龙六个分测验。

1.2 取样方法

样本年龄45至98岁。以汉族为主(只包括2位长期和汉族混居,完全习惯了汉文化的少数民族老人)。男女尽量无太大悬殊。样本来自华东、中南和西南三个大行政区在经济和教育状况方面比较有代表性的城镇和近郊。

有明显疾病的,包括脑部疾病史和现有明显的神经精神疾病者除外。

取样时间2000年5月至2001年1月。全部样本由三名具有一定临床经验的专业人员收集完成。

2 结果

*现在中科院心理所读博士后

2.1 常模样本

2.1.1 总人数、性别、民族、地域和年龄分布 从45岁至98岁共取样本443例,有效样本427例,男244例,女183例,结果见表1。其中完成了编码测验的只有151例,男82人,女69人。从45~75岁之间每10岁一组,75岁以上共设立一组。

表1 各年龄组的样本人数

| | 45岁~ | 55岁~ | 65岁~ | 75岁~ |
|----|------|------|------|------|
| 人数 | 95 | 115 | 127 | 90 |
| % | 22.3 | 26.9 | 29.7 | 21.1 |
| 男 | 50 | 66 | 80 | 50 |
| 女 | 45 | 49 | 47 | 40 |

2.1.2 样本的受教育程度 被试的教育程度以低文化程度者为多,总体来看,男性的教育程度要高于女性,男性除文盲的人数少于女性外,在其余各种教育程度上均多于女性,见表2。

表2 样本的受教育程度分布

| | 文盲 | 半文盲 | 小学 | 初中 | 高中 或中专 | 大专 | 本科 及以上 | 总数 |
|----|------|------|------|------|-----------|-----|-----------|-----|
| 人数 | 132 | 46 | 55 | 58 | 58 | 34 | 44 | 427 |
| 男 | 45 | 34 | 36 | 36 | 40 | 20 | 35 | 246 |
| 女 | 87 | 12 | 19 | 22 | 18 | 14 | 9 | 181 |
| % | 30.9 | 10.8 | 12.9 | 13.6 | 13.6 | 8.0 | 10.2 | 100 |

2.1.3 样本职业分布 样本的职业是以农民居多,占39.1%,其余依次为工人(18.5%),科技人员(11.2%),干部(10.3%),无业人员(7.3%),教师(5.1%),服务人员(4.9%),医务人员(4.0%),商人(0.4%)。

2.2 各年龄组测验粗分

测验结果分原始分(测验直接得分),量表分(由原始分换算标准分均数为10,标准差为3)和总标准分(由量表分换算而来,均数为100,标准差为15)三种。在此只比较了年龄组的原始分(见表3)。从结果来看,各分测验成绩随着年龄增长而下降。除45岁~组和55岁~组之间差别不显著外,其余各组均有显著差异。

2.3 项目通过率

GNIT中的分类、填图、接龙和填数四个分测验的项目顺序是按照少数民族样本的通过率高低排序。已证明汉族城市儿童常模样本的项目难度顺序与此相似。汉族中老年人的项目难度顺序是否相似?采用与汉族儿童中所用的相同计算方法来比较本样本与五个少数民族样本的项目难度顺序。结果两组各分测验难度顺序相差大多在1~2级之间,等级相关系数在0.959~0.998之间。结果见表4。

表3 各年龄组分测验的原始分比较($\bar{x} \pm s$)

| | 45岁~ | 55岁~ | 65岁~ | 75岁~ | P | 组间比较 |
|----|-------------------|-------------------|-------------------|-------------------|------|-------|
| 认辨 | 17.44 \pm 3.09 | 17.17 \pm 3.21 | 15.88 \pm 3.11 | 13.84 \pm 2.92 | .005 | , > > |
| 分类 | 21.95 \pm 2.10 | 21.67 \pm 2.27 | 20.96 \pm 3.07 | 18.96 \pm 2.77 | .003 | , > > |
| 填图 | 17.12 \pm 4.37 | 16.52 \pm 4.81 | 15.21 \pm 4.86 | 12.46 \pm 3.71 | .000 | , > > |
| 填数 | 19.27 \pm 3.64 | 19.12 \pm 3.94 | 17.45 \pm 4.21 | 14.35 \pm 3.84 | .006 | , > > |
| 接龙 | 16.02 \pm 4.04 | 15.49 \pm 4.58 | 13.16 \pm 4.59 | 9.96 \pm 3.04 | .000 | , > > |
| 编码 | 41.31 \pm 13.47 | 39.52 \pm 12.11 | 33.14 \pm 12.96 | 21.55 \pm 12.81 | .000 | , > > |

表4 两个样本在四个分测验中项目通过率比较和等级相关系数()

| 分测验 | 分类(30项) | | | 填图(32项) | | | 填数(30项) | | | 接龙(26项) | | |
|-----------|---------|------|------|---------|------|------|---------|------|------|---------|------|------|
| | 0 | + | - | 0 | + | - | 0 | + | - | 0 | + | - |
| 相等或者相差项目数 | 9 | 12 | 9 | 17 | 7 | 8 | 16 | 6 | 8 | 12 | 4 | 10 |
| 占总项目的% | 30 | 40 | 30 | 53.1 | 25 | 21.9 | 53.3 | 20 | 26.7 | 46.2 | 15.4 | 38.5 |
| 等级相差平均数 | 0 | 1.67 | 2.22 | 0 | 1.29 | 1.13 | 0 | 1.50 | 1.13 | 0 | 3.75 | 1.50 |
| 值 | 0.975 | | | 0.998 | | | 0.993 | | | 0.959 | | |

2.4 信度

2.4.1 重测信度 对30名成人在平均间隔30天

(28~32天)后进行龚氏非文字智力测验重测。结果表明:第二次的标准分平均高2.4分,标准分的重测

信度为.85,认数辨色、分类、填图、填数和接龙测验的重测信度分别为.88、.43、.58、.76和.62。

2.4.2 分半信度 分类、填图、填数和接龙四个分测验,采用奇偶分半的方法计算分半相关系数(用 Spearman - Brown 公式校正)。按顺序依次为.75、.92、.92、.90。

2.4.3 各分测验之间的相关 表5表明,将各分测验原始分作两两相关,发现它们之间的相关系数均在中等或以上,其中认数辨色测验之间的相关系数较高。

表5 各分测验量表分之间的相关

| | 编码 | 认数1 | 认数2 | 认数3 | 分类 | 填图 | 填数 |
|-----|------|------|------|------|------|------|------|
| 认数1 | .456 | | | | | | |
| 认数2 | .672 | .670 | | | | | |
| 认数3 | .614 | .572 | .815 | | | | |
| 分类 | .526 | .437 | .528 | .503 | | | |
| 填图 | .541 | .292 | .383 | .334 | .452 | | |
| 填数 | .548 | .363 | .559 | .509 | .587 | .529 | |
| 接龙 | .528 | .359 | .441 | .423 | .405 | .492 | .572 |

2.4.4 概化系数 计算了分类、填图、填数和接龙4个分测验的概化系数,依次为.79、.88、.89和.89。

2.5 效度

2.5.1 GNIT与WAIS-RC的相关 所有被试均同时完成WAIS-RC五合一(知识、相似性、数字广度、图画填充和木块图)的简式。结果表明GNIT的标准分与WAIS-RC的言语智商、操作智商和全智商的相关系数分别为.56、.58和.62;GNIT的标准分与知识、相似性、数字广度、图画填充和木块图等WAIS-RC各分测验量表分的相关系数依次为.65、.56、.49、.64和.57。

2.5.2 因素分析 采用方差极大正交旋转对常模样本进行因子分析,结果只提取1个因子,该因素能够解释总变异的51.8%,各分测验的因素负荷在.67~.77之间。另外还进行了二、三因子的模型分析,见表6。

表6 各分测验的因素负荷

| 分测验 | 单因子 | | 二因子 | | 三因子 | |
|------|------|------|------|------|------|------|
| | G负荷 | 因子1 | 因子2 | 因子1 | 因子2 | 因子3 |
| 分类 | .683 | .197 | .763 | .132 | .287 | .889 |
| 填图 | .705 | .753 | .251 | .728 | .083 | .380 |
| 填数 | .778 | .654 | .448 | .618 | .253 | .473 |
| 接龙 | .669 | .849 | .106 | .846 | .269 | .079 |
| 认数辨色 | .705 | .145 | .844 | .097 | .882 | .255 |
| 编码 | .770 | .423 | .663 | .386 | .744 | .169 |

3 讨论

3.1 GNIT能否用于中老年人

GNIT原为居住在农村,不懂或者不全懂汉语的少数民族而编制的^[1,3]。测验方法、内容和形式上与传统的智力量表不同,自然存在着能否适用于老年人的问题。不过从以下几个方面来看,该量表应用于老年人是可行的:第一、从测验的实施过程看,本量表根据非文字测验的编制原则,在实施中减少了言语指导和言语回答;增加了演示、练习和回答方法;操作简便,测验内容直观;总的测验时间短,容易使被试保持对测验的注意^[3]。第二、该量表在其它几个特殊人群和汉族儿童中应用的效果也较满意。如在汉族儿童常模编制中发现测验分数能够反映儿童的智力发展特点,具有与同类测验相似的信度和效度^[2]。将分类、填图、填数、接龙和编码分测验用于聋人智力量表的编制中发现信、效度都达到了能力测验的要求;第三、曾经有人把它用于老年人及其特殊人群,也表明它能够反映出他们的智力特点^[5-6]。老年精神分裂症和老年痴呆患者的GNIT的成绩能够反映出老年痴呆患者的智力比正常老年人低,老年精神分裂症患者的智力介于正常老年人和老年痴呆患者之间的特点;这些研究已经初步说明它可以用于汉人和某些群体的汉族老人。

能力测验的项目顺序是按照由易到难排列的。当它用于其它人群时,如果某些条目的难度发生改变,那么测验的项目顺序也应该做相应改变。因为一些适应年龄广和测验难度范围宽的测验,往往设有起始点和终止点,以便适应不同的受试。某一人群的项目难度顺序与另一人群的顺序不一定相同,GNIT在少数民族样本的项目难度顺序与汉族儿童的稍有不同^[2],与本组老年人样本相比,绝大部分项目在两个样本的等级完全相等或相差不大,如各分测验中项目顺序完全相同的条目占30~53.3%,不同的条目顺序相差也大部分在1~2级之间。由于本测验各分测验不长,均没有设立起始点和终止点,所以不必将项目作出加调整,不影响结果。龚耀先在分析GNIT能否应用于汉族儿童样本时指出,该量表建立其它特殊群体的常模时,如果出现项目顺序不一致时,也可以按照原来的分测验顺序进行测验^[2]。另外,对测验条目的难度和区分度在老年人群中的分析,也表明基本符合测量学的要求。根据心理测量学的原则,非多选题形式的测验条目难度

在.50时有最好的鉴别能力,而多选题的测验条目难度水平可比非多选题形式低些,如.69左右。由于智力测验条目的编制是由易到难,理论上,每一个分测验应具备最容易的条目(难度为1.0)和最难的条目(难度为0.0)^[7,8],因此,编制者希望一个分测验的平均难度为.50左右。而GNIT的四个图案分测验的平均难度为.57,其中分类测验的难度较低,只有.69。鉴别指数是重要的区分度指标。因为区分度有相对性,所以很难确定一个绝对的水平作为判断条目好坏的标准,一般要求鉴别指数在.20以上。美国测量学家 Ebel 根据自己编制测验的经验提出了鉴别指数评价测验的标准:.19以下,属于劣;.20~.29,尚可;.30~.39,良好;.40以上,非常优良^[7]。而本次四个分测验的平均鉴别指数为.30,已属良好,这也说明该套测验的项目,无论是难度还是区分度,都比较适合老年人用。

3.2 GNIT在中老年人样本中的信度和效度

GNIT无论是总的标准分还是各个分测验的重测相关系数都在中等水平,尽管对一个能力测验来说仍然不算高,不过在韦氏量表中操作分测验的重测信度也不高,而且国外许多著名的非文字测验或操作测验的复测相关较标准智力测验低^[3,9,10],如French的图画测验的重测相关在.70~.80^[9],Brown报道他编制的非言语智力测验的重测相关为.86^[9]。说明GNIT的这一结果可以接受。GNIT在编制时为了减少非文字测验不便测量言语智力的限制,增加了测量思维活动的分测验和设立测量言语流畅性的分测验。并且在儿童常模中通过与韦氏儿童智力量表的相关发现,填图和接龙两个分测验也负荷了言语理解因素,因此作者认为该测验确实测量了言语理解因素。曹颂尧也发现了类似的结果^[11]。本研究发现各分测验与言语智商、操作智商都有中等程度的相关,而且差别不大,进一步证明了上述措施确实起到了一定作用。

对常模样本的因素分析发现,本研究对常模样本进行因素分析,发现只能提取一个因子,该因子能够解释方差的51.8%,进一步说明较好地测量了智力的g因素。对二、三因子的因素分析结果显示,经旋转后二因子模型与三因子模型存在很多共同特征,三因子的功能比较明确。其中,填图、填数和接龙分测验都以知觉和推理为主,所以将该因子叫做知觉和推理因子。认数辨色和编码测验,前者需要视知觉与言语配合的速度,后者需要视知觉与手的运动速度的结合,因此可以命名为视觉心理运动速度因子。分类测验单独负荷一个因子,因为分类测验主要测量分析综合,抽象概括的能力,所以被命名为抽象概括因子。(在取样过程中,得到了邹涛、郑华、马建东的帮助,深表谢意!)

参 考 文 献

- 1 龚耀先,张致祥,程灶火,等.非文字智力量表的少数民族常模.中国临床心理学杂志,1995,3(3):129~133
- 2 龚耀先.非文字智力测验的汉族儿童常模.中国临床心理学杂志,1999,7(3):129~134
- 3 龚耀先.非文字智力测验手册.内部资料.1997.1~11;20~21
- 4 孟宪章,龚耀先.聋人智力量表的编制.中国临床心理学杂志,1995,3(2):75~78
- 5 杨蕴萍,李冬梅,陈统献,等.龚氏非文字智力量表用于老年痴呆患者的智力评估.中国临床心理学杂志,1996,4(3):214~216
- 6 杨蕴萍,张巍.龚氏非文字智力量表用于老年精神分裂症患者的智力评估.中国临床心理学杂志,1998,6(3):187~188
- 7 陈英豪,吴裕益.测验与评量.台湾复文图书出版社,1980.457~462
- 8 漆水清,戴海崎,丁树良.现代教育与心理测量学原理.江西教育出版社,1998.60~92
- 9 Brown L, Sherbenou RJ, Johnson SK. Test of Nonverbal Intelligence 2th, Pre-ed. Inc, 29~50
- 10 Leiter RG. Leiter International Performance Scale. Chicago: Stoelting, 1948
- 11 曹颂尧,卢原,唐细容,等. GINI与C-WISC在一组汉族儿童中测验结果的比较研究.中国临床心理学杂志,2000,8(1):45~47

(收稿日期:2002-04-11)

《临床精神医学杂志》征订

《临床精神医学杂志》是经国家科委批准的国内外公开发行的刊物(刊号:ISSN1005-3220 CN32-1391/R),系中国科技论文统计源期刊。本刊坚持突出临床、注重实用、内容新颖的宗旨,内容涉及儿童、老年、司法、戒毒、妇女、诊断、药物、心理、管理、社区、生物、护理等精神医学及医学心理学,主要栏目有论著、学术交流、临床研究、病例报告、专题讨论、讲座、综述、述评、心理治疗、国外动态、笔谈综合、读者园地等。适用于精神卫生工作者、临床医务人员、司法鉴定工作者及医学心理工作者等。本刊为双月刊,每册5.5元,全年33元。全国邮局订阅,邮发代号28-224。亦可与本刊编辑部联系,地址:南京广州路264号《临床精神医学杂志》编辑部收,邮编:210029

《临床精神医学杂志》编辑部