



常用心理评估量表 手册

主 编 戴晓阳

CHANGYONG
XINLI PINGGU LIANGBIAO
SHOUCE



人民军医出版社

PEOPLE'S MILITARY MEDICAL PRESS

常用心理评估量表手册



CHANGYONG XINLI PINGGU LIANGBIAO SHOUCHE

主 编 戴晓阳

副 主 编 张进辅 程灶火



人民军医出版社

PEOPLE'S MILITARY MEDICAL PRESS

北京

图书在版编目(CIP)数据

常用心理评估量表手册/戴晓阳主编. —北京:人民军医出版社,2010.7

ISBN 978-7-5091-3964-6

I. ①常… II. ①戴… III. ①心理测验-手册 IV. ①B841.7-62

中国版本图书馆 CIP 数据核字(2010)第 132267 号

策划编辑:张 田 张忠丽 文字编辑:赵晶辉 责任审读:余满松

出版人:齐学进

出版发行:人民军医出版社

经销:新华书店

通信地址:北京市 100036 信箱 188 分箱 邮编:100036

质量反馈电话:(010)51927290;(010)51927283

邮购电话:(010)51927252

策划编辑电话:(010)51927300-8225

网址:[www. pmmp. com. cn](http://www.pmmp.com.cn)

印、装:中国农业出版社印刷厂

开本:710mm×1010mm 1/16

印张:22.5 字数:417千字

版、印次:2010年7月第1版第1次印刷

印数:0001~3000

定价:56.00元

版权所有 侵权必究

购买本社图书,凡有缺、倒、脱页者,本社负责调换

编著者名单

主 编	戴晓阳	深圳大学			
副 主 编	张进辅	西南大学			
	程灶火	无锡精神卫生中心			
编 者	(以章节先后为序)				
	戴晓阳	李春波	吴文源	程灶火	苏林雁
	杜亚松	郑全全	骆 一	郑 涌	雷 辉
	徐银芳	杨 洁	陈 坡	张 炜	吴明霞
	解亚宁	肖计划	肖水源	叶悦妹	吴 艳
	安 芹	雷 雳	杨 洋	刘炳伦	王明辉
	李宗波	杨 洪	王孟成	李淑然	罗 杰
	孟宪璋	刘秀菊	潘素珍	刘文华	丛 中
	吴薇莉	刘协和	方 莉	崔汉卿	杨 东
	朱相华	赵后锋	吴汉荣	蒙 衡	张兴贵
	何立国	刘仁刚	朱丽霞	田 澜	李捷华
	马惠霞	静 进	黄 旭	姚 彬	江 琦
	张大均	葛明贵	余益兵	任 芬	郭文斌
	姚树桥	陈小莉	董 琴	卢国华	梁宝勇
	杨宏飞	张进辅	赵永萍	杜建政	王 君
	訾 非	周 旭	梁海霞	李超平	蔡太生
	阳锡洲	胡 建	王志平		
学术秘书	王孟成	深圳大学			
	李捷华	深圳大学			

前 言

在指导学生做研究实践或毕业论文的过程中,他们常会询问从哪里能够找到所需要的心理评估量表。多数时候他们难以找到所需量表的完整条目和计分键,以至放弃原先的设想。这就是我们编写《常用心理评估量表手册》的始因,希望将一些经过实践检验的信度和效度较好的量表汇编在一起,尽可能完整地提供量表的相关资料。

我们通过中国期刊网检索 1993 年以来相关学术期刊上发表的有关心理评估量表的论文,获得 410 个自编或修订的量表,这一检索结果构成了本书所选量表的基本库。但这些量表仅仅反映十几年来我国心理工作者和其他研究者所编制和修订心理评估量表的一部分,因为早期有些专业期刊并没有被中国期刊网收录,还有许多量表没有公开发表。因而,我们将张明园主编的《精神科评定量表手册》和王向东主编的《心理卫生评定量表手册》等书也作为重要参考。

《常用心理评估量表手册》选录量表基本遵循以下 3 条原则:

1. 量表的研制或修订过程符合心理测验编制程序,并且研究结果显示该量表具有较好的信度和效度。
2. 所选量表已经在国内心理学专业期刊上公开发表或已被接受待发表(极少数例外)。
3. 量表的编制者或修订者除愿意提供与量表相关的理论、结构、信度与效度资料外,同时也愿意提供量表的全部条目、计分键和评分方法,以保证使用者可以方便地应用。

需要特别说明的是,本书没有纳入的量表并不说明其质量不好,可能有许多原因:一是量表版权或专利的原因,编制者或修订者不愿意或不能公开量表的条目和计分键;二是由量表的性质所决定,一旦公开了其条目和(或)计分键则可能导致该量表不再有效(如成就测验);三是同一类型的量表太多,限于本书的篇幅不能全部收入。

为了能够准确地介绍量表的性质、功能和结构,以及实施、评分、结果分析和解释方法,我们尽可能邀请量表的编制者或修订者亲自撰写,只有少量量表由于原编制或修订者出国、工作太忙等原因由我们代笔。因此一本不太厚的手册竟由100多位作者共同努力了近2年的时间才成稿。但是,编写这本手册所花费的精力与本书所含量表的研制或修订所花费的精力相比,是不可同日而语的。

有鉴于此,我要对本书量表的研制者和修订者表示崇高的敬意和衷心的感谢,是他们的慷慨和奉献精神才使得这本书得以完成,为使用者提供获得这些量表的方便途径。希望使用者在运用这些量表进行心理评估或研究时,感念量表的作者在研制和修订过程中所付出的努力,并在研究成果或论文中能给出他们的名字。量表的编制和修订者当然希望自己的研究成果得到更广泛的应用,所以在此我也代表本书所有量表的编制者和修订者对使用者表示感谢!

戴晓阳

2010年3月于深圳大学

目 录

第 1 章 心理量表概论	(1)
一、心理量表与心理测验	(1)
二、心理态度量表的形式	(1)
三、标准化心理量表的基本特征	(4)
四、量表使用中的注意事项	(9)
第 2 章 一般心理健康与行为问题量表	(13)
一、90 项症状清单(SCL-90)	(13)
二、儿少心理健康量表(MHS-CA)	(19)
三、Achenbach 儿童行为量表(CBCL)	(28)
四、长处和困难问卷(SDQ)	(43)
五、初中学生攻击行为的主要心理特征问卷	(48)
六、青春期性心理健康量表	(52)
七、匹兹堡睡眠质量指数(PSQI)	(56)
第 3 章 应激及相关行为量表	(60)
一、生活事件量表(LES)	(60)
二、中学生生活应激评定量表	(65)
三、中学生应激源量表(SSMSS)	(72)
四、大学生恋爱心理压力源量表	(76)
五、中小学教师工作-家庭冲突问卷(WFCS)	(80)
六、简易应对方式问卷	(82)
七、应付方式问卷(CSQ)	(85)
八、社会支持评定量表	(90)

九、青少年社会支持量表	94
十、青少年学习倦怠量表	96
十一、中国心理咨询与治疗专业人员枯竭量表	99
十二、青少年病理性互联网使用量表(APIUS)	103
十三、青少年上网成瘾自评量表	107
十四、中学生网络成瘾诊断量表	109
十五、IT 企业员工工作成瘾问卷	111
十六、自杀态度问卷(SAQ)	113
十七、创伤后应激障碍检查表(PCL)	117
第 4 章 精神病学临床量表	122
一、老年精神评定量表(PAS)	122
二、抑郁自评量表(SDS)	133
三、流调用抑郁自评量表(CES-D)	137
四、贝克抑郁自评量表(BDI)	141
五、抑郁体验问卷(DEQ)	144
六、医院焦虑抑郁量表(HAD)	150
七、焦虑自评量表(SAS)	153
八、状态-特质焦虑问卷(STAI)	156
九、大学生强迫症状问卷	160
第 5 章 家庭与人际关系量表	163
一、家庭亲密度和适应性量表中文版(FACES II -CV)	163
二、父母教养方式评价量表(EMBU)	167
三、安全感量表(SQ)	174
四、修订版成人依恋量表(AAS)	177
五、分离体验量表第 2 版(DES- II)	180
六、人际信任量表(ITS)	186
七、社交回避及苦恼量表(SAD)	189
八、青少年学生疏离感量表(ASAS)	191
九、汉区少数民族学生文化疏离感量表	196
十、UCLA 孤独量表	200
十一、儿童期虐待史自评量表(PRCA)	203

第 6 章 生活质量与满意度量表	(209)
一、儿童少年生活质量量表(QLSCA)	(209)
二、儿少主观生活质量问卷(ISLQ)	(218)
三、青少年学生生活满意度量表	(222)
四、纽芬兰纪念大学幸福度量表(MUNSH)	(226)
五、中国人婚姻质量问卷	(231)
六、舒适状况量表(GCQ)	(238)
第 7 章 自我意识与自尊量表	(242)
一、Piers-Harris 儿童自我意识量表(PHCSS)	(242)
二、核心自我评价量表(CSES)	(249)
三、Rosenberg 自尊量表(RSES)	(251)
第 8 章 学习与教育量表	(254)
一、大学生学习动机问卷	(254)
二、中学生自我导向学习倾向性量表(SDLRS)	(257)
三、大学生一般学业情绪问卷	(261)
四、学习障碍儿童筛查量表(PRS)	(267)
五、儿童汉语阅读障碍量表(DCCC)	(273)
六、中学生考试心理和行为问题症状自评量表(EMP) ..	(277)
七、中学生感知的学校气氛问卷(PSCI-M)	(282)
第 9 章 行为方式量表	(285)
一、思维风格量表(TSI)	(285)
二、归因方式问卷(ASQ)	(293)
三、Aitken 拖延问卷(API)	(299)
四、大学生坚韧人格评定量表	(301)
五、学龄前儿童活动调查表(PSAI)	(304)
第 10 章 人生价值、态度量表	(307)
一、中文人生意义问卷(C-MLQ)	(307)
二、目标追求的人世、出世量表	(310)

三、中学生与其父母价值观差异问卷·····	(313)
四、大学生公正世界信念量表·····	(318)
五、多维完美主义问卷(MPS)·····	(320)
六、中文 Frost 多维度完美主义问卷(CFMPS)·····	(325)
七、消极完美主义问卷(ZNPQ)·····	(328)
第 11 章 职业价值、态度量表·····	(334)
一、职业延迟满足量表(ODGS)·····	(334)
二、变革型领导问卷(TLQ)·····	(338)
三、心理授权问卷·····	(341)
四、大学生文科兴趣量表·····	(344)

第 1 章

心理量表概论

一、心理量表与心理测验

在测量学领域,量表(scales)是用于测量的准尺。朱智贤主编的《心理学大词典》将量表解释为“一个具有单位和参照点的连续体”,将被测量的事物置于该连续体的适当位置,看它离开参照点多少单位的计数便得到一个测值。量表在心理与教育测量中专指测验量表。因此,心理量表(psychological scales)与心理测验(psychological tests)的概念是一致的,指用来测量不同心理特征(能力、人格、成就、兴趣等)的工具。

Hopkins KD 根据测量的目标将心理测验分为两类:一类是认知测验(cognitive tests),用来评价个体某项作业能力的最高水平(此人“能做什么?做得怎么样?”),例如智力测验、特殊能力测验、成就测验就属于这类测验;另一类测验是情感测量(affective measures),用于评价个体的典型作业特征(此人“常做些什么?”或“感觉怎么样?”),通常包括评价人格、兴趣等心理特征的测验。

态度量表(attitude scales)是情感目标测验中最主要的一类,通常采用纸-笔测验方式,包括自评和他评两种评价方式。前者由被测验者自己对照测验条目的情景根据感受作答;后者则由知情人对照测验条目的情景根据对被测者的了解做出评价,他评方法实际上是从自然观察延伸而来的。

在国内,许多研究者习惯将评价心理卫生、学校和社会行为、价值观点(一般不包括人格测验、兴趣测验)等态度量表统称为评定量表(rating scales),其依据是将英文“rate”的动词不定式翻译成“评定”。但是在心理测量学中“rating scales”特指等级量表,是一种以等级评分方式为特征的心理量表。

二、心理态度量表的形式

1. 瑟斯顿态度量表(Thurstone attitude scales) 该量表是 20 世纪 30 年代,

瑟斯顿在芝加哥大学做研究时发展的,用于评价被测验者对待某些事情的态度、雇员的道德水平等。它采用7点评分方式,非常赞成某个题目的观点给6分,非常不赞成给0分,中性(两者都不是)则给3分,被测者得分越高,表明越同意题目所示的观点。

瑟斯顿态度量表一般由20~45个项目组成,可以包含几个不同的维度,各维度的项目经过随机排序组成正式量表。许多研究证明这种量表形成是有效的,但其所能测量的因子受到一定限制。

2. 里克特量表(Likert scales) 这是一种最常用的心理量表评价方式,最常见的是5点量表(如大学生一般学业情绪问卷),分别是:非常同意、同意、不确定、不同意、非常不同意。也可以使用3点量表方式,如同意、不确定、不同意,对于年龄较小的儿童(如小学生)使用3点量表,他们可能更容易做决策。有时也有研究者采用5点以上的评价方式。测试时被测试者可以根据自己的实际情况选择最适合自己想法的一个进行标记。在实际应用中,这种形式的量表分半信度和重测信度可以达到0.8或以上。

里克特量表评价方式具有很好的灵活性,适用于各种态度测量,所有典型的描述肯定或否定态度的问题都可以采用这种方式。在测验中将肯定态度和否定态度的问题混合编排有利于减少被测者没看清楚题目就随意选答案的倾向。但是,在编制否定态度的题目时应当注意自然而然,不要露出太明显的人为痕迹,办法是尽量使用简单的语言和词汇。

3. 等级量表(rating scales) 等级量表本质上与里克特量表没有什么不同,通常用来描述同一个因素影响程度的差异。例如:无影响、轻度、中度、偏重、严重。这种方式在评估被测者的精神症状和问题的临床量表(如SCL-90)中普遍使用。在教师评估学生的非认知功能(如自我导向学习倾向性量表)或学生对教师教学质量的评价时也常会使用。

等级量表的评估效果有时候容易受评价者所掌握的评价标准的准确性、评价者熟练程度影响,研究显示有经验的评价者得到的结果其信效度比缺乏经验者高。

4. 语义区分量表(semantic differential scales) 这是一种测量心理意义的有用工具。这种量表最早由Osgood于1957年研创,其经典方法是在一个三维语义空间中评价被测者对某个具体或抽象事物或概念的态度。研究表明通常用3个维度就能够解释大部分变异,这3个维度分别是:性质(例如好—坏)、力量(例如强—弱)和活动(例如快—慢)。在设计量表时需要为被评估的某个内容或概念设计3~15组意义相反的形容词,每组都要包含这3个维度(分量表),最少也要设计3组。下面是一个对教师进行评价的举例:

教 师		
1. 快	____:____:____:____:____:____:____	慢 (活动)
2. 好	____:____:____:____:____:____:____	坏 (性质)
3. 被动	____:____:____:____:____:____:____	主动 (活动)
4. 小	____:____:____:____:____:____:____	大 (力量)
5. 无价值	____:____:____:____:____:____:____	有价值 (性质)
6. 强	____:____:____:____:____:____:____	弱 (力量)
7. 轻松	____:____:____:____:____:____:____	严厉 (力量)
8. 有帮助	____:____:____:____:____:____:____	无帮助 (性质)
9. 激动	____:____:____:____:____:____:____	平静 (活动)

一般采用7点评分,同一维度的项目根据性质的方向给予1~7分评分,通常给予正性评价高分,负性评价低分。例如在性质维度给“好”“有价值”和“有帮助”这3个端点7分,而在“坏”“无价值”和“无帮助”这3个端点给1分;同理,在力量维度给“大”“强”和“严厉”这端7分,给“小”“弱”和“轻松”这端1分;而在活动维度给“快”“主动”和“平静”这端7分,给“慢”“被动”和“激动”这端1分。设计问卷时注意不要将性质方向相同的形容词放在同一边,以避免被试者形成定势,不加思考地随意作答。

研究表明:当评估的内容或概念性质比较清楚、容易界定时,运用语义区分技术编制量表相对比较容易;而且语义区分量表较等级量表更容易实施,适合低年级学生或文化程度偏低的被试者;其信度也比较好,常可达0.8或以上。但是这种量表的可塑性较差,对于许多概念很难找到适合的形容词来描述它,或者难以满足3个维度的要求。另外,结果解释时给出的信息较少,许多测验者就直接用题目(形容词)解释被试者的特征。

5. 自我报告调查表(self-report inventories) 自我报告调查表与里克特量表或等级量表没有本质上的差别。典型的自我报告调查表指那些要求被测者以“是”或“否”方式回答,或者以“是”“不知道”或“否”方式回答的量表。这种方式被广泛地用于评估个体的习性、态度和技能。

6. Q-分类技术(Q-sort technique) Q-分类技术由 Stephenson 于 1953 年发明,是常用于测量态度、兴趣和其他情感变量的方法。被试者将一组附有问题、特征、图片或其他内容的卡片对照自己的情况分成9个等级(也可以分为5个、7个或11个等级)。每次只能遵循一个分等原则,如从“最重要”到“最不重要”,或从“最喜欢”到“最不喜欢”等。通常每个等级里卡片的数量接近正态分布,即“最重要”和

“最不重要”等级的卡片较少,中间等级的较多。卡片的内容通常设计成与某个特殊目的有关,分类结果反映了被测者自己的倾向,Q-分类技术多用于心理咨询和心理治疗实践。

Q-分类技术实施和计分相对较烦琐,项目较多、测验时间较长,而且每次只能评估一个维度,因此不如里克特量表或语义区分量表使用得广泛。

7. 问卷(questionnaires) 将一些问题印制在纸上,邮寄给特定的人群,请其填写完毕后寄回。问卷通常是从一些有关被测者背景的问题开始,例如性别、出生日期、受教育程度等。问卷多采用等级量表的方式调查被测者对某些问题的态度和看法。

问卷除了具有各种自我报告工具所共同存在的缺陷外,还存在一些特殊的问题。对于通过邮寄回收的问卷,研究者不知道被测者是在什么时候和何种情况下完成的,事实上从许多回收的问卷中可以发现被测者完成得非常草率,漏填或填错的情况很常见。一般情况下都有一定比例的问卷收不回来,这种情况可能使研究结果产生偏差。比较科学的做法是在问卷研究计划中至少设置1~2个随访组,调查没有邮回问卷的被试者,由此可以判断这些未邮回问卷者的态度,修正影响调查结果的误差。即使是百分之百回收的问卷,研究者在做结论时也要谨慎,被调查者所报告的感受,不一定是他的实际感受,他们在填写时的疏忽、记忆错误、故意隐瞒都可能在一定程度上影响结果。

8. 形容词核查表(adjective checklists) 形容词核查表的编制、实施和计分都不难。在编制这种测验时,研究者需要将能够描述所要测量的心理特征的相关形容词尽可能找出来,经过仔细筛选后确定要保留的形容词。

三、标准化心理量表的基本特征

(一) 取样

目前,人们尚没有办法直接测量心理特征,一般是通过测量人的各种行为表现间接地评估他们心理活动和特征。采用测量某种心理特征的测验对一组有代表性的人群(样本)进行测试,将他们在测验中的行为反应量化,并以某种分数的形式表达,这就构成了可供比较的标准。因此,样本的代表性决定了一个心理测验的适用范围,也会影响到它的功能。取样的方法很多,有方便取样、随机取样、整群取样、分层按比例取样等,不同方法所得到的样本的代表性可能相差很大。研究者在量表编制过程中是根据自己的设想和实际条件来选择测试对象的,因此使用者在准备应用某个量表前必须充分考虑该标准用于自己将要评价的对象是否合适?

从理论上讲,量表和问卷都应当是标准化了的测量工具,它们的实施和计分一般都有标准化的说明。但实际上除了人格测验、部分兴趣测验外,大部分评估态度的心理量表并不是标准化的,即它们的实际样本并不能很好地代表研制者想要测

量的目标群体。因此,在使用这类量表时要注意它们的局限性,特别是在解释测验结果和进行推论时要十分谨慎。

(二)原始分与标准分

心理测验的结果通常都以分数的方式来呈现。根据一定的评分原则,对被测者的回答进行评判并给予相应的分数,这个分数就是被测者得到的原始分。未经过标准化的心理量表一般都是等级量表,由此所得到的原始分仅能表示相对高低、优次等差异,不能够进行加减处理,因为分数之间不是等距的;同一个测验不同分测验的原始分之间也不能进行比较,因为不同分测验的原始分不等距,相互之间没有可比性。比如一名被试在智力量表算术分测验中获得15分(最高分为20分),而在词汇分测验中获得40分(最高分为80分),我们无法判断该被试者哪方面的成绩更好,因为这两个测验分数的全距不同,互相没有可比性。在另一种情况下,被试者甲为16岁青年,在词汇分测验中获得40分,被试者乙为40岁中年人,在同一分测验中获得50分,我们也无法判断两名被试者谁的词汇水平更高,因为不同年龄的语言发展水平不同,他们也不具备可比性。解决这个问题的办法就是将原始分转换成标准分,使该测验由等级量表变为等距量表。

然而,在测量态度的心理量表中除了人格测验、部分兴趣测验等少数量表外,大多数心理量表都没有采用标准分来表示结果,因为发展一个标准化心理测验需要耗费大量的人力和物力。目前,国内外大多数心理量表都属于等级量表,一般用原始分的平均值和标准差来表示某群体的心理特征。当然,如果有条件,研究者应尽量发展标准化的心理测验,并且用标准分代替原始分(图1-1)。

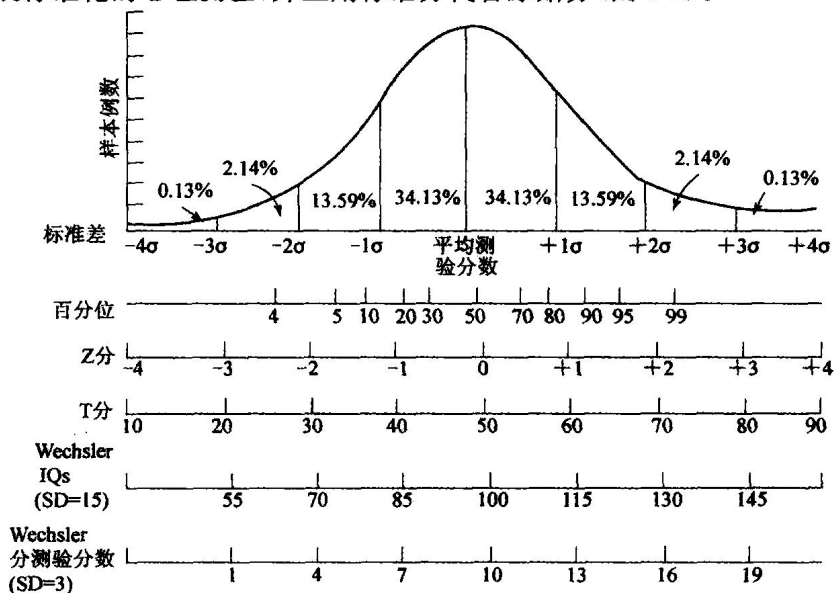


图 1-1 正态分布与标准化的关系

心理测验的标准分是参照常模——即标准化样本在测验中的操作数量化结果。因此测验常模是对一个具有代表性的样本在该测验中实际的操作结果(分数)进行标准化转换而建立起来的。

标准分的形式有很多,其共同点都是基于统计学的正态分布理论衍化而来的。

1. Z分 Z分是最基本的标准分,其他形式的标准分均是在Z分基础上转化而成。其公式是:

$$Z = \frac{X - \bar{X}}{SD} \quad (\text{公式 1})$$

上式中, X 为某被试者在测验中所获得的原始分

\bar{X} 为标准化样本在该测验的平均原始分

SD 为标准化样本在该测验中所获原始分之标准差

从公式 1 中可以知道, Z 分实际上是某一被试者在测验中的得分与标准化样本平均分的差距,在以标准化样本的标准差为单位这杆标尺上的位置。它不但可以说明被试者的操作水平在平均水平之上(Z 分为正)还是之下(Z 分为负),也能表明被试者与平均水平的相差程度(以相差多少个标准差来表示)。但 Z 分的缺点是存在负分,这在某些测验如能力测验、成就测验中使用起来不方便。下面是经过改良后的常用标准分计算公式:

$$\text{标准分} = M + S(Z) \quad (\text{公式 2})$$

上式中: M 为设计量表分的平均值

S 为设计量表分的标准差

Z 为 Z 分

我们在智力测验中常用的离差智商便是由公式 2 衍变而来。一般人们将智商的平均值定在 100,标准差定为 15(韦氏智力量表等)或 16(斯坦福一比内量表等)。如韦氏智力量表的智商公式为:

$$IQ = 100 + \frac{15(X - \bar{X})}{SD} \quad (\text{公式 3})$$

2. T 分 由 Z 分衍化而来的一种标准分。

$$T = 50 + \frac{10(X - \bar{X})}{SD} \quad (\text{公式 4})$$

T 分是一种使用得非常广泛的标准分,许多心理测验如明尼苏达多相人格调查表(MMPI)、艾森克个性问卷(EPQ)等都用 T 分来表示结果。

3. 标准 10 分和标准 20 分 也是由 Z 分衍化而来的一种标准分,卡特尔 16 种人格因素问卷(16PF)就是采用标准 10 分。

$$\text{标准 10 分} = 5 + \frac{1.5(X - \bar{X})}{SD} \quad (\text{公式 5})$$

$$\text{标准 20 分} = 10 + \frac{3(X - \bar{X})}{SD} \quad (\text{公式 6})$$

4. 百分位 百分位是临床心理评估中常用的一种非标准分表示结果的方法,它比标准分更容易理解,即使没有统计学基础的人也很容易理解它的含义,许多测验特别是成就测验经常用百分位来表示被试者的操作成绩。另外,临床心理学家在一些用标准分表示操作水平的测验报告中,也常用百分位来解释被试者的操作水平,如:“某某在韦氏成人智力量表中国修订本(WAIS-RC)获得全量表智商(FIQ)为117,相当于85%的百分位……”,说明该被试者的智力水平比85%的同龄的人好。

5. 划界分(cut-off score) 划界分是一种用来对被测者做出“全有”或“全无”分类的分数。在许多情况下要求个体的某些能力或心理特征必须达到一个最低资格,例如考取获得驾驶证、选拔飞行员,出于安全需要设立作业的临界点。在临床和咨询实践中,在进行治疗和行为建议的决策时,也要求同样的判断。使用这种分数的理由是:某种作业要求一些关键性的变量,这些变量的能力的缺失不能由其他变量的优秀能力来补偿;在临床和咨询实践中,这些具有特殊意义的心理特征对于诊断和治疗具有决定性的作用。

但是,运用划界分对被测者做出全或无的分类也容易产生误判,特别是那些得分接近划界分的个体容易被误判。Anastasi认为应努力减少单次测验所产生的误差,如果有可能,应当根据多次施测的一组分数来决定,而不是根据一次测验所得出的单一分数。另外,测验编制者确定划界分时通常会依据相应的理论和自己的目的选择一个适当的划界值,在减少一类误差(如假阳性)的同时,必然会增大另一类误差(如假阴性)的发生。

(三) 信度(reliability)

信度即可靠性,是对测验分数测量误差的估计。通俗地讲,测验的信度指施测分数能在多大程度上反映个体的“真实分数”。其专业的定义是:测验信度的测量即评价误差(error variance)在测验分数总方差中所占的比例。

在心理测验编制和使用过程中,影响测验分数的误差通常来源于以下几个方面。

1. 内容抽样误差 在编制一个测验时,测量同一功能的测验条目往往有很多,人们不可能全用,只能在其中抽取具有代表性的少数条目来构成测验,这样产生的误差称为内容抽样误差。

2. 时间抽样误差 同一名主试给同一名被试在不同时间实施测验,也会产生误差,这就是时间抽样误差。

3. 评分者误差 同一份测验结果让不同的人评分,由于个人掌握的标准有差异而产生的误差。

用信度系数来表示误差的大小,信度系数在 $-1 \sim +1$,绝对值越大(接近1.0),表明误差越小;绝对值越小(接近0),表明误差越大。不同的测量误差可采