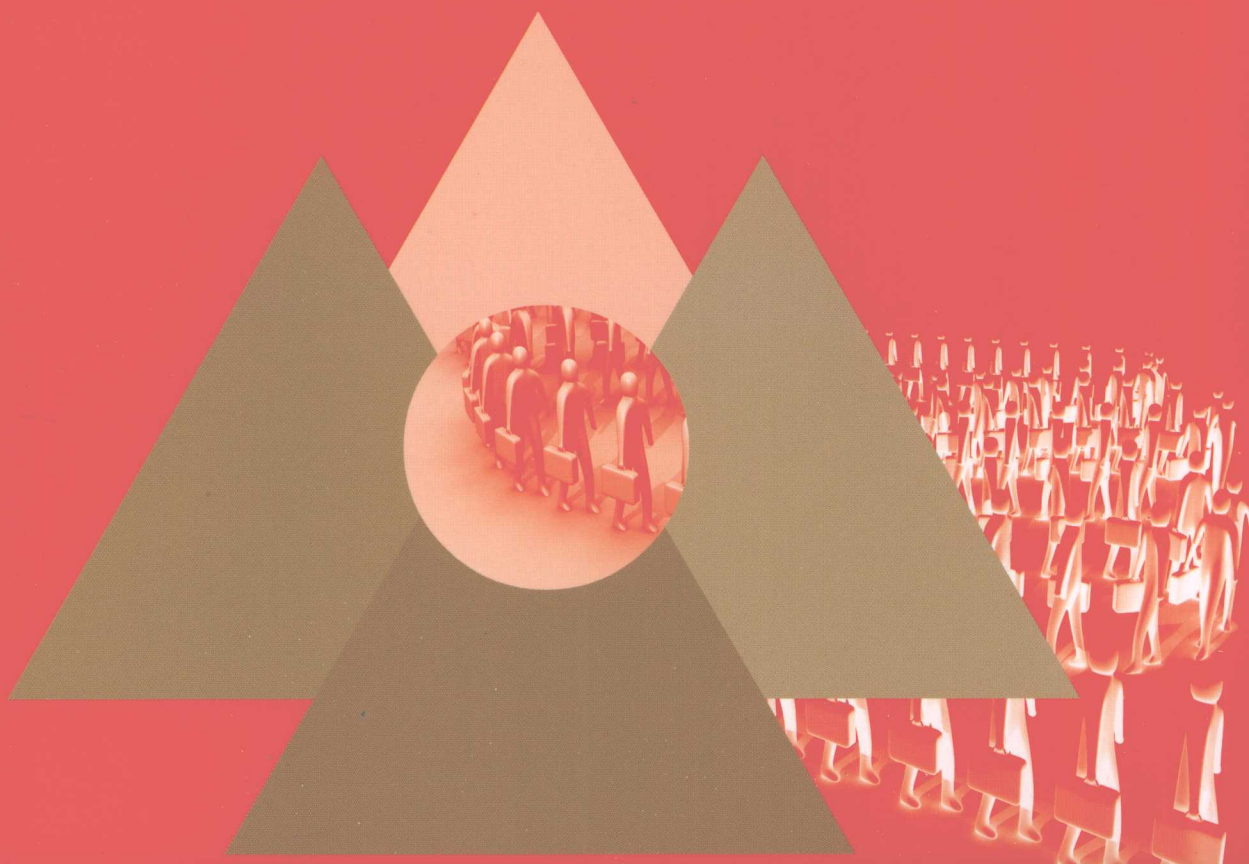


21 世纪应用心理学教材



实用心理与人事测量

王 垒 施俊琦 童佳瑾 编著



北京大学出版社
PEKING UNIVERSITY PRESS

21 世纪应用心理学教材

实用心理与人事测量

王 垒 施俊琦 童佳瑾 编著



北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

实用心理与人事测量/王垒,施俊琦,童佳瑾编著. —北京:北京大学出版社,2008.11
(21世纪应用心理学教材)

ISBN 978-7-301-14399-5

I. 实… II. ①王… ②施… ③童… III. 心理测量学-应用-企业管理-人事管理-高等学校-教材 IV. F272.92

中国版本图书馆CIP数据核字(2008)第167200号

书 名: 实用心理与人事测量

著作责任者: 王 垒 施俊琦 童佳瑾 编著

责任编辑: 陈小红

封面设计: 薛 磊

标准书号: ISBN 978-7-301-14399-5/B·0758

出版发行: 北京大学出版社

地 址: 北京市海淀区成府路205号 100871

网 址: <http://www.pup.cn>

电 话: 邮购部 62752015 发行部 62750672 编辑部 62752021 出版部 62754962

电子邮箱: zpup@pup.pku.edu.cn

印 刷 者: 北京大学印刷厂

经 销 者: 新华书店

787mm×1092mm 16开本 20印张 462千字

2008年11月第1版 2008年11月第1次印刷

定 价: 42.00元

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话:010-62752024 电子邮箱:fd@pup.pku.edu.cn

前 言

随着中国的改革开放,随着经济发展、全球化,工作社会的要求也越来越高、越来越复杂,竞争也越来越激烈。无论是企业,还是其他类型的组织,如医院、学校,都面对着竞争。如今,全球范围内的一个共识是:竞争首先是人才的竞争。一个组织要想在激烈的竞争中取得关键性优势并最终胜出,首要任务之一就是获得适用的人才。人才是竞争优势的基础和保障。而要获得人才,就必须首先懂得如何识别人才。这就好比需要千里马,就必须知道如何识别千里马。因此,客观、可靠而有效地甄别、评价人才,就成为竞争中的一个关键。这就需要有科学的心理与人事测量。

不仅从组织的角度,从个人的角度而言,正确地认识自己,了解、把握和发掘自己的优势、潜能、特长,对于在今天的工作社会中有效地生存,也是十分重要的。不了解自己的人,很难为自己找到适合自身特长和发展的工作。而且,不能很好地了解自己,也就难以很好地管理自己;而不能很好地管理自己,也就只能被别人管。所以,从个人职业生涯规划、发展而言,通过科学的心理与人事测量系统全面地了解自己,也是十分重要的。

值得一提的是,市场经济造成的结果之一是人员流动。竞争为人才提供了机会,使人力资源在社会中形成动态调配,从而应了一句老话,“人往高处走,水往低处流”,人力资本也趋于流向其价值最大化的地方。如今,“跳槽”在中国早已成为见惯不怪的事;合理的人员流动已经被人们普遍接受,并且成为各类组织所需人才的巨大供应渠道。人员流动促成了活跃的人事机制,使组织有可能根据自己的需要灵活地配置人才。然而,要想使这种配置科学化,真正实现人一事匹配,人适所能、才尽其用的理想,就需要了解人,就需要对人的特性做出科学的诊断,也就需要心理与人事测量。

随着现代职业社会的发展、进步,每一个职业人获得了更大的发挥空间。人们有机会不断评价自己,调整自己,不断地发展自己。人们开始接受个人职业生涯规划与开发的概念,并逐渐用于自己的职业指导。过去那种“吃老本”、“吃大锅饭”的日子一去不复返了,取而代之的是不断“充电”,追求自我提升、自我完善。然而,要想最大限度地开发自己,最大限度地发挥自己,就必须首先全面、准确地了解自己。这同样需要运用科学的心理与人事测量。

心理与人事测量有许多功能,企业需要它,各种组织需要它,个人也需要它。它能服务于人力资源规划,能为招聘、安置、考核、晋升提供依据,也是个人择业的参考,是职业生涯规划与开发的基础。科学的心理与人事测量能提供客观的评价依据,能消除人为主观的臆断,能避除任人唯亲的不公;它为组织人事决策的最大准确性和效率提供了科学基础,使组织和个人的利益都得到保证。

在发达工业国家,心理测验很早就被开发出来,并得到广泛应用,涉及的领域包括军事、企业、教育、政府机构等。例如,美国很早就使用人格测验参与士官的选拔;法国政府

推动将心理测验为义务教育的实施服务;美国政府很早开始在公务员的选拔中使用心理测验。欧美的很多大企业很早就开始使用心理测验于人事甄选。一些企业资助研究机构进行心理测验的开发,有的企业干脆自己设立专门的机构研发心理测验服务于专门的人事管理目的。在国内,早在上世纪初,就有包括北京大学在内的一些机构运用心理测验提供职业指导。这些事例都充分体现了心理学技术在人事评价中的重要作用。

自从上世纪80年代我开始从事心理测验与人事评价有关工作以来,接触了企业、教育、政府、军队各类机构。在为这些机构提供研究与咨询服务的过程中,我深切感受到心理与人事测量在我国现代化发展中起到的作用,经历的坎坷,也看到这一事业的不断发展和成熟。随着各类机构越来越多地应用心理与人事测量,这方面的专业知识、技术、方法的推广传播也越来越重要。这些年来,也有少数相关的专业书籍出版,但在系统性、理论与应用的结合特别是操作性方面仍不能满足业内需要。1989年,我编著出版了《心理测验程序库》一书和配套软件。90年代,我在北京大学出版社出版了多种心理与人事测量软件。1999年,为了配合北京大学的课程教学,也为了给业内人士提供一个专业化的工具书,我和我的学生们撰写出版了《实用人事测量》一书。如今,又有很多年过去了,心理与人事测量在国内又有了许多发展,许多新的实用测量工具被开发出来,各类机构对心理与人事测量的需求与日俱增,人力资源管理的需求与实践越来越复杂,各类职业人士对自我认识的要求越来越迫切。为此,我和我的同事们在原来《实用人事测量》一书的基础上进行了大量的修订,编著成这本《实用心理与人事测量》,还是当时那句话,将此书献给读者,献给那些诚于探索自我、勇于开发自我的人,特别要献给那些锐意改革、勇敢面对竞争的人和组织。

本书是我与同事二十余年的研究与实践经验积累的结果,书中绝大多数案例都出自我们的亲身经验。其中既有著名西方企业的经典案例,有国内民营企业的大胆尝试,也有国营企业的谨慎试验。从这里我们也看到,涉入市场经济越深的企业,越对科学的人事测量予以重视,因为他们的经验证明了科学人事测量的价值。实际上,我认为,人事测量有“四两拨千斤”的作用,人事测量上的少许投入,能取得人力资源的重要保障,换得组织的巨大效益。有的企业不肯在科学人事测量上投资,却肯把整个企业交给一些并不了解的人,造成千百万甚至数以亿计的损失。这样的交换怎么会有人算而不清呢?我希望通过本书引发读者和企业的反思,使人们肯于把管理交给科学。

科学的心理与人事测量是由一系列经过科学方法研制成的测量工具,以科学的方法使用而完成的,也就是说,是由工具和过程两个环节来保证的。因此,要了解科学的心理与人事测量,就要了解这些工具是怎样构成的,以及它们是如何工作的。两者都很重要,不可偏废。为此,本书在编著时紧紧围绕这两个方面展开。全书共分三大部分。第一部分是“原理篇”,可以说是心理与人事测量的基础知识篇,共有10章,系统介绍心理与人事测量的基本概念、原理,尤其重点说明心理与人事测量的科学基础。作者尽力使该部分严谨之余不失可读性。第二部分“工具篇”,是心理与人事测量的实用技术篇,详细介绍了各种类型的心理与人事测量工具近40种,使读者能对各种不同功能用途的心理与人事测量工具有较直观的感性认识。第三部分是“实务篇”,是心理与人事测量的高级综合应用,介绍如何在实际工作中组合、运用各种不同的心理与人事测量工具以达到不同的目的,包括如何针对不同规模的企业、不同类型的岗位、不同层次的员工、不同的管理

目的、不同的企业特殊需要如企业文化等进行心理与人事测量操作。该篇还介绍了如何运用心理与人事测量得到的结果进行总体的人事评价。为了进一步加强该篇的实务性,还特意安排了各类型企业实务案例一章,以及如何撰写测评报告一章。

本书由我和北京大学心理学系人力资源评价与开发中心的同事施俊琦博士、童佳瑾博士共同编著。本书的编著也与此前很多同仁、学生、朋友的支持、合作、帮助是分不开的,他们是:肖敏、姚宏、廖芳怡、王欢、李林、陈祎,在此特别表示感谢。还要感谢北京泰来猎头事务所、北京泰来人力资源评价中心、北京欣智成长科技有限公司,它们提供了许多测量实务实施与案例调研的机会,而本书提到的绝大部分测量工具也由北京欣智成长科技有限公司实现了网上在线实施(参见 www.wisejoy.com.cn)。

最后,要感谢北京大学出版社的陈小红编辑,她为本书出版计划的立项与实施,对本书的文字和专业性,都倾注了极大的精力和热忱,提出了不少宝贵的意见和建议。

今年恰逢改革开放三十年。这三十年也是心理与人事测量重新获得各界重视和逐渐发挥作用的三十年。谨以此书作为特别的纪念。

王 垒

2008年10月

于北京大学

目 录

• 原 理 篇 •

1. 心理测量概论	3
1.1 什么是心理测量	3
1.2 心理测量的应用领域	7
重要概念和术语	8
讨论题	8
2. 心理测量的历史	9
2.1 最早的智力测验:比奈的智力测验	9
2.2 心理测量在欧美的发展	11
2.3 心理测量在中国的发展	15
2.4 人事测量技术的革新	16
重要概念和术语	17
讨论题	18
3. 心理测量的基本类型	19
3.1 标准化的纸笔测验	19
3.2 投射测验	20
3.3 行为模拟与观察类测量	22
3.4 人事评估中基于工作情境的综合类测量	24
重要概念和术语	26
讨论题	26
4. 心理测量工具的建立	27
4.1 针对不同目标的测量工具的设计	27

4.2 题目的设计和测量的编制	29
4.3 题目的编排	35
重要概念和术语	36
讨论题	36
5. 心理测量的实施与计分	38
5.1 实施测量操作的要领与误差控制	38
5.2 实施测量计分的要领与误差控制	43
重要概念和术语	45
讨论题	45
6. 心理测量结果的解释	46
6.1 测验结果解释的概念	46
6.2 常模与基于常模的解释	47
6.3 效标与基于效标的解释	55
重要概念和术语	56
讨论题	56
7. 心理测量的关键技术 1:信度	58
7.1 心理测量的可靠性——信度	58
7.2 如何评估心理测量的信度	59
7.3 影响信度的因素	65
重要概念和术语	66
讨论题	67
8. 心理测量的关键技术 2:效度	68
8.1 心理测量的有效性——效度	68
8.2 如何评估心理测量的效度	69
8.3 影响效度的因素	75
重要概念和术语	77
讨论题	77
9. 心理测量的关键技术 3:项目分析	79
9.1 测量的难度分析	79
9.2 测量的项目鉴别度分析	82
9.3 测量的诱答分析	86

9.4 其他类型的项目分析及要点	88
重要概念和术语	89
讨论题	89
10. 人事测量在现实中的应用	91
10.1 现实工作中的人事测量	91
10.2 人事测评的工具和方法	97
重要概念和术语	104
讨论题	104
• 工 具 篇 •	
11. 基本测验 1: 个性品质测验	107
11.1 卡特尔 16 因素人格测验	107
11.2 DISC 个性测验	114
11.3 管理人员人格测验	119
11.4 控制源取向测验	126
12. 基本测验 2: 职业适应性测验	129
12.1 生活特性问卷	129
12.2 需求测试	133
12.3 职业兴趣测验	135
12.4 道德成熟度测验	140
13. 基本测验 3: 能力测验	144
13.1 多项能力与职业意向咨询	144
13.2 数量分析能力测验	151
13.3 管理人员逻辑推理测验	153
13.4 敏感性与沟通能力测验	155
13.5 销售技能测验	159
13.6 情绪智力测验	162
14. 基本调查	164
14.1 个体行为评估	164
14.2 领导行为评估	170

14.3 团体行为评估	178
15. 基于情境的测验	181
15.1 公文筐测验	181
15.2 无领导小组讨论	189
15.3 结构化面试	196
16. 面向高绩效管理的测验	214
16.1 人际敏感能力测验	214
16.2 管理变革测验	215
16.3 团队指导技能测验	216
16.4 自我实现测验	217
16.5 人际关系管理测验	217
16.6 沟通技能测验	218
16.7 XYZ 管理方式测验	219
16.8 基本管理风格测验	220
16.9 管理情境技巧测验	222
16.10 组织绩效测验	225

• 实 务 篇 •

17. 人事测量工具的组合原理	229
17.1 测量工具组合设计的含义	229
17.2 企业组织行为和人力资源评价系统回顾	229
17.3 测量工具的组合和设计论证	232
18. 针对不同管理目的的人事测量	236
18.1 用于招聘的测量组合设计	236
18.2 用于培训和开发的测量组合设计	241
18.3 用于晋升的测量组合设计	243
18.4 用于考核的测量组合设计	247
18.5 用于激励的测量组合设计	252

19. 针对不同对象的人事测量	254
19.1 用于不同层次的人事测量	254
19.2 用于不同职务层次的人事测量	258
19.3 用于不同岗位系列的人事测量	264
20. 针对企业特征和需求的人事测量	267
20.1 针对不同行业特征的测量组合设计	267
20.2 针对企业其他特征的测量组合设计	270
20.3 其他出于具体需要的人事测量	271
21. 人事测量案例	274
案例 1:某英国跨国公司的人事测量	274
案例 2:某美国跨国公司的人事测量	277
案例 3:某瑞士跨国公司的人事测量	280
案例 4:某国有企业的人事测量	287
案例 5:某民营企业的人事测量	289
22. 人事测量结果的报告	290
22.1 个人报告的撰写	290
22.2 总体报告的撰写	299
参考文献	301
重要概念和术语汇总	302

原理篇

在“企业竞争就是人才竞争”的今天,人才的素质越来越被企业所重视。随着人才素质要求多元化的发展,过去单凭经验选人、评价人的方法已不再适合。对人才素质的考察需要科学、客观的方法。因此,人事测量对于企业也就变得越来越重要。而人事测量学的基础是心理测量,本书首先介绍心理测量学的基本知识和概念,然后将心理测量学的知识运动到人事测量中。

本书也将有大量的篇幅向读者介绍人事测量技术在人力资源管理与开发领域如何实际操作和运用。但要真正做到对人事测量灵活地运用,还应首先对心理测量的原理有一个基本的了解。所以这开卷第一篇的目的就是向读者介绍心理测量的基本原理。在本篇中,你将能够了解到下列内容:

- 什么是心理测量,它的功能、用途、基本程序是什么;
- 心理测量的历史和现况;
- 如何编制一项成熟的心理测验;
- 如何建立测验计分体系和解释体系;
- 如何建立一个稳定的实用测量数学模型;
- 实施测验时应注意哪些问题;
- 怎样对测验的可靠性和有效性进行评估;
- 如何对测验的各个项目进行分析;
- 心理测量的方法如何运用到人事测量中去。

1

心理测量概论

1.1 什么是心理测量

心理测量、心理测验与人事测量

心理测量 (psychological measurement) 是通过科学、客观、标准的测量手段对人的特定素质进行测量、分析和评价。研究心理测量的学科叫做 **心理测量学 (psychometrics)**。这里的所谓素质,是指那些完成特定工作或活动所需要或与之相关的感知、技能、能力、气质、性格、兴趣和动机等个人特征,它们是以一定的质量和速度完成工作或活动的基础。

人事测量 (personnel assessment) 是心理测量技术在人事管理领域的应用,它以心理测量为基础,针对特定的人事管理目的如招聘、安置、考核、晋升、培训等,对人的素质进行多方面的系统评价,从而为人事管理、开发提供参考依据。因此,相对来说,人事测量是一个更广泛的实用概念,它通常要求运用多种心理测量工具来完成。

心理测验 (psychological test) 是心理测量的一种具体方法和手段,它是结合行为科学和统计学以评价特定个体在特定素质上相对于特定群体所处的水平的手段。心理测验是人事测量中最常采用的方法之一。除测验法之外,人事测量还经常采用观察法、访谈法、调查法等。以上几个概念之间的关系可由图 1.1 来表示:

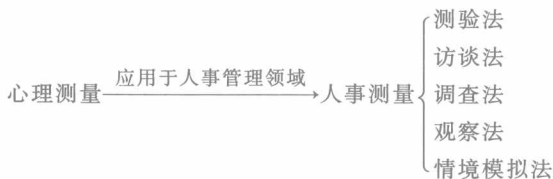


图 1.1 心理测量、心理测验与人事测量的关系

就具体方法而言,心理测量是通过对特定个体的有限的具有代表性的行为进行观察,依据事先确定的原则,对贯穿在行为活动中的心理特征进行推论和数量化分析。也就是说,由外部行为推及内在特征。个体的心理活动和心理特征是很难用直接测量的手段来度量的,只有通过对心理特征的外显结果——行为进行测量,才能推知个体内部的

心理活动状态和心理特征。因此,心理测量的对象实际上是行为样本,而不是心理状态。这种由行为表现到心理状态的推论并不是主观随意的,它必须在成熟的心理测量学理论的基础上,采用客观、科学的方法进行推断。

心理测验的五个基本要素

心理测验是人事测量的重要手段,心理测验是对行为样本的客观的和标准化的测量。它包含以下五个基本要素:

1. 行为样本

人们在对一类事物的某种特性进行考察时往往无法对这类事物的每一特定情况进行逐个观测,而总是抽取这类事物中有典型代表性的一部分进行观测,进而推论该类事物的普遍特性。

对个体行为或行为特征的测查同样无法做到面面俱到。在进行心理测量时,往往只能对经过科学选择的少数行为样本进行测查,借以推测个体的心理特征。例如,要考察个体的数学运算能力,可以选择一定数量、有代表性的运算题进行测量。个体对所抽选出来的问题的解决行为就叫做**行为样本 (behavioral sample)**。通过对这个行为样本的测量,可推测个体对所考察问题的解决行为,进一步推断个体解决问题能力的优劣。

由于实际上只考察了有限的行为样本,由此去推论总体行为,结论是否可靠,取决于所抽选出的行为样本能否很好地代表总体行为。为此,必须获取一组能够提供足够有用的信息、能恰当反映个体行为特征的行为,即要保证行为样本具备一定数量和具有代表性。

首先,行为样本的数量必须足够多,以使从样本到总体的推测错误率尽可能低。例如,从个体对一两道运算题的解决情况来推测个体的运算能力,这种推测的风险是极大的,因为猜对的概率很大。也就是说,从少数一两个行为,无法判断被评价者是在猜测结果,还是真正具备了解决问题的能力。

其次,要保证所选样本具有对总体样本的代表性。例如,不能只靠检测个位数加法题目来代替多位数加法能力的检测,因为个位数加法只是多位数加法的一个特例。又比如,不能只通过测查个体的计算能力来判定其财会技能。因为基本的计算能力只是财会技能的一部分,不能代表所有财会行为。因此,在选择行为样本时一定要十分慎重,尽可能使行为样本有较强的代表性。

2. 标准化

标准化 (standardization) 是指测验编制、实施、计分和解释必须严格遵循统一的科学程序,保证对所有受测者来说,施测的内容、条件、计分过程和解释系统都相同。只有这样才能保证测验的客观性和准确性。

要达到测验的标准化,应做到:

(1) 测验题目的标准化

对所有属于同一群体的受测者来说,测验题目本身不存在差异。对同一项心理测验,测验题目所引发的受测者心理状态、行为反应可以不同,但呈现在受测者面前的测验题目本身应该是同样的。例如,一个数理逻辑能力测验题可能引起不同受测者不同的推

理过程,但题目本身对所有人来说是一样的。只有保证题目标准化,即对所有受测者给定的条件一样,我们才能通过考察不同受测者的不同反应来判断其内在素质特征。这里,真正的输入(自)变量是受测者的内在素质,输出(因)变量是所观测到的受测者的反应,即应答行为,包括行为的过程和结果。

(2) 实施过程和计分的标准化

标准化的人事测量在实施时必须保证所有受测者在相同的环境中接受施测并得到标准化的测量指导说明。只有这样才能确保测验结果不受其他无关因素的影响。比如,如果同样的测验,有的人在炎热、潮湿、昏暗的环境接受测量,有的人在舒适明亮的环境接受测量,或者有的考场主考人口齿不清、说明含糊,而有的考场主考人口齿清晰、解释明确,由此造成的受测者的反应差异,无法判定究竟是由于受测者本人真实素质的差异所致,还是由考场环境的差异所致。

同理,测验的计分也必须有标准化的程序。要做到不同的计分者对同一受测者同一测验答案的计分方法和结果是相同的。人事测量中有一部分测验是较为依赖于评分者的主观判断的,如“无领导小组讨论”测验和“公文筐测验”的部分内容,这种类型的测验要做到计分系统的完全标准化相对较为困难。在这种情况下,虽然很难要求不同计分者计分结果完全一致,但通常要求有较高的相关性,一般相关系数要达到 0.80 甚至 0.90。目前,由于专业机构的努力,这类测验在计分方法和程序上已达到相当客观的程度。

(3) 选用有代表性的常模

常模(norm)是一组具有代表性的受测者样本的测验成绩的分布结构,包括它的集中趋势(通常用平均数表明)和离散度(通常用标准差表示)。常模是用以比较不同受测者测验分数的标准,它能够说明某一测验结果分数相对于同类受测者所处的水平。因此,心理测验的常模实际上起到度量衡的度量定制的作用,就如同确定如何计量长度(厘米、尺)或重量(克、两)一样,心理测验的常模确定如何表达能力的水平、工作动机的高低、职业兴趣的强弱或某种行为倾向(如内向和外向)的大小。

平均数(mean)是常模的重要内容之一,也是一个最常用的统计概念,它的数学定义是一个数据序列所有数据之和除以数据个数的商,即:

$$M = \frac{\sum X_i}{N} \quad (i = 1, 2, \dots, N) \quad (1.1)$$

其中, M 为平均数; N 为数据个数; X_i 为各数据值。例如,如果一组大学生能够在 30 分钟内平均正确回答 30 道逻辑推理题的三分之二,即 20 道题。则这 20 题的分数为一般大学生的平均值。平均分(数)描述了相应受测者群体的平均水平。

只有平均数还不够。因为不能把一个受测者的原始测验分数和平均数直接比较来判断它的高低。比如,一个人在上述的逻辑推理测验中得了 17 分,比平均分差了 3 分,但这 3 分究竟意味着什么?如果所有大学生的逻辑推理测验得分分布在 16—24 分,那么,这个 17 分可能就相当差,因为它几乎是在所有人成绩的最低端。但如果所有大学生该项测验的分数分布在 5—30 分,那么,这个 17 分就可能是还不错的成绩了。所以,只有平均分还不能反应一个测验分数的真实水平。这就需要常模中的另一个概念“离散度”,两者合在一起才能共同确定心理测验的度量定制。

离散度最常用的表示方法是标准差。**标准差(standard deviation)**的数学定义为一

个数据序列各数据与平均数之差的平方和除以数据个数之商的平方根,通常记作 S ,即:

$$S = \sqrt{\frac{\sum (X_i - M)^2}{N}} \quad (i = 1, 2, \dots, N) \quad (1.2)$$

标准差值越大,说明数据分布越分散,离散度越大。

一般来说,测验的原始分数并没有多大的意义,而是要通过以离散度为基准与平均值的比较来体现受测者的特点。在智力测验中,人们关心的往往是得分是高于一般水平还是低于一般水平。这就需要将测验的原始分数转换为标准分数,如将智力测验的原始分数转换为标准分数。通常这个转换公式设定为:

$$IQ = 100 + 15 \times \frac{X - M}{S} \quad (1.3)$$

其中, IQ 为标准分数; X 为实得原始分; M 为平均分; S 为标准差。

这个公式很好地体现了常模的作用,其意义是,把智力分数转化为以 100 为平均分、以 15 为标准差的标准(智商)分数。举例来说,如果一个人的实得测验分数刚好等于平均分,那么代入公式,末项分式取值为 0,最后结果刚好为 100 分。如果一个人的实得分数刚好高于平均分一个标准差,那么代入公式,末项分式取值刚好为 15,则最后得分为 115 分。

常模是否可靠,关键是有没有一个具有代表性的受测者样本,即建立常模的这些受测者样本要有足够的数量,而且是依据随机抽样或分层随机抽样原则抽选出来的。例如,如果要建立一般人的智力测验的常模,就不能用大学生作为智力测验的受测者样本,因为他们的分数往往高于一般水平,而无法代表一般人的水平。

3. 难度的客观测量

测验题目的难度水平影响到测验的客观性。无论是测验题目太容易,所有人都成功地完成所有项目,得分都很高,还是测验题目太难,使得大部分人的得分很低,都会使所有受测者的反应趋于一致,抹消了个体差异,无法就某一素质将不同的受测者很好地区别开来,致使测验的目的无法达到。通常将测验题目过于容易,致使大部分个体得分普遍较高的现象,称为**天花板效应(ceiling effect)**;而将测验题目过难,致使大部分个体得分普遍较低的现象,称为**地板效应(floor effect)**。在编制心理测验题目时要力求避免出现这两种情况。选择题目时,通常以通过某一题目的人数的百分数来作为难度的指标。有关的具体方法我们将在第 9 章详细介绍。

4. 信度

信度(credibility)是指测验结果的可靠性、稳定性,即测验结果是否反映了受测者的稳定的、一贯性的真实特征。这种可靠性体现在:

- 测验结果跨时间的一致性——受测者在不同时间所测结果一致;
- 测验内容的一致性——同一测验内的各部分题目所测的是同一种行为或行为特征;
- 不同评分者之间的一致性——不同评分者对同一测验结果的评分一致。

一项成熟的心理测验应具有比较理想的信度。信度较低的测验其测验分数的稳定性较差,往往不能反映受测者的真实水平。有关信度的具体内容我们将在第 7 章详细