

张杉杉 魏华颖 编著

ZHAOPIN ZHONG DE
XINLI CELIANG

招聘中的心理测量
原理与操作



经济管理出版社

ECONOMY & MANAGEMENT PUBLISHING HOUSE

**ZHAOPIN ZHONG DE
XINLI CELIANG**

招聘 中的心理测量
原理与操作



图书在版编目 (CIP) 数据

招聘中的心理测量: 原理与操作/张杉杉, 魏华颖编著. —北京: 经济管理出版社, 2011.8

ISBN 978-7-5096-1581-2

I. 招… II. ①张… ②魏… III. ①人力资源—招聘—考试—应用心理学—高等学校—教材 IV. ①F241.32

中国版本图书馆 CIP 数据核字 (2011) 第 168430 号

出版发行: **经济管理出版社**

北京市海淀区北蜂窝 8 号中雅大厦 11 层

电话: (010) 51915602 邮编: 100038

印刷: 北京晨旭印刷厂

经销: 新华书店

组稿编辑: 房宪鹏

责任编辑: 任爱清

责任印制: 杨国强

责任校对: 蒋 方

720mm×1000mm/16

13 印张

175 千字

2011 年 10 月第 1 版

2011 年 10 月第 1 次印刷

定价: 28.00 元

书号: ISBN 978-7-5096-1581-2

·版权所有 翻印必究·

凡购本社图书, 如有印装错误, 由本社读者服务部
负责调换。联系地址: 北京阜外月坛北小街 2 号

电话: (010) 68022974 邮编: 100836

前 言

有些人力资源管理人员认为心理测量过于专业，在人力资源管理工作中的应用极为有限，对人力资源管理的整体工作无关紧要。但事实上，如果不通晓心理测量学知识，任何人想要在人力资源管理中达到相当职业水准，只能祈祷自己是个天才。在每一项任务中，“现成的”测量工具要么不合适，要么根本不能用。如果外包测量专家，采用过于理想化的测量可能会无法体现实际工作的现实需要；如果尊重自己的直觉，测量方法又过于通俗、随意，无法满足专业要求；要想获得理想的效果，人力资源管理者学习心理测量势在必行。

心理测量学中的基本概念及其理论让大多数人望而生畏，如果不是心理测量学专业人员的，可能会被信度、效度等心理测量概念吓倒！但事实上，大多数人都能通过接受教育理解并使用心理测量的相关概念。当然，我们也不否认：传统的心理测量是围绕一系列听起来技术性很强的概念构建的，为了直观地解释这些概念，本书将选取具体的组织、公司的招聘案例为引导，讨论招聘所需要的各种心理测量学知识，而尽量避免相关公式的讲解，同时，我们也在每章后面提供了相应的建议阅读材料，为那些有可能希望了解这些原理的读者提供相关信息。

本书内容可以划分为两大部分，第一部分为第一至第六章，从心理测量出发讲解相关原理，并就招聘中的相关环节进行对接；第二部分为第七至第十章，从招聘出发，在现实或者模拟招聘案例中挖掘心理测量的藏身之处，让您在信息较为充分的情况下做出自己



招聘中的心理测量：原理与操作

的判断：心理测量能否、又是如何以及在怎样的背景和条件下使用，使得人力资源管理的招聘活动开展得更为有效。

本书第一至第六章由张杉杉老师主笔，第七至第十章由魏华颖老师主笔。感谢我们的学生，尤其感谢李娟娜、段红真、唐琳、李秋丹、李静雅、裴书涛同学在本书校稿中提供的帮助。由于作者能力有限，写作中疏漏和偏颇之处难免，恳请读者批评指正。

张杉杉 魏华颖

2011年5月29日

目 录

第一章 心理测量概述	1
一、测量对象	2
二、测量数字	3
三、测量法则	5
四、对心理测量的基本态度	9
五、本章案例：“国考”试题的测量质量评估	10
第二章 潜变量和因素分析	17
一、经典测量理论的测量原理	18
二、因素分析	21
三、本章案例：香满园胜任力结构的确定	29
第三章 信度和效度	37
一、信度	37
二、效度	40
三、概化理论	47
四、本章案例：自编沟通能力问卷的信度、效度检验	53
第四章 项目分析	59
一、难度	60
二、区分度	60



招聘中的心理测量：原理与操作

三、难度与区分度的关系	62
四、项目分析和信效度的关系	64
五、本章案例 1：托福的加试题	67
六、本章案例 2：题目的项目分析举例	70
第五章 项目反应理论	75
一、项目反应理论的简单历史回顾	76
二、项目反应理论的基本理论体系	77
三、项目反应理论的应用	84
四、对项目反应理论的评价	87
五、本章阅读材料：计算机自适应测验	88
第六章 心理测验	93
一、确定测量目的	93
二、选择测量的方式、方法	96
三、明确测验项目	100
四、收集数据	100
五、确定主题和项目	102
六、测量模型的试测和分析	104
七、验证、完善模型	106
八、本章案例：ETS 如何编制其测验题目与问卷	107
第七章 招聘流程	111
一、招聘的一般流程	111
二、招聘流程的测量学意义	117
三、本章招聘案例	121
第八章 校园招聘	137
一、校园招聘简介	138

二、校园招聘的特点	138
三、校园招聘的具体形式	140
四、校园招聘形式的测量学意义	142
五、本章案例——联合利华的“商业夏令营计划”	144
第九章 网络招聘	159
一、网络招聘简介	159
二、网络招聘的特点	161
三、网络招聘的测量学意义	163
四、本章招聘案例：欧莱雅的“全球在线商业策略竞赛”	167
第十章 面试及其发展	177
一、面试的基本概念	177
二、面试的类型	180
三、影响信度的要素	182
四、对主考官的培训	184
五、面试的发展趋势	187
六、本章招聘案例：宝洁面试的“经典八问”	194

第一章 心理测量概述

本章要点：

1. 心理测量的主要成分
2. 心理测量中常模的意义
3. 心理测量在测量法则上的特点
4. 测验标准化的内容

在选拔招聘过程中，测量是一个焦点。以下面的情境为例：

(1) 公司老总打过招呼：要在此次公司招聘中接收一位应届毕业生，应聘者的基本条件，诸如毕业院校、专业、本科成绩以及英语、计算机证书等都符合公司基本要求。现在面试已经结束，结果尚未公布，依据面试表现，老总的关照对象在淘汰之列。作为人力资源管理部门招聘组组长，你如何回复上司？

(2) 炙手可热的“国考”：时下，媒体已把国家公务员招录笔试称为“国考”。2010年度中央机关及其直属机构公务员招录笔试开考。根据统计，2010年报名并确认参加考试的人数为103万人，这一数字比上年少了1万人，而录取的人数是1.6万人，比上年增加了1000人，这也就意味着，2011年度国家公务员考试的录取率将有所提高，报考与录取比是63.6：1（2010年12月5日新华网）。

(3) 百度贴吧 <http://tieba.baidu.com/f?kz=737932959> 描述了武汉新东方总部人力资源部的招聘程序。版主针对新东方在各个高校不断举行类似的、以淘汰为主要手段的招聘会提出质疑：这样的招聘能否留得住人才？



在以上任何一种条件下，对特定领域感兴趣的个体在工作的时候都遇到了测量问题，虽然他们最初没有一个人对测量本身感兴趣，但是，想要完成上述任何一项任务都必须清楚这个情境与测量概念之间的关系。

测量是一个基本的科学活动。我们通过观察人类、物体、事件和过程而获得相关知识，要弄清楚这些观察结果常常需要对事物、活动本身进行量化。最早的社会测量程序，如投票、人口普查以及工作提升系统等“似乎都是为了满足大众的需要，而不仅仅是为了合乎科学好奇心而进行的实验”，同样的程序“可以从物理学的历史中得出：古代的人类在解决社会和实践问题的过程中，成功地实现了对长度或距离、面积、数量和时间的测量，物理科学就是建立在这些成就基础之上的”。通俗地说，测量是指利用合适的工具（Instrument），确定某个给定对象在某个给定属性上的量的程序或过程。从该定义可以看出，测量包括测量对象、测量数字、测量法则三个主要元素。

一、测量对象

心理测量的对象是人的行为。为了考察个体的特定行为，如个体的英语词汇掌握情况，只要使用一组有代表性的词汇测量即可，并不需要拿着英语词典，将词典中所有的词汇逐一进行测试。这种抽样的方法与其他学科的科学家用观察的方法是一样的，就像医生检验血液，只需分析一个或几个血液样本。也就是说，人们对一类事物的某种特征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察，而是对少量的、经过科学选择的个体行为样本进行测量，上例抽取出来的英语词汇就叫做行为样本。通过对这个行为样本的测量，可以推测出个体对所考察问题的解决行为，从而由点



及面，推测个体英语词汇水平。显然，少量的行为样本能否准确测量个体行为或者行为特征，依赖于行为样本是否具备一定数量、具有足够代表性。然而，由于所取得的行为样本只是代表某些心理功能，并不能反映这种功能的全部，因此，总免不了某种程度的偏差，因此，只有在全部了解行为样本的意义以后，才能正确使用心理测量。

美国著名心理测量学家 Anastasi 指出，测量题目不必与测量所预测的行为非常相似，但需要证实两者之间的经验性对应。从形式而言，测验样本与所预测的行为之间的相似程度并无强制规定。有时，测验行为样本与所预测的行为之间相当吻合，如英语词汇测验，而大部分的行为样本与实际中完成任务之间只有适度的相似性。一个测量是否具有应用价值，必须在经验上证实：被试在测量中的表现与在其他情境中的行为具有相当的一致性。

二、测量数字

数字代表某一事物或属性的量。在测量中，我们根据事物的属性和属性的大小分派数字，而事物的属性决定了数字是否具有自然数系统的功能（如区分性、等级性、等距性和可加性等）。当客观事物经测量在数字上显示出差异时，人们就可以把握事物的个别差异性，并针对客体对象的水平采取不同的对待方法。比如在招聘中录用高分者而拒绝低分者。

使用心理测量，对潜在变量的行为样本进行测量获取测量原始分数。一般而言，个人测量中获得的原始分数并不具有实际意义。例如，一位同学在某次测验中获得 90 分的成绩，是高还是低不得而知，只有将它与其他人的分数相比较之后才能做出客观判断。也就是说，测验的原始分数往往不具备独立解释的意义，同样一个分



数，比如 80 分，在语文、数学、政治等不同学科测验中有不同的意义，因为不同学科测验的难度不同，平均分也不同，80 分在一个较难的测验中是高分，在容易的测验中只算是中等的成绩。因此一个分数只有与总体水平相比较才能决定其高低优劣。常模就是为了这一目的而制作出来的。

常模是一组具有代表性的受测者样本的测验分数的分布结构，一般包括集中趋势（通常使用算数平均数表示）和离中趋势（通常使用标准差表示）。常模是用以比较不同的受测者测验分数的标准，说明某一测验结果分数相对于同类受测者所处位置，从而显示出心理测量与物理测量的一个显著差异——心理测量的相对性。显然，心理测量的常模实际上起到度量衡的度量定制作用，依据个体在其所在团体中的相对位置获得个体表达能力水平、工作动机高低、职业兴趣强弱、内向或外向的倾向性的大小判断，从而对测验分数加以解释。常模是否可靠，关键是有无一个具有代表性的被试样本。建立常模的通行方法是：在将要使用测验的全体对象中，选择有代表性的一部分人（称标准化样本），对此样本施测并将所得的分数加以统计整理，得出一个具有代表性的测验分数的分布结构。这就意味着建立常模需要受测者样本的数量足够多，并依据随机抽样或分层随机抽样原则抽取样本个体，确保样本对总体的代表性。建立常模的难度可想而知。所幸，在招聘中，我们并不强求常模。作为一种典型的选拔性测验，我们只要将应聘者的得分依次排列，选取适合特定招聘目的的应聘者即可。也就是说，在一般组织招聘中都不会特意制作常模，但是将被试的测验分数由高向低排列，选择高端的、相应的百分等级以上被试的选拔方式就是来自于常模思路。

三、测量法则

法则就是测量所依据的规则和方法。例如，用杆秤测量物体重量，依据的是杠杆的原理；用温度计测量物体的温度，依据的是热胀冷缩的规律；而对人的心理特征的测量，就是根据智力理论编制，以得分多少衡量智力水平。法则有好坏之分，使用较好的法则可以得到比较理想的结果，而较差的法则所获得的结果则不令人满意。例如，招聘启事中涉及的性别、毕业院校等内容易于测量，这种法则作为录用刚性条件易于建立和遵守，但使用不当就会成为性别歧视、地域歧视的把柄。而其他的很多胜任力特征难以测量的重要原因就是因为我们很难设计清晰而良好的法则。

制定法则、依据法则测量获得客观指标是对一切测量的基本要求，物理测量对象个体差异性小，测量中需要控制的误差变量较少，客观测量相对较为容易。作为生物科学，心理测量的个体差异显著高于物理科学中的研究对象，因此，心理测量更为复杂和抽象，更难以满足客观性的要求。测验法为了避免无关变量的影响，会对测验条件进行控制，以尽量保证所有被试接受的刺激条件相同。这里所谓共同的刺激就是测验的技术和方法。如果在严格控制无关变量的情况下对被试施测同样的刺激，被试的反应不同，就有理由认为是由于其心理特质的不同导致的。

心理测量中对误差控制的过程称为测验的标准化。测验的标准化包括测验内容的标准化、测验过程的标准化和评分的标准化。测验的标准化是测验编制的一个重要环节，它是在测验实施之前，对测验的内容、测验实施的情境、测验的时间、主试、测验指导语及评分等做出明确的统一的标准，以保证测验的全过程都能按照这一标准严格进行。



(一) 测量内容的标准化

即对所有被试施以同样的测验内容。测验内容就像心理实验中创设的实验条件（刺激），通过同样的刺激作用而引发被试不同的反应。测验内容在一张试卷（量表）上全部显示出来，包括同样的指导语、同样的答题要求和同样的测题。另外，测验内容的印刷要统一、工整，没有错误和遗漏。

(二) 测量实施过程的标准化

指测验实施中一切作用于被试的外界条件都应该相同，如主试、指导语、测验的时限、外部环境等。①主试。主试的年龄、性别、态度等因素都可能对被试产生重要的影响。有的主试能使被试产生轻松和安定的情绪，有的主试却会使被试感到紧张、焦虑。测验实施时，主试的各方面条件要保持一致，主试也要尽量创造一个宽松的气氛。主试的知识、经验也是必须考虑的因素，对于一般的学科成就测验，主试只要具备本学科的知识就行了；而对于智力测验和人格测验，主试则需要由受过心理测量学专业训练的人来担任，而投射测验的主试必须是使用投射测验的专家。②指导语。指导语分两种，一是给主试的，一是给被试的。给主试的指导语就是要告诉主试如何实施测验，包括测验要求、测验时限及对出现的意外情况如何处理等。测验开始前和测验实施的过程中，主试要用简明的、平和的语言介绍测验的情况，说明答题的要求，解释被试可能产生的疑问，以便被试迅速掌握答题的方法和要求。主试应说哪些话、不应说哪些话，都应有明确的规定。给被试的指导语应介绍测验目的、答题要求等内容，以指导被试完成测验。③测验实施的外部环境。测验实施的外部环境条件也应该相同，这要求测验应在同一时刻和相同环境条件下施测。在测验实施时，要求环境完全相同是不可能的，但要保证将环境的变异对测验结果的影响控制到最小限度。④测验的时限。测验的时限以大多数被试能完成所有测题

的时间为宜，这在测验的试测阶段就应该确定。大多数人格测验并没有严格时间限制，但也要对测验时间提出大致的要求，让被试尽快如实地答完测题，而不应在部分测题上停留过多时间。

（三）评分的标准化

评分的标准化，是指评分时使用完全客观的标准，使不同评分者对同一测题或一个评分者在几个等值的测题评分上得到一致的结果。一般而言，客观性试题都有唯一的正确答案，即标准答案，无论是人工评分还是计算机评分，评分误差都是很少发生的。而对论文式试题，由于试题不存在唯一的正确答案，在测试前就应想象出所有被试可能做出的反应，然后将所有答案依其水平和层次的不同划分等级，每一等级内包括属于一定范围的答案。评分时先将答案归入某一等级范围之内，再根据其等级给予相应的评分，评分尽量做到客观。

（四）解释的标准化

没有判断标准时，对测验的解释往往是主观的和任意的，所谓测验结果解释的标准化是指对测验分数高低优劣的判断要以一定的标准为依据，这一标准即是常模。常模的概念在常模参照测验中是必需掌握的。要知道招聘往往是在常模参照测验背景下进行。因此，我们后续的测量概念介绍均以所谓的常模参照测验为主，除非特别说明，否则基本不会涉及标准参照测验的内容。

在心理测量中，纸笔测验很好地反映了标准化的上述特点。标准的纸笔测验一般由事先确定好的测验题目和答案组成（测验内容的标准化）；有详细的答题说明，统一时限、统一印制格式、统一座次要求（实施过程的客观性）；往往以客观题居多，即使是主观论述题也有较为详细的评分要点（评分标准的标准化）；得分往往不是局限在一个班、一个组织下属单位的小范围内的纵向比较，而是将分数进行大规模的统一排名，避免“自相残杀”和“矮子里拔将军”（分数解释的客观化）。标准化的纸笔测验之所以得以广泛使



用，是因为它具有很多其他类型测量无法替代的优点，如方便性（很容易实施，非专业人士也可以很好地使用），经济性（较短时间内获得大量被试信息），客观性（如上所述，纸笔测验往往具有标准化的实施说明、计分系统和解释系统，测试受到主观因素影响较小，其结果容易被接受和信服）。

除了标准化要求之外，心理测量与物理测量的有关法则问题还涉及大量的理论问题，依旧以纸笔测验为例，尽管由于纸笔测验的方便性、经济性和客观性使其无所不在，但其不足之处也不容忽视：①纸笔测验形式制约其无法对受测者的实际行为表现进行测量。②程式化的实施过程，导致测试者只能搜集测验中考察的结果得分，而对于测验之外的信息一无所知、信息流于零散。③应试技巧和猜测往往成为受测者提高测验得分的捷径，标准化答案的筛选往往阻碍个体创新能力的展示。

依据上述定义可以看出，符合定义的测量在科学出现之前已经得到了发展。我国始于汉代、兴于隋唐的科举取士制度被中外学者公认为世界上最早的心理测量实践；而在古希腊，测验作为教育的附加物，同时测量智力技能和运动才干。当然，科学的心理测量则是以心理测验的形式在 19 世纪的欧洲兴起。随着时间的推移，心理测量的影响力逐步扩大。邓肯^①指出，测量学本身作为一种方法学的范例出现，它的影响主要表现在以下三个方面：①心理测量学的信度和效度概念得到广泛使用。②社会科学研究中的因素分析备受欢迎。③运用社会科学方法编制的量表所包含的变量数量远远多于心理测量学最初关注的变量数量。也正是在这个背景下，借助于心理测量提升人力资源管理招聘效率才成为顺理成章的事情。

总之，邓肯的上述评价内容涉及心理测量中的诸多重要概念，我们将逐一讨论。因为只有理解这些概念，才能更好地使用心理测量这一通用工具。

^① Duncan (1984), Notes on Social Measurement: Historical and Critical. New York: Russell Sage.



四、对心理测量的基本态度

在人力资源管理的职业生涯中，估计从业者中的大多数或早或晚、或多或少都会不同程度地接触过心理测验。心理测验使用的国际化趋势明显，越来越多的心理测验促使你成为其施测对象、施测者甚至设计者，我们需要在此强调：必须要对心理测量持客观的态度。

自测量问世以来，人们对其评价毁誉参半，在 20 世纪 20 年代，测量如旭日东升引来万人瞩目，这种过热状态甚至导致一种极端的观点：“测验万能论”，认为心理测验可以解决一切问题。人们迷信测验，将测验分数绝对化，甚至对 IQ（智商）的 1 分之差耿耿于怀，认为这种差异具有意义。随着心理测验的逐步应用，人们逐渐认识到：测验这种粗糙的测量工具具有局限性，当测验结果和那些毫无根据的期望大相径庭的时候，这种顶礼膜拜、奉若神明的极端表现带来的反面效应是：失望、怀疑，乃至敌视情绪，“测验无用论”成为另一种极端观点。有人开始提出以下两种观点：人格测验侵犯个人隐私、违背民主原则；测验产生“宿命论”和种族歧视的偏见。即使是现在，我们也还是可以从心理测验的使用中看到上述观点的各种表现：“我想选一个聪明人给我当秘书，给候选人做一下智力测验，录用得分最高的！”“我昨天做了一本杂志上的人格测验，结果我是神经症，气死我了！一天都无法安心工作，就是休息也不舒服。”“给候选人做一个自杀倾向测验，凡是得分高的一律不要录用！”“测量专家就会故弄玄虚，不要花这些冤枉钱，让各部门老总自己去招自己想要的人。”“测量专家又不是我们公司的，他们怎么会知道我们公司需要什么样的人，对他们的话不要太当真。”“找个咨询公司来给我们做个绩效考核吧，我们别自己直接出