

SHIYONG RENSHI CELIANG

王 垒 / 著

实用人事测量

(简明版)



经济科学出版社

SHIYONG RENSHI CELIANG

王 垒 / 著

实用人事测量

(简明版)

BAS07/04



经济科学出版社

图书在版编目 (CIP) 数据

实用人事测量：简明版 / 王垒著 .—北京：经济科学出版社，2002.3

ISBN 7-5058-2965-3

I. 实… II. 王… III. 企业管理：人事管理
IV. F272.92

中国版本图书馆 CIP 数据核字 (2002) 第 015193 号

责任编辑：边 江

责任校对：徐领弟

版式设计：代小卫

技术编辑：邱 天

实用人事测量

(简明版)

王 垒 著

经济科学出版社出版、发行 新华书店经销

社址：北京海淀区阜成路甲 28 号 邮编：100036

总编室电话：88191217 发行部电话：88191540

网址：www.esp.com.cn

电子邮件：esp@esp.com.cn

中国科学院印刷厂印刷

河北三河新路装订厂装订

787×1092 16 开 18.5 印张 350000 字

2002 年 3 月第一版 2002 年 3 月第一次印刷

印数：0001—5000 册

ISBN 7-5058-2965-3/F·2333 定价：33.00 元

(图书出现印装问题，本社负责调换)

(版权所有 翻印必究)

序 言

进入 21 世纪，对中国的经济领域影响最大的环境变化莫过于中国加入 WTO。这使中国面对更广泛、更激烈的全球性竞争与较量。在这场竞争、较量中，制胜的核心之一是人才的竞争。一个组织要想在激烈的竞争中胜出，就必须首先获得高质量的一流人才。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、可靠而有效地甄别人才就成为竞争中的一个关键，这就需要有科学的人事测量。

市场经济造成的结果之一是人员流动。由于竞争提供了机会，使人力资源在社会中形成动态调配，这正应了一句古话，“人往高处走，水往低处流”。如今，人们对“跳槽”已经见惯不怪，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，就需要对人的特性进行科学的诊断，也就需要人事测量。

改革开放促成了生产力的解放，也使每一个人的能力获得了巨大的发挥空间。尤其是竞争机制的引入，使人们有机会不断评价自己，调整自己，按需求不断地发展自己。人们开始接受个人职业生涯规划与开发的概念，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”、追求自我提升的需要。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人事测量。

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,能避免任人唯亲的隐患。它保证了组织人事决策的最大准确性和效率,使组织和个人的利益都得到保证。

科学的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。

然而,尽管人事测量如此重要,但很长一个时期里,国内却没有一部系统、全面介绍这些内容的论著,这与我国的改革进程是不相适应的,尤其不能满足激烈竞争下人力资源管理与开发的需要。为此,笔者1999年编著了《实用人事测量》一书,献给读者,献给那些诚于探索自我、勇于开发自我的人,特别是献给那些锐意改革、勇敢面对竞争的人和组织。该书是笔者十余年的研究与实践经验积累的结果,其中案例有的出自著名西方企业的经典经验,有的出自国内民营企业的大胆尝试,有的出自对国有企业的谨慎观察。从该书中可以看到,涉入市场经济越深的企业,越对科学的人事测量青睐,因为它们的经验证明了科学人事测量的价值。实际上,依我的观察,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证,换得企业的巨大效益。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,造成千百万甚至数亿的损失,这样的交换怎么会有人算而不清呢?出版该书是希望引发读者和企业的反思,使人们肯于把管理交给科学。

该书出版不久,就受到学术界尤其是业界的关注,许多高校的管理学院、商学院、心理学系把它选为教材,很多企业参考了该书并选用了该书推荐的人事测量软件。为了更广泛地推动人事测量的学习和应用,为了方便广大读者,有必要出一个简明版本。特别是根据社会实践的需要,“人事测量”已被北京市列入成人高等教育自学考试课程,为了适应自学考试的特点和要求,应主考单位的要求,在出版社的同意和支持下,笔者简化、浓缩原来的《实用人事测量》,形成现在的这个简明版本,希望为读者和各界朋友提供学习和应用上的方便。

本书仍保持了原有的大致结构,共分三大部分。第一部分是“原理篇”,共有10章,系统介绍人事测量的基本概念、原理,尤其重点说明人事测量的科学基础,尽力使该部分严谨之余不失可读性。第二部分“工具篇”,共有6章,详细介绍各种类型的人事测量工具,总计超过30种,使读者能对人事测量工具有较直观的感性认识。第三部分是“实务篇”,共有4章,介绍如何在实际工作中组合、运用各种不同的人事测量工具达到

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等进行人事测量操作。该篇还介绍了如何运用人事测量得到的结果进行总体的人事评价。

最后，特别要感谢经济科学出版社第一编辑中心的朱丹和边江同志，她们对本书出版的立项与实施给予了极大帮助和支持，尤其是在出版任务极其繁重的情况下，她们仍悉心致力于本书的出版，令我十分感动。

因编著时间仓促，加之经验有限，本书难免有不足之处，谨请读者不吝赐教。

王 垒

2002 年 2 月于北京大学

目 录

第一部分 原理篇

第一章 人事测量概论	3
第一节 什么是人事测量	3
第二节 人事测量的功能与用途	8
第三节 人事测量的基本程序	10
重要概念和术语	13
讨论题	14
第二章 人事测量的历史	15
第一节 最早的智力测验：法国比奈的智力测验	15
第二节 心理测量在欧美的发展	18
第三节 心理测量在中国落户	24
第四节 人事测量技术的革新	26
重要概念和术语	27
讨论题	27
第三章 人事测量的基本类型	29
第一节 标准化的纸笔测验	29
第二节 投射测验	31
第三节 行为模拟与观察类测量	33
第四节 基于工作情境的综合类测量	37
重要概念和术语	38
讨论题	39

第四章 如何建立人事测量工具	40
第一节 针对不同目的的测量工具的设计	40
第二节 题目的设计和测量的编制	43
第三节 题目的编排	52
重要概念和术语	54
讨论题	54
第五章 人事测量的实施与计分	56
第一节 实施测量操作的要领与误差控制	56
第二节 实施测量计分的要领与误差控制	64
重要概念和术语	66
讨论题	66
第六章 人事测量结果的解释	68
第一节 解释的概念：为什么测量的实施与解释相分离	68
第二节 常模与基于常模的解释	69
第三节 效标与基于效标的解释	79
重要概念和术语	81
讨论题	81
第七章 人事测量的关键技术：信度	83
第一节 人事测量的可靠性——信度	83
第二节 如何评估人事测量的信度	85
第三节 影响信度的因素	92
重要概念和术语	94
讨论题	94
第八章 人事测量的关键技术：效度	96
第一节 人事测量的有效性	96
第二节 如何评估人事测量的效度	98
第三节 影响效度的因素	105
重要概念和术语	108
讨论题	108
第九章 人事测量的关键技术：项目分析	109
第一节 测量的难度分析	109
第二节 测量的项目鉴别度分析	114

第三节 测量的诱答分析	118
第四节 其他类型的项目分析及要点	121
重要概念和术语	123
讨论题	123
第十章 人事测量在现实中的应用	124
第一节 现实工作中的人事测量	125
第二节 人才测评的工具和方法	130
重要概念和术语	137
讨论题	138
第二部分 工具篇	
第十一章 基本测验：个性品质测验	141
第一节 卡特尔 16 种人格因素测验 (16PF)	141
第二节 DISC 个性测验	149
第三节 管理人员人格测验	154
第十二章 基本测验：职业适应性测验	162
第一节 生活特性问卷	162
第二节 需求测试	166
第三节 职业兴趣测验	168
第十三章 基本测验：能力测验	174
第一节 多项能力与职业意向咨询	174
第二节 数量分析能力测验	178
第三节 管理人员逻辑推理测验	180
第四节 敏感性与沟通能力测验	181
第十四章 基本组织人事调查	186
第一节 个体行为评估	186
第二节 领导行为评估	189
第三节 团体行为评估	196
第十五章 基于情境的测验	199
第一节 公文筐测验	199
第二节 无领导小组讨论	204
第三节 结构化面试	210

第十六章 面向高绩效管理的测验 226

第三部分 实务篇

第十七章 人事测量工具的组合原理 241

- | | | |
|-----|-------------------|-----|
| 第一节 | 测量工具组合设计的含义 | 241 |
| 第二节 | 企业组织行为和人力资源评价系统回顾 | 242 |
| 第三节 | 测量工具的组合和设计论证 | 245 |

第十八章 针对不同管理目的的人事测量 250

- | | | |
|-----|----------------|-----|
| 第一节 | 用于招聘的测验组合设计 | 250 |
| 第二节 | 用于培训和开发的测验组合设计 | 253 |
| 第三节 | 用于晋升的测验组合设计 | 256 |
| 第四节 | 用于考核的测验组合设计 | 261 |
| 第五节 | 用于激励的测验组合设计 | 266 |

第十九章 针对不同对象的人事测量 268

- | | | |
|-----|---------------|-----|
| 第一节 | 用于不同层次的人事测量 | 268 |
| 第二节 | 用于不同职务层次的人事测量 | 272 |
| 第三节 | 用于不同岗位系列的人事测量 | 276 |

第二十章 针对企业特征和需求的人事测量 279

- | | | |
|-----|-----------------|-----|
| 第一节 | 针对不同行业特征的测验组合设计 | 279 |
| 第二节 | 针对企业其他特征的测验组合设计 | 284 |
| 第三节 | 其他出于具体需要的人事测量 | 285 |

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，人事测量也就变得越来越重要。

本书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，还应首先对人事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人事测量的基本原理。在本篇中，你将能够了解到以下内容：

- 什么是人事测量，它的功能、用途、基本程序；
- 人事测量的历史和现况；
- 一项成熟的人事测验如何编制；
- 如何建立测验记分体系和解释体系；
- 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。

第一章

人事测量概论

第一节 什么是人事测量

一、人事测量、心理测量与心理测验

关于人事测量，首先要澄清一些相关的概念。

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、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、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必要基础。

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系统评价，从而为人事管理、开发提供参考依据。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，它是结合心理学和统计学方法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、情景模拟法等等。

心理测量是人事测量的基础。人事测量是以人为对象，是对人的测量，这正是以心理测量为基础进行的。绝大多数人事决策都需要以人事测量为基础。例如，招聘人员时，要测量应聘者的能力、工作动机；考虑晋升时要评价能力是否胜任新的职位；计划培训时要考察对象在哪些方面对于完成规定的工作任务有不足。所有这一切都要运用心理测量。所以，要从事人事测量，首先要掌握心理测量。

具体来说，心理测量是通过对人的有限的具有代表性的行为进行观察，依据预先确定的原则，对贯穿在行为活动中的心理特征进行推论和数量化分析，由外部行为推及内在特征。这是因为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，需要通过对心理特征的外显结果——行为进行测量，来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，它必须在心理学有关理论的基础上，采用客观、科学的方法进行推断。

二、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，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。例如，在进行产品的废品率检测时，无需对每一出产产品进行逐个检测，而只需随机抽取一定比率的产品进行检测。

对个体行为或行为特征的测查同样无法做到对所有情况面面俱到。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。如果是检查小学生是否具备多位数加法的技能，我们不需要穷尽所有可能的加法算题，而是通过有限个算题举例，来判定其技能。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)**。我们通过对这个行为样本的测量推测个体对所有运算问题的解决行为，进而推断个体的数学运算能力的优劣。

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，其结论

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最低。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,因为应答者猜对的概率很大。也就是说,从少数一两个行为,我们无法判断被评价者是靠猜还是靠真正的能力解决问题。

其次,要保证所选样本具有对总体样本的代表性。例如,我们不能只靠检测个位数加法题目来代表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,我们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。

2. 标准化

标准化 (standardization) 是指测验编制、实施、记分和测验分数解释必须遵循严格的统一的科学程序,保证对所有被测者来说施测的内容、条件、记分过程、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被测者来说,对同一项心理测验,测验题目所引发的被测者的心理状态、行为反应可以不同,但呈现在被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被测者不同的推理过程,但题目本身对所有人来说是一样的。又例如,一个测查工作动机的题目可能引发不同被测者不同的联想和反应,但题目本身对所有的人都是一样的。只有保证题目标准化,即对所有被测者给定的条件一样,才能通过考察不同被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)变量是假定的被测者的内在素质,输出(因)变量是所观测到的被测者的反应即应答行为,包括行为的过程和结果。

2) 实施过程和记分的标准化

标准化的人事测量在实施时必须保证所有被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、解释含糊,而有的考场主考人口齿清晰,解释明确,由此造成的被测者的反应差异,我们无法判定究竟是由于被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,测验的记分也必须有标准化的程序。要做到不同的记分者对同

一被测者同一测验答案的记分方法和结果是相同的。人事测量中有一部分方法是相对较依赖主观记分的,如“无领导小组讨论”和“公文筐测验”的部分内容,这种类型的测验要做到记分系统的完全标准化相对较为困难。在这种情况下,虽然很难要求不同记分者记分结果完全一致,但通常要求有较高的相关,一般相关系数要达到 0.80 甚至 0.90。目前,这类测验在记分方法、程序上已达到相当客观化、量化的程度。

3) 选用有代表性的常模

常模 (norm) 是一组具有代表性的被试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被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被测者所处的水平。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 (如内外向) 的大小。

平均数 (mean) 是常模的重要内容之一。这也是一个最通俗的统计概念,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:

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

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 2/3,即 20 道题,则这 20 道题的记分分数为一般大学生的平均值。平均分 (数) 描述了相应被测者群体平均的水平。

只有平均数还不够,因为我们还不能把一个被测者的原始测验分数和平均分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为 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)** 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:

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

其中, M 为平均数; N 为数据个数; X_i 为各数据值。

标准差取值越大,说明数据分布越分散,离散度越大。

标准差的实用意义可以从以下这个例子体现出来 (见表 1.1)。5 个考

官给甲、乙两个应聘者就某项能力打分。5个考官给甲的分数分别是1、2、3、4、5分，而给乙的分数都是3分。甲、乙两个人的平均分一样，都是3分。但谁更好？因为通常是一个职位两个甚至更多的人应聘，我们总希望能把应聘者区分开来。针对这个例子，显然乙得到了一致的评价，而对甲的评价则众说纷纭，褒贬不一。也就是说，两个同样的平均分，“含金量”不同。相对来说，乙要更好一些。实际上，乙的得分的标准差为“0”，而甲的得分的标准差为“1.41”。

表 1.1 不同考官对不同应聘者的评分

	考官 1	考官 2	考官 3	考官 4	考官 5	平均分	标准差
应聘者甲	1	2	3	4	5	3	1.41
应聘者乙	3	3	3	3	3	3	0.00

从以上说明可以看到，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被测者的特点。在智力测验中，人们关心的往往是自己的测验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 IQ 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， 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，后端分子式取值为“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，后端分子式取值刚好为“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被试样本，即建立常模的这些被试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被试样本，因为他们的分数无法代表一般人的水平，而往往是高于一般水平。

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。测验的题目太容易，所有人都成功地完成所有项目，得分都很高，或测验题目太难，使得大部分人的得分很低，都会使所有被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被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