

应用型高等院校经管类系列实验教材·心理学

员工素质测评实验

李 红 / 主编

Yuan Gong Su Zhi Ce Ping Shi Yan



经济科学出版社
Economic Science Press

应用型高等院校经管类系列实验教材·心理学

员工素质测评实验

李 红 / 主编

Yuan Gong Su Zhi Ce Ping Shi Yan

图书在版编目 (CIP) 数据

员工素质测评实验 / 李红主编. —北京: 经济科学出版社,
2010. 1

(应用型高等院校经管类系列实验教材·心理学)

ISBN 978 - 7 - 5058 - 8987 - 3

I. 员… II. 李… III. 企业管理: 人事管理—考核—
实验—高等学校—教材 IV. F272. 92 - 33

中国版本图书馆 CIP 数据核字 (2010) 第 006040 号

责任编辑: 白留杰 白 炜

责任校对: 杨 海 王肖楠

版式设计: 代小卫

技术编辑: 李长建

员工素质测评实验

李 红 主编

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: 100142

教材编辑中心电话: 88191354 发行部电话: 88191540

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

电子邮件: [esbj3@ esp. com. cn](mailto:esbj3@esp.com.cn)

北京密兴印刷厂印装

787 × 1092 16 开 8.75 印张 200000 字

2010 年 3 月第 1 版 2010 年 3 月第 1 次印刷

ISBN 978 - 7 - 5058 - 8987 - 3 定价: 18.00 元

(图书出现印装问题, 本社负责调换)

(版权所有 翻印必究)

总 序

实践教学是高等教育本质的必然要求,是践行应用性人才培养的必经之路,是地方行业性教学型本科院校办学的重要特征。近几年来,各高校经济与管理类专业实验教学已经逐步开展,把实验教学作为教学改革的抓手、知识融合的平台以及联系社会的桥梁,然而如何进一步完善实验教学体系、提高实验实践教学水平与质量已经成为各高校亟待解决的问题。应用型高等院校经管类实验教材系列以提高高等院校经济与管理类专业实验教学的建设水平为目的,以实验教材建设为突破口,探讨高等院校经济与管理类实验教材的新方向、新思路、新内容、新模式。

本系列实验材的编写紧紧围绕“知行合一,能力为尚,积淀特色,共享协作”的地方行业性教学型经济与管理类实验教学理念,贯彻以现代教育技术为基本手段,以实验资源共享与应用为条件,强化理论教学与实践教学互动与互补,“实践与理论相结合”和在“做中学”的指导思想,强调实验教材建设与实验课程建设、实验项目建设、实验教师队伍建设以及深化实验教学改革相结合,力图通过系列教材建设规范实验教学内容和实验项目,促进实验教学质量的提高。

(一) 本系列实验教材内容与教学方式符合实验教学规律和要求。具体表现在以下几个方面:

1. 实验教材以实验项目为章节,按如下体例编写:实验目的和实验要求;实验的基本原理;实验仪器、软件 and 材料或实验环境;实验方法和操作步骤;实验注意事项;数据处理和实验结果分析;实验报告。当然,对于不同的课程,根据其本身的学科特点,实验教材的编写体例并不完全一致。

2. 增加综合性、设计性、创新性实验项目的比例,并逐步将科研成果项目转化为教材的实验项目。

3. 与当前流行的实验平台软件或硬件及教材内容紧密结合,符合一般软件要求。

4. 充分体现以学生为主体,明确实验教学的内涵。实验教学过程体现以学生操作为主,教师辅导为辅,少量时间教师讲解,大部分时间学生操作的特点。

5. 按实验教学规律分配学时,并且有多余的实验项目供学生利用开放实验室自主学习。

6. 内容精练,主次分明,详略得当,文字通俗易懂,图表与正文密切配合。

(二) 本系列实验教材遵循实验教学规律,体现时代特色,总体来说,具有以下四个特点:

1. 与现代典型案例相结合。以培养应用型人才为原则,根据实验教学大纲,注重理论联系实际,教材具有较强的实践性、新颖性、启发性和适用性,有利于培养学生的实践能力和创新能力。

2. 建设形式新颖。实验教材分为纸质实验教材和网络资源的形式;纸质教材实验报告

尝试做成活页形式，或做成可撕下的带切割线形式；在纸质教材出版，配套建有供学生实验前和实验后学习使用的网络资源。

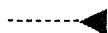
3. 实验内容创新。对于实验教材编写内容上的创新，一是凸显应用型人才培养特色实验项目，提高了综合性、设计性、创新性实验项目的比例；二是将教师的科研成果转化为本科学生实验教学项目。

4. 编写程序严格。对实验教材的申请立项的实验教材经由学院领导及专家进行立项审查；实验教材初稿经由相关同行专家给出鉴定，最终审核后，送交出版社评审出版。

本系列教材得到各方面人士的指导、支持和帮助，尤其是得到中国经济信息学会实验经济学与经济管理实验室专业委员会的专家，广东金电集团等多家业界人士，以及各高校同行老师们的支持和帮助，我们在此表示由衷地感谢。本系列实验教材尚处于探索阶段，作为一种努力和尝试，存在诸多不足之处，竭诚希望得到广大同行及相关专家的批评指正。

应用型高等院校经管类系列实验教材编委会

2009年12月



前言

员工素质测评是以心理测量理论为基础,人事管理领域中应用性、操作性强的新兴学科,相应心理测量工具在人员招聘录用、绩效考核、职业生涯管理和干部选拔与任用方面使用广泛。

本实验教材编写目的在于整合员工测评领域中最新进展,将员工素质测评实验课程与人力资源管理专业、应用心理专业、社会保障专业等相关专业学生实验操作能力培养结合起来,掌握心理测评技术技能;同时本教材为职业生涯规划提供理论依据和实操技术;本教材也是各专业大学生就业指导的良师益友。

本教材的特色是实验基本原理与操作技术相结合,既阐明实验的心理学原理,又明确具体实验操作步骤,结合相应的实验分析和案例。具体特色如下:

1. 实验内容以实验项目为单位。每个实验项目包括:实验目的、实验原理、实验内容、实验步骤、实验结果分析、实验报告、案例、实验延伸使用思考等。

2. 突出内容实务性、操作模块性和环境真实性。本教材的实验部分为员工素质测评常用基础内容,操作上既独立为模块,也可综合使用设计,真实实用。

3. 实验软件与教材内容紧密结合。采用北森测评软件,此软件购买为中央与地方共建项目,2007年购买,2008年升级。

4. 设计性实验项目和特色性试验项目结合。

本教材分为三个部分:第一部分,员工素质测评基础部分,简要地介绍相关的概念术语以及北森软件;第二部分,各个实验的详细介绍;第三部分,是测量的综合性组合设计和特色实验项目。其中第二部分是本教材的主体部分,除了测验的理论解释以及维度解释,还有实际应用价值的案例,为了帮助学生更好地理解和使用测验。

本教材是广东金融学院出资,由李红、李夏妍和杨洋合作编写,其中李红主持教材编写工作,负责前言、员工素质测评基础、性格测验、智维测验、北森软件简介、特色实验项目部分;李夏妍编写能力测验、动机测验、职业兴趣测验、职业锚测验和领导力测验部分;杨洋编写样章,负责16PF测验、团队角色测验、评价中心、SCL-90心理健康测验、综合组合设计和课件部分;李夏妍和杨洋负责全书校对。

该书编写过程中,华南师范大学莫雷教授不吝赐教,百忙之中,给予学术指导和技术资源支持,并且积极肯定,在此表示深深的谢意!

在此感谢广东金融学院同仁的帮助支持,吴晓义教授中肯的建议是我们学术的有力支持!感谢劳动经济与人力资源管理系热诚相助,感谢北森公司提供的资料与帮助。

本教材为第一版,不当之处敬请各位读者提出宝贵的意见及建议。

广东金融学院应用心理教研室

2009年12月

目 录

第一部分

员工素质测评基础	(1)
北森人力资源测评系统简要介绍	(11)

第二部分

实验一 16PF 测验	(13)
实验二 性格测验	(23)
实验三 智维测验	(32)
实验四 能力测验	(38)
实验五 动机测验	(44)
实验六 职业兴趣测验	(50)
实验七 职业锚测验	(56)
实验八 领导力测验	(61)
实验九 团队角色测验	(69)
实验十 评价中心 (综合性实验)	(77)
实验十一 SCL - 90 心理健康测验 (学生自主实验)	(98)

第三部分

测评的组合设计	(107)
特色实验项目: 大学生职业生涯规划	(117)
附录 实验报告样例	(125)
参考文献	(129)

第一部分

员工素质测评基础

随着我国经济改革开放的发展，员工素质和人力资源的开发问题越来越受到国内企业家和研究者的关注，在借鉴国外人力资源管理经验的同时，也在不断尝试开展本土化的、符合我国国情的员工开发措施和管理理论。可见，员工素质和人力资源的开发是经济发展的决定因素，也是当今世界各国都十分重视的领域。

所谓人力资源开发，简单地说，就是以人力资源为中心，培养、发掘和充分利用人的能力，以实现社会、经济、文化的可持续发展的全过程。人力资源开发的内容主要包括对人力资源的能力培养、能力发挥和充分利用等方面。那么，如何有效地开发和使用人力资源，也正是目前国内一些专家、学者以及管理人士不断探索和研究的课题，员工素质预测实际上是推动社会重视和充分利用员工资源的手段之一，也是人力资源开发的重要环节。

一、员工素质测评

员工素质测评就是心理测试技术在人力资源管理领域的应用，它以心理测试为基础，针对特定的管理目的如招聘、安置、考核、晋升、培训等，对员工的素质进行多方面的系统评价，从而为人力资源管理工作提供参考依据。素质包括那些完成特定工作所需的或与之相关

的技能、能力、性格、兴趣、动机等个人特征，它们是以一定的速度和质量完成工作的必要基础。例如，在招聘时，用人单位希望了解应聘者多方面的素质，他喜欢与人打交道吗？他的思维逻辑性如何？他善于影响带动他人吗？他是否具备创新精神？他能把繁琐的工作处理得井井有条吗？以上这些都能在人事测量的结果中找到答案。人们常说的心理测验（Psychological Test）则是心理测试的一种工具和手段，阿娜斯塔西（A. Anastasi）是著名的心理测验学家，她认为，心理测验实质上是行为样本的客观和标准化的测量。通俗地说，就是通过观察人的少数有代表性的行为，对于贯穿在人的行为活动中的心理特征，依据一定的原则进行推论和数量化的一种科学手段，比如对一些标准化测验的运用过程，本书所介绍的能力测验、性格测验、职业兴趣测验等都属于标准化测验。

具体来说，员工测评是指运用先进的科学方法，对社会各类人员的知识水平、能力及其倾向、工作技能、个性特征和发展潜力实施测量和评鉴的人事管理活动。它是一门融现代心理学、测量学、社会学、统计学、行为学及计算机技术于一体的综合性学科。

二、员工素质测评的功能和作用

（一）员工素质测评的功能

1. 甄别和评定功能。甄别是测量个体之间的素质差异，评定是衡量受测者素质构成及其成熟程度，看其是否具备规定的资格条件和常模标准。

2. 诊断和反馈功能。诊断是指通过测评找出被测者素质构成及发展上的不足。反馈是指根据测评结果，提供调整和改进测评对象素质缺点的信息，分析缺点和不足及其产生原因，提出诊断意见和素质优化方案，帮助其克服缺点、发扬优势，推动其素质全面发展。

3. 预测和激励功能。预测功能是指通过对员工素质现在状态的鉴别评定，可以预知、推测其素质发展的趋向。

（二）员工测评的作用

1. 配置员工资源的作用。员工资源的配置包括人与事的配置及人与人的配置两个方面。员工资源的配置是员工资源管理的基础工程，员工测评的作用是由测评的功能派生出来的。如前所述，员工测评可以测查人的素质状况，优点及缺点，了解其特长、兴趣和爱好。掌握了这些信息，可以为人事部门任用、安置人员提供客观依据，为领导进行人事决策提供科学依据。

2. 推动员工开发的作用。人事管理的实质是一个不断开发人力资源的过程。无论是发现、选拔、任用新的员工，还是对已有员工根据新的工作任务和目标要求重新进行调整、优化组合，都必须首先了解人，即“知人”。“知人”必须借助员工测评的科学技术和手段，才能对不同人的德、智、能、绩的实际水平有较为客观准确的了解，从而达到“人尽其才”的最佳人事管理状态。

3. 调节员工市场的作用。在社会主义市场经济条件下，员工流动一方面要靠计划配置，

另一方面主要是靠基础性的市场机制来调节。这种调节既要有政府宏观调控及法规保障，同时也离不开员工测评的信息咨询和作用。

三、人事测量的发展和应用

虽然心理测验只是世界上最年轻的学科之一，但其思想和实践源远流长。我国始于汉代、兴于隋唐的科举制度被中外学者公认为最早的心理测试的实践。

在心理测验的发展史上，法国心理学家比奈（A. Binet）编制的智力测验是世界上第一个具有应用价值的心理测验，比奈也因此成为本领域内的鼻祖。他的突出贡献在于建立了相对客观、量化的度量工具，并将智力测验成功地运用于教育领域。

在第一次世界大战期间，团体施测的陆军甲种测验和陆军乙种测验用于大规模选拔军官，心理测验应用的普遍性大大提高。在第二次世界大战期间，美国中央情报局的前身——美国战略情报局开始大量使用评价中心技术来选拔情报人员。自1956年开始，美国电话电报公司（AT&T）率先将评价中心引入商业组织中，并证实了这种测评技术的有效性。

在中国，从20世纪的80年代开始，随着改革开放外资企业进入国内，人事测量进入迅速发展时期，其在组织管理中的价值受到了越来越广泛的关注，同时也出现了一批专业的机构和公司，从事相关的实务。

四、心理测验的分类

1. 对测验类型的选择是人事测量过程中不可避免的一环，心理测验的分类方法多种多样，在人事测量领域，按照测验的内容分类，可分为两种。

能力测验。测量人的一般能力或者某方面的能力，例如，逻辑推理能力、领导技能、机械操作能力、英语能力、计算机编程能力等，都属于能力测验。能力测验的特点是最高作为，题目有一定难度，有正确答案，要求受测者尽量表现出最高水平。

人格测验。广义的人格测验用来测量受测者的性格、兴趣、动机、价值观、态度等特点。例如，本书介绍的卡特尔16人格因素问卷（16 Personality Factor Questionnaire, 16PF）、职业兴趣测验、职业锚测验等。狭义的人格测验仅仅指性格测验。与能力测验不同的是，人格测验一般没有所谓的答案，而是要求受测者按照自己的典型行为作答。

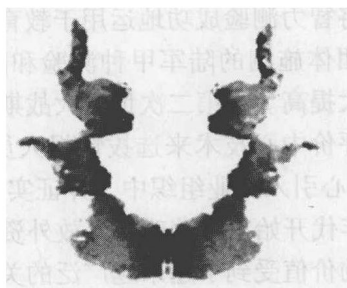
现在流行的胜任力（Competency）测验，则是综合了上述两类测验的内容，对与工作绩效相关的心理特质加以考察。

2. 从测验的不同形式来分类，可以分为以下两种：

（1）标准化测验。一个标准化测验系统包括客观的计分系统、解释系统、良好的常模，以及信度、效度、项目分析等数据。大多数能力测验、人格测验、成就测验等都采用了标准化测验的形式。标准化测验具有方便、经济、客观等其他测验无法替代的优点，但同时也存在不足，如不能对受测者实际的行为表现进行测量，很多人格测验采用自陈式的题目，也就是由受测者自己报告自己的行为。例如，判断“我在接人待物方面总是不太成功”这

句陈述是否符合自己，有的受测者可能会掩饰自己的真实情况，按照社会认可的方式去作出自我评价，使自己看起来更“好”，这就是社会赞许性作答，多数测验往往无法完全解决这个问题。

(2) 投射测验。投射测验主要对于人格、动机等内容的测量。要求受测者对一些模糊不清或结构不明确的刺激作出描述或反应，这些刺激包括图片、未完成的句子，或者让受测者创造一些图画、故事等，由专业的评价者作出解释。例如，罗夏墨迹测验（见下图）、主题统觉测验（Thematic Apperception Test, TAT）等。投射测验因为掩蔽性高，受测者一般不知道测验的目的和自己的反应作何解释，所以伪装的可能性大大降低。但是由于投射测验的计分和解释相对缺乏客观标准，分析结果费时费力，对评价者的要求很高，使得这类测验在人事领域中的广泛应用受到了限制。



“罗夏墨迹测验”的墨迹图

行为模拟与观察。这类测量通过安排一定的情境，在其中观察受测者的特定行为，来从中分析所要考察的素质或特征。如角色扮演、无领导小组讨论、公文筐测验等形式。行为观察法能够尽可能接近自然地观察受测者的真实行为反应，因此备受关注，其弊端是操作较为困难，对评价者的要求很高，测量的成本也很高，因此对于高层管理人员应用较多。

在实际的应用中，要根据希望对受测者的哪些方面加以考察和评估，再综合考虑人力成本、经济成本、时间成本等多方面因素，权衡利弊，选择合适的测验内容和测验方式。

五、心理测验的信度和效度

企业在选择心理测验工具的时候，一个重要的标准就是测验的准确性，但这一点对于很多人力资源工作者来说很难判断，因为心理测验有很强的专业性，从表面上很难看出测验的“质量”如何，其实，从心理测量学的角度来讲，衡量其准确性有两个重要的指标——信度和效度。

1. 信度——什么样的测验才“可信”？信度（Reliability），主要是指测量结果的可靠性或一致性。也就是说，我们希望在相同时间、不同测验条件下所得分数之间保持一致，而要避免这样的情况出现：一位应届生招聘时的测验结果显示责任心非常强，但录用进来半年以后再测，发现责任心又变得非常差了，这就是测验的信度不足导致的。此外，信度还反映了测验中得到的分数与他的“真实分数”之间的接近程度，这也就是很多人担心的一类问题：“测验显示他的人际影响力很强，可他的人际影响力真的有这么强吗？”因为心理测试与物

理测量一样，都存在着无法消除的误差，不可能达到百分之百的准确无误，而且人的行为会在不同的时间、不同的情境下变动，所以，心理测试比物理测量的误差来源要更多。信度就是衡量误差大小的指标，有内部一致性信度、重测信度、复本信度等。信度越高，误差越小。那么，一个测验的信度达到多少，企业才可以放心地使用呢？从专业的角度，一般来说，能力测验的信度应在 0.90 以上，人格（性格）、兴趣等测验的信度通常在 0.75 以上，如果测验整体的信度在 0.65 以下，就不适宜使用了。

2. 效度——什么样的测验才“有效”？接下来再看衡量测验的第二个指标——效度（Validity），效度的定义是：与测量目标有关的真实分数方差与总分方差的比率，通俗地说，也就是测验是否精确地测量了想要测的东西。例如，一个测量影响力的测验，由于测验题目编制得不好，实际上测量的是受测者的性格外向性型，而我们知道，尽管二者之间有一定关联，但并不是外向的人就善于影响他人，内向的人就不善于影响他人，这就是测验效度不足导致的。再有像前面提到的社会赞许性作答的问题，在招聘选拔中很常见，如果测验的题目编制得不好，有很多明显的正面描述的题目，就会导致测验测量的不是工作所必需的素质，而仅仅是应聘者的“装好”倾向，这样的测验效度就很低。显而易见，如果在实际工作中运用了这类效度不足的测评工具，会产生严重的影响。避免这种社会赞许性作答的方法之一是在测验中加入专门的“测谎量表”来对受测者是否真实作答进行鉴别，从而提高效度，像明尼苏达多相人格测验（Minnesota Multiphasic Personality Inventory, MMPI）、艾森克人格问卷（Eysenk Personality Questionnaire, EPQ）这些经典的测量工具中都包含“测谎量表”。

反映测验效度高低的指标有多种，例如内容效度、构想效度、效标关联效度等，与信度不同，效度由于指标多样复杂，并没有一个简明的标准。但标准化的测验系统应包含效度的指标。

3. “洋为中用”——测验的中文版修订。现在国内应用源于西方的测验非常普遍，将西方的心理测验拿到国内使用，仅仅将题目翻译成中文是远远不够的，还需要进行严谨的中文版修订工作，重新进行信度、效度研究。这是由于中西方的文化背景有别，中国人和西方人的思维习惯、归因方式、人格特征、行为模式等都有重大的差异，导致国外的测验原样照搬并不适用，从测验的理论基础、题目陈述到计分解释都可能发生偏差。所以，在选择国外的测验工具时，更要特别留意是否经过了中文版的修订，也就是“本土化”的过程。这个过程包括题目的翻译、回译和修改、国内常模（参照人群）数据的收集以及项目分析、信效度研究等一系列工作，才能保证测验的准确性。中文版修订的过程需要大量的人力、物力和时间，而且需要由专业人员完成。

总之，需要慎之又慎的选择可信、有效的测验。否则，一旦人力资源管理中使用了测量不准确的测验工具，例如，把不恰当的人安置到了公司的岗位上，对组织不仅无益，而且有害。

六、心理测验的实施、计分和解释

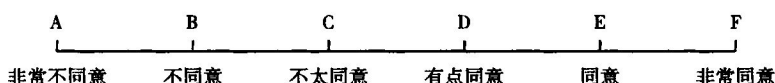
人事测量的目的是对受测者作出尽可能准确和公平的评估，这在很大程度上依赖于心理测验的公平实施、正确计分和合理解释。任何一个步骤出现问题，都会影响到结果的准确性。

1. 测验的实施。

(1) 标准化指导语。测验的实施过程中，标准化的第一步是指导语的标准化，即测验实施过程中应该使用统一的指导语。对受测者的指导语应力求清晰简单，说明他应该做什么，如何对题目作反应，可以写在测验的开头部分让受测者自己阅读，也可以由测验的施测者朗读。例如“北森职业锚测验”的指导语为：

本问卷的目的在于帮助您思索自己的能力、动机及价值观从而找出影响您作出职业决策的最重要的因素。下面给出了一系列描述，每个题目有6级判断，分别表示“非常不同意、不同意、不太同意、有点同意、同意、非常同意”。请对每个选项进行6级评判，并选择相应选项。

例题：我希望我的工作能为我带来很高的收入。



如果你对此描述“有点同意”，则请点选“有点同意”。

注意：所有选项均没有对错、好坏、高低之分，与您的生活和工作没有任何利害关系，请放心根据您的实际情况和真实想法作答，不要考虑应该选择什么或别人希望自己选择什么。

(2) 标准时限。关于测验时限的确定，最重要的考虑因素是测量的目标，例如以典型行为作为主要考察方式的人格测验，受测者的反应速度就不十分重要，但像能力测验属于最高作为测验，速度是需要考虑的重要因素之一。

(3) 测验的环境条件。测验环境也会影响测验的结果，良好的物理环境包括安静而宽敞的地点、适当的光线和通风条件，还要防止干扰。如果测验环境有任何意外因素，需要记录在案，以备解释结果时把意外因素考虑在内。

(4) 计算机辅助的测验实施。计算机实施测验，指导语可以通过视觉或录音呈现，并自动记录测验反应。计算机辅助测验控制更严格，更标准化，优势有以下几点：

- 严格遵守计划和时间。
- 能按照受测者的表现选择题目，例如回答正确后呈现更高难度的题目。
- 不受施测人员疲劳、厌烦、注意涣散和记忆误差的影响，计算机能公正、一致地对待受测者。
- 能迅速准确地计分。

(5) 施测者的职责。施测者需要在测验前做好准备工作，包括通知受测者测验的时间、地点、大致内容等，以便受测者做好必要的准备，还要熟悉测验指导语和测验的具体程序，准备测验材料，并确保合适的测验条件，防止作弊。在测验实施过程中，施测者还需要保证公平、一致地对待所有受测者，不作任何暗示和提示，并和受测者建立起一种友好、合作的、能促进受测者最大限度做好测验的关系。

(6) 影响受测者反应的因素。受测者对测验题目作答时，很多因素会影响受测者的反应，这些因素需要在选择和使用测验时加以注意和避免，或者在参考测验结果作出相关决策时把它们考虑在内。

• 测验的技巧与练习因素。综合各种研究成果，练习的影响仅限于第一次和第二次重测，以后的影响就微不足道了，平均而言，练习的影响幅度约在 0.2 个标准差以下。

• 动机因素。动机过高或过低都可能影响受测者的反应。在测量人格、兴趣、态度、价值观等典型行为的测验中，动机高的受测者为了给别人留下好印象，会产生社会赞许性作答的倾向，会降低测量的有效性。如果受测者的动机不强，在能力测验中，可能不会尽力作答，能力就会被低估，在人格类测验中可导致结果准确性降低。告知受测者测验的目的，以及测验结果对受测者的影响有助于引起受测者的兴趣。

• 反应定势。反应定势指的是受测者在回答问题时有特定的习惯，比如大多数人偏好正面的叙述，偏好较长的选项等。好的测验应该在题目编制上尽量避免受到反应定势的影响。

2. 测验的计分。准确测量的另一个要求是客观计分，可理解为两个或两个以上受过训练的合格评分者之间所评结果具有一致性，如果能达到 90% 以上就是较为客观的。只有计分是客观的，才能把分数的差异归因于受测者的差异。

客观题的计分可以使用计分板或机器计分，能够很好地保证有效客观。计算总分时也可以根据评价者所认为的重要性，在不同的题目上赋予不同的权重。例如，一套测验测量了创新意识、严谨性、团队合作、计划性等多项素质，在计算总分时可以根据对绩效影响的大小给各项素质赋予不同的权重。

对于主观题的计分，首先要记录受测者的反应，有些情境类的测量方式，如无领导小组讨论，可用录像记录，以避免记忆困难和记忆错误。然后将受测者的反应和标准答案加以比较，将受测者的作答归类或赋予分值。当评分者的判断可能成为影响分数的一个因素时（例如：问答题、论文题），就需要对评分规则作详细的说明，可能的话由两位以上的评分者共同评分，取其平均值。

3. 测量结果的解释。人事测量的解释不同于测量的施测和计分，它们是测量中两个相对独立的部分，测量过程和结果解释是分离的。测量的解释是一个相当复杂的系统程序，以一系列的统计工作为基础，并依赖于解释者的专业性。

从测验中直接获得的分数称为原始分数（Row Score），是通过受测者的反应与标准答案比较获得的，其本身的意义并不大。例如，某位受测者的逻辑推理能力是 25 分，我们并不知道他这项能力优秀与否，与平均水平相比是低还是高。我们必须具有可比性的分数标准，也就是常模，常模来自于测验的目标人群总体的分数分布及其特征。原始分数通过与常模的比较，可以转换成等值的分数，叫做标准分（Standard Score），标准分是有意义的，也是具有可比性的。例如，智商 IQ 就是标准分，所以才能够从智商得分判断出人们的智力水平在人群中的位置。如果 IQ 分为 100，就意味着智力中等，高于人群中 50% 的人，如果 IQ 得分为 115，就意味着智力水平高于人群中 84% 的人。这种由测量的原始分数通过与常模的对照得到导出分数的过程，就是常模参照解释（Norm Reference Explanation）的方法，另外一种方法是效标参照解释（Criterion Reference Explanation），它是根据外在的标准来对受测者的分数进行解释。二者的区别在于，前者是将受测者的分数与同类群体的其他人进行比较，例如，实施一项机械技术考试后，如果将受测者的成绩与其他人的成绩进行比较，可以知道受测者在人群当中的位置，则是常模参照解释，如果将受测者的成绩与机械师等级标准进行比较，可以知道是否达到了某一级机械师的水平，就是效标参照解释。

对于常模参照解释的测验，常模的选取是结果解释准确与否的重要因素。常模的选择必须能够代表测验目标人群的总体，取样具有代表性，而且达到一定的数量，常模的大小并没有明确的指标，一般是从经济或实用的可能性和减小误差这两方面来综合考虑常模大小的。常模的时间性也是需要注意的，需要定期更新，一般来说，需要5年更新一次，否则可能会给结果的解释带来偏差。在选择测验时，常模的情况也是衡量测验优劣的一个重要标准。

常用的常模有百分位常模和标准分位常模。使用百分位常模的测验分数用百分位等级 (Percentile Rank) 来表示，它表示人群中低于这个分数的人数百分比。例如，一位受测者的情绪稳定性的百分位等级为85，则表示他的情绪稳定性比85%的人要高。百分位等级的分数非常简单明了，但其缺点是单位不等，如果原始分数的分布接近正态分布，^① 则平均数附近的分数转换为百分位等级时，差异将被夸大，相反的，处于高分和低分两端的原始分数之间的差异将被缩小。例如：在某个测验中，原始分数为80和85的两名高分受测者，百分位等级得分相差3，而原始分数为55和60的两名分数居中的受测者，同样是原始分数相差5分，百分位等级得分相差了20。在实际工作中，这个问题是在应用百分位等级时需要加以注意和考虑的。

标准分的特点是有相等的单位，有确定的平均数和标准差。标准分位常模有好几种，比如常用的韦氏智力量表使用的智商 IQ 就是一种标准分数，平均数是100，标准差是15，又因为智商分数呈正态分布，所以从 IQ 分数可以清楚地看出一个人的分数在人群中的位置。例如，IQ 分数在85~115之间的人，都处于平均数附近1个标准差的范围之内，基本上属于正常水平，如果 IQ 得分70，则对鉴别智力落后是有意义的。

效标参照的测验也有多种表示方法，比如使用受测者答对题目的百分比。另外，在有些管理技能的测验中，可以根据测验分数判断受测者达到某一水平的绩效有多大的可能性，例如，获得65~69分的受测者有94%的可能性达到绩效C等。

七、应用人事测评的方法步骤

如何能将人事测评为我所用，为组织的人力资源管理提供辅助和参考呢？应用人事测评的方法和步骤依据测量目的而有所不同，一般可以按照以下四个步骤进行。

1. 明确需求。要应用测评，第一步是明确测评的需求，做到有的放矢。比如，是将测评应用到哪项工作？是选拔高层管理者、招聘应届毕业生、还是员工职业发展？澄清了应用目标后，下一步还要明确需要应用测验评估哪些内容？例如，如果是用于应届毕业生招聘，就需要确定哪些素质对做好招聘岗位的工作是必须具备的？敏锐的洞察力？出色的人际沟通协调能力？还是关注并完善细节的能力？更进一步的，还要考虑“度”的问题：这些素质要达到什么程度才能胜任这个职位？是越高越好，还是中等即可？这些问题的答案可以从公司的职位要求、工作说明书、胜任力素质模型等渠道中获得，也可以从绩优员工，或者对职

^① 正态分布的函数为 $\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ 在自然界和人类社会中大量现象均按正态形式分布，例如，能力高低、

行为表现、社会态度等。正态分布的基本特征是对称的，而且平均数点最高，然后逐渐向两侧下降。

位有充分了解的上级主管处获取这些信息。如果是以员工普查、员工发展为目的，确定内容就相对简单，不一定需要设定可接受的标准。

2. 选择可信有效的测评工具。在确定了考察的需求点之后，就可以有针对性地去选择相应的测评方式和工具了。衡量测验的准确有效性如何，可以从信度、效度资料着手进行判断，同时，还要注意测验的常模是否有代表性，是否经过了本土化的修订等。

3. 分析和应用测验结果。在应用测验结果中需要注意的是，心理测验中测量的很多素质，并不是一般所认为的越高越好。例如，“挫折承受”指的是人们面临已知的或可能存在的困难与障碍、压力与失败时的心理感受。如果有位应聘者“挫折承受”的百分位等级达到 98（表明他的挫折承受水平比人群中 98% 的人都高），是不是就表明他这方面的素质非常优秀呢？答案并非如此，这样的高分意味着他可能行事有莽撞的倾向，不顾后果，这对于一些需要谨慎周全的职位是不太适合的——例如财务部门。同样，像成功愿望、影响愿望、计划性、外向型等很多看上去“好”的特质，分数过高并不适合某些职位的要求。所以，在根据测验结果对人员作出评价的时候，关键要看得分是否与职位要求相匹配，而不是一味追求越高越好。

在参考测量结果作出相关决策时应该注意，测量结果只是决策信息的一部分，在利用人事测量辅助决策时，例如，人员选拔任命等，也要通盘考虑其他因素和其他渠道获得的信息。另外，决策应该由企业中的相关人员作出，而不能由测验中给出的分数或者等级代替决策者作出决定，因为每个组织有其自身的特点，而且决策所带来的任何结果的承担者也是组织本身，而不是心理测验。

4. 跟踪检验和反馈。人事测评的应用是一项长期的工作，在多数情况下，还需要对测量结果及相关的决策结果进行跟踪，例如，对于招聘中运用心理测试，要根据录用人员的工作绩效进行检验，进行测评的效果研究，比如如果有位员工工作绩效不高，或者几个月后就离职了，那么这位员工的问题出在哪里？是否由于测评中考察的内容不够全面，遗漏了某些重要的素质？还是对某项素质的要求标准有所偏差？这就为前面的工作提供了重要的反馈。这个步骤是非常重要的，可以不断修正测评的应用过程，逐步提升应用人事测评辅助相关的人力资源工作的准确性，使测评在组织中发挥更大价值。

八、心理测验中的伦理道德

心理测验的伦理道德是在应用心理测验中不容忽视的问题，与心理测验所发挥的效能关系重大，同时如果被滥用或者误用将会给社会带来不同程度的危害。中国心理学会在 1992 年颁布了《心理测验工作者的道德准则》，对国内测验使用人员的行为加以规范。美国心理学会（American Psychological Association, APA）在心理测验的伦理道德方面也有详细的规定。心理测验中的职业伦理的基本原则有以下几点：

1. 善行。指负起促进受试者福祉的责任，在测验过程中顾及受测者的利益。

2. 避免伤人。指不要作出伤害别人的事情，例如，在利益冲突情况下，不要从事（甚至是在无意间）可能伤害到受测者的活动。例如，心理测验工作者对测量中获得的个人信息要加以保密，除非在对个人或社会可能造成危害的情况下，才能告知有关方面。

3. 自主权。指维护受测者自我决定的权利。例如，参加心理测验是自愿性的行为，不得强迫他人参加。

4. 公正。指以公平公正的态度去对待所有的受测者，不受其年龄、性别、种族、文化、残障、社会经济地位、生活方式、地域、宗教等因素的影响。也意味着在顾及当事人利益的同时，也必须考虑其他人的利益。

5. 忠实。指作诚实的承诺以及忠于自己的专业工作。创造一种互信的气氛，使受测者在此气氛下与施测者共同合作。本原则也意味着不能欺骗或剥削受测者。