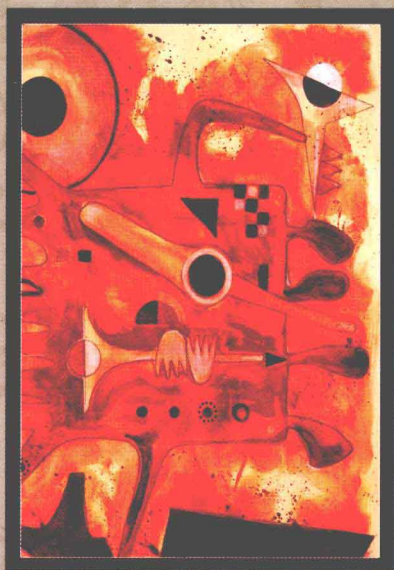


全国高等学校心理学系列教材

总主编 乐国安



Psychological Measurement

# 心理测量学

主编 杜林致

本书重点介绍心理测量学基本概念、基本理论、基本技术和工具等方面的专业知识，兼及这些专业技术在实际生活，特别是人才测评、心理健康评估以及心理健康诊断等领域的应用。

南开大学出版社

# 序 言

由于社会的迫切需要，近二十年我国心理学专业的学生和从业人员数量急剧增长，设有心理学专业的教学科研单位从20世纪80年代末的4个发展到当前的200多个。心理学科不论在政治、经济、文化、教育、体育、管理、健康服务、社区服务、危机处理等领域，还是在学校、企业、医院、行政、司法、军队等部门都发挥着越来越重要的作用。而从学科内部来看，当前不论国外还是国内的心理学研究均在迅速发展，各种新的理论和思想此起彼伏，各种新的研究方法和技术手段不断涌现，使心理学各个领域在宏观的行为层面以及微观的脑基础层面都取得了丰富的新成果与长足进步，从而使心理学的面貌发生了极大的改变。

因此，为了反映当前国内外心理学各个领域的变化与发展，进一步深化高等院校心理学教学改革，加强心理学专业学生的理论素养以及能更好地培养适应新时期社会需要的专业技能，促进我国心理学学科建设和发展，我们组织了目前活跃在心理学教学、科研和实践工作第一线的中青年专家、学者编写了这套反映当前心理学科发展和成果的“全国高等学校心理学系列教材”。

本套系列教材包括《普通心理学》、《实验心理学》、《认知心理学》、《心理统计学》、《心理测量学》、《教育心理学》、《发展心理学》、《社会心理学》、《管理心理学》、《咨询心理学》、《人格心理学》、《西方心理学史》、《中国心理思想史》共13部，其内容选择和结构编排本着专业课程细化兼顾学科交叉的原则，切合当前心理学研究发展的主流方向。

我们在编写本套教材时力图体现以下特色：

第一，科学性与实用性的结合。一方面，在内容的选择上，既确保知识的科学性、正确性，注重科学研究、科学数据对心理现象的说明作用，强调理性对感性的超越，同时，也注重科学原理对日常经验、生活事实的解释作用，体现教材内容对“活生生”社会、生活实际的实用性。另一方面，在材料的组织上，注意处理好学科科学性和教材科学性的关系，既强调学科体系的科学性、

系统性、完整性，同时也从有利于学生学习的角度出发，注重学科的基本结构，注意把握学科体系与教材体系的关系，突出有利于学生学习与掌握的实用性。

第二，前沿性与经典性的结合。虽然科学的心理学至今不过只有一百二十余年的历史，但在这短短的一百二十余年中，心理学家们已从事过数不胜数的研究，获得了无法计量的数据和结果。因此，作为主要面向大学生的教材，需要在科学性、系统性的原则指导下，突出各领域的经典性研究、经典性方法与核心概念和原理，用经典或权威的研究、数据阐述学者们的核心思想与代表性研究。而由于最近十余年心理学界的研究和思想都正在和已经发生了巨大的变化，因此，本套教材在继承历史的基础上，更希望面向现在和未来，强调尽可能多地吸收和反映当前各学科领域的最新成果和进展，力图做到前沿与经典、历史与现在甚至未来相结合。

第三，国际化与本土化的结合。科学的心理学起源于欧洲，成长和壮大于北美，直到今天，欧美心理学仍在当今国际心理学界占据着主导地位。但中国国内外的华人心理学工作者在过去的近百年中，也在学习和借鉴西方心理学研究成果的基础上探索着自己的生存和发展之路，取得了不少重要和有影响的成果。因此，本套教材一方面注重较全面反映国际心理学各领域研究和发展的轨迹、前沿，同时也尽可能结合中国（华人）心理学界的研究与成果，注意反映中国及华人社会特有的心理现象与特点。

第四，学术性与可读性的结合。作为主要面向 21 世纪新时代大学生的教材，在编写过程中，我们既注重专业教材的学术性和科学性，同时也尽量顾及当代大学生学习和阅读的心理特点，不论在内容编选还是在写作风格、编排体例上，均强调教材的易读性、生动性和形象性，力图做到学术性与易读性的结合，希望使这套教材能成为一套教师认为好用、学生认为好学的专业教材。

本系列教材汇集了集体的智慧，是大家精诚合作的产物。虽然在写作过程中我们尽心尽力，力求完善，但由于时间和学识的限制，书中难免存在这样或那样的缺陷和不足，敬请广大读者指正。

本系列教材编写过程中，参考和引用了国内外大量的研究资料，在此向这些作者表达诚挚的谢意！同时，也要感谢南开大学出版社有关领导的大力支持和诸位编辑的精心工作，尤其要衷心感谢策划编辑莫建来同志长期以来对出版心理学专著与教材的热忱和远见卓识。

乐国安 谨识

2010年11月3日于南开园

# 目 录

第一章 心理测量和测验概述 .....	1
第一节 心理测量和心理测验的概念 .....	1
一、测量的基本问题 .....	1
二、心理测量的概念 .....	3
三、心理测验的概念 .....	4
四、测量、测验、评价的关系 .....	5
第二节 心理测验的类型 .....	6
一、按照测验的功能分类 .....	6
二、按测验对象分类 .....	7
三、按测验方式分类 .....	7
四、按测验目的分类 .....	7
五、按测验难度分类 .....	8
六、按测验要求分类 .....	8
七、按测验性质分类 .....	8
八、按测验解释分类 .....	9
九、按测验应用分类 .....	9
第三节 心理测验的用途 .....	10
一、心理测验在教育工作中的用途 .....	10
二、心理测验在人事决策中的用途 .....	11
三、心理诊断 .....	12
四、促进自我了解和发展 .....	13
五、科学研究 .....	13
第二章 心理测量和测验的历史 .....	14
第一节 西方早期的心理测量尝试 .....	14

一、对智力落后儿童的分类和训练的早期关注	14
二、心理测验的先驱	15
第二节 西方心理测量的诞生及其发展	17
一、比纳和世界上第一个智力测验	17
二、西方心理测验的发展	18
第三节 中国心理测量的历史与现状	22
一、中国古代心理测量思想和方法	22
二、近现代中国心理测验的发展	26
第三章 心理测量的流程	29
第一节 心理测量的步骤	29
一、心理测量的一般步骤	29
二、确定测量目标	30
三、测验工具的建立	32
第二节 心理测量的实施、计分和解释	36
一、心理测量的实施	36
二、心理测量的计分	41
三、心理测量结果的解释	43
第四章 心理测量的信度	45
第一节 信度的理论	45
一、信度的理论含义	45
二、测验分数不一致性的来源	47
第二节 信度估计的方法	50
一、重测信度	51
二、复本信度	52
三、内部一致性信度	52
四、评分者信度	55
第三节 影响信度系数的因素	56
一、受测团体的范围	56
二、测验的长度	57
三、测验的难度	58
第四节 测量的标准误差	59

二、项目分析的目的	92
第二节 项目分析的方法	93
一、项目难度	93
二、题目的鉴别力	97
三、选择题的选项分析	99
四、项目难度、鉴别力和迷惑项之间的关系	100
第三节 不同类型测题的项目分析	102
一、常模参照测验的项目分析	102
二、标准参照测验的项目分析	104
三、速度测验的项目分析	107
四、项目分析举例	107
第四节 项目反应理论	109
一、经典测验理论的局限性及项目反应理论的诞生	109
二、IRT 的概念及其假设	111
三、项目反应理论的基础模型	114
四、IRT 的实际应用	115
第七章 量表、转化和常模	119
第一节 测验分数及其意义	119
一、原始分数和导出分数	119
二、常模和标准化样组	120
第二节 发展性常模和发展量表	126
一、年龄常模	126
二、年级当量	126
三、顺序量表	127
四、发展商数	128
第三节 组内常模和量表	130
一、百分等级与百分位数	130
二、四分位数和十分位数	131
三、标准分数	131
第四节 标准参照分数	135
一、内容参照分数	135
二、结果参照分数	137

三、教育成就测验的作用 .....	182
第二节 标准化学业成就测验 .....	184
一、标准化学业成就测验的编制步骤 .....	184
二、标准化学业成就测验的题目类型 .....	187
三、常见的标准化成就测验 .....	191
第三节 教师自编测验 .....	195
一、教师自编测验的特点 .....	195
二、教师自编测验的步骤与方法 .....	196
第十一章 人格测验 .....	199
第一节 人格测验概述 .....	199
一、人格与人格测量 .....	199
二、人格测量的历史 .....	202
三、人格测验的类型 .....	204
第二节 自陈人格量表 .....	208
一、卡特尔 16 种人格因素问卷(16PF) .....	209
二、明尼苏达多相人格问卷(MMPI) .....	218
三、艾森克人格问卷(EPQ) .....	228
四、加州心理调查表(CPI) .....	232
五、迈尔斯—布里格斯性格类型测验 ( MBTI ) .....	238
六、其他问卷式自陈人格测验 .....	243
第三节 投射测验 .....	253
一、投射测验概述 .....	253
二、常见的投射测验 .....	255
三、对投射类测验的评价 .....	268
第四节 人格测量中的评定量表和情境测验 .....	268
一、人格评定量表 .....	268
二、情境测验 .....	271
第五节 人格测验中存在的问题 .....	273
一、人格测验中的误差——反应偏向 .....	273
二、人格测验中存在的其他问题 .....	277
第十二章 其他常用的心理测验 .....	279

第一节 需要和动机的测验	279
一、需要测验	279
二、动机测验	281
第二节 兴趣测验	283
一、兴趣概述	283
二、兴趣测验	283
第三节 态度和价值观测验	292
一、态度测验	292
二、价值观的测量	297
三、民意测验和问卷编制	300
第十三章 心理测量的应用	302
第一节 心理测量在工业组织中的应用	302
一、人力资源测评的含义和特征	302
二、人力资源测评的种类	303
三、人力资源测评的方法	305
第二节 心理测量在临床和咨询心理学中的应用	318
一、心理健康调查问卷	318
二、专项心理疾病诊断问卷	321
三、应激及相关问题评定	326
四、神经心理测验	328
第三节 心理测量在政治和军事领域中的应用	332
一、公务员考试	332
二、在军事领域的应用	334
附录一 心理测验管理条例	340
附录二 心理测验工作者职业道德规范	347
主要参考文献	349
后记	351



测量的工具（某种法则）：温度计。读数时温度计与眼睛平齐。

测量的结果（数字）： $\times\times^{\circ}\text{C}$

例 2：测量学生的数学水平

测量的对象（事物的属性）：学生的数学水平。

测量的工具（某种法则）：用预先编制好的数学试卷，按照测验的要求进行，包括规定测验的时间、不准作弊等等。

测量的结果（数字）：测验的分数。

### （三）测量的量表及其类型

要测量某个事物，必须有一个定有单位和参照点的连续体，将所要测量的每个事物放在这个连续体的适当位置上，看它距离参照点的远近，以此得到一个测量值，这个连续体就叫量表（measurement scale）。在这里，参照点是计算的起点，分为两种：绝对参照点，即以绝对的零点作为测量的起点，如长度测量和重量测量；相对参照点，即以人为确定的零点为测量的起点，如对地势高度的测量。单位是测量的基本要求，理想的测量单位应当具备两个条件：第一，有确定的意义，即对同一单位，所有的人的理解都是相同的，不能做出不同的解释。第二，有相等的价值，即第一个单位与第二个单位之间的距离等于第二个单位与第三个单位之间的距离。

根据测量的精确程度不同，斯蒂文斯（S. S. Stevens）将测量从低级到高级分成四种水平，即名称量表、等级量表、等距量表和比率量表。

#### 1. 名称量表

也称类别量表，是指用数字或符号对事物属性进行分类，如性别、国籍等。名称量表是测量水平最低的量表，所得的数据或符号只有区分性，而没有序列性、等距性和可加性。因此，运用名称量表时不能作常用的数量化分析，只能计算频数、百分比、卡方等。

#### 2. 等级量表

也称顺序量表，是用数字将事物的属性分成等级，如优良中差、名次等。它是次低水平的量表，所得的数据或符号，能区分不同类别，并能排出等级或顺序，但没有相等的单位，也没有绝对的零点。也就是说数据具备了序列性，但是仍然没有等距性和可加性，因此不能进行加、减、乘、除等代数运算，仅限于次数统计，以及中位数、百分位数、等级相关系数、肯德尔和谐系数、秩次变差分析等。

#### 3. 等距量表

也称区间量表，是指将事物属性等距划分，其单位等值，但没有绝对零点，

### （一）测量与测验的关系

测量是根据一定的法则用数字对事物加以确定，凡是涉及人的心理活动和心理属性的测量都可以称为心理测量。而心理测验是心理测量的一种，是高级形式的测量法，它不仅具备心理测量的所有属性，还具备其特定的含义。如教师将学生操行品德分为“甲、乙、丙、丁”四个等级，根据每个学生的表现（外显行为）打分，就属于心理测量，而不属于心理测验。心理测量的研究范围比测验要广泛得多。

心理测量是一种实践性的活动，主要在“动词”意义上使用；而心理测验是了解人的心理与行为的工具，主要在“名词”意义上使用。概括地说，心理测验是心理测量的一种工具和手段。

### （二）测量和评价的关系

测量和评价是紧密相连的两个问题。所谓评价，就是对测量的结果进行价值判定。测量是评价的基础，若没有测量，就没有可进行评价的素材。评价是对测量结果的进一步判定，若仅有测量，而不进行评价，则无法确定该事物属性的价值是优是劣，因此也可把评价的过程看作是一个总结的过程。

评价的内容一般多于测量。比如，评价某个学生的智力，是为了了解该学生在常规学习中的能力。这样的评价不仅应显示学生的智力强项和弱点，也应有社会技能的判断。相反的，测验“可以不用直接回答特定问题，甚至没有测验者直接面向客户或被试而进行”（Maloney & Ward, 1976, P9）。另外，在测量中，测验者可能“将正确答案的数量或特定类型回答的数量相加……很少注意内容是如何运作的”（Maloney & Ward, 1976, P39）。评估更适用于集中于个体的过程是如何进行的，而不仅仅是过程的结果。

## 第二节 心理测验的类型

按照不同的标准，可以把心理测验分成不同的类型。

### 一、按照测验的功能分类

可以把心理测验划分为能力测验、成就测验和人格测验。

能力测验可分为实际能力测验与潜在能力测验。实际能力是指个人已有的知识、经验与技能，潜在能力则是指个人将来可能达到的水平。能力测验还可分为普通能力测验与特殊能力测验，前者即通常所说的智力测验，后者多用于

可以预测人们从事各种活动的适宜性，向人们提供职业选择的科学化数据，可以大大减少职业选择的误差。

#### （四）保护学生的身心健康

不同身心特点的学生其学习会面临不同的问题，根据学生的身心特点进行相应的教学、管理方式、方法的调整，可以有效地保护学生的身心健康。比如根据临床经验，高级神经活动有两种极端不平衡的类型，即不可遏止的强兴奋型和过分抑郁的弱抑制型，它们往往是精神病的主要候补者；强烈的愿望，过度的紧张与不知疲倦的劳累，常常使胆汁质者的抑制过程更加减弱，于是容易出现精神衰弱或发展为时而狂暴时而忧闷的躁郁性精神病；困难的任务，社会的冲突，个人的不幸遭遇，都会使神经过程本来脆弱的抑郁质，感到无法忍受，易于转入慢性抑制状态，于是容易出现受暗示和富于情绪的歇斯底里，或发展为精神分裂症。为了保护学生的心身健康，教师知晓学生的气质类型，诊断出这种极端不平衡类型者，就可以采取一些特殊措施，如使第一类学生多得到工作与休息的交替机会，使后一类学生在集体中获得真正的友谊和生活的乐趣。心理测验具有诊断功能，能帮助教师做科学而有效的鉴别。

## 二、心理测验在人事决策中的用途

在西方国家，心理测验在人事决策中得到广泛的应用。在企业员工招聘、管理人员选拔与晋升、公务员选拔、法官考试，以及参军、升学等领域，心理测验成为必不可少的手段。以美国为例，每年仅人才测评服务的直接收入多达十几亿美元，如果包括与测评服务相关的咨询和培训费用，多达 100 亿美元。

### （一）人才选拔

各行各业的人事部门经常面临选拔人才的问题，即要辨认和挑选出那些具有极大成功可能性的人。心理测量学家根据对各种工作的性质和特点的分析，寻找出适应特定工作要求的心理模式，然后根据这种模式编制测验，借此识别适合从事这种工作的人。用测验法甄别人才不仅大大提高了选才的效率，而且可以避免选才过程中的各种人为因素的影响，从而提高选才的科学性和客观性。

#### 例一：飞行员选拔

据估计，培养一个飞行员所需的费用如果按黄金计算，那么所需要的黄金就相当于飞行员的体重。并不是一般人都可以当飞行员的。在当代快速飞行的情境中，实践已对飞行员的反应提出极其严格的要求。汽车驾驶员把脚从加速器移到制动器踏板上的时间约为 500 毫秒。按这种数据计算，如果飞行员需要突然转换飞行方向，那么必须飞行几百公尺，神经冲动才能从大脑传到手或脚。

如再加上感觉过程中可能遇到的时滞，那么飞行员的反应就可能在飞行了一两公里后才能发出，这显然是不行的。所以必须选拔那些在反应速度上有特殊能力的人参加训练。不仅如此，据资料，在飞行着陆的短短 5 分钟时间内，一个飞行员需做出一百多个操纵动作。在这一过程中，飞行员平均注视仪表一百多次，每次注视时间只有 0.4 至 0.6 秒，但操纵准确度的要求却很高。比如着陆俯仰角度只要相差几度，就有可能造成严重事故。因此，飞行员的素质选拔成为决定飞行员能否合格的至关重要的前提。如果我们事先对飞行员的候选者进行注意广度、视觉聪度、动作准确度、情绪稳定性以及图形识记、运算能力等方面的测验，就可以大大降低淘汰率。我国使用心理测验方法选拔飞行员，使飞行员淘汰率由 65% 下降到 36%。

#### 例二：美国电话电报公司选拔管理人员

1956 年，美国电话电报公司选拔管理人员，分别运用了主管自己选拔、部门经理推荐和评价中心技术测评等三种方法。在把评价结果对公司的高层管理者保密了 8 年之后，将员工实际发展情况与评价结果进行核对发现，评价中心技术测评正确率达到 76%，而部门经理提名的正确率只有 35%，主管自己选拔的正确率仅有 15%。

#### （二）人员安置

随着社会化大生产的发展，人们之间的分工越来越精细，不同的工作对人的素质要求有所不同，这就要求在人员和工作之间选择最佳匹配。通过人事测量，可以对个体的兴趣、人格、能力、技能等多方面进行分析，为实现人才的合理安置提供信息。

#### （三）绩效考评与培训计划

在人事管理领域，对员工的能力水平、工作满意度水平及可供开发的潜力等方面进行评价，是对员工进行考核或培训时应了解的信息；而人事测量能够提供关于个体的行为的描述，形成对被测者的全面的评价，从而为人事考核及培训提供依据。

### 三、心理诊断

对于智力缺陷者和心理障碍者的识别是推动心理测验发展的重要动力。直到现在，对各种智力落后、精神疾病和脑功能障碍等应用心理测验来诊断仍然是一种重要的途径。心理测验获得的资料还可以作为从事心理咨询工作的依据。例如，综合成就测验、智力测验、能力倾向测验、职业兴趣测验和性格测验的资料，可以为一个人的未来职业方向提供咨询意见，以帮助来访者做出正确的

职业选择。利用人格测验和临床精神障碍测验的资料，还可以帮助来访者改善心理环境，提高心理适应的能力。

心理测验的诊断功能不只限于临床，在教育工作中同样可以发挥作用。比如，对特殊儿童（天才儿童、智力低下儿童、问题儿童）进行心理诊断，可为教育措施的制定提供依据。

#### 四、促进自我了解和发展

“认识你自己”是古希腊神庙石碑上镌刻的箴言，是人类亘古以来永恒的话题，但认识自己并不是一件简单的事情。特别是人的性格、气质、需要、动机、价值观等内在心理活动，犹如深藏大海的冰山底层，无法直接透视，必须借助其他手段来间接感知。应用心理测验，就可以对人们在智力水平、人格特点等心理特质上的优势和劣势做出描述和评价，使一个人知道自己的长处和短处，以便扬长避短，更好地学习、工作和生活。

#### 五、科学研究

进行心理学研究的心理学工作者，为了提出或检验自己的某种理论，或解决某种现实问题，会编制和施测心理测验问卷。这类研究是纯学术性的，但正是这些研究成果的累积，才形成今日如此丰富的心理学成果的应用。

## 第二章 心理测量和测验的历史

### 第一节 西方早期的心理测量尝试

#### 一、对智力落后儿童的分类和训练的早期关注

促使产生心理与教育测量技术的最重要的因素是社会发展的需要。

在西方一些国家，工业革命成功后，对劳动力的需要急剧增加，且分工日益精细，因而有了专门人才的训练、人员选拔与职业指导的需要，这是促使测验发展的重要因素。

在19世纪以前，智力落后者（mentallyretard）和精神异常者（insane）遭到与精神病人同样的待遇，常常被忽视、禁闭甚至拷打。随着对精神病人了解的增加，欧美对智力落后者的态度有所好转，有些国家还建立了特殊的医疗机构来收容他们。出于对智力落后者和精神病人治疗和帮助的需要，急需建立一种客观的分类标准和鉴别方法。法国医生艾斯克瑞尔（Esquril）首次提出了区别智力落后者和精神病者的方法：用观察个体使用语言的能力来判断个体的智力水平。这是促使测验发展的另一个重要因素。

工业革命成功后，西方国家对劳动力的需求急剧增加，工厂大量雇用童工。为了使低能者能寻找到维持生计的职业，一些地方官员与工厂主定约：每招收20名童工，必须兼带一名低能者。为了设法使低能者尽可能适应工厂的技术要求，法国医生沈干开始训练智力落后儿童，并于1837年创办了第一所专门教育智力落后儿童的学校。1846年，沈干出版了《白痴：用生理学方法诊断与治疗》一书，专门研究从感觉辨别力和运动控制力方面来训练智力落后儿童，其中的一些方法如形式板为后来的非言语智力测验所采用。

科学家最初发现人的心理具有个体差异是起因于天文学的一个事件。1796

年，英国格林威治天文台的皇家天文学家 N. 马斯基林因为助手金内布鲁克观察星体通过的时间比自己迟 0.8 秒，认为他私心自用、不依法行事而将他辞退。此事在 20 年后引起另一天文学家贝塞尔的注意，他通过研究认为，这是一种不可避免的个人观察误差，于是引起了学者们对个体差异的研究。

## 二、心理测验的先驱

心理测量学几乎与实验心理学同步并举，其先驱者首推冯特、高尔顿和卡特尔。美国著名的心理史学家波林曾指出：在测验领域中，“19 世纪 80 年代是高尔顿的 10 年，90 年代是卡特尔的 10 年，20 世纪头 10 年则是比纳的 10 年”。

### （一）冯特实验心理学的影响

19 世纪中叶以后，由于社会生产力和科学技术的发展，心理学逐步和哲学分开。1879 年，德国生理学家冯特（W. Wundt）在莱比锡大学建立了第一所心理实验室。由此，心理学开始成为一门独立的实验性科学。冯特试图用生理学和物理学的仪器来发现人类心理的一般规律，但在研究中发现不同被试对同一刺激的反应常常不同。最初他以为这是实验设计程序上的问题，经过长时间的实验才认识到，这种差异并不是偶然的错误，而是由于个人能力上的真正差异所导致，这就为日后开展对个别差异的测量学研究提出了课题。

### （二）高尔顿的个别差异测量

最先倡导测验运动的是优生学创始人、英国生物学家和心理学家高尔顿（F. Galton）。高尔顿开创了个别差异心理学研究，并采用了定量研究方法。1869 年，他出版了《遗传的天才》一书，提出人的能力是由遗传决定的，每个人的遗传是有差异的，而这种差异又是可以测量的。高尔顿做了著名的双生子实验来证明他的假说。他于 1884 年在国际卫生展览会内创设“人类测量实验室”以测试人的身高、体重、阔度、呼吸力、听力、视力、色觉等项目，参观者只要付三便士就可以测量。后来，这个实验室迁移到伦敦的肯辛顿博物馆，继续开办 6 年，共测量 9337 人。他在 1883 年出版的《人的能力研究》一书中说到：“外部世界的信息是通过我们的感觉到达我们的大脑的。我们的感觉越敏锐，获得的信息便越多；获得的信息越多，我们的判断与思维便越有用武之地。”高尔顿还注意到，白痴对于冷、热、痛的鉴别能力较低。这一观察结果使他进一步确信，感觉辨别能力基本上是心智能力中最高的能力。虽然他试图从这些生理的、感知觉的材料里推出心理差异的一般规律的努力失败了，但他在研究中所提出的相关概念，却构成了心理测量学的重要思想和研究方法。高尔顿是应用等级评定量表、问卷法及自由联想法的先驱。高尔顿在 1893 年出版的《人类

能力及其发展的研究》一书中，首次提出了“心理测量”和“测验”这两个术语，堪称直接推动心理测验产生的第一人。他的另一个重要贡献是把统计方法应用于对个体差异资料的研究。他将以前数学家们研究出来的统计技术改造为简单形式，使那些未经专门训练的调查者也能使用。他不但扩充了古特莱特（Gmtelet）的百分位法，还创造了一种浅显易用的计算相关系数的方法。他的学生皮尔逊（K. Pearson）推进其事业，创立了积差相关法，成为测量学重要的工具。

### （三）卡特尔的早期“心理测验”

美国心理学家卡特尔是心理测量史上另一位重要的代表人物。他早年留学德国，是实验心理学创始人冯特的弟子。他继承了冯特和高尔顿二人的学说，在哥伦比亚大学实验室里编制了 50 个测验，测量哥伦比亚大学学生的肌肉力量、运动速度、反应时、痛感受性、视听敏感度、记忆力等，使测验走出了实验室。

卡特尔于 1890 年在《心理》杂志上发表了《心理测验与测量》一文，这是心理测量第一次出现于心理学文摘中。文中说道：“心理学若不立足于实验与测量上，决不能有自然科学的精确”，“如果我们规定一个统一的手续在异时、异地得出的结果可以比较、综合，则测验的科学性和实用性都可以增加”。他当时就极力主张测验手续和考试方法应有统一规定，并要有常模以便比较。他的所有这些观点都是测验编制的重要指导思想。

卡特尔认为生理能量与心理能量密切相关，因此他对智力的测量主要是对感觉辨别力、反应时、动作过程等的测量。卡特尔的学生维斯勒（C. Wissler, 1901）进行了一项开创性的研究工作。他收集了哥伦比亚大学和伯纳德大学 300 名学生的智力测量分数和学习成绩，试图以学生的智力测量分数预测学习成绩。但研究结果令人失望，因为二者的相关极低。维斯勒的研究证明用简单的心理过程测量智力的方法是错误的，这启发后来的研究者通过其他途径来研究智力的差异，并为后来比纳—西蒙智力测验的制定提供了借鉴。

卡特尔的重要贡献还在于他为美国心理学界培养了很多优秀的学生，如桑代克（动物实验室创始人、教育测验的鼻祖）、弗朗兹（以研究大脑皮层机能定位而著称）等。



验心理学的鼻祖，我们不得不称比纳为心理智力测量的鼻祖。”

## 二、西方心理测验的发展

心理测验运动自 20 世纪初兴起，20 年代进入狂热，40 年代达到顶峰，50 年代后转向稳步发展。在此期间测验主要有以下几个方面的发展。

### （一）智力测验的发展

比纳—西蒙量表问世之后，立刻传到世界各地，被广泛修订、应用。其中最著名的是美国斯坦福大学推孟（L. M. Terman）教授于 1916 年修订的斯坦福—比纳量表（简称为斯比量表）。其总测题为 90 题，适应的年龄组为 3~14 岁儿童，另加普通成人组和优秀成人组。其最大的改变是采用了智商（intelligence quotient, IQ）的概念，从此智商一词便为全世界所熟悉。此后，它分别在 1937、1960、1972、1986 和 2003 年五次被修订。

1939 年，韦克斯勒（D. Wechsler）发表了第一个用于测量 16~60 岁成人智力的韦克斯勒—贝尔韦量表（W-BI），此后该量表又发展成为韦克斯勒成人智力量表（WAIS, 1955；1981 年修订本 WAIS-R）和儿童智力量表（WISC, 1949，测量对象为 6~16 岁儿童；1974 年修订本 WISC-R；1991 年 WISC 第三版 WISC-III 出版；2003 年第三版 WISC-IV 发行）。1967 年韦氏幼儿智力量表（WPPSI）出版。韦克斯勒智力量表的特点，一是在 1949 年出版的 WISC 中用离差智商代替比率智商；二是由各个分测验结果可以得到言语、操作和总体等三个分数，既可以区分个别间差异，又可以评定个别内差异。在对人的智力的描述方面，从笼统地划分聪明不聪明，转向区分智力的不同侧面，说明人人皆有所长和所短。

1938 年，英国心理学家瑞文（J. C. Raven）出版了瑞文标准推理测验。这是一个著名的非文字智力测验，既可弥补语言文字量表在理论上的缺陷，又可用于测试文盲和有语言障碍的人；既可用于个别测验，又可用于团体测验。1947 年，瑞文又出版了彩色推理测验和高级推理测验。

推孟的研究生奥蒂斯（A. S. Otis）编制出的团体智力测验扩大了测验的应用范围。

心理测验在第一次世界大战期间大显身手，得到了迅猛的发展。以叶克斯为首的一些美国心理学家在原有心理测验的基础上编制出了“陆军甲种测验”（文字测验）和“陆军乙种测验”（非文字测验），并对 200 多万官兵进行了智力检查，取得了巨大的成功。战后，这些测验经过改造，广泛运用于民间，如教育界和工商界等。