

# 第 1 章

## 心理测量总论

人类历经亿万年的演化过程，其认识世界、改造世界的本领达到令人叹为观止的程度，人类对自然界的全面征服已是指日可待。然而，与此形成鲜明对照的是人类对自身认识的贫乏。“人心不同，各如其面”，说的是个体间人格特点的独特性和丰富多样性，“人心难测”又道出了对人的内部心理世界难以了解的几分无奈。随着科学技术的飞速进步和心理科学的日益发展，人们不但对物体的长度、重量、温度以及时间、空间、运动等物理特性作出了越来越精确的测量，而且对人的能力、气质、性格等诸多的心理特性也能够在不同程度上进行测量。尽管人心是难测的，但是人心却绝非不可测量。美国心理学家桑代克在 1918 年曾说：“凡物之存在必有其数量。”仅隔四年之后，美国测量学家麦柯尔（1922）又进一步指出：“凡有数量的东西都可以测量。”人类对自身心理的认识已经从黑箱状态转入到灰箱状态，心理测量与诊断这门学科将为人们提供一把认识自我与他人的金钥匙。

## 第一节 心理测量概述

### 一、测量及其量表

测量是人类生产和生活中普遍存在的现象。例如，农业生产要丈量土地面积，工业生产要测定产品的技术指标，医疗工作要测定人体的生理指标，教育工作要测定学生的学业成绩，等等。根据测量对象的性质和特点，可以将测量大致分为物理测量（如长度测量、重量测量等）、生理测量（如对人体各种生理机能的测量等）、社会测量（如人口普查等）和心理测量（如智力测量、人格测量等）四种类型。

那么，究竟什么是测量呢？史蒂文斯（Stevens S S）曾说：“广义而言，测量是根据法则给事物赋予数量。”也就是说，依据一定的规则对事物的特征进行定量描述的过程就是测量。在此，所谓“一定的规则”是指任何测量都要建立在科学规则和科学原理基础之上，并通过科学的方法和程序完成测量过程。如用杆秤测量物体的重量，依据的是物理学上的杠杆原理；用温度计测量温度，依据的是热胀冷缩

原理。有的测量依据的规则比较科学和完善，测量的结果比较准确、可靠，而有的测量依据的规则比较粗糙和欠成熟，测量结果的准确性和可靠性较差。

任何测量都必须具备两个要素——测量的参照点（计算的起点）和测量的单位。

### 1. 参照点

参照点是计算的起点。参照点不同，测量结果就会因其所代表的意义不同而无法进行比较。两个测量结果如果要相互比较，这两个测量就必须建立在同一个参照点上。参照点有两种：

(1) 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就是建立在以绝对的零点为参照起点的基础上的测量。这个绝对的零点的意义就是“无”，即没有重量或没有长度。

(2) 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，就是以海平面为测量的起点。此时，人们假定海平面的高度为“零”，然后去确定陆地高出海平面多少。

最为理想的测量参照点是绝对参照点，因为它的意义最为明确。但在许多情况下，人们难以找到绝对参照点，所以必须改用相对参照点。物理测量的量具常用绝对零点（如各种度、量、衡器具上的零点）作为参照点。但是，在心理测量中，由于所要测量的特质常常很模糊，所以，很难找到绝对零点，参照点一般随其所参照的标准而变化。心理测量的参照标准主要有两种：常模参照和标准参照。采用相对参照点为测量起点的测量结果只能进行加减运算，而不能进行乘除运算。

### 2. 单 位

单位是测量的基本要求。例如在物理测量中，测量长度以米、厘米等为单位，测量重量以千克、克等为单位，测量时间则以分、秒等为单位。不同测量所用的单位是不同的，甚至同一测量也可用不同的单位。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，不允许做出不同的解释

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。

一般来说，具有绝对零点的物理测量的单位皆符合上述两个条件。但是，心理测量的单位却往往不等即（如 90 分与 80 分之差和 60 分与 50 分之差）或等距不等值（如在难易不同的两道题上同样得 3 分）。

要测量某一特定事物的特定特征的数量，必须首先选择一个具有确定单位和测量参照点的数字连续体，将所测量的特征与这个连续体相比照，确定它的位置，看它距参照点的远近，就会得到该特征的一个测量值。这种能够使事物的特征数量化的数字的连续体就是量表，制定量表的单位和参照点不同，就会编制出不同的量表。不同的量表具有不同的测量水平，相应的测量的精确度也不同。根据测量的不同水平以及测量中使用的不同单位和参照点，我们把测量量表分为以下四种类型：

### 1. 命名量表

亦称类别量表。它是测量中最简单粗糙、水平最低的量表。命名量表只是用数字代表事物或用数字对事物进行分类，数字只是事物的符号，而没有任何数量的意义。也就是说，命名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，运用命名量表时不能作常用的数量化分析，只适用于次数的统计因此，有人认为运用命名量表进行的测量不能算是真正意义上的测量。命名量表可细分为两种形式：一是名称量表，即用数字指代个别事物（如运动员的号码）。二是类别量表，即用数字指代事物的种类（如用 0 代表教师，用 1 代表学生）。

### 2. 顺序量表

亦称等级量表。它是次低水平的量表，它不仅能够指代事物类别，而且能够表明不同类别的大小、等级或事物具有某种特征的程度。也就是说，顺序量表所得的数据或符号，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。各种体育比赛中的名次排列或学生学习成绩的名次排列就是典型的运

用顺序量表进行的测量。在顺序量表中，数字只表示等级、大小和程度的顺序。它既没有相等的单位，也没有绝对的零点。换言之，它既不表示事物特征的真正的数量，也不表示绝对的数值。

### 3. 等距量表

亦称区间量表。它是较高水平的量表，因为它不仅能够指代事物的类别、等级，而且具有相等的单位。等距量表所得的数据既具备区分性、序列性，也具备等距性和可加性。也就是说，等距量表的数字是一个真正的数量，这个数量中各个部分的单位是相等的，因此就可以对其进行加减运算。但是由于等距量表没有绝对的零点，它的零点是人们假定的相对零点，所以对于等距量表中的两个数量不能进行乘除运算，它们之间不存在倍数关系。

### 4. 比率量表

它是最高水平的量表，也是一种理想的量表，因为它除了具有类别、等级、等距的特征外，还具有绝对的零点。使用比率量表，不仅可以知道测量对象之间相差的程度，而且可以知道它们之间的比例。在长度测量和重量测量等物理测量中人们广泛使用比率量表。而心理测量由于难以确定绝对零点，因此很难达到这一量表水平。

一般来说，心理测量是在顺序量表上进行的。但是，由于等距量表适合于大量的统计分析方法，所以人们总是试图将心理测量的结果放在等距量表的背景下去解释。为了避免使分数的解释和推论发生错误，人们常采用的统计方法是把原始分数（测验分数）转换为标准分数。

## 二、心理测量与心理测验

心理测量，就是根据一定的法则用数字对人的行为加以确定。即根据一定的心理学理论，使用一定的操作程序，给人的行为确定出一种数量化的价值。心理测量与物理测量相比，具有以下两点性质：

### 1. 心理测量的间接性

由于心理现象的非实体性和动力性等性质，心理测量比物理测量要复杂得多。科学发展到今天，我们还无法直接测量人的心理实体，

而只能根据人的外显行为，间接地测量人的心理活动的特点和水平，也就是说，我们只能通过一个人对测验项目的反应来推论出他的心理特质。所谓特质，即个体独有的、稳定的、可辨别的特征，它是描述一组内部相关或有内在联系的行为时所使用的术语，是在遗传与环境影响下，个人对刺激作反应的一种内在倾向。但特质又是一个抽象的产物，一个构想，而不是一个能被直接测量到的有实体的个人特点。由于特质是从行为模式中推论出来的，所以心理测量永远是间接的。

## 2. 心理测量的相对性

任何测量都要有参照点和单位，否则，测量就是无意义的。如果参照点不统一，量数所代表的意义就不同，测量的结果也就无法相比较，理想的参照点是绝对零点。测量单位也是进行测量必不可少的，没有单位就无法得到量数，测量结果就无法表述出来，理想的单位应该是有确定的意义和相等的价值。而我们目前使用的心理量表既无绝对零点，也无相等单位。因为我们无法确定一个人的智力、性格、兴趣、态度等的零点是一个什么状态，我们只能人为地确定它们的相对零点。如智力年龄为零，实际上指的是零岁儿童的一般智力，而不能说没有智力。同时，我们还无法用等距的单位将测量结果表达出来。例如，一个测验由 20 个难易不同的题目组成，每题为 1 分，那么我们就不能说 5 分与 8 分之差同 15 分与 18 分之差是相等的。尽管从表面上看二者的差异量均为 3，但实际上这两个 3 分的含义却并不相同。因为从 5 分提高到 8 分与从 15 分提高到 18 分的难度是不一样的，后者比前者要难得多。既无绝对零点又无相等单位的心理量表所测得的结果，只能是从等级顺序上反映出个体间的差异。因此，心理测验所测得的结果只是一个相对的量数，是与被试所在团体某种行为常模相比较而言的。也就是说，在对人的行为进行比较时，没有绝对的标准，我们有的只是一个连续的行为序列。所谓心理测量就是看每个人处在这个序列的什么位置上。

此外，心理测量和其他所有的测量一样，均具有客观性的性质或特点。客观性是对一切测量的基本要求。因此，测量中所使用的测验也必须具备客观性。测验的客观性实际上就是测验的标准化问题。

第一，测验用的项目或作业、施测说明、施测者的言语、态度及施测时的物理环境等，均经过标准化，测验的刺激是客观的，特别是对测验项目的选择不是随意的，而是在预测基础上，通过实证分析确定的。

第二，评分计分的原则和手续经过了标准化，对反应的量化是客观的。评分方面的客观性随测验种类和项目类型而异，一般说来，投射测验的客观性差些，而选择题的客观性较好。

第三，分数的转换和解释经过了标准化，对结果的推论是客观的。

在心理测量中要控制的因素很多，因此，要保证客观性是很困难的，尽管心理测验的客观性有待于进一步提高，但它毕竟是测量人的心理特性的较为客观、较为科学的方法。目前，还没有更有效、更实用的方法能够取代它。

心理测量工作是在测验的基础上完成的，而测验是由有关领域的专家、学者经过长期的编制、试用、修订、完善而逐渐形成的标准化测量工具。在许多场合，心理测量和心理测验常被作为同义词来使用。的确，这两个概念的内涵在很大程度上是重叠的，但它们又存在明显的区别，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。相对而言，心理测量的意义范围更广泛一些。因此，简单地说，心理测量与心理测验的关系是：心理测验是心理测量的一种工具和手段。

心理测验在编制、施测、评分和解释方面依据一套系统的程序，按照严格的科学程序去编制和使用的测验称之为标准化测验。标准化测验有三点好处：第一，可以减少无关因素对测验目的的影响，使测量准确、客观。第二，有统一标准，便于对不同人的测验成绩进行比较和交流。第三，同一份测验可用于许多人，并可反复使用，较为经济。

## 第二节 心理测验的种类和作用

### 一、心理测验的种类

心理测验是多种多样的，可以按照不同的标准加以分类。

首先，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，前者即通常说的智力测验，后者多用于测量个人在音乐、美术、体育、机械等方面的特殊才能。成就测验主要用于测量个人（或团体）经过某种正式教育或训练之后对知识和技能掌握的程度。因为所测得的主要是学习成就，所以称为成就测验，最常见的是学校中的学科测验。人格测验主要用于测量性格、气质、态度等方面的个性特点。

其次，按测验对象分类，可以把心理测验划分为个别测验和团体测验。个别测验每次仅以一位被试为测验对象，通常是由一位主试与一位被试在面对面情形下进行。其优点在于主试对被试的行为反应有较多的观察与控制机会，主要缺点是不能在短时间内收集到大量的资料。团体测验是在同一时间内由一位主试（必要时可配几名助手）对大量的被试施测。此类测验的优点主要是可以在短时间内收集到大量资料，因此，在教育上被广泛采用。团体测验的缺点是被试的行为不易控制，容易产生测量误差。

第三，按测验方式分类，可以把心理测验划分为纸笔测验、操作测验、口头测验和电脑测验。纸笔测验所用的是文字或图形材料，实施方便，团体测验多采用此种方式编制。文字材料易受被试文化程度的影响，因而对不同教育背景下的人使用时，其有效性将降低，甚至无法使用。操作测验项目多属于对图片、实物、工具、模型的辨认和操作，无需使用文字作答，所以不受文化因素的限制。操作测验的缺点是要花费大量的时间，一般不宜团体实施。口头测验项目为言语材

料，主试口头提问，被试口头作答。电脑测验项目可为文字或图形，在电脑上显示，被试按键作答。

第四，按测验目的分类，可以把心理测验划分为描述性测验、诊断性测验和预示性测验。描述性测验的目的在于对个人或团体的能力、性格等进行描述。诊断性测验的目的在于对个人或团体的某种行为问题进行诊断。预示性测验的目的则在于通过测验分数预示一个人将来的表现和所能达到的水平。

第五，按测验难度分类，可以把心理测验划分为速度测验和难度测验。速度测验题目较为容易，一般不超出被试的能力水平，但数量较多，而且时限较短，几乎每个被试都不能做完所有题目。难度测验则包含各种不同难度的题目，由易到难排列，其中有一些极难的题目，几乎所有被试都解答不了，但作答时间较为充裕，因此测量的是解答难题的最高能力。

第六，按测验要求分类，可以把心理测验划分为最高作为测验和典型作为测验。最高作为测验要求被试尽可能作出最好的回答，主要与认知过程有关，有正确答案。能力测验、成就测验均属最高作为测验。典型作为测验要求被试按通常的习惯方式作出反应，没有正确答案。一般说来，人格测验测量的均属典型作为。

第七，按测验性质分类，可以把心理测验划分为构造性测验和投射性测验。在构造性测验中，所呈现的刺激和被试的任务是明确的。而投射性测验的刺激则没有明确意义，问题模糊，对被试的反应也无明确规定。

第八，按测验解释分类，可以把心理测验划分为常模参照测验和标准参照测验。常模参照测验是将一个人的分数与其他人比较，看其在某一团体中所处的位置。标准参照测验则是将被试的分数与某种标准进行比较来解释。

第九，按测验应用分类，可以把心理测验划分为教育测验、职业测验和临床测验。教育部门是测验应用最广的领域，许多能力和人格测验都可在学校中应用，但用得最多的是成就测验，平时说的教育测验，主要指后者。职业测验主要用于人员选拔和职业指导，可以是能

力和成就测验，也可以是人格测验。

临床测验主要用于医务部门。除感觉运动和神经心理测验外，许多能力和人格测验也可以用来检查智力障碍或精神疾病，为临床诊断和心理治疗工作服务。

## 二、心理测验的功能

### (一) 心理测验在实际工作中的应用

#### 1. 选拔人才

现代社会需要大量不同类型、不同层次的人才，心理测验的发展为大规模地选拔人才提供了可能。心理测量学家根据对各种工作的性质和特点的分析，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。

#### 2. 人员安置

随着社会化大生产的发展，人事分工越来越细，不同的工种需要不同的人来做，借助于心理测验可以做到人尽其才，从而提高劳动生产率。在教育领域，可以借助于心理测验的资料，作为按能力和成绩分班的依据，从而为分类教育、因材施教提供条件。

#### 3. 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应用心理测验来诊断仍然是一种重要的途径。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。

#### 4. 描述评价

应用心理测验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作出描述和评价，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。

#### 5. 心理咨询

心理测验获得的资料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，以帮助来访者作出正确的职业选择。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。

## （二）心理测验在理论研究中的应用

### 1. 收集研究资料

在心理学和教育学的许多研究工作中，都需要通过测验来获得第一手资料。例如，为了查明影响学生学业成绩的心理因素，我们需要运用智力测验、学习能力倾向测验、成就动机测验、学习兴趣测验、人格测验和学业成就测验，对各种心理因素的测量分数与学业成就测验的分数之间的相关系数进行回归分析，然后根据测验所获得的实证资料做出科学结论。

### 2. 建立和检验理论假设

在心理学的研究中，通常需要根据已有的测验研究成果提出理论假设，然后通过测验进一步检验这个假设。斯皮尔曼的智力二因素理论，瑟斯顿的智力群因素理论，以及吉尔福特的智力三维结构理论都是建立在对智力测验结果的因素分析基础上的。这些智力理论来源于智力测验，反过来又为进一步编制智力测验提供理论基础。

### 3. 实验分组

心理测验还可以和实验方法结合起来运用于研究工作中。在一些实验心理学的研究课题中，为了考察不同自变量对被试因变量的不同影响，通常选择两组被试进行比较研究，这时需要控制与实验变量无关的被试的其他心理变量，使两组被试实现等组化。心理测验（如智力测验）可以满足实验设计中的上述要求。有时，我们需要研究具有不同心理特征的被试在完成心理实验任务过程中的差异。在这种情况下，我们首先需要通过心理测验识别不同心理特征的被试，然后分成两个极端组进行比较实验。

## 万舰梦者

### 阅读 I

对于许多工作，测验是用于选拔过程的一部分。显然，不只是测验的使用，连同整个选拔过程及其所有的组成部分，都要关心“机会均等”。从机会均等观点评估选拔过程的要点列在下面。它们适用于选拔过程的所有方面，而不仅仅是使用测验这一方面。

**监测不同团体的成绩** 除非收集的数据是用来记录申请者的主要事实，否则进一步分析是不可能的。希望把自己称为机会均等的雇主的人需要常规地收集数据以监测选拔过程的方方面面，包括内在的及外在的方面。这些数据可用于查看选拔过程的不同阶段、不同团体的成功率。

**设定决定分数的策略** 在英国，不准许优惠某些团体（如妇女或少数民族）的限制或正性歧视。选拔过程的目的是对所有团体一视同仁。这意味着即使会产生副作用，时所有团体也要使用同一决断分数。

**副作用** 如果某些团体比另一团体更少成功，那么选拔过程可被视为有副作用而应该进行反省，选拔规则对所有团体都有效吗？要收集证据证明选拔过程是与工作有关的，是有效的。在英国，选拔过程没有法律约束，而无法表明选拔过程是与工作有关的。如果某一测验受工作中不需要的技能影响，那么使用这个测验是不恰当的。

如果发现选拔过程有副作用，这可能反映真正的团体差异，但也应该探究其他的解释。团体是否在其他方面有差异，例如受教育的资历？当前的选拔过程是否对某些申请者有不利方面？这些能否改变而不减弱选拔的效度？另一不同的选拔过程会更公平而同样有效吗？（摘自《了解心理测验过程》〔英〕查尔斯·杰克逊著，妇七萍译，万传文审校，北京大学出版社，P103-104）

### 第三节 心理测量的历史

地球上自从有了人类，就产生了研究人的心理这一课题。人们想加强相互了解，从而发生了测度心理的兴趣。有史记载的探心术，最早也许可以追溯到伏羲和女娲“结绳记事”的时代。我国后人不断完备的“八卦”就是从那时开始的。而至今久盛不衰的“星相学”，一般被认为是公元前 3000 年在美索不达米亚确立的。

在我国几千年文明的发展过程中，存在着丰富的心理测量的思想和实践。如早在 2500 多年前，孔子就曾根据自己的观察来评定学生的个体差异，把人分为中人、中人以上和中人以下三类。孟子也说：“权然后知轻重，度然后知长短。物皆然，心为甚。”这反映了人们在很久以前，就认识到了心理测量的重要性。在我国实行了 1300 多年的科举制度，则是心理测量在人才选拔运用上的一种雏形。清朝后期出现的七巧板可以看做最早的创造力测验之一。尽管我国有着丰富的心理测量的宝藏，但遗憾的是，科学的心理测量的诞生，则要追溯到 19 世纪的西方。

#### 一、国外心理测量发展的历史

##### （一）心理测量的产生是社会的需要

承认人的个别差异及其对个人行为的重要影响是进行心理测量工作的基本前提。在西方，科学家最初发现人的心理的个别差异的重要性，是起因于 18 世纪天文学上的一个偶然事件。1796 年，英国格林尼治天文台的皇家天文学家 N. 马斯林基因为其助手金内布鲁克观察星体通过的时间比自己晚了 0.8 秒钟而将他辞退。20 年后，另一天文学家贝塞尔对这一事件作了研究，认为这不是金内布鲁克的过错，而是一种不可避免的个人观察的误差。贝塞尔的这一发现引起了学者们对个别差异的重视和研究。但在当时并未引起心理学家的注意。1879 年，德国心理学家冯特在莱比锡大学建立了世界上第一个心理实验室。在研究中他发现，不同被试对同一刺激的反应常常不同。研

究者最初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。当时的实验心理学所研究的内容主要集中在感知觉等低级心理现象上面，而对诸如能力、人格等高级心理特征还无能为力，这就为日后开展对个别差异的测量学研究提出了课题。

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。19世纪，在欧洲和美洲开设了一些护理精神病人的特别医院，因而急需确定收护标准和客观化的分类方法，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，工厂大量雇佣童工。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订约，每雇佣20名童工，必须同时带雇1名低能者。为了设法使低能者尽可能适应工厂技术的要求，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，介绍了在感觉和肌肉运动方面训练智力落后儿童的方法。他的著作中的一些内容现已转化为能力操作测验的组成部分。

19世纪，由于科学的发展和欧洲人道主义思想的广泛传播，人们对智力缺陷者和精神病人的态度发生了重大变化，开设了一些医院专门护理和医治精神病人。这就在客观上要求确定鉴别各种心理疾病的统一标准。法国医生艾斯科罗尔首次对智力落后与精神病作了区分。他还认为，智力落后从接近正常到最严重的白痴之间有一系列等级，而诊断智力落后程度的最可靠的方法是观察儿童运用语言的能力。他的这一思想至今还体现在智力测验之中。

科学的心理测量的产生，既有着深刻的社会时代背景，又与科学技术的发展水平相适应，同时还与科学家个人的学术贡献密不可分。

## （二）心理测量的先驱

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”。

### 1. 高尔顿的 10 年

第一个直接推动心理测量运动的学者是优生学创始人，英国生物学家、心理学家弗兰西斯·高尔顿。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。为了测量个体间的差异，他于 1884 年在国际博览会上设立了“人类测量实验室”，参观者只要付三便士就可以测量到自己的某些身体素质状况，如视觉或听觉的敏锐性程度、肌肉的力量、反应速度等。博览会闭幕后，这个实验室迁移到伦敦的南圣顿博物院，在这里继续开办了六年。此实验室共积累了 9337 人的材料，是第一个大量系统地测量个别差异的尝试。高尔顿还出版了《人类才能及其发展的研究》一书，在这本著作中，他首先提出了“测验”、“心理测量”这两个术语。

高尔顿的另一个重要的贡献是把数学中的统计方法加以改造，变成比较简单的形式，使那些没有经过训练的调查者也可以使用。这对于推动心理测量的科学化是一个很了不起的贡献。因为有了这种数学统计方法，人们对心理现象的研究就可以从定性的分析上升到定量的分析了。

### 2. 卡特尔的 10 年

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。他将新兴的实验心理学与测验运动结合起来，完成了题为“反应时的个别差异”的博士论文。从 19 世纪 90 年代起，卡特尔的研究重心开始从反应时转向心理测验，以各种心理测验来研究个别差异。他在自己的实验室（哥伦比亚大学实验室）里编制了 50 个测验，分别用来测量肌肉力量、运动速度、反应时、痛感受性、视听敏度、记忆力等，这些研究为美国心理测量学的发展开了先河。

卡特尔于 1890 年在《心理》杂志上发表了《心理测量与测量》

一文，这是心理测量第一次出现于心理学文摘中。文中主张心理学应建立于实验和测量上，心理测量应有一个普遍标准。

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。有人说，卡特尔在为美国心理学界培养后继人才方面所作的贡献，并不逊色于他对美国心理测量学发展所作的贡献。

### 3. 比纳的 10 年

在心理测量领域，高尔顿、卡特尔、克雷正林、艾宾浩斯、皮尔逊、斯皮尔曼等都是早期心理测量先驱者的代表人物，而比纳则被认为是心理测量的鼻祖，他是发明智力测验常模量表的第一人。

比纳在 1886 年发表了第一部著作——《推理心理学》。他认为当时的心理测验偏重于简单的感觉测量，而在他看来，智力是一个复杂的概念，决不仅仅限于感觉的能力。1904 年，法国教育部组织一个委员会，专门研究公立学校中低能班的管理办法，比纳即为该委员会的委员之一。比纳主张用测验法去辨别有心理缺陷的儿童，比纳根据自己对智力的理解，与其助手合作，在 1905 年编制了世界上第一个智力测验量表——比纳—西蒙量表。同年，他在《心理学报》上发表《诊断异常儿童的新方法》一文介绍此量表，因而历史上称其为“1905 年量表”。这个量表包含 30 个由易到难排列的项目，可用来测量各种能力，特别侧重于判断、推理、理解能力，这些也是比纳认为的智力的基本组成部分。

1908 年，比纳对量表作了修订，采用智力年龄的方法计算成绩，并建立了常模，这是心理测量史上的一个创新。1911 年，比纳对量表进行了第二次修订。不幸的是，就在这一年，比纳英年早逝，终年只有 54 岁。比纳在心理测量这个领域获得成功并不是偶然的，他曾经测量过人的头盖骨，研究过面相、手相和字相，在走了许多弯路后才终于取得了成功。目前世界上的智力测验为数众多，其基本原理和主要方法都是由比纳奠定的。在心理测量的发展史上，比纳的贡献不可磨灭。

比纳—西蒙量表问世之后，立刻传到世界各地，美国斯坦福大学的推孟教授于 1916 年对比纳—西蒙量表进行修订（修订后的量表简称为斯比量表），其中最大的改变是采用了智商的概念。

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

### （三）心理测量的蓬勃发展与应用

心理测量运动自 20 世纪初兴起后，20 年代进入狂热，40 年代达到顶峰，50 年代转为稳步发展。在此期间，心理测量主要有以下几方面的发展：

1. 编制出了一批操作测验，既弥补了语言文字量表在理论上的缺陷，也适用于文盲和有言语障碍的人，以宾特纳所编的非文字量表为代表。

2. 编制出团体智力测验，扩大了测验的应用范围。

3. 多重能力倾向测验逐渐受到重视。随着因素分析理论的发展，多重能力倾向测验在第二次世界大战后编制出来，它适合进行个体内部心理结构的分析和个别差异的诊断，也包括单个能力倾向测验所包含的信息。

4. 传统的学校考试方法也取得了技术上的突破。20 世纪初，桑代克编制了第一个标准化的教育成就测验，该测验第一次利用了心理测量的原则，编制出评定学生书写、作文、拼读、算术、计算和推理的量表。桑代克由此被公认为教育测验的鼻祖。此后，第一个成套测验——斯坦福成就测验于 1923 年发表。

5. 心理测量的另一领域是对情感适应、人际关系、动机、兴趣和态度等人格特点的测量。人格测验的先驱是克雷正林，他最早用自由联想测验来诊断精神病人。最早的人格问卷则是武得沃斯编制的“个性资料调查表”。而 1921 年问世的罗夏墨迹测验则是投射测验的开端。人格测验后来也发展到要求数量化，如明尼苏达多相个性调查