

第五章 常用的心理测验



一、能力测验

(一) 智力测验

1. 个体智力测验

(1) 比内智力量表

比内智力量表源于斯皮尔曼的智力二因素论。

① 比内-西蒙量表

这是比内和西蒙出于诊断异常儿童的需要而编制并不断修订的智力量表，包括：

A. 1905年版

a. 包括 30 道题，项目难度逐渐上升，以通过项目多少作为区分智力的标准； b. 适用于智障儿童； c. 是世界上第一个智力测验。

B. 1908年版

a.包括 59 道题，项目以年龄分组；b.适用于 3~13 岁儿童；c.提出第一个年龄量表，即智龄（智力的绝对水平）。

C.1911 年版

增设成人题目，扩大适用范围。

②斯坦福——比内量表

戈达德第一次将比内——西蒙量表介绍到美国，以斯坦福大学推孟教授的修订最负盛名，包括以下版本：

A.1916 年版

a.在原来基础上增加 39 道题；b.适用于 3~13 岁儿童；c.首次引入比率智商，有详细的指导语和计分标准以保证标准化。

B.1937 年版

a.由两个复本（L 型和 M 型量表）构成；b.适用于 1.5~18 岁；c.拥有代表性样本，但仅限于白人被试。

C.1960 年版

a.合并两个复本形成 LM 型单一量表；b.适用于 2 岁~成人；c.舍弃比率智商，引入离差智商（100，16）。

（2）韦氏智力量表

韦氏智力量表基于能力的结构理论，采用因素分析法得来。

①韦氏成人智力量表（WAIS）

A.韦克斯勒——贝尔韦量表（W-B）

a. 题目难度逐渐增加，分为 I 型和 II 型，II 型信度和代表性较高；b.适合 10~60 岁被试；c.第一个成人智力量表。

B.修订版（1955）

a.共有 11 个分测验，包括 6 个言语量表和 5 个操作量表。分测验内题目难度逐渐增加，并且言语测验和操作测验交替施测；b.适合 16~74 岁的成人；c.先将原始分转为均值为 10、标准差为 3 的标准分数（标准 20 分），再将言语分测验或操作分测验的标准 20 分相加得到言语量表分或操作量表分，最后将量表分转换为均值为 100、标准差为 15 的智商分数（离差智商），可得到言语智商（VIQ）、操作智商（PIQ）和总智商（FIQ）；d.是标准化程度很高的测验，信效度很高，以斯坦福——比内量表作为校标，其相关达到 0.8 以上。

②韦氏儿童智力量表（WISC）英文版

a.12 个分测验，5 个言语测验，5 个操作测验，2 个备用测验（背数和迷津），每 4 个月为一个年龄组；b.适合 6~16 岁儿童，可用于鉴别智力发展水平，鉴别发展性障碍、学习障碍等；C.同 WAIS，先将原始分数转换为量表分（标准 20 分），再将量表分转换为智商分数（离差智商），d.首次提出离差智商，是当今世界上应用最为广泛的儿童智力量表。

③韦氏幼儿智力量表（WPPSI）

A.共 11 个分测验，其中，3 个分测验是为了适应幼儿特点而新编的，另外 8 个则与 WISC 相同；B.适用于 4~6.5 岁的儿童；C.多种证据表明 WPPSI 的信效度都较好。

（3）戴斯认知测验（CAS）

戴斯、纳格里尔于 1977 年基于智力的 PASS 模型编制，又称认知评估系统。

A.12 种任务，4 个分测验；B.适用于 5~17 岁 11 个月的人；C.分辨儿童在特定认知成分上存在的缺陷及个体差异，是可以动态评估的智力测验，有理论和实证基础；D.该测验以认知心理学的相关研究成果为基础。

(4) 考夫曼儿童成套评价测验 (K-ABC)

美国心理学家考夫曼夫妇于 1983 年基于卡特尔的流体智力和晶体智力理论编制。

A.共有 16 个分测验，按年龄选用 7~13 个分测验进行实施；

B.适用于 2.5~12 岁半的正常儿童及特殊儿童；

C.用于评价智力和成就水平。其中，智力评估包括评估受测者处理那些既要同时性加工也要继时性加工的能力；成就评估包括评估受测者在阅读和算术上习得的技能；

D.该测验以认知心理学和神经心理学的相关研究为成果基础，特别为学前、少数民族和特殊儿童编制。

2. 团体智力测验

(1) 陆军测验

测验由 Otis，推孟的学生编制，用来招募和选拔士兵。

① 陆军甲种

A.由 8 个分测验组成，以英文呈现；

B.第一个团体智力测验，由斯——比量表改为纸笔测验。

② 陆军乙种

A.由 7 个分测验组成，属于非文字测验；

B.适用于母语非英语和有英文阅读障碍的人群；

C.与甲种测验的相关达到 0.8。

(2) 瑞文推理测验

又称瑞文渐进图阵。它是基于智力二因素中 G 因素理论的非文字型的图形测验，可个别施测，结果先转化为百分等级，再转化为 IQ 分数。测量一般智力，如推理能力。

瑞文将智力的 G 因素分为两种独立的能力：再生性能力（经过教育后达到的水平，适合词汇测验）；推断性能力（不受教育影响的理性判断能力，适合非语言的图形推理测验）。

瑞文测验分为三个水平：

① 瑞文标准推理测验 (SPM)

A.1938 年编制，5 个黑白系列，60 个项目；B.适用于 5.5 岁以上智力正常者；C.属于中等水平的瑞文测验。

② 瑞文彩图推理测验 (CPM)

A.1947 年编制，3 个系列，36 个项目；B.适用于幼儿和智力低于平均水平的人；C.属于最低水平的瑞文测验。

③ 瑞文高级推理测验 (APM)

A.1941 年编制；B.适用于高智力成人；C.是最高水平的瑞文推理测验。

(3) 认知能力测验 (CAT)

由桑代克等美国心理学家于 1968—1972 年编制成功，由 4 个部分组成，分别是：①初

级测验；②文字测验；③数量测验；④非文字测验。

(4) 联合瑞文测验 (CRT)

①72幅图、72个题目，二级评分，答对1分、答错0分；②适用于5~75岁；③原始分转换为百分等级，再转化为IQ分数；④可用于有言语障碍的智力测验，也可作为不同民族、不同语种间的跨文化研究工具，每次施测团体不应超过50人，属于文化公平测验。

(5) 其他团体测验

①文化公平测验 (CFT)

R.B.卡特尔与A.K.卡特尔基于流体智力和晶体智力的理论编制，属于非文字测验；可个体施测也可团体施测，测量受文化影响较小的智力的一般因素，如发现联系和规律的能力；结果以百分等级表示。

②画人测验：

由Goodenough编制，适用于4~13岁被试，评估非言语智力，属于文化公平测验。

③希——内学习能力测验

由M.Hiskey编制，基于智力的二因素理论，适用于3~16岁聋哑或听力正常儿童的智力评估。在分数解释上，主要采用学习年龄评价聋哑儿童的能力，采用智力年龄评价正常儿童的能力。

总结：

以智力二因素理论为基础的智力理论：比内智力量表、瑞文推理测验（基于G因素）、希——内学习能力测验、多水平团体智力测验；

以卡特尔流体智力与晶体智力理论为基础的智力量表：考夫曼儿童成套评价测验（K-ABC）、文化公平测验（CFIT）；

以智力PASS模型为基础的智力量表：认知能力测验（CAT）；

适用于文化公平的智力量表：陆军乙种测验、联合瑞文测验（CRT）、文化公平测验（CFTT）、画人测验。

3. 智商的发展

智商概念的发展经过智龄、比率智商、离差智商三个阶段。

(1) 智龄 (MA)

智力的绝对水平，说明智力实际达到了什么程度。但是智龄相同的儿童若实际年龄不同，他们的智力仍不同，即不能够比较实际年龄不同的儿童智力的高低，所以要考虑智龄与实际年龄的关系。

(2) 比率智商 (IQ)

施特恩 (Stern) 首次提出比率智商 ($IQ = \text{智力年龄} / \text{实际年龄} \times 100$)。比率智商可以比较实际年龄不同的儿童智力的高低，但是相同智商分数在不同年龄具有不同意义。因为智力的发展速率先快后慢，与年龄增长不是直线关系，因而随着年龄的增长，比率智商将逐渐下降，所以比率智商的概念对成人不适用。

(3) 离差智商 (IQ)

1949年，韦克斯勒首创离差智商 ($IQ = 15Z + 100$)。离差智商表示一个人的智力偏离本年龄组平均水平的方向和程度。其优点在于同样的离差智商在任何年龄水平上都代表同样的

相对位置。

4.智力测验的局限性

- (1) 主要是对智力进行了量的分析，质的分析过少；
- (2) 只测查了某个时段的智力，没有考虑智力发展的速度及趋势；
- (3) 只是部分反映了一个人的智力水平，不同智力量表都只是从某个侧面反映了智力；
- (4) 对于创造性思维考查得较少。

(二) 其他能力测验

1.能力倾向测验

(1) 概述

①产生原因

- A.智力测验的局限。
- B.个体能力倾向的差异。
- C.来自实际的需要。

②特点

- A.在测验目的上，能力倾向测验的预测目的更强。
- B.在测验编制上，能力倾向测验一般同时测量几种能力因素，以分测验形式组成，每个分测验针对一种能力，每个测验应该是独立的，各分测验间的相关要尽可能低。测验内容广泛，且较少涉及与学校习得知识有关的内容。
- C.在测验结果解释上，被试在能力倾向测验上可以得到若干测验分数，既可表明不同被试在每一分测验所测能力上的相对位置水平，也可以表明同一被试在所测各能力上的相对优劣状况。针对不同工作，不同能力因素的权重应有变化。一般采用多重回归模式解决这一问题。

(2) 学术能力倾向测验 (SAT)

美国教育测验服务中心编制，预测学生是否具备大学学习和研究的能力以及倾向于在哪些专业领域更具优势；分为语文、数学两部分，均采用多选题方式，相当于中国的高考。

(3) 分辨能力倾向测验 (DAT)

美国心理公司编制，8个分测验，7个能力测验和1个速度测验（文书）；适用于美国8~12年级学生的职业咨询和教育咨询；是运用最广泛的成套能力倾向测验之；结果用百分常模表示，转换后画剖面图。

(4) 一般能力倾向成套测验 (GATB)

美国联邦劳工部用因素分析法编制，12个分测验，8个纸笔测验，4个仪器测验，可确定9种能力倾向，其中的装配测验属于速度测验，有36个职业群常模；用于职业咨询，即可团体施测又可个体施测；结果采用百分等级和标准分数（100，20），用剖面图表示；与DAT相比更注重实际操作。

2.特殊能力测验

针对一种特殊能力所包含的各方面因素进行测量，测验性质介于成就测验和能力倾向测验之间。其内容与相应的专业或职业训练的重点是一致的，而测量的目的既想了解个体在此专业领域的既有水平，又想预测个体今后在此专业领域成功的可能性。常见测验包括：

(1) 音乐能力测验：西肖尔音乐才能测验（第一个标准化的音乐能力测验）、戈登的音乐能力测验图 MAP、Drake 音乐能力倾向测验等。

(2) 美术能力测验：梅尔美术判断力测验（能力与美术成就）、格雷福斯图案判断测验（美学知觉和判断）和霍恩美术能力问卷等（操作型测验）。

(3) 机械能力测验：空间关系测验（机械拼合、空间关系、书面形式拼板）和 Bennett 机械理解能力测验等。

(4) 文书测验：一般文书测验（GCT）、明尼苏达文书测验（MCT），属于速度测验。

3. 创造力测验

(1) 吉尔福特发散思维测验（南加利福尼亚大学测验）

① 吉尔福特基于其三维智力理论编制，14 个分测验，针对其中 11 种能力因素进行测量；② 一般适用于初中水平以上的人；③ 考查思维的流畅性、变通性和独特性三方面；④ 分半信度在 0.6~0.9 之间，结果以百分位数和标准分数表示。

(2) 托伦斯创造性思维测验

① 是托伦斯在吉尔福特的智力理论及其发散思维测验基础上编制而成的，由两套测验组成，每套有两个复本；② 适合幼儿园到成人被试；③ 考察思维的流畅性、变通性、独特性和精确性四个方面。

(3) 芝加哥大学创造力测验

① 由盖泽尔斯和杰克逊编制；② 适用于青年人；③ 考察思维的流畅性、变通性和独特性。

二、人格测验

（一）概述

1. 定义

人格测验就是通过一定的方法，对在人的行为中起稳定调节作用的心理特质和行为倾向进行定量分析，以便进一步预测个人未来的行为。

2. 历史

(1) 高尔顿

首先提倡用科学方法测量人格；通过记录心率和脉搏的变化来测量情绪；通过观察社会情境中人的活动来评估人的性情、脾气等特征；他还初步尝试编制了一个人格的评定量表，这一切标志着科学人格测量的开始。

(2) 克雷培林

人格测验的先驱。1892 年将联想测验用于临床诊断，其基本做法是要求受测者在听到或看到刺激词后说出他最先想到的词汇，然后，通过分析反应词的内容来判断受测者的人格特征。这种方法后来被广泛地应用于人格测验项目的编制。

(3) 武德沃斯

1919 年，美国武德沃斯发表了第一个自陈人格量表一个人资料调查表，开创人格问卷测量之先河。

(4) 罗夏克

1920 年，罗夏克墨迹测验问世，投射测验由此诞生。

3. 存在的问题

存在的主要问题是反应偏差。反应偏差有两类，一类是反应定势，一类是反应形态（文称反应风格）。最常见的反应定势就是社会赞许倾向。反应形态主要表现为默认倾向（只做肯定回答或只做否定回答）和由于错误记忆和模糊记忆产生的潜在“合理化”加工。

除此之外，过度谨慎、装好或装坏、求“快”或求“精准”、喜好正面评述、喜好特殊位置、喜好较长选项、猜测、回避、极端化等一些反应定势也在很大程度上影响测验结果。这就使得人格测量比能力测量更不稳定，加上人格特征的定义和题目设计困难，其信度常低于能力或成就测验。

（二）自陈量表

自陈人格测量就是根据要测量的人格特质，编制许多有关的问题，要求受测者根据自己的实际情况逐一回答这些问题，然后根据受测者的答案，去衡量受测者在这种人格特质上表现的程度。为完成自陈人格测量而编制的测量工具叫自陈测验或自陈量表。

1. 自陈量表的特点

- （1）自陈量表的题量比较大，多数用于测量人格的若干特质；
- （2）自陈量表通常采用纸笔测验的形式，可以团体施测；
- （3）自陈量表的项目形式一般采用是非式或选择式，它的计分规则比较客观，施测手续比较简便，测量分数容易解释，应用广泛；
- （4）信度指标通常采用重测信度和内部一致性信度，信度系数不低于 0.6；效度常采用构想效度，较少使用效标效度。

2. 自陈量表的编制方法

（1）逻辑分析法

①基本程序：首先由专家依据某种人格理论，确定要测量的特质，用逻辑分析的方法编写和选择一些能测出这些特质的题目，最后组卷编排成问卷。

②举例：爱德华个人偏好量表（EPPS）、詹金斯活动调查表、显性焦虑量表等。

（2）因素分析法

①基本程序：首先对标准化大样本施测大量题目，然后通过被试在各题上的得分进行因素分析或其他相关分析，把相关题目构成一个因素并命名，便可以得到若干个同质量表来测量对应于这若干个因素的若干个人格特征。

②举例：16PF、EPQ 等。

（3）经验法

①基本程序：先分组，即选取具有某一特征的效标组和对照组，然后，把一系列的测试题给各组施测，选出能把两组分开的题目构成测验。

②举例：MMPI 等。

（4）综合法

①基本程序：将上述三种编制方法综合起来使用。首先，采用逻辑分析法经推理获得一大批题目，同时用经验法确定效标组特征也获得一大批题目；然后采用因素分析法编出若干同质量表；最后，将同质量表中没有效标效度的题目删掉。同时，将表面效度太高的题目删掉，最好保留效标效度高，但表面效度不高的题目。

②举例：CPAI（中国人个性测量表）、CPI（加州心理调查表）。

3.题目编制方法

包括：(1) 是否式；(2) 二择一式；(3) 是否折中式；(4) 文字等级式；(5) 数字等级式。

(三) 常用的自陈量表

1.明尼苏达人格调查表 (MMPI)

(1) 哈撒韦, 麦金利根据经验法编制, 可个体施测 (卡片、问卷), 也可团体施测 (问卷);

(2) 共 566 题, 16 个重复项目, 用于检验受测者反应的一致性, 共 13 个分量表 (10 个临床量表和 3 个效度量表), 其中效度量表包括说谎量表 (L)、校正量表 (K)、诈病量表 (F), 也可加上疑问量表 (Q), 通常在临床诊断中, 只使用前 399 个项目;

(3) 适合 16 岁以上具有初中文化的人;

(4) 原始分数应转换为 T 分数, 对测量结果解释的最好方法是分析 MMPI 剖析图。临床量表需加 K 分数校正;

(5) MMPI 可以全面测量人格的各个方面, 可以作为编制新量表的参考, 信度和效度比一般的人格测验高, 但题目过多且测试时间过长会影响被试的情绪。

2.卡特尔 16 种人格因素量表 (16PF)

(1) 美国伊利诺伊州立大学教授卡特尔根据根源特质采用因素分析法编制, 适用于团体施测;

(2) 有 5 个版本: A、B 版 (等值, 各 187 题, 16 种因素), C、D 版 (缩减版, 各 106 题), E 版 (适用于文化程度低的人, 128 题), 每个题目有 a、b、c 三个选项。受测者根据自己的情况选一个合适的选项, 计分方式是 (0, 1, 2) 或 (2, 1, 0), 只有聪慧性因素采用的是 (0, 1) 计分;

(3) 适用于 16 岁以上受测者;

(4) 原始分数须转换为标准 10 分 (5.5, 1.5), 另外, 需要根据剖面图的描述解释人格特征, 根据公式进行综合描述, 预测某些情境中的行为特征;

(5) 各因素间相互独立 (相关小), 重测信度高, 表面效度不高;

(6) 卡特尔在 16 种人格因素基础上进行了二阶因素分析, 得到四个二阶公共因素, 分别为: 适应—焦虑、内向—外向、感情用事—安详机警、怯懦—果断。

3.艾森克人格问卷 (EPQ)

(1) 艾森克夫妇根据艾森克人格三维度理论采用因素分析法编制, 适用于团体施测;

(2) EPQ 由四个分量表组成 (P、E、N 和 L), 用于测量受测者在精神质 (P)、外倾性 (E) 和神经质 (N) 三个人格维度上的特征。L 是说谎量表, 用于识别受测者回答问题时的诚实程度。EPQ 分为儿童版和成人版两种。儿童问卷共有 97 个项目, 成人问卷共有 101 个项目, 均采用是否 (0, 1) 计分, 注意反向计分的情况。

(3) 儿童版适合 7~15 岁的受测者, 成人版适合 16 岁以上受测者。

(4) 原始分数须转换为 T 分数。

(5) 可用来分析受测者的个性特点或气质类型。

4.爱德华个人偏好量表 (EPPS)

(1) 爱德华以默瑞的 15 种人类需要理论为基础采用逻辑分析法编制。

(2) 量表由 15 种需要量表和 1 个稳定性量表所组成，有 225 对由陈述句组成的题目，15 个题目重复两次。每个题目包括两个自我描述性的陈述。采用强迫选择法控制社会赞许，一定程度上也提高了测验的信度和效度。

(3) 适用于大学生和成年人。

(4) 用于测查 15 种不同心理需要的反应倾向的相对强度并画出人格剖面图；参照框架是个体而不是常模团体，分数具有自比性质。

5.其他

(1) 加州心理问卷 (CPI)

①高夫 1948 年以 MMPI 为基础采用综合法编制，可个体施测也可团体施测；②共 480 题 (178 题源于 MMPI；另外的题目反映正常青年和成人的人格)，18 个分量表，包括 3 个效度量表；③适用于 13 岁以上的正常被试；④测验结果用 T 分数表示；⑤用于正常个人人格特征的测查，可预测人们在特定情境下的行为，是目前美国使用最广泛的测查正常人格特点的量表之一。

(2) 詹金斯活动性调查表 (JAS)

①詹金斯 1964 年依据弗里德曼和罗斯曼的 AB 人格类型采用因素分析和叛变函数编制。②共 52 个题目，4 个量表，无时限。其中，A 型量表作为从整体评价个体的 A 型行为程度的指标，另外三个分量表分别从 S 因素、J 因素、H 因素上对 A 型行为的三种特征因素进行描述，S 因素是速度和性急因素，J 因素是对工作献身的因素，H 因素是刻苦和竞争的因素。③用于评价 A 型人格特征。

(3) 梅耶尔斯-布雷格斯类型指标 (MBTI)

①梅耶尔斯和布雷格斯依据荣格性格理论编制；②4 个分量表 (内倾-外倾，感觉-直觉，思维-情感，判断-知觉)；③用于衡量和描述人们在获取信息、做出决策、对待生活方面的心理活动规律和性格类型。

(4) NEO-PI 五因素调查表

①科斯塔、麦克雷依据大五人格因素模型采用因素分析法编制；②共 300 个题目，采用 5 点量表，有自我报告形式和观察报告形式的两个版本；③适用于成年人；④观察者报告形式目前被广泛接受。

(四) 投射测验

1. 投射测验的定义

投射测验是向被试提供一些未经组织的刺激情境，让其在不受限制的情境下自由表现出反应，然后分析反应的结果，便可推断人格结构。

2. 投射测验的假设

(1) 人们对于外界刺激的解释性反应都是有其原因且可以预测的。

(2) 反应者过去形成的人格特征、当时的心理状态以及对未来的期望等心理因素也会渗透在他对刺激的反应过程及其结果之中。

(3) 人格结构的大部分处于潜意识中，当被试面对一种不明确的刺激情境时，就可以使隐藏在潜意识中的欲望、需求、动机等“泄露”出来，即把一个反映他的人格特点的结构

加到刺激上，通过分析，就有可能获得对受测者自身人格特征的认识。

3.投射测验的特点

(1) 测验材料没有明确的结构和确切的意义，为受测者提供了针对测验材料进行广阔自由联想的机会和空间。

(2) 受测者对测验材料的反应不受限制。

(3) 测验目的具有隐蔽性，避免受测者的伪装和防卫，使结果更能反映真实的人格特征。

(4) 测验结果的解释重在对受测者的人格特征获得整体了解。

(5) 不受语言文字的限制。

(6) 计分困难，难以对测验结果进行定量分析。

4.投射测验的分类

林德赛 (G. Lindzey) 根据受测者的反应方式，将投射测验分为：

(1) 联想测验：让受测者说出某种刺激所引起的联想内容。如罗夏克墨迹测验 (RIT) 和荣格的词语联想测验。

(2) 构造测验：让受测者根据所看到的刺激编出一套含有过去、现在、将来等发展过程的故事。如主题统觉测验 (TAT)、图片故事测验等。

(3) 完成测验：据供一些不完整的句子、故事等材料，让受测者自由补充，使之完整。如句子完成测验。

(4) 选排测验：让受测者在一些故事中进行选择，成按照偏好进行各种排列。如汤姆金斯-福恩图片排列测验、内田测验等。

(5) 表露测验：

让受测者利用某种媒介 (绘画、游戏、心理剧等) 自由表露心理状态。如画人测验等。

5.常用的投射测验

(1) 罗夏克墨迹测验 (RIT)

测验材料采用 10 张左右对称但形状模糊的墨迹图，通过指导让受测者自由联想并表达对图案的想法。

①由瑞士精神病学家罗夏克 1921 年采用联想法创制的一种投射测验，适用于个别施测；

②包括 10 张墨迹图 (5 张黑白图、3 张彩色图、2 张“黑加红”图)，包括自由反映阶段 (不交谈)、提问阶段 (有针对地提问题)、类比阶段 (进一步商讨)、极限测验阶段 (确定受测者是否看到了具体事物) 四个阶段；

③适用于成人，3 或 5 年级以上受测者；

④一般从反应的部位，反应的决定因素 (形状、运动、浓淡、色彩)，反应的内容，反应的独创性这四个方面进行计分和解释；可预测人们在特定情境下的行为；

⑤用于精神病人人格分析或分析精神障碍对知觉的影响，是当今世界使用范围最广的投射测验，但是其评分很困难，极为费时费力，只有经验丰富的人才能掌握这种方法。

(2) 主题统觉测验 (TAT)

通过呈现图片让受测者自由联想编造故事，即向受测者呈现一系列意义相对模糊的图卡，并鼓励他按照图卡不假思索地编述故事。

①莫瑞和摩尔根在 1935 年根据“需要——压力”理论采用构造法编制；

②包括 30 张黑白卡，1 张空白卡，分 4 组（成年男性、成年女性、儿童男性、儿童女性），各组均为 19 张黑白卡加 1 张空卡；

③评分分为两部分：需要变量和情绪变量上的分数以及每一种压力变量上的分数。解释分数时，假设主人公的归因代表受测者的人格倾向性，并且受测者统觉的环境压力代表着过去、现在和将来；

④TAT 假设对模糊情境的解释倾向与自己的经历和愿望一致，讲述故事时利用了自己的经历，并在故事中表达了自己的情感与需要。

（3）句子完成测验

句子完成测验比罗夏墨迹测验和 TAT 实施起来更为简便，当受测者完成个句子时，他的潜意识就会投射出来，可以判断其人格特征。测验用的句子都是第一人称，可以团体施测也可以个体施测，主要包括两种形式：

①限制选择式：在一个未完成的句子后面列有数个短句，要求受测者从其中选择一个最合适的短句来完成句子。

②自由完成式：要求受测者将未完成的句子补充成完整的句子，而对受测者不加以任何限制。主要有以下两种：

A. 未完成语句测验：罗特编制，包括 40 道题，采用 7 级评分。

B. 萨式语句完成测验：萨克斯编制，包括 60 道题，分为 4 类，每一类中都有 4 道题目与日常生活密切联系。

（4）绘画测验

①画人测验

先画一个人，再画一个与先前性别不同的人，再根据这两个人讲一个故事。

画人测验是绘画测验中最重要的一种。古德伊纳夫（F. L. Goderoge）首先将其用于测验智力，麦科维尔（K. Machover）将其用于测验人格。画人测验评分注重质的方面，对各个项目的解释都有自己的标准，有许多评分手册。

②画树测验

随意画一棵树，然后将画的树与定的标准相匹配，看与哪一棵标准树最接近，便可发现受测者的人格特征。画树测验是最简单的投射测验，由瑞士心理学家卡尔柯契所创，有 20 种事先订好的标准。

③房一树一人测验

测验要求受测者随意画出一间房子、一棵树和一个人，然后主试对每幅画大约提 25 分钟的问题，让受测者对所画图形及背景作描述、解释以及联想。

A. 房一树一人测验源于画树测验，与 RIT 和 TAT 涉及言语表达及联想不同，房一树一人测验是非言语性的，主要描绘的是画，它涉及被测者人格特征中的感受性、成熟性、灵活性、效率性和综合性，并且还具有一定的创造性。

B. 分数的解释实质上是定性分析，大多是根据单个目标进行广泛概括，此测验使用方便，为临床治疗师所青睐。

C. 有假设如下：房屋作为人居住的地方，可以引起对于家庭及亲人的联想；树的图面可

以反映受测者的无意识；人则反映受测者的自我形象以及与人相处的情形。

(5) 逆境对话

①罗森博格根据其挫折攻击理论编制；

②测验有成人（18岁以上）、少年（12~18岁）、儿童（4~13岁）三种，每种测验有24个图，每张图有两个人，左边的人说一些挫伤右边的人的话，假定受测者是受挫者，他会代替图中的受挫者做出反应；

③反应方式有3种（为自己辩护、强调障碍所在、建设性地解决问题），攻击方向也有3种（外向攻击、内向攻击、想象攻击或无攻击），所以反应一共有9种类型（3×3）；

④测验用于研究挫折，预测受测者在受挫折时的反应倾向，此测验项目少，有较客观的评分标准，可以建立常模，是投射测验中标准化最好的一个。

(6) 其他

词语联想测验、卡斯尔等编制的团体人格测验（CPPT）、贝拉克编制的儿童统觉测验（CAT）、里茨勒编制的密西西比主题统觉测验、哈特等编制的密执安图片测验。

(五) 其他人格测验

1. 评定量表

(1) 莱式品质评定量表：莱德编制，评定人格的内外向，肖孝嵘修订。

(2) 猜人测验：哈特松、梅等编制；同班同学互相评价人格特质，属于教育情境品质评定。

(3) 形容词检核表（ACL）：高夫编制，共24题（15题源于爱德华15种需要理论，9题源于不同团体的反应偏差），用于评价他人人格特征。

(4) 儿童人格问卷：鉴别需要心理治疗的儿童。

(5) Q分类技术：了解“自我”理解及一致性。

2. 情境测验（评价者中心技术）

评价者中心技术是在工作情境模拟测评的基础上逐步发展起来的，以测量评价被试的管理素质为核心的套标准化、程序化的评价活动。在20世纪90年代以来，该技术被引入我国人事选拔领域，并被广泛运用。

(1) 品格教育测验（CEI）

①哈特松、梅和雅克布松编制的诚实测验，属于直接测验、真实活动情境测验、公正情境测验。

②科尔伯格采用两难故事法对道德认知进行测量，属于间接测验、假象问题情景。

(2) 情景压力测验

包括：军事情境，无领导讨论（测量进取心、说服力、领导力），文件筐（公文）测验，预算计划小组（理财、合作、领导），管理游戏（组织适应能力），角色扮演（获取信息、洞察力），佩内愤怒的客户电话测验（控制情绪），情景作业，模拟经营，小组辩论，推销，演讲，面试，能力及人格测验等。

其中，无领导小组和文件筐测验是用得较多的情境测验。

3. 客观测量法

例如威特金编制的镶嵌图形测验，测查场独立和场依存性。

4.社会计量法

例如莫里洛编制的社会关系图解和社会距离量表,分别测查团体中人际关系地位以及团体成员的亲疏。

三、态度测验

(一) 传统态度测验 (直接测量)

1.瑟斯顿量表 (等距量表法)

围绕某一态度主题,选取能代表该方面的态度语或项目若干,由专家对这些项目进行等级排列,并把专家排列的结果进行项目分析,保留有效的项目以及根据专家的反应确定项目等级。

要了解某个受调查者的某方面态度,只需看其对该量表的反应,最后运用对全部项目反应结果的等级求中位数,以中位数表示该受调查者的态度状况,信度为 0.8~0.9。

编制过程复杂,中位数代表态度等级不定合适。

2.李克特量表 (总加量表法)

李克特量表假定每一项目或态度语都具有同等的量值,项目之间没有差别量值。受测者按每个项目 (正面、负面描述) 态度强弱按五级或六级反应,最后,以受测者在所有项目中评定等级的总和来估计受测者的态度。比瑟斯顿量表简易,但无法对态度差异作进一步解释。

3.哥特曼量表 (累计量表法)

试图确立一个单向性量表,单向性好坏指标采用复制系数 $Crep=1-\text{误答数}/\text{总反应数}$,当 $Crep>0.9$,单向性得以保证。

单向性使态度分数与结构一致,但编制困难。

4.奥斯古德编制的语义分化量表 (多维度量表法)

其前提是:态度是由人们对所给概念 (刺激) 的含义 (语义) 所组成的,这个含义可以通过对关键词的反应来加以确定。

测查时,先给受测者提出一个关键词 (态度对象),要求受测者按自己想法在两极形容词间的 7 个数字上圈一个数字,各系列分值的综合代表他对有关对象态度的总分,即总态度。

采用因素分析法编制,包括三个维度:评价、强度、活动。

5.博格达斯社会距离量表

测量人际关系亲疏。

(二) 间接测量法

1.内隐联想测验 (IAT)

(1) 格林沃德根据威尔逊提出的双重态度模型编制;(2) 包括 4 类概念 (2 种客体和 2 种属性)、2 种任务 (相容和不相容);(3) 相对态度= $RT(\text{相容})-RT(\text{不相容})$;(4) 用于调量内隐态度。

2.其他

包括投射法、生理反应法、启动测验等。

四、兴趣测验

(一) 概述

兴趣测验通常要考虑两个基本问题：一是兴趣的客观表现；二是兴趣的主观表现。

（二）常见兴趣测验

1. 斯特朗职业兴趣问卷（SVIB）

最早的兴趣问卷，经验法编制；用于确定某人的职业兴趣。

2. 库德职业兴趣调查表（KOIS）

由迫选项目组构成，三择一的强迫选择题，用于确定某人的职业兴趣。

3. 自我指导问卷（SDS，又称自我定向调查表）

霍兰德以其人格类型学说（现实、研究、艺术、社会、企业、传统）为基础而编制。霍兰德 1953 年编制了职业偏好调查表（VPI），其中受测者要求对 160 个不同的职业判断自己喜欢与否；霍兰德于 1969 年又编制了 SDS，由职业类型测验和职业搜索表两部分构成。

4. 斯特朗—坎贝尔兴趣问卷（SCII）

坎贝尔编制；包括四个个人风格量表；从 3 个层次[霍兰德一般职业主题（GOT）、相互异质的同质性量表（BIS）、职业量表]，5 类量表（管理指标、一般职业主题、基本兴趣量表、职业量表、个人风格量表）上进行解释；多用于帮助高中、大学生进行职业生涯规划。

五、心理健康测验

（一）症状自评量表 SCL-90

（1）德罗盖提斯于 1975 年编制；

（2）共 90 个项目，10 个因子，采用 0-4 的 5 级评分；

（3）适用于 16 岁以上受测者，评定范围为最近一周；

（4）总分超过 160 分，或阳性项目超过 43 项，或任一因子分超过 2 分，需进一步检查；

（5）是最广泛的心理障碍门诊检查表，不适合躁狂和精神分裂患者。

（二）焦虑量表

1. 显性焦虑量表（MAS）

由泰勒依据卡默龙关于显性焦虑反应所描述的显性焦虑概念，运用逻辑分析法编制；研究焦虑对学习的动机或驱动作用。

2. 状态—特质焦虑量表（STAT）

（1）由施皮尔伯格等根据其状态—特质理论编制而成；

（2）包括状态焦虑和特征焦虑（卡特尔和塞欧提出的两种焦虑形式）两个部分，采用 4 级计分，注意反向计分；

（3）适用于初中文化水平以上的受测者；

（4）最好先做状态焦虑的题（会形成一定的测试气氛），后做特征焦虑的题。

3. 测验焦虑量表（TAI）

（1）施皮尔伯格根据状态—特质理论编制；

（2）测验共 20 题，采用 4 级评分，包括两部分，测量 W 因素（对失败结果的认知）和 E 因素（紧张引起的神经系统反应）；

（3）将测验焦虑看作特质，测查焦虑的倾向性。

4. 焦虑自评量表（SAS）

（1）由扎格（William W. K. Zung）编制；

(2) 包括 20 个反应焦虑主观感受的项目，每个项目按症状出现的频率进行 1~4 的四级评分，评定范围为最近一周；

(3) 适用于具有焦虑症状的成年人；

(4) 用于评定焦虑症状出现的轻重程度及其在治疗中的变化。

5.其他

(1) 贝克焦虑量表 (BAI)

21 个自评项目，每个项目是一种焦虑症状，采用 4 级评分，总分 $Y(总) = INT(1.19X)$ 取整；评定范围为近一周内，测查主观感到的焦虑程度。

(2) 汉密尔顿焦虑量表 (HAMA)

最经典的焦虑量表之一，14 个症状，0~4 的 5 级评分；评定病人焦虑严重程度：得分 >29 分表示有严重焦虑，>21 分表示焦虑肯定明显，>14 分表示肯定有焦虑，>7 分表示可能有焦虑，<7 分表示没有焦虑。

此外还包括：抑郁自评量表 SDS (扎格，20 题，四级评分，评定范围为最近一周)、焦虑自评量表 SAS (扎格，20 题，评定范围为最近一周)、康奈尔医学指数 CPA (布鲁德曼)、大学生心理健康调查表 UPI (樊富珉从日本引进)、生活事件量表 (LES)、社会支持评定量表 (SSRS)、应对方式问卷 (CSQ) 等。

六、发育量表

(一) 格赛尔发展量表 (GDS)

(1) 格赛尔编制，又称耶鲁量表；(2) 包括动作能力、应物能力、言语能力和应人能力等四个方面的能力，包括 0~5 岁的各种行为的发展常模；(3) 适用于 4 周至 3 岁的婴儿；(4) 格赛尔是婴儿量表创始人，GDS 是第一个婴儿量表，主要用来诊断孩子发展是否正常，测得结果以发育商数 (DQ) 表示， $DQ = (\text{测得的发育成熟年龄} + \text{实际年龄}) \times 100$ 。

发展商数如低于 65-75，则表明发育有严重的落后；若婴幼儿阶段应物能力发展商数低于 85，则表明机体存在损伤。

(二) 丹佛发展筛选测验 (DDST)

(1) 弗兰肯堡、多兹编制的筛选性测验，并非测定智商，对婴幼儿目前和将来的适应能力和智力高低无预测作用，只是筛选出可能的智商落后者；(2) 分为 4 个方面 (应人能力、细动作-应物能力、言语能力、粗动作能力)，采用限定式描述法，家长从四个答案 (通过、不通过、拒绝、无条件) 中选择一个；(3) 适用于 0~6 岁受测者；(4) 用于快速发现幼儿的智能发展差异。

(三) 新生儿评定表 (NBAS)

(1) 布雷泽尔顿编制；(2) 测验包括 27 个项目，分为习惯化、朝向反应、运动控制的成熟性、易变特点、自我安静下来的能力、社会行为 6 大类，按 9 级评分，中间等级为正常，两端为偏离正常的反应；(3) 是目前适用于年龄最小 (0~30 天) 的婴儿使用的行为量表；(4) 目的在于诊断和预测婴儿的行为发展情况。

(四) 贝利婴儿发展量表 (BSID)

(1) 贝利编制；(2) 包括三个分量表：智能量表、运动量表、婴儿行为记录表；(3) 适用于 2 个月-2 岁半的受测者；(4) 该量表信效度都较高，被认为是最好的婴儿测验。其

成绩用智能发展指数和心理活动发展指数表示，分别评定智能水平和运动水平，均值 100，标准差为 16。

七、成就测验

1. 定义

对个体在一阶段的学习或训练之后知识、技能的发展水平的测定，也称学绩测验。与其他心理测验不同的是，成就测验希望看到练习和发展的效应；与能力测验一样，成就测验属于最佳行为测验。成就测验一般是团体测验，所测量的是认知性心理能力，一方面是认知内容的多寡，另一方面是认知能力的高低。

2. 成就测验的编制

由学科与测验编制专家共同完成，其步骤与一般心理测验的编制程序相同。

3. 成就测验的标准化

与一般心理测验的标准化程序相同，只是在分数解释的标准化时，不是要编制好常模，而是要求命题组认真研究教材和教学大纲，分析合格标准的确切含义，经过调整试题难度结构准确划分合格分数线。

4. 常用成就测验

(1) 斯坦福成就测验

①加德纳编制；②适用于 1-9 年级；③是最早的综合成就测验。

(2) 其他

镇术学科测验、GRE、韦氏个别成就测验（WIAT）、考夫曼教育成就测验（K-TEA）、大都会成就测验（MAT）、艾奥瓦基本技能测验（TBS）等。

八、测验中常见的常模

1. T 分数：50 为平均数，10 为标准差（MMPI、EPQ）。

2. 标准九分：5 为平均数，2 为标准差。

3. 标准 10 分：5.5 为平均数，1.5 为标准差（16PF）。

4. 标准 20 分：10 为平均数，3 为标准差（韦氏分测验的量表分）。

5. 离差智商：100 为平均数，15 为标准差（韦氏）。

6. 离差智商：100 为平均数，16 为标准差（比内一西蒙 1960 版，斯坦福一比内）。

九、心理测验的应用

（一）心理测验在心理咨询中的应用

心理测验在心理咨询中的作用主要是诊断与效果评估，尤其以诊断用得最多。

1. 在自我认识、人生规划咨询中的应用

心理测验可以帮助个体了解自己的性格、智力、价值观、气质类型等心理特性，从而客观地对待自己的优缺点，对自己有好处，对社会也有价值。常用的量表有《卡特尔 16 种人格因素测验》《YG 性格测验》《艾森克人格问卷》等。

2.在神经症、人格障碍等咨询中的应用

心理测验可以帮助咨询师对来访者做出咨询诊断和进行咨询效果评估。常用的量表有《明尼苏达多相人格调查表》《艾森克人格问卷》《症状自评量表(SCL-90)》，以及部分焦虑测验量表。

(二) 心理测验在人事测评中的应用

对于在岗人员，测验的应用主要有两个方面：一是做出在岗人员是否合格的诊断。二是对不合格者重新安置及培训的效果评估。对于要挑选的不在岗人员而言，主要是其选拔作用。如果把这两类人合在一起，心理测量的应用主要有3个方面：一是人员的心理特点评估。二是人员培训后的心理特点评估。三是工作人员的绩效评估。

1.在人的心理特点评估中的应用

此时主要作用有两大方面：是一般心理品质的测量，主要有智力和个性测量等；二是专业知识和特殊能力测验，包括专业知识技能测验以及一些特殊能力测验，如音乐能力、美术能力、文书能力、机械能力以及管理能力等。

2.人员培训后的心理特点评估

这类测量的对象可以是知识或技能水平的提高，也可以是工作态度或工作兴趣的改变。可以使用成就测验、兴趣或态度测验。

3.工作人员的绩效评估

用于评估工作人员的绩效。如使用对领导者行为效果及员工心态评估的PM量表，PM量表包括P因素量表、M因素量表和情境因素量表。由PM量表的得分区分领导的类型，用情境量表的得分高低作为部下士气、态度和满意度的反映，也是领导效果的表征。

(三) 测量在教育评价中的应用

1.在测量学生的学习与发展状况中的应用

- (1) 弄清学生的学习和发展状况，是因材施教的前提；
- (2) 弄清学生的学习和发展状况，是评价教育过程中不同阶段成效的依据；
- (3) 弄清学生的学习和发展状况，是评价一种新的教育思想、教育措施或教育技术等有效与否的重要指标。

2.在测量教师与管理者评价中的应用

对教师的评价主要有四个方面的内容：一是教师的资格评定，二是教师的教学水平的评定，三是教师的管理水平的评定，四是教师的个性评定。

(四) 测量在研究中的应用

- 1.搜集资料得到大量有关受测者能力、性格、与群体的关系方面的资料，有助于发现问题、探讨规律；
- 2.建立和检验假说可以通过分析测验结果，提出和检验心理与教育理论；
- 3.实验分组，用测验对受测者进行实验分组。