

心理测量学

第一章 心理测量基本理论

1. 施瑞奥克把三国时期刘副的《人物质》翻译成《人类能力的研究》在美国出版 统 10 选
2. 七巧板是最早的创造力测验 统 13 选
3. 心理测量的基本要素: 2 点 参照点 (2 类 绝对 相对) 单位 (2 要求 确定意义 相等价值) 选择
4. 心理测量的分类: 9 类 选择

功能: 能力测验 (一般能力: 智力测验 特殊能力: 音乐、体育) 成就测验 (achievement) 人格测验

对象: 纸笔 口头 操作 电脑

目的: 描述 诊断性 (临床 学习) 预测性

要求: 最高作为 Performance 典型行为

速度难度: 速度 难度 结合

性质: 构造性 (结构性) 投射性 project

解释: 常模 标准参考

应用: 教育 职业 临床
5. 心理测量的作用: 2 类 理论 (收集数据 建立和检验理论假设 实验分组) 实际 (人才选拔 人员安置 心理诊断 评价描述 心理咨询) 选择 统 11 选 (分班考试是安置性)
6. 心理测量中的行为样本不能有偏差, 要有代表性, 能反映被试行为的一组特征, 要是被试出现的可观察的行为 统 17 选
7. 2008 年中国心理学会: 《心理测验管理条例》选择 统 14 选

第二章 经典测量理论

1. 测量的误差种类: 2 类 随机误差 (影响稳定性, 也影响准确性) 系统误差 (影响稳定性, 不影响稳定性) 选择 18 选 (心理测验中的误差是指) 统 08 选
2. 控制测量误差的方法: 4 个 提高编制测验的科学性 施测过程的标准化 评分与计分的科学性与标准化 主试与被试的配合及操作规范 选择 16 选 (测验标准化是为了控制测量误差)
3. 心理科学的调查研究中, 不可避免的误差有: 3 个 随机误差 组间误差 抽样误差 统 14 多
4. 经典测量的数学模型: $X = T + E$ X -观察分数 T -真分数 E -随机误差 选择 16 选 (无数次测量获得的平均值为真实分数) 统 07 选
5. CTT 的假设: 2 点 X 与 T 间线性关系 X 与 T 间只相差一个随机误差 选择 18 多
6. CTT 模型 Gulliken 引申的三个假设公理: $\lim_{n \rightarrow \infty} E(X) = X$ $\rho_{ET} = 0$ $\rho_{E_1 E_2} = 0$ 选择 统 10 选
7. CTT 模型下变异的关系: $S_X^2 = S_T^2 + S_E^2 = S_V^2 + S_I^2 + S_E^2$ S_V^2 -与测量目的有关的变异数 S_I^2 -与测量无关的变异数 S_E^2 -和测量误差的变异数 选择 统 13 选
8. 信度的含义: 同一测量工具反复测量某人的同一种心理特质, 多次结果间的一致性就是信度 选择
9. 信度的 3 个等价定义: $r_{XX} = S_T^2 / S_X^2$ $r_{XX} = \rho_{XT}^2$ $r_{XX} = \rho_{XX'}$ 选择
10. 信度的作用: 3 个 随机误差大小的反映 解释个人测验分数的意义 帮助进行不同测验分数的比较 统 12 选 (信度系数可以解释随机误差的变异和真分数的变异)
11. 一个测验的标准误: $S_E = S_X \sqrt{1 - r_{XX'}}$ 选择 统 15 选 ($r \rightarrow 1$ 时 $S_E \rightarrow 0$) 统 10 选
12. 真分数的置信区间: $[X - Z_{(\alpha/2)} S_E, X + Z_{(\alpha/2)} S_E]$ 统 15 论述
13. 测验结果精确程序的指标是: 2 个 测验信度 测验标准误 多选 统 11 多
14. 信度的估计方法: 5 个 选择

- ① 重测信度=再测信度：两次测试的皮尔逊积差相关 条件：3个 心理特性稳定 遗忘和联系抵消 两次时间间隔 内心理特质没有获得更多的学习和训练 16选=统 09选（重测信度的主要误差源是时间不同）统 08选（概念）
- ② 复本信度：两次平行测验测量同一批被试的皮尔逊积差相关 条件：2个 构造真正的平行测验 被试有条件 接受两个测验
- ③ 分半信度：斯皮尔曼-布朗公式校正： $r_{xx} = \frac{2r_{hh}}{1+r_{hh}}$ 适用：两半测试分数的变异系数相等 不相等用弗朗那根 公式 or 卢仑公式 18选（奇偶两半测验 $r=0.5$, 求信度）统 2007选（计算）统 15选（幼儿快速发展，适用其的信度资料方法是分半法）
- ④ 同质性信度：测验所有题目间的一致程度 库-理系数（K-R20 只适用 0 1 记分 K-R21 只适用题目难度接近）
 克龙 α 巴赫系数 $\alpha = \frac{K}{K-1} \left(1 - \frac{\sum s_i^2}{s_x^2} \right)$ 统 15选（克龙的含义）统 15论述
- ⑤ 评分者信度：多个评分者给同一批人的答卷进行评分的一致性程度 两个评分者（积差相关 or 等级相关） 多个（肯德尔和谐系数）
15. 克龙巴赫系数是所有可能的分半信度的平均值 它只是测量信度下界的一个估计值 α 大测量信度高， α 小不能断定测量信度不高 选择 统 16选
16. 影响信度的因素：5类 被试方面 主试方面 施测情境方面 测量工具方面 两次施测的间隔时间 简答 统 07简答
17. 提高测量信度的方法：5点 适当增加测验的长度 使测验中试题难度接近正态分布 努力提高测验试题的区分度 选取恰当的被试团体 主试严格执行实测规程 选择、论述 统 15论述
18. 效度的定义：效度 r_{xy} ，效度系数： $r_{xy}^2 = \frac{S_y^2}{S_x^2}$ 得到效度 a ，则可以解释变异的 a^2 选择 统 07选 统 12选（不能解释的是 $1-a^2$ ）
19. 效度与信度的关系：信度高不一定效度高，但想要效度高，信度必须高 $r_{XY}^2 = r_{XX}^2 - \frac{S_I^2}{S_X^2} < r_{XX}^2$ 一个测验的效度总是受它的信度所制约 选择 统 09选（信度系数恒大于或等于效度系数）
20. 效度的估计：3个
- ① 内容效度：成就测验、选拔和分类的职业测验 不能用于能力倾向测验和人格测验 逻辑分析法 16选（内容效度主要考查成绩技能测验）
- ② 结构效度=构想效度：5个逻辑和实证方法 ①测验内部寻找证据法 ②测验之间寻找证据法 ③考察测验的实证效度法 ④多种特质-多种方法矩阵法（相容效度：极不同方法测同一特质相关高则相容效度高 区分效度：相似方法测不同特质相关高则区分效度高 相似方法测同一特质相关则信度高）⑤因素分析法 18选（因素分析法用于检验结构效度）统 12选（构想效度含义）统 14选（MTMM 考察的是构想效度）
- ③ 实证效度=效标关联效度：同时效度（诊断现状） 预测效度（预测某个个体将来行为） 效标（测验是否有效的外在标准） 效标的估计方法（3个 相关法 区分法 命中率） 16选（高考分数和高校学习的成绩的相关是预测效度） 18选（为了维护测验公平性应重视总命中率） 18多（实证效度的确定方法） 统 08选（选了 50 人，10 人不理想，预测效度是 0.8） 统 15选（预测效度含义）
21. 常用的效标：7个 学业成就 等级评定 临床诊断 专门的训练成绩 实际的工作表现（绩效） 对团体的区分能力 其他现成的有效测验 单选、多选 统 12多 统 15选（人员选拔最恰当的指标是绩效）
22. 影响效度的因素：5个 测验的构成 测验的实施程序 接受测验的被试 所选效标的性质 测量的信度 选择
23. 提高效度的方法：精心编制量表 妥善组织测验 创设标准的应试情境 选好正确的效标 简答
24. 项目难度的计算：2类 二分法记分（通过率 $P = \frac{R}{N}$ 极端分组法 $P = \frac{P_H + P_L}{2}$ ） 非二分（ $P = \frac{\bar{X}}{X_{\max}}$ ） 注意 P 和难度成反比 选择
25. 测验难度水平的确定：测验项目的难度应接近 0.5 一个测验的大多数项目的难度范围在 0.3~0.7 间，测验就能够最大限度地获得有关个体间差异的信息 若测验为了被试对某方面知识技能的掌握情况，也不用过多考虑

难度选择

26. 难度的等距变换：将 P 作为正态曲线下的面积（注意是从正无穷开始积分的）转换成 $Z \Delta = 13 + 4Z$
 $P \uparrow Z \downarrow \Delta \downarrow$ Δ 越大越难 选择 统 16 选
27. 难度对测验的影响：2 个 分布形态（难-正偏态 简单-负偏态）分数的离散程度（过难过易使全距变小 $P=0.5$ 分布较广）**联系统计中偏态中三个集中量数的关系** 选择 16 选（一组二分计分的题中，平均正确率为 0.5 表明题目难度适中）统 08 选 统 07 多（若 P 大问分布情况）
28. 项目区分度的计算：2 类 项目鉴别指数法 ($D = P_H - P_L$) 0.4 0.2 一般用 27% 正态理论上最好的是 33% 团体人数小于 100 用 50% 相关法（点二列 二列 pho 相关积差相关）**联系系统中相关的适用** 选择 17 选（项目区分度的取值）统 17 选
29. 区分度与难度的关系：难度越接近 0.5，项目的潜在区分度越大；难度越接近 0 或 1 项目的潜在区分度越小 选择 16 选（难度为 1 或 0 时区分度最小）18 选（ $P=0.5$ 区分度最大）
30. 猜测校正的公式：2 个 已知 P (对于全体而言) $CP = \frac{KP-1}{K-1}$ 未知 P (对于一个被试) $S = R - \frac{W}{K-1}$ (100 题 4 重选择题，错了 18，则猜了 $\frac{18}{3/4} = 24$ 题，猜对了 $24/4=6$ ，则做对了 $82-6=76$ 题) 选择 统 14 选 (计算)
31. 多重选择题目的：更有效控制随机误差导致的测量误差 选择

第三章 测验常模

1. 某一标准化样本的**平均分**数为常模分数 16 选
2. **常用的导出分数**：4 个 百分等级分数（缺点：单位不等 只具有顺序性）标准分数 T 分数 ($T = 10Z + 50$) 标准九分数 (0.5 1.75) 选择 统 16 多（韦氏 IQ=115 则百分等级、Z、T、标准九分数分别是多少）统 13 选=统 17 选（导出分数中不属于等距变量的）统 07 选（Z 到 T 的变换是线性变换）
统 15 选 施测某标准化测验获得的原始分数均值为 80，标准差为 16；导出分数均值为 50，标准差为 10. 某被试参加测试所得导出分数为 70，则他的原始分数为： $50+10Z=70$ $Z=2$ $(X-80)/16=2$ $X=112$
3. **标准分数的变式**：T 分数 CEEB 分数= $100Z+500$ 韦氏 IQ= $15Z+100$ EPT= $20Z+90$ 选择 (计算) 统 17 选 (Z=1，韦氏 IQ 是多少)
4. **分数合成的种类**：3 个 项目的合成 分测验或量表的组合 测验或预测验的合成 多选 18 多 统 16 选 (若可以补偿最好方法是多元回归)
5. 分数合成的方法：4 个 直觉合成 加权求和合成 多重回归 多重划分=连续栅栏法 (特质具有某种程度的互偿性) 单选、多选
6. **确定常模的注意事项**：6 个 群体构成的界线必须明确 常团体有代表性 取样过程有明确详尽的描述 样本大小适当 常团体必须近似 一般常模与特殊常的结合 多选 18 选 (常模样本大小适当的关键是[没有选项我也不知道是什么])
7. **制定常模的步骤**：5 步 确定测验群体 对测验群体进行抽样得到常模团体 对常模团体进行施测 确定常模分数类型制作常模表 撰写抽样说明和常分数的解释 简答 统 16 简答
8. 主要的常模参照分数：5 类 发展量表（年龄常模 年级常模）商数（教育商数 成就商数）百分等级常模 标准分数常模 剖析图 选择 统 09 选 (剖析图概念)

第四章 心理测验的编制与实施与测验等值

1. 编制测验的基本程序：7 步 确定测验目的 制定测验计划 编辑测验项目 预测与项目分析 合成测验 测验标准化 鉴定测验 简答 17 选 (设计一个测量自杀态度的测验，现已经收集了 800 个题目，这时需要的处理技术是预测试) 统 14 简答 (简述心理测验编制的基本过程)
2. 测验前的准备：4 个 主试在测试前熟悉测验的结构和内容及其使用方法 准备好测验材料 准备测验的环境 提前预告测验 简答 17 简答
3. 我们通常希望一个测验的**特异性高，敏感性高** 17 选

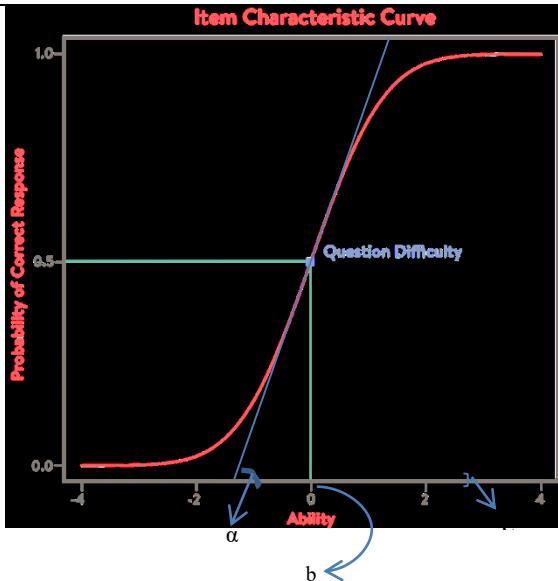
4. 测验标准化的方法: 4 个方面 测验内容 施测过程 (相同的测验情境 相同的指导语 相同的测验时限) 测验评分 测验分数的解释 简答 统 08 简答 (简述测验标准化的含义及提高测验标准化程度的方法)
5. 解释心理测验分数时应该注意哪些问题? 5 点 ①充分了解测验的性质和功能 ②对导致测验结果的原因的解释应慎重谨防片面极端 ③必须充分估计测验的常模和效度的局限性 ④解释分数还应参考被人口统计学变量等其他有关资料 ⑤应将测验分数视为分数段或范围, 而不是一个确定的点值 简答 统 12 简答
6. 测量等值的条件: 6 个 同质性 等信度 公平性 可递推性 对称性 样本不变性 选择 统 13 选
7. 经典测量理论等值关系的计算: 2 个 等百分位等值法 线性等值法 单选、多选 统 08 多

第五章 标准参照测验 (目标参照实验)

1. 常模参照测试与目标参照适用: 常 (选拔性考试 总结性考试) 目 (诊断和发现学生不足 广泛用于学生的各种基础水平测试及各种证书考试) 选择 统 13 选 (日常教学中教师用来测评学生学习效果的单元小测试属于非标准化测验)
2. 目标参照提出者: 2 个 戈莱塞 Glaser 克劳斯 Klaus 选择
3. 目标参照的两个重要内容: 内容范围 行为标准 了解
4. 目标参照内容范围的确定: 包括所有欲测量特质中蕴涵的全部行为 双向细目表 (考察目标 考察内容 比重) 选择
5. 目标参照测验项目的内容效度分析: 专家评定法 选择
6. 目标参照项目的难度和区分度分析: 测验的预测方法 (3 个 前测-后测方法 已接受教学组-未接受教学组方法 对照组方法) 难度分析 (通过率 也有人说不重要) 区分度 (2 类 ①难度差值 鉴别指数 $D = P_A - P_B$ 个人获得指数 ②相关系数)
7. 目标参照测验的信度: 2 个 分类一致性信度 $p_0 = \frac{a+d}{a+b+c+d}$ 方差分析方法 (荷伊特信度) 选择
8. 目标参照测验的效度: 2 个 内容效度 效标关联效度 (决策效度) 选择
9. 目标参照测验分数的解释=分数界点的确定: 2 类 专家判定法 (3 个 Nedelsky 方法 Angoff 方法 Bookmark 方法) 效标组预测法 (临界组法 对照组法 单选、多选)

第六章 项目反应理论 IRT

1. 项目反应理论的假设: 2 点 特质空间的单维性假设 项目特征曲线假设 选择
2. 项目特征曲线: $P(\theta) = C + (1-C) \int_{-\infty}^{\theta} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ 选择 统 11 选 (那条线叫项目特征曲线) 统 12 选=统 16 选 (a 参数意义) 统 15 选 (多条线比较区分度 看 a)
 - ① a: 拐点的斜率-区分度
 - ② b: 中等难度-难度
 - ③ c: 猜测率
 - ④ P: 通过率
 - ⑤ θ: 能力
3. 参数模型: 单参数模型 (难度系数 b) 双参数模型 (区分度参数 c) 三参数模型 (a b c)
4. 项目反应理论优良性: 4 个 题目参数的跨群体不变性 潜在特质量表的可选择性 参数的科学性 信息函数概念的引进 不慌不忙信息函数的可加性 (信息函数越大, 测试的误差越小 $SE = (\text{项目信息总量})^{\frac{1}{2}}$) 单选、多选 统 17 选 (项目信息总和为 25, 间测量精度: $1/\sqrt{25}=0.2$)
5. 项目反应理论的应用: 5 个 题库建设 常模参照测验的编制 目标参照测验编制 计算机自适应测验编制 认知诊断测量模型发展基础 多选 统 08 选 (计算机化自适应测验设计的主要理论依据是 IRT)



第七章 概化理论 GT

1. GT 的测验情境相关说：2 点 一个测验可以有不同的水平 测验的真分数（全域分数）不止一个 选择 统 09 选（强调对情境进行考察的心理测量理论是 GT）统 11 选（CTT 的 T 类似 GT 的全域分数）
2. GT 数据采集的方法：4 类 全交叉采集 相互嵌套采集 嵌套混合采集 其他（三侧面交叉设计 三油布嵌套设计 三侧面混合设计）选择
3. GT 的统计分析：2 阶段 G 研究 D 研究=决策研究 选择 统 15 选（D 研究是决策研究）
4. GT 研究对总方差的分类：3 类 目标主效应方差 侧面主效应方差 交互效应方差（各测量侧面与测量目标形成的各级交互方差 纯由各测量侧面自己形成的各级交互方差）多选 统 13 多
5. G 研究的结果：对测验观察领域中各种效应期望均方的估计 选择
6. 概化理论中最具特色的计量分析手段：D 研究 选择
7. D 研究的目的：利用 G 研究的结果数据，提供各种测验方案下的测验误差估计值 选择
8. D 研究的误差指标：2 个 相对误差方差（所有与测量目标有关的交互效应方差之和） 绝对误差方差（除测量目标效应方差之外的所有方差之和）（总方差=侧目主+侧面主+交互） 选择 统 12 选
 $\frac{\text{相对误差}}{\text{绝对误差}}$
9. D 研究的综合指标：2 个 概化系数 (G 系数 $G = \frac{\text{测量目标效应方差}}{\text{测量目标效应方差} + \text{相对误差方差}}$)
 依存系数 (Ψ 系数 $\Psi = \frac{\text{测量目标效应方差}}{\text{总效应方差}}$) 选择

$$G = \frac{\hat{\sigma}^2(P)}{\hat{\sigma}^2(P) + \hat{\sigma}^2(\xi)}$$

$$\hat{\sigma}^2(P) \text{ 侧目 } \hat{\sigma}^2(\xi) \text{ 相对 } \hat{\sigma}^2(\Delta) \text{ 绝对 } \text{ 相对+绝对=总 } \text{ 相对=侧目 } \text{ 侧目+绝对=总}$$

统 16 选 下列关于概化理论的表述，错误的是 C

- A 属于随机抽样理论 B 关于行为测量可靠性的理论
 C 关于被试能力估计的理论 D 全域分数类似于经典测量理论的真分数

第八章 常用心理测验

1. 学绩测验所测的东西：认知性心理品质 选择
2. 学绩测验的作用：3 个 测定获得了多少学科知识 鉴别学习困难的学生，诊断学生困难原因 在教学管理中学生的升学、毕业、升级、留级、划分班级 单选、多选
3. 学绩测验的分类：5 类 按测验的编制方法（教师自编课堂测验 标准化学绩测验）按测验的内容（单科测验 多

- 科测验) 按测验的用途(考察性/诊断性) 按测验评分系统的参照系(参照测验/目标测验) 按测验的题型(定向反应型=客观型/自由反应型=论文式) 单选、多选
4. 标准化学绩测验的基本要求: 4个命题组卷标准化 施测标准化 评分标准化 测验分数解释标准化 多选、简答
 5. 常见的标准化学绩测验: 3个 斯坦福成就测验 关键数学算术诊断测验 高考 选择
 6. **个体智力测验:** 3类 比内量表(3个 比内-西蒙量表 斯坦福-比内量表 中国比内量表) 韦克斯勒量表(韦氏成人智力量表 韦氏儿童智力量表 韦氏儿童智力量表) 戴斯的认知测验 CAS(创新性的认知状态测量工具)
 7. **团体智力测验:** 3类 陆军测验(甲 乙-非文字测验) 瑞文推理测验(非文字型的图形测验 文化公平) 认知能力测验(Thorndike)
 8. 能力倾向测验的目的: 不在于总结过去, 而在于预测将来, 预测个体在将来的学习或工作中可能达到的成功程度 选择
 9. **常见能力倾向测验:** 4个 学术能力倾向测验 SAT 分辨能力倾向测验 DAT(适用于初中和高中生的教育咨询及就业问题 广泛用于教育领域中对学生将来学术成就的预测) 一般 肋条倾向成套测验 GATB(综合式职业性向测验 重于实际操作) 行政职业能力倾向测验 AAT(国家行政机关招考主任科员以下非领导职务公务员) 选择
 10. 创造性思维常见量表见心导 选择
 11. **常见人格测验量表:** 2类 自陈量表(7个 明尼苏达多项人格调查表 MMPI 卡特尔16种人格因素量表 16PF 艾森克人格问卷 EPQ 爱德华个人偏好量表 EPPS Y-G 性格测试 加州心理调查表 CPI 儿童心理 14种人格因素问卷 CPQ) 投射测验(2个 罗夏克墨迹测验 主题统觉测验) 选择 18 选(EPQ采用自陈量表)
 12. 为了确认被试在**人格测验中的偏差**, 可以使用的方法是: **信度量表** 选择 17 选
 13. 兴趣测验: 3个 斯特朗职业兴趣问卷 SVIB 库德职业兴趣调查表 KOIS 霍兰德自我指导问卷 SDS 选择
 14. 常见态度量表: 两类 量表测量(等距量表法=瑟斯夫量表 利克特量表法 哥特曼量表法) 内隐联想测验 选择
 15. 瑟斯夫的等距量表测项目一致性的指标是: 中位数 18 选(不一定有这个题目)
 16. 心理健康量表: 6个 SCL-90 心理状况自测表(检测各类神经症状的工具) 显性焦虑量表 MAS 状态-特质焦虑量表 STAI 测验焦虑量表 TAI 贝克焦虑量表 BAS 汉密顿焦虑量表 HAMA 选择
 17. 发育量表: 格塞尔发育量表(发育商数 DQ=(测得的发育成熟年龄 ÷ 实际年龄) × 100) 选择
 18. 评价中心技术 AC: 文件筐测验 无领导小组讨论 管理游戏 角色假伴演
17 选 如果你要是用一个人格测验来帮助进行**人职匹配**, 你会使用: **Myer-Briggs 类型指标**

• 常见量表

学绩测验 (成就测验)	斯坦福成就测验				
	关键数学算术诊断测验				
	高考				
一般能力测验 (智力测验)	个体智力测验	比内量表	比内-西蒙量表		
			斯坦福-比内量表		
			中国比内量表		
		韦克斯勒量表	韦氏成人智力量表		
			韦氏儿童智力量表		
			韦氏幼儿智力量表		
	戴斯的认知测验 CAS				
	团体智力测验	陆军测试	陆军甲种		
			陆军乙种		
			瑞文推理测验		
			认知能力测验		
能力倾向测验	学术能力倾向测验 SAT				
	分辨能力倾向测验 DAT				
	一般能力倾向成套测验 GATB				
	行政职业倾向测验 AAT				
特殊能力测验	音乐能力测验	西肖尔音乐才能测验			
		音乐能力测验图			
	美术能力测验	梅尔美术判断力测验			
		格雷福斯图案判断测验			
	机械能力测验	霍恩美术能力问卷			
		空间关系测验			
	创造力测验	机械理解能力测验			
		吉尔福特发散思维测验			
人格测验	自陈量表	托伦斯创造性思维测验			
		卡特尔 16 种人格因素量表 16PF			
		艾森克人格问卷 EPQ			
	投射测验	明尼苏达多项人格调查表 MMPI			
		罗夏克墨迹测验			
		主题统觉测验			
兴趣测验	常见职业兴趣测验	斯特朗职业兴趣问卷 SII			
		库德职业兴趣调查表 KOIS			
		自我指导问卷 (霍兰德) SDS			
态度测量	量表测量	等距量表法=瑟斯顿量表			
		利克特量表法			
		哥特曼量表法			
	内隐联想测量 IAT				
	情境测验法	真实活动情境测验			
		假想的问题情境			
		问卷测量法			
心理健康量表	SCL-90 心理状况自测表				
	显性焦虑量表 MAS				
	状态-特质焦虑量表 STAI				

	贝克焦虑量表 BAS
	汉密尔顿焦虑量表 HAMA
发育量表	格塞尔智能发育量表

@想搞学术的余赛赛