

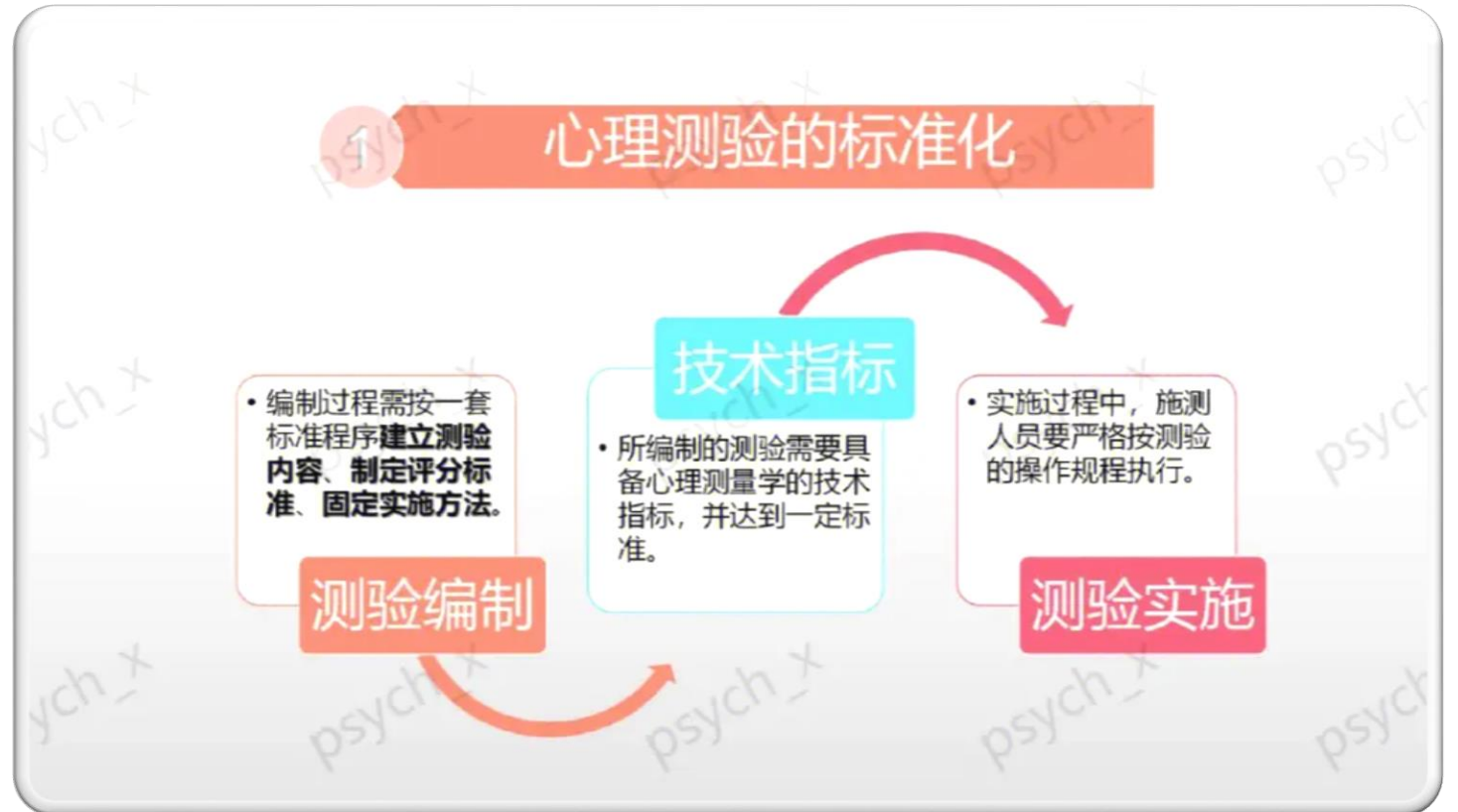
第九讲 心理测验的编制

——大学生疾病感知量表的编制及其信效度分析

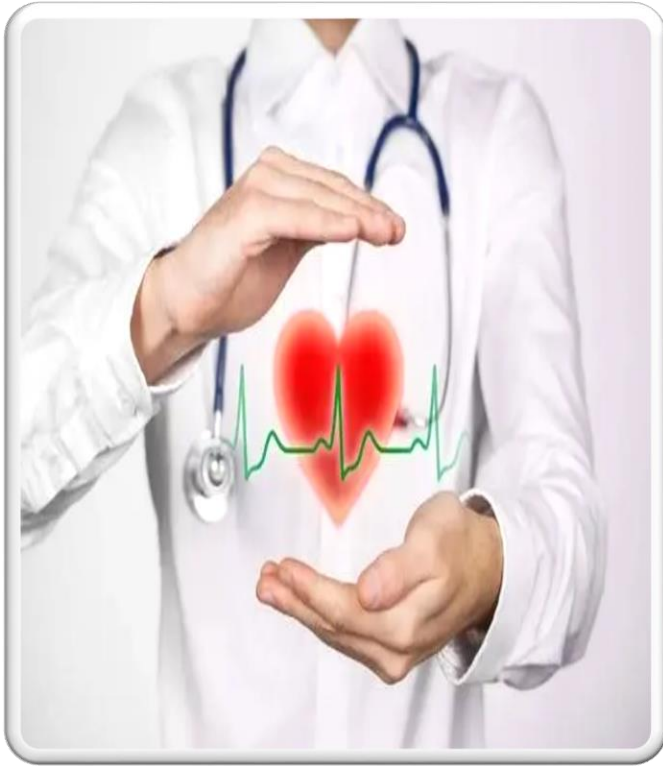


编制测验的一般程序

- (1) 确定测验编制的目的
- (2) 编辑测验项目
- (3) 预测
- (4) 测验的项目分析
- (5) 合成测验
- (6) 常模的编制
- (7) 测验标准化
- (8) 鉴定测验
- (9) 编写测验说明书



第一节 编制测验的目的



□ 待测属性、对象、用途。

□ 一、待测属性

□ 疾病感知，又称疾病认知，是指是个体在疾病状态或健康受到威胁时，利用已有的疾病知识经验分析、解释当前症状（或疾病），进而形成对疾病的理解及情绪反应，导致应对行为的产生，之后经由评价过程修正自身对疾病的认知，进而影响患者预后的过程。

信息来源

疾病感知 疾病认知 疾病体验 疾病感受等关键词

CNKI首页 手机版 网站地图 帮助中心 旧版 使用手册 | 检索设置 会员 我的CNKI 欢迎来自 广州中医药... 的您, 个人账户 登录

cnki中国知网 www.cnki.net

主题 疾病感知 问答 结果中检索 高级检索 知识元检索 > 引文检索 >

总库 621 中文 外文 学术期刊 459 学位论文 156 会议 2 报纸 1 年鉴 图书 0 专利 标准 0 成果 0

科技 社科 检索范围: 总库 主题: 疾病感知 主题定制 检索历史 共找到 621 条结果 1/32

全选 已选 0 清除 批量下载 导出与分析 排序: 相关度 发表时间↓ 被引 下载 综合 显示 20

	题名	作者	来源	发表时间	数据库	被引	下载	操作
<input type="checkbox"/>	1 克罗恩病患者应对方式现状及其相关影响因素分析	郑丽梅	罕少疾病杂志	2023-11-25	期刊	19		
<input type="checkbox"/>	2 疾病感知和认知情绪调节在造血干细胞移植患者社会限制与癌症复发恐惧间的中介效应	沈志莹;李成媛;石双姣;阮春红	护理学杂志	2023-11-25	期刊			
<input type="checkbox"/>	3 缺血性脑卒中病人疾病感知在社会支持与恐惧疾病进展间的中介效应	黄显义;刘霜;岳洁;刘艳;李玉 >	循证护理	2023-11-22	期刊	121		
<input type="checkbox"/>	4 自我调节理论指导健康教育在老年动脉瘤患者护理中的应用及对疾病感知的影响	杨磊	吉林医学	2023-11-15	期刊	38		
<input type="checkbox"/>	5 基于自我调节模式的疾病感知护理结合预见性化疗反应护理对白血病化疗患者毒副反应、应对方式及疾病感知的影响	朱苗苗;郝婷婷;董昌虎	临床医学研究与实践	2023-11-11	期刊	30		
<input type="checkbox"/>	6 基于压力与适应理论的护理干预结合“三三法则”对经尿道前列腺电切术后膀胱过度活动症患者疼痛程度、排尿情况的影响	杨磊;曹辉茹;周影;周莉	临床医学研究与实践	2023-11-11	期刊	23		

主题: 疾病感知 (281), 影响因素 (72), 生活质量 (48), 现状及影响因素 (48), 相关性研究 (45), 影响因素分析 (36), 乳腺癌 (36), 自我管理 (35), 冠心病 (28), 脑卒中 (27)

学科: 肿瘤学 (32), 中医学 (3)

第一节 编制测验的目的

二、对象

- 通常以年龄、性别、职业、受教育程度、经济状况、民族、文化背景等指标来区分测量对象。



第一节 编制测验的目的

□二、用途

□ 描述测验

- 又称诊断测验，在于测验待测属性的当前情况，用于了解情况、处理当前问题、鉴别、分类等。

□ 预测测验

- 多用于选拔和预测，从当前的行为样本来预测被试的将来情况。
- 测验用于“预测”时应该谨慎，只有确定现在的行为样本能够用来预测将来的行为时才可称为预测性测验。



第二节 编辑测验项目

首先，采用半开放式问卷调查，向 116 名广州中医药大学大一至大五在读学生发放问卷（问题 1：在您的生命历程中，是否患过疾病？（普通感冒，或者重大疾病等）选项 1：有，选项 2：无；问题 2：当提到“疾病”一词时，您脑海中联想到的词语有哪些？（必答 5 个）；问题 3：请您用 5 个词语描述一下当您患病时，有什么样的感受？（必答 5 个）），收回有效问卷共 116 份。问卷收集后，结合疾病意义相关资料，与导师和 4 名本科生进行词汇的整理，得出 128 个关键词。接下来与导师和 4 名本科生一起将近义词和同义词进行合并与归类，形成三个维度，66 个条目。

一、收集测验资料

资料要丰富

- 无论是能力或人格，均是十分复杂的复合性心理结构，不能仅凭一两种简单的项目去推断一个人的智力或人格结构，必须包含不同类型的材料。

资料要有普遍性

资料要有趣味性

- 趣味性可以减少受测者由于缺乏足够的动机而引起的测量误差。

第二节 编辑测验项目

二、选择测验项目的形式

（一）选择题：

1、定义：

题干

选项

17、我觉得自己是个有用的人，有人需要我。

- 没有或很少时间。
- 小部分时间。
- 相当多时间。
- 绝大部分或全部时间。

18、我的生活过得很有意思。

- 没有或很少时间。
- 小部分时间。
- 相当多时间。
- 绝大部分或全部时间。

19、我认为如果我死了，别人会生活得好些。

- 没有或很少时间。
- 小部分时间。
- 相当多时间。
- 绝大部分或全部时间。

指导语：以下列出了有些人可能会有的问题，请仔细地阅读每一条，然后根据最近一星期以内下述情况影响您的实际感觉，在每个问题后标明该题的程度得分。其中，“没有”选1，“很轻”选2，“中等”选3，“偏重”选4，“严重”选5。

题目	选择
1.头痛。	1-2-3-4-5
2.神经过敏，心中不踏实。	1-2-3-4-5
3.头脑中有不必要的想法或字句盘旋。	1-2-3-4-5
4.头昏或昏倒。	1-2-3-4-5
5.对异性的兴趣减退。	1-2-3-4-5
6.对旁人责备求全。	1-2-3-4-5
7.感到别人能控制您的思想。	1-2-3-4-5

第二节 编辑测验项目

□ 2、注意事项

- 题干所提的问题必须明确，尽量使用简单而且明晰的词语。做到题干意义完整，即使被试不看选项亦能完全理解。
- 选项切忌冗长，要简明扼要。
- 能力测验：
 - 每题只给一个正确答案，其它属诱答。若是找最合适的答案，则应用这样的问句“下列答案中哪个最合适?”，以免引起困惑。
 - 各选项长度应相等，尽量不要有长有短。同时，选项与题干的联系要非常密切。诱答也必须一致，以免被试很容易就排除了诱答项目。

第二节 编辑测验项目

□ (二) 是非题

□ 1、定义

- 是非题又叫正误题，是指出一个论点要被试判断是否正确，或是从是非两个答案做出选择，因此可以把是非题看作是两个备选答案的选择题。

例：(1)你常常会主动地去做一些有意义的习题吗？

是 否

(2)你常常主动给朋友写信或打电话吗？

是 否

第二节 编辑测验项目

□2、编制是非题应注意：

- 内容应以有意义的概念、事实或基本原则为基础，不要在叙述中出现琐碎的细节或无关的话语。
- 每道题只能包含一个概念，避免两个或两个以上的概念出现在同一个题目中，造成“半对半错”或“似是而非”的情况。
- 尽量避免否定的叙述，尤其是要避免用双重否定的叙述
 - 因为否定的叙述常会被人误认为肯定的叙述，将“不”字忽略，双重否定尤其容易使人困惑，不如直接采用肯定的叙述为佳。
- 若是表达意见的题目，最好说明意见的来源和根据，以便测出被试是否了解某个人或某些人的意见、信念或价值观念等。
- “是”、“非”题的选项应有适应比例，基本相等，且要随机排列。“是”、“非”题目的编写在长度和复杂性上应尽量保持一致。

第二节 编辑测验项目

□ (三) 简答题

□ 1、定义

- 要求被试用一个正确的词或句子来完成或填充一个未完成句子的空白，或者是提供一个正确的答案。
- 有时将前者称之为填充题，后者称之为简答题。

了解心理的钥匙是_____。

指导语：“现在我来问你一些问题，请你回答，这些问题有的很容易，你很快便可回答；有的可能也要想一想才能回答，你一想好了便回答。懂吗？好现在我们开始”。

- 1、钟表有什么作用？
- 2、球是什么形状？
- 3、一年有多少月？
- 4、国庆节是哪一天？
- 5、一年中哪个季节白天最长？

第二节 编辑测验项目

□（三）简答题

□2、编制简答题有三条原则：

- 宜用问句形式。如果使用未完成的句子，则空格尽量放在最后。
- 如果是填充形式，空格不宜太多，过多空格会使题意不明确，并且所空出的应该是关键词句。
- 每题应只有一个正确答案，而且答案要简短而具体。对不完整的答案，应事先规定评分标准。

第二节 编辑测验项目

□（四）操作题

□1、定义

- 操作题是介于一般认知结果的纸笔测验和在未来真实情境的实际活动之间的测验，是让被试实际操作，如画图、走迷津、拼配物体等，可作为纸笔测验题的补充。



第二节 编辑测验项目

□（四）操作题

□2、编制操作测验有四条原则：

□明确所要测量的目标，并将其操作化。

□即要进行工作分析，辨认出操作中最重要特殊活动，找出具有代表性的工作样本；

□还要建立作业标准，指出通过此项作业的最低标准。

□指导语要简明扼要，主要让被试明白要他们做什么和在什么条件下做，如使用什么工具、时间限制以及评价的依据等。

□制订评分标准，确定计分方法。

□有些操作项目可根据完成的数量和错误次数客观记分，有些项目的评分则较为困难。

第二节 编辑测验项目

□ 三、编写测验项目

□ 1、原则

- (1) 测验项目的取样应当对欲测的心理特质具有代表性。
- (2) 测验项目的难度应有一定的分布范围。
- (3) 编写测验项目的用语要力求精炼简单，浅显明了。
- (4) 初编题目的数量要多于最终所需要的数量，以便筛选或编制复本。

第二节 编辑测验项目

□ 2、材料来源

□ (1) 已有的测验

- 最简单、最直接的方法是从已经出版的各种标准测验中选择合适的题目。

□ (2) 访谈

□ (3) 学术资料

□ (4) 专家经验

- (5) 临床观察和记录：临床的观察也可以作为题目的来源，各种观察量表或检核表很多都是来源于观察到的行为表现。

第三节 测验预测

一、预测目的

- 预测的目的在于获得被试对测验项目做何反应的资料。
- 它既能提供那些题目意义不清、容易引起误解等质的信息，又能提供测验项目优劣的量的指标。

样本二：通过问卷星平台，面对广州中医药大学大一至大五在读学生进行预测量表的发放，共收回 369 份问卷，有效问卷为 338 份。（其中女性共 210 名，约占 62%，男性共 159 名，约占 38%；其中医学生 138 名，非医学生（接受过基础的医学课程）200 名；其中大一的学生 21 名，大二的学生 30 名，大三的学生 146 名，大四的学生 105 名，大五的学生 36 名）。样本二为填写预测问卷的被试。

第二节 编辑测验项目

□ 二、预测的注意事项

- (1) 预测对象应取自将来正式测验时准备施用的群体。
- (2) 预测的情境应力求同正式测验的情境一致。
- (3) 预测的时限可以适当延长，以便每一受测者都能将题目做完。
- (4) 施测者应对受测者的反应加以记录。如在同一时间内，受测者所完成的题数，以及受测者反应的题意不清等，以便修改项目时参考。
- (5) 预测人数：一般而言，预测人数为题目数的0.5-2倍。

第二节 编辑测验项目

三、数据录入

尽量使用SPSS录入

编号	服刑人员姓名	关系	本人姓名	性别字符	性别	年龄	职业	V1	V2	V3	V4	V5	V6	V7
1	樊照波	兄弟	樊少波	男	1.00	45	个体户	3	2	2	2	1	3	
2	袁剑威	兄弟	袁剑锋	男	1.00	41	个体户	3	2	2	2	1	3	
3	冯翠怡	母亲	冯玉英	女	.042		散工	2	2	3	2	3	3	
4	陈锦华	父子	陈及森	男	1.00	62	无	1	3	3	1	3	1	
5	陈福根	爱人	冯添弟	女	.035		种植	3	2	2	1	3	2	
6	周健能	夫妻	冯可君	女	.035		务农	2	2	3	1	3	1	
7	郭洪军	夫妻	吴清霞	女	.039			2	2	2	1	1	1	
8	黄福全	夫妻	冯少洋	女	.053		打工	2	2	2	2	1	1	
9	陈志平	父子	陈镜华	男	1.00	49	务农	2	2	3	1	3	3	
10	李瑞华	兄妹	刘桂清	女	.023		待业	2	2	3	2	3	3	
11	梁连锋	爱人	梁焕婷	女	.023		自由职业	2	2	3	3	3	3	
12	孙亚宁	夫妻	郭小凡	女	.029		葵园	2	2	3	1	1	3	
13	张泉星	夫妻	张泽洪	女	.040		务农	3	2	3	2	3	3	
14	何文光	父子	何锡金	男	1.00	57	保安	3	2	3	2	3	3	
15	何志明	夫妻	郭桂莲	女	.042		务农	3	2	3	2	3	3	
16	李邵华	夫妻	袁叠云	女	.058		退休	3	2	3	2	3	3	
17	郭满堂	夫妻	卢桂英	女	.037		个体户	3	2	3	2	3	3	
18	陈志洪	夫妻	郭艳梅	女	.037		家庭主妇	3	3	3	2	3	3	
19	陈国辉	兄弟	陈锋	男	1.00	29	物业管理员	3	2	3	2	3	3	
20	何锡洪	夫妻	樊维嘉	女	.048		务农	3	1	3	2	3	3	
21	陈胜楠	夫妻	张少泉	女	.041		务农	3	3	3	2	3	3	
22	刘神义	父子	刘映山	男	1.00	59	个体户	3	2	3	2	3	3	
23	张海萍	父子	张执金	男	1.00	59	务农	3	2	3	2	3	3	
24	郑达	夫妻	郑春	女	.051		家庭妇女	3	3	3	2	3	3	
25	陈锐番	父子	陈均良	男	1.00	52	务农	3	3	3	2	3	3	
26	梁柏辉	父子	梁贵照	男	1.00	29	务农	3	3	3	2	3	3	
27	曾伟苹	母子	郭文锋	男	1.00	18	学生	3	3	3	2	3	3	
28	袁均涛	兄弟	袁锐尧	男	1.00	35	雇员	3	3	3	2	3	3	
29	郭杰财	夫妻	范孝凤	女	.027		工人	3	3	3	2	3	3	
30	张志勋	夫妻	黄瑞娟	女	.037		医生	3	2	3	2	3	3	
31	陈锦洪	夫妻	黎桂月	女	.042		务农	3	2	3	2	3	3	
32	刘锡良	夫妻	陈妹	女	.051		务农	3	2	3	2	3	3	

第四节 项目分析

- (一) 项目的难度
- (二) 项目区分度
- (三) 项目难度与区分度的关系
- (四) 项目分析的特殊问题



第四节 项目分析

□一、项目分析的定义：

□ 指根据被试的反应，对组成测验的各个题目（项目）进行分析，包括定性分析和定量分析，帮助筛选和修订项目。

- 定性分析：涉及项目编写的恰当性、有效性以及表达是否清楚。
- 定量分析：主要指对项目的难度和区分度的分析。

第四节 项目分析

□ 二、项目的难度（多针对能力测验而言）

□ （一）定义

- 难度就是指项目的难易程度。通常以通过率表示，即以答对或通过该题目的人数的百分比来表示。



第四节 项目分析

□ (二) 难度的计算

- 1、二分法记分项目的直接算法
- 2、多项选择项目的校正算法
- 3、非二分法计分项目的平均分算法
- 4、用极端分组法计算难度

第四节 项目分析

□ 1、二分法记分项目的直接算法

□ 以答对的百分比（或称通过率）来计算难度

□ 对0、1两级记分的测题：

$$P=R/N \times 100\%$$

P代表难度，N代表全体被试人数，R为答对或通过该题的人数。

□ 例，100人参加某一测验，其中第2题和第6题通过的人数分别为20人和30人，求第2题和第6题的难度。

第四节 项目分析

□ 2、多项选择项目的校正算法

$$\square \quad CP = (KP - 1) / (K - 1)$$

CP是校正后的通过率、P为实得通过率、K为备选答案的数目。

□ 比如P为75%

□ 5个备选时，CP为69%；

□ 4个备选时，CP为67%；

□ 3个备选时，CP为63%；

□ 2个备选时，CP为54%。

□ 例，一个五择一的测题，未校正前的难度指数为0.50，一个四择一的测题，未校正前的难度指数为0.53，请问校正后哪一题的难度大？

第四节 项目分析

□ 3、非二分法计分项目的平均分计算法

□ $P = X / X_{\max} \times 100\%$

□ X 为全体受试者在该题目上的平均分数， X_{\max} 为该题的满分。

□ 例，某语文测验第五题的满分为12分，这道题考生的平均得分为8.5分，问该题的难度是多少？

第四节 项目分析

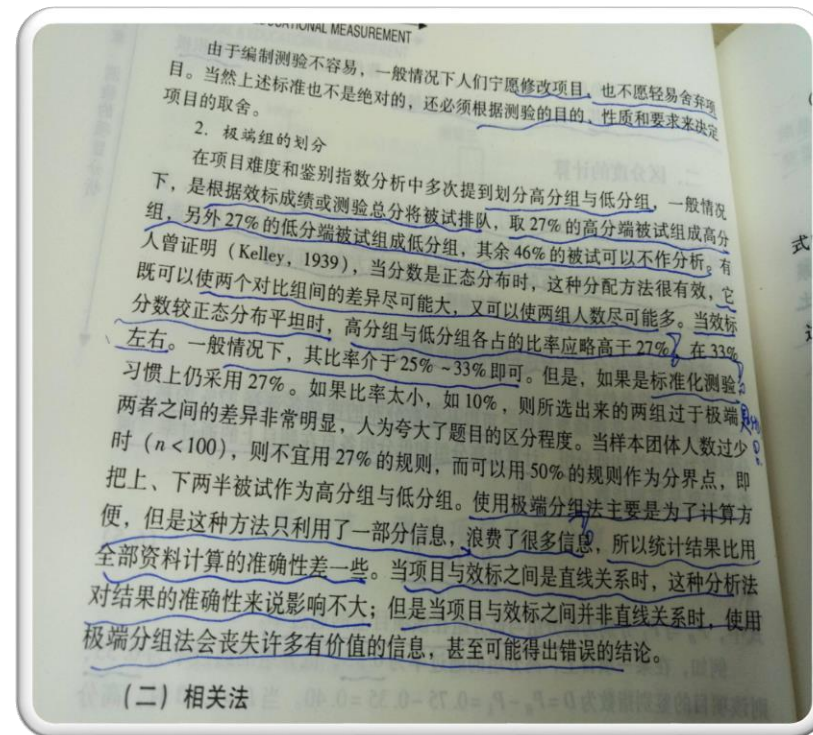
□ 4、用极端分组法计算难度

□ 具体步骤如下：

- 第一步：先依测验总分的高低，按由高到低的顺序，依次排列试卷；
- 第二步：当效标分数比正态分布平坦时，从得分最高的一份试卷开始依次向下选出全部试卷的27%，作为高分组
- 第三步：从得分最低的一份试卷开始依次向上选出全部试卷的27%，作为低分组
- 第四步：按下列公式计算难度

$$p = \frac{P_H + P_L}{2}$$

- P 为难度； P_H 为高分组在该项目上的通过率； P_L 为低分组在该项目上的通过率。



第四节 项目分析

- (三) 难度水平的确定
- 根据测验的目的、项目的形式和测验的性质来确定。
- 1、学业考试：
 - 目的是教学内容掌握的情况，所以只要认为内容是重要的，就可以选择，而不论难度。



第四节 项目分析

□ 2、人员选择:

- 采用难度值接近录取率的项目。录取率为40%，则理想的难度在0.40左右。
- 0.50左右难度的题目，可以对全体被试作最大程度的区分。

□ 3、对于选择题

- 难度值要大于概率。
- 如难度值等于或小于概率，如鉴别力好，通过者都是水平最高的，也可保留。

第四节 项目分析

□ (四) 难度水平与分数的分布关系

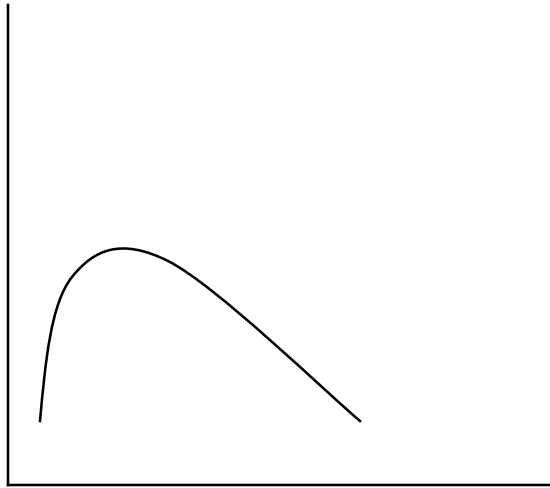
□ 1、正态分布：平均难度为0.50，即大多数被试通过50%的题目。

□ 2、偏态分布

➤ 正偏态：分数密集在低分区段，表明整体难度下限太高，缺乏足够的较易的项目。

➤ 负偏态：分数集中在高分区段，表明难度上限太低，缺乏足够的难题，所以多数人都得高分。

□ **注：并非所有测验都要求正态分布，有些允许出现偏态，如标准参照测验的分数就常常是偏态，可根据实际需要来确定。**



- 正偏 负偏
- 当测验的分数（预试结果）呈偏态分布时，可以通过改变项目难度的比例对测验的难度加以调整，通常，如果被试的取样具有代表性，对于中等难度的测验，其分数的分布呈现正态分布。

第四节 项目分析

□（五）项目难度的等距量表（难度转换）

□1、使用项目难度等距量表的理由

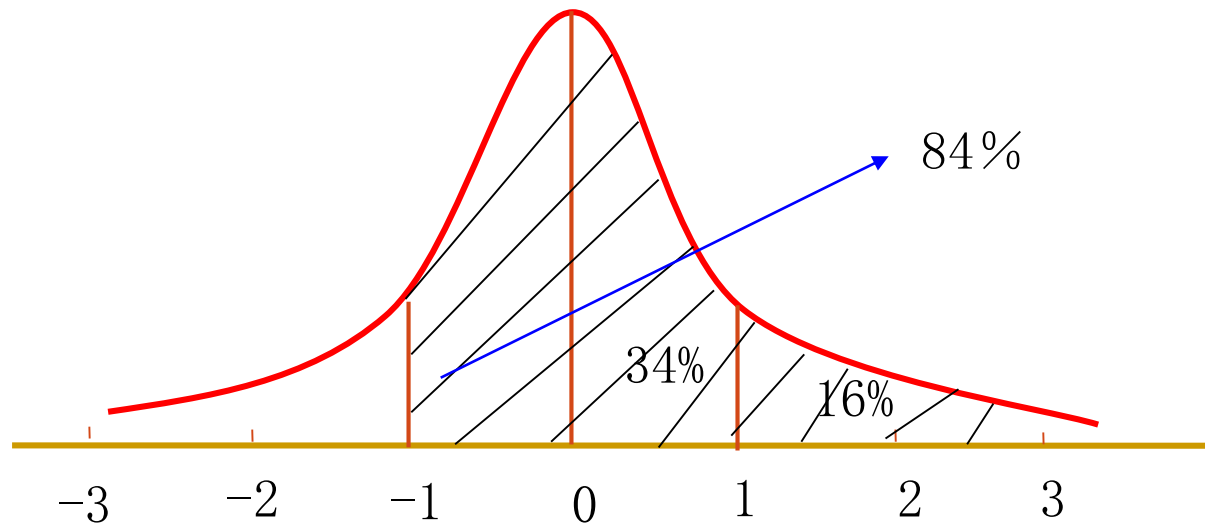
- 前面我们计算的难度是用百分比来表示难度。
- 百分量表是等级量表，不是等距量表，缺乏相等的单位，只能表示事物之间位次的关系，不能用来对项目之间的难度作精确的比较，也不能由各试卷的难度来计算整份试卷的平均难度。
- 例如，1、2、3三个项目的难度指数分别为0.50、0.40、0.30

第四节 项目分析

□2、方法

- 先将某题的通过率视作正态曲线下的面积，然后查正态分布表找出正态曲线右侧的面积所对应的 z 分数。这一 z 分数就可视为该题的标准难度。

例，某题的通过率为84%，求该题的标准难度。



第四节 项目分析

□ 3、标准难度指数

□ Δ (delta) 量表标准难度 (美国教育测验服务中心 (ETS))。 Δ 值可以用以下公式计算: $\Delta = 13 + 4Z$

□ 希腊字母 Δ 表示标准化难度指数, Z 为标准分数。 Δ 值是一个均数为 13, 标准差为 4 的正态分布。

□ 例 求难度指数为 0.50、0.40、0.30 的标准难度指数 (Δ 值)

□ Δ 的取值范围:

□ 标准正态分布的全距一般包括 6 个标准差的距离, 即从 -3 到 +3, 当 $Z = -3$ 时, 表明通过率达到 99.87%, 这种情况极少发生, 此时 $\Delta = 1$; 当 $Z = 3$ 时, 表明通过率为 0.13%, 也是不大可能出现的情况, 此时 $\Delta = 25$ 。 Δ 的取值范围是 1~25。

第四节 项目分析

□ (六) 测验难度与信度的关系

- 过难和过易的测验，会使测验分数的分布相对集中于高分段和低分段，从而缩小分数的分布范围。
- 分数分布范围影响信度，范围越广，信度越高，分数分布范围越小，信度越小。
- 1965年，艾伯尔用三套测验进行研究，发现当难度集中在0.50附近时，分数分布的范围较广。

第四节 项目分析

□ 总结：

- 一般能力测验和成就测验的平均难度在0.50左右为宜。出现偏态情况时，宜对项目进行调整，以使测验分数的分布接近正态。
- 但项目难度还与测验的目的有关，正偏态分布适合于筛选性测验（如选拔性测验、竞争性测验），达标考试属于负偏态分布。

第四节 项目分析

□ 三、区分度

□ (一) 定义:

□ 指测验项目对所测量的心理特性的区分程度或鉴别能力。

□ 确定区分度的方法包括鉴别指数法和相关法。

□ 项目鉴别指数法

□ 相关分析法

□ 点二列相关(Point biserial correlation)

□ 二列相关(biserial correlation)

□ 积差相关

积差相关法	连续变量	连续变量
点二列相关	连续变量	二分称名变量
二列相关	连续变量	人为划分的两个类别

第四节 项目分析

□ (二) 鉴别指数法

□ 1、定义

□ 当效标测量是一个连续变量时，可从分布的两端选出高分组（H）和低分组（L），然后将高分组和低分组通过每个项目的人数用百分比来表示，这两个百分数之间的差异就提供了题目区分度的指标。

□ $D = P_H - P_L$

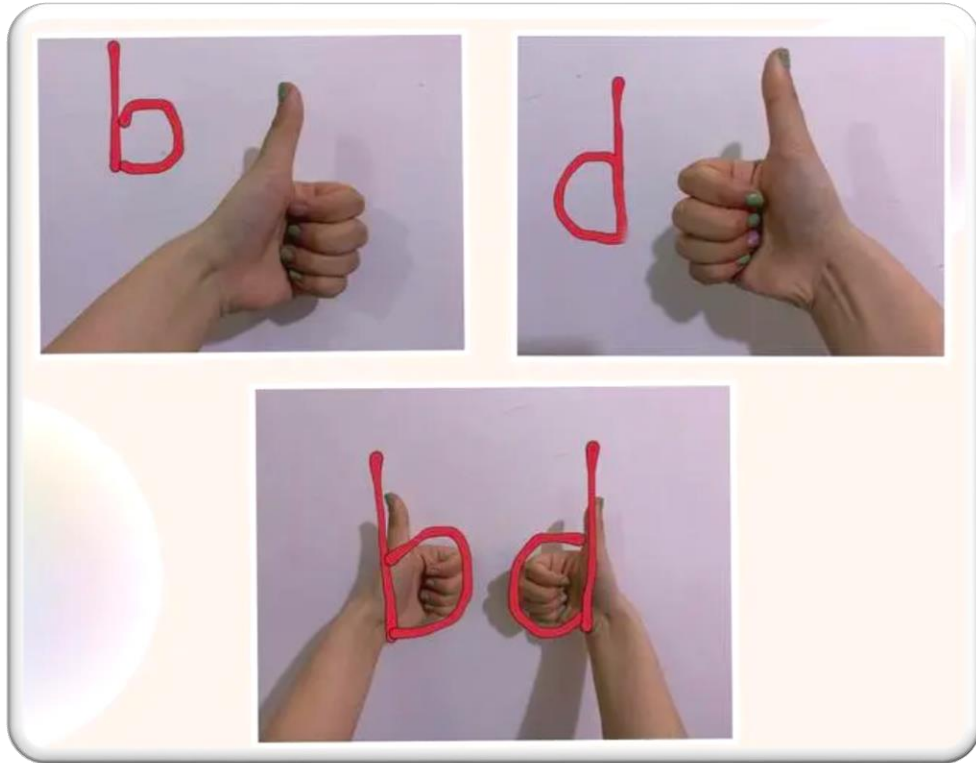
□ D为鉴别指数

□ P_H 为高分组某项目通过人数百分比

□ P_L 低分组某项目通过人数百分比。

□ D值越大，项目区分度越高，项目越有效。

第四节 项目分析



- 例，某高中物理测验，被试共18人，高分组和低分组各取总人数的27%，则两组各为5人，第五题高分组5人全部答对，低分组只有1人答对，计算该题的鉴别指数。

第四节 项目分析

□2、L. Ebel提出的鉴别指数标准

□鉴别指数（D） 题目评价

- 0.40 非常优良
- 0.30-0.39 良好，如能改进更佳
- 0.20-0.29 尚可，仍须修改
- 0.19以下 劣等，必须淘汰

□ 口诀：见（鉴）到妖怪，一舅逃（淘），二舅伤（尚），三舅凉。



第四节 项目分析

□ (三) 项目与效标或总分的相关

□ 1、二列相关

□ 测验总分或效标分数、某个测验项目的分数都是连续变量，其中一个变量被人为地分成两类，可以是测验总分或效标分数被人为地分成两类，也可以是某个项目的分数被人为地分成两类。

- (1) 项目分数连续，效标或测验总分，分为高低或及格不及格。
- (2) 效标连续，项目分数分为对、错或通过、不通过。

第四节 项目分析

$$r_b = (\overline{X_p} - \overline{X_q}) \times pq / (S_t \times y)$$

$$r_b = (\overline{X_p} - \overline{X_t}) \times p / (S_t \times y)$$

- $\overline{X_p}$ 为二分变量中高级组x值的平均数
- $\overline{X_q}$ 二分变量低级组x值的平均数
- $\overline{X_t}$ 为所有x的平均数
- S_t 所有x的标准差
- p为高分组人数与总人数之比
- q为低分组人数与总人数之比
- y是p与q交界处正态曲线的高度

第四节 项目分析

- 例题：20个学生语文测验总分和一个作文题、一个选择题上的得分情况。如果作文分37分以上（包括37）算通过，计算作文分与总分的相关。

❖ 学生	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
❖ 总分	86	52	94	72	65	22	76	83	80	75	76	73	62	91	47	74	81	88	62	58
❖ 作文	47	37	55	27	22	10	35	42	46	39	40	41	38	52	21	39	42	48	29	27
❖	1	1	1	0	0	0	0	1	1	1	1	1	1	1	0	1	1	1	0	0
❖ 选题	1	0	0	1	1	0	0	1	1	1	1	0	1	1	0	1	1	0	0	0

- $\bar{X}_p = (86 + 52 + 83 \dots + 74 + 81 + 88) / 13 = 78.08$
- $\bar{X}_q = (72 + 65 + 22 + 76 + 47 + 62 + 58) / 7 = 57.43$
- $p = 13 / 20 = 0.65$
- $q = 1 - 0.65(7 / 20) = 0.35$
- $\sum X = 1417 \quad \sum X^2 = 105947$
- $S_t^2 = (\sum X^2 / N) - (\sum X / N)^2 = 105947 / 20 - (1417 / 20)^2 = 277.63$
- $S_t = 16.66$
- 当 $p = 0.65$ 时，查常态曲线表得
- $y = 0.3704$
- $r_b = ((78.08 - 57.43) / 16.66) \times (0.65 \times 0.35) / 0.3704 = 0.76$

□ 二列相关系数 r_b 的显著性考验用下面公式:

□ $Z = r_b / \sqrt{((1/y) \times pq/N)}$ (7.8)

□ 代入公式 (7.8) 得:

□ $Z = 0.76 / \sqrt{(1/0.37) \times 0.65 \times 0.35 / 20} = 2.6$

□ 查正态曲线表, 如果 Z 值大于1.96, 即为相关显著。

□ 作文与测验总分有显著相关。

第四节 项目分析

□ 注意：

- 计算二列相关系数时，要求二分变量的分布在连续测量时必须为正态的，即使样本分布不太正态，总体分布也应是正态。
- 连续变量的分布虽不要求是正态，但须是单峰和相当对称。
- 当两个变量均为连续变量时，计算积差相关更好。二列相关可靠性较积差相关差，并要求大样本。
- 当一个连续变量分布不够规则，如在量表值的某一极限外没有个案或组别甚少，而它们在量表上的分布距离是否相等值得怀疑时，不能计算积差相关。此时可以将这个变量分为两个类别计算二列相关作积差相关的估计值。

第四节 项目分析

□ 2、点二列相关

□ (1) 适用性：适用于一个变量是连续变量，另一个是二分变量的资料。

□ 例如，某项目答对记1分，答错记0分，则此项目为二分变量，而总分为连续变量。

□ (2) 计算公式：

$$\square r_{pb} = ((\bar{X}_p - \bar{X}_q) / S_t) \times \sqrt{pq}$$

第四节 项目分析

学生	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
总分	86	52	94	72	65	22	76	83	80	75	76	73	62	91	47	74	81	88	62	58
作文	47	37	55	27	22	10	35	42	46	39	40	41	38	52	21	39	42	48	29	27
选题	1	0	0	1	1	0	0	1	1	1	1	0	1	1	0	1	1	0	0	0

$$\square X_p = (86 + 72 + 65 + 83 + 80 + 75 + 76 + 62 + 91 + 74 + 81) / 11 = 76.82$$

$$\square X_q = (52 + 94 + 22 + 76 + 73 + 47 + 88 + 62 + 58) / 9 = 63.56$$

$$\square p = 11 / 20 = 0.55 \quad q = 1 - 0.55 = 0.45 \quad St = 16.66$$

$$\square r_{pb} = ((76.82 - 63.56) / 16.66) \times \sqrt{0.55 \times 0.45} = 0.396$$

第四节 项目分析

□ 3、积差相关

- 适用资料：两列数据都是测量的数据；两列变量各自总体的分布都呈正态，即正态双变量。
- 例：下表是30个学生一次语文测验的分数，试计算作文题的区分度。

第四节 项目分析

学生	1	2	3	4	5	6	7	8	9	10
测验总分	78	65	66	81	55	48	82	79	83	88
作文题得分	41	34	30	38	26	20	43	44	41	42
学生	11	12	13	14	15	16	17	18	19	20
测验总分	55	57	87	68	55	46	67	77	72	73
作文题得分	23	26	41	31	30	25	26	31	32	33
学生	21	22	23	24	25	26	27	28	29	30
测验总分	78	77	74	84	42	61	45	63	66	61
作文题得分	39	36	32	43	19	33	18	33	35	30

第四节 项目分析

❖ 4、 ϕ 相关

- ❖ (1) 适用性：适用两个变量都是二分变量。一些连续变量也可以用此方法计算相关系数。 ϕ 相关不要求变量呈正态分布。
- ❖ 用该系数作为区分度指标时，要求题目反应和效标量都是二分变量。
- ❖ (2) 公式：

$$r_{\phi} = \frac{AD - BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

第四节 项目分析

❖ 例题：45名考生考取大学的人数和在某题上通过的人数见下表：

❖ 升学情况

❖

	考取	未考取	合计
通过	13 (A)	7 (B)	20
未通过	5 (C)	20 (D)	25
合计	18	27	45

第四节 项目分析

❖ 代入公式:

$$\text{❖ } r_{\phi} = (13 \times 20 - 7 \times 5) / \sqrt{20 \times 25 \times 18 \times 27} = 0.456$$

❖ 通过 r_{ϕ} 和 X^2 的关系计算: $X^2 = N \times r_{\phi}^2 = 45 \times 0.456^2 = 9.375$

❖ $\alpha = 0.05$, $X^2 = 7.88$, 所以显著。

- 练习1：下表是30个学生一次测验的分数，试计算第三题的难度；若第三题是是非题，答案分对和错，请计算其区分度；若第三题是作文题，人为地将其分成通过和不通过，请计算其区分度。

学生	1	2	3	4	5	6	7	8	9	10
测验总分	85	73	62	98	76	35	42	76	84	85
第3题得分	1	0	0	1	1	1	0	0	1	1
学生	11	12	13	14	15	16	17	18	19	20
测验总分	96	94	53	82	79	64	53	84	47	65
第3题得分	1	1	0	1	0	1	1	1	0	1
学生	21	22	23	24	25	26	27	28	29	30
测验总分	51	64	97	32	46	88	81	76	53	32
第3题得分	0	1	1	1	0	1	1	0	1	1

- 练习2：下表是某大学学生的一次普通心理学期末考试的成绩总分和某一论述题的得分（满分为15分），请计算这一论述题的难度（分别用基本公式和极端分组法计算。）和区分度。

学生	1	2	3	4	5	6	7	8	9	10
测验总分	88	86	85	81	77	62	54	53	56	62
论述题得分	12	11	10	9	8	7	6	5	5	5
学生	11	12	13	14	15	16	17	18	19	20
测验总分	68	91	52	68	64	63	78	77	74	79
论述题得分	6	12	5	6	6	6	7	7	7	7
学生	21	22	23	24	25	26	27	28	29	30
测验总分	71	58	65	36	56	69	84	85	87	88
论述题得分	7	5	5	4	5	6	8	8	8	8

第四节 项目分析

□ 四、区分度和难度的关系

- 1、依据连续变量求得的相关值 (r_t 、 r_b) (二列相关, 积差相关), 这些指标不受难度影响。
- 2、 r_Φ (Φ 相关)、 r_{pb} (点二列相关)和鉴别指数D偏好中等难度水平, 即中等难度时这些指标最好, 区分度最好。

第四节 项目分析

□ 3、项目难度与鉴别指数D的关系

□ 项目通过率P 鉴别指数D

□ 1.00	0.00
□ .90	.20
□ .70	.60
□ .50	1.00
□ .30	.60
□ .10	.20
□ 0.0	0.00

□ 4、并非要全部项目难度都是.50，实际需要测验的平均难度在.50就最好。

第四节 项目分析

□ 五、项目分析的特殊问题

□ （一）备选答案的反应模式

□ 多重选择题依然需要分析被试对各个备选答案的反应模式。

□ 1、正确的备选答案被所有受试选择，说明题目太容易或题目可能提供了某种暗示。

□ 2、错误答案没有一个被试选择，说明该选项不具迷惑性，错得过于明显。最少要有2%以上的人选择，否则这个备选答案要去掉。

第四节 项目分析

- 3、如果被试都选择了同一错误答案，可能是编制测验时把正确答案定错了
- 4、高分组被试的选择集中在两个答案上，二者选择率相近，说明该题可能有两个正确答案或在某种意义上另一个选择也有一定的意义。
- 5、高分组对正确答案的选择与低分组相等或低于后者，说明该题所考察的东西与被试水平无关，即不具有辨别力。
- 6、一个题目被试未作答的人数多（速度测验除外）或选择各个备选答案的人数几乎相等，说明该题可能过难或题意不清。

□例，表3—4是一张2道4重选择题的项目分析表。被试共100人，高低分组共有27人。表中两道题的正确答案都是B。

表3—4 2道4重选择题的项目分析表

题号	分组	选项A	选项B	选项C	选项D	未答	P	D
1	高分组	2	24	0	4	0	0.72	0.26
	低分组	1	16	0	8	1		
2	高分组	9	12	4	0	5	0.56	-0.04
	低分组	3	18	4	0	5		

第四节 项目分析

□ (二)标准参照测验的项目分析

□ 以上介绍的是常模参照测验的项目分析方法，不完全适用标准参照测验，需要注意以下方面。

□ 1、难度分析

□ 标准参照测验可以采用常模参照测验的方法计算难度，但对难度水平的要求不同，只要认为项目重要，就可包括在测验中，而不管其通过率和鉴别力如何。

□ 2、区分度分析

□ 标准参照测验不适合用相关法计算区分度，但可以采用类似鉴别指数的方法计算，即比较两组通过率。

第四节 项目分析

□ (三)速度测验的项目分析

□ 1、项目特点：

- 项目容易，多数人都可以通过。
- 项目多，有时限，很少有人做完全部项。

□ 2、通过率和区分度的特征

- 测验前部项目通过率高，后部低。
- 前部区分度低，后部区分度被高估。
- 从速度测验得到的项目分析的指标，主要反映了项目的位置，而不反映项目的难度和区分度。

□ 3、前面介绍的分析方法不适用速度测验。

第五节 项目的合成

□ 一、合成测验

- 合成测验就是把经过预测后证明有价值的项目排成有组织的测验。
- 测验题目选出之后，必须根据测验的目的与性质，并考虑被试作答时的心理反应方式，加以合理安排。如下几点应当是测验编排的一般原则：
 - 1. 测题的难度排列宜逐步上升。
 - 开始是十分容易的题目，以使受测者熟悉作答程序，解除紧张情绪，建立信心，进入测验状态。对试题的总的编排原则是要由易到难。

第五节 项目的合成

□一、合成测验

□2. 尽可能将同类型的测题组合在一起。

□ 这样使每一类型的试题仅需作一次答题说明，也使被试可用相同的反应方式来回答，同时可以简化记分工作和对测验结果的统计分析。

□3. 各种类型测题本身的特点。

□ 如在是非题或选择题中必须避免将选择相同选项的测题安排在一起，以免引起被试的定势反应。

第五节 项目的合成

□ 二、项目排列方式

- **并列直进式**：将整个测验按测验项目材料的性质归为若干分测验，对于同一分测验的测验项目，则依其难度由易到难排列。
- **混合螺旋式**：将各类测验项目依难度分成若干不同的层次，再将不同性质的测验项目予以组合，作交叉式的排列，其难度则渐次上升。

第五节 项目的合成

□ 三、编制复本

□ 为增加实际的效用，有时测验要编制等值的副本。复本要符合下列条件：

- (1) 测量的是同一种心理特质。
- (2) 具有相同的内容和形式。
- (3) 两份测验中不应有重复的项目。
- (4) 两份测验的题目数量要相等，并且有大体相同的难度和区分度。

❖ **A本： 1、 4、 5、 8**

❖ **B本： 2、 3、 6、 7**

第五节 项目的合成

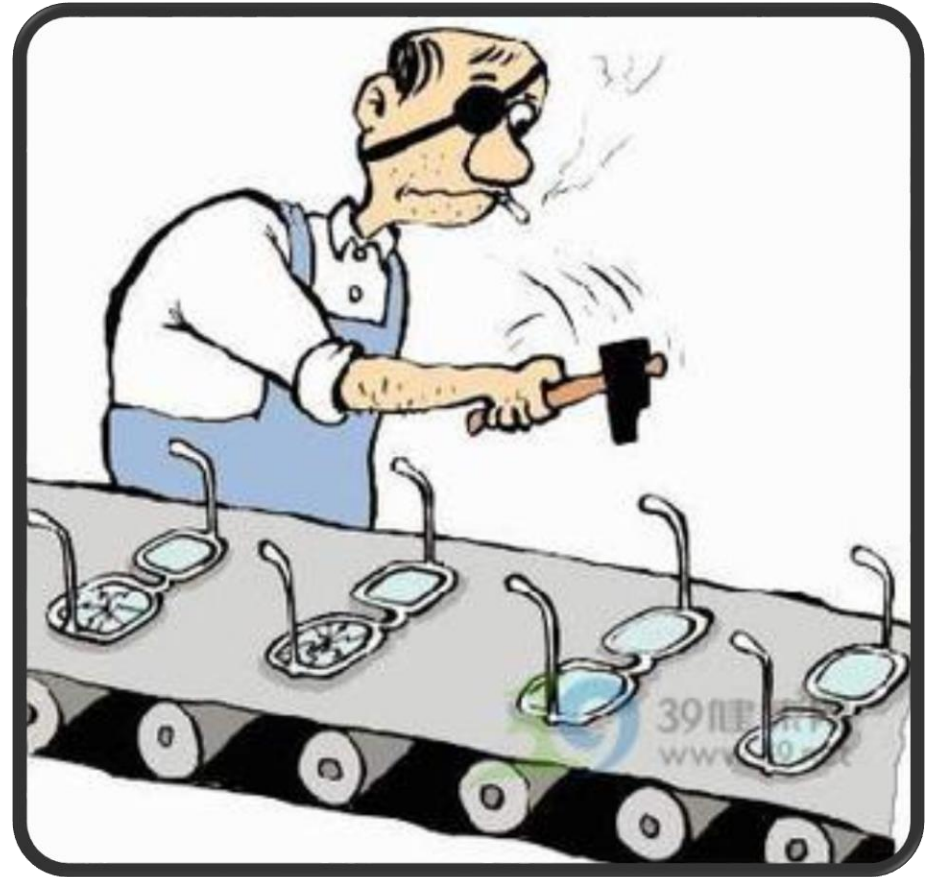


四、测验的解释

- 测验分数必须与某种参照系统比较，方能显出它所代表的意义。
- 多数心理测验是把个人所得的分数与代表一般人同类行为的分数相比较，以判别其所得分数的高低。此处的代表一般人同类行为的分数，即为“常模”。

第六节 常模的制定

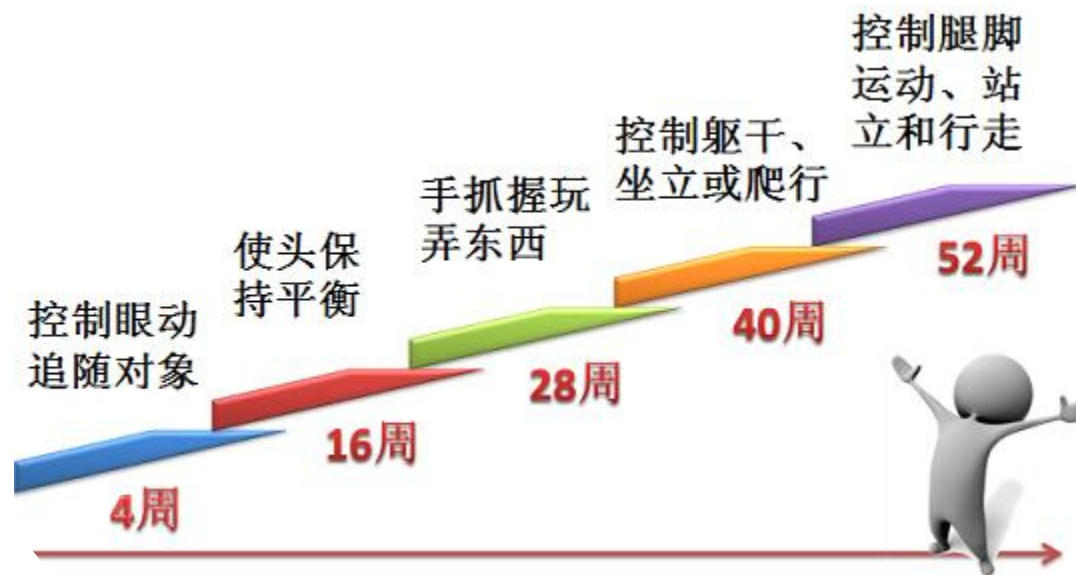
- 建立常模的方法是，在将来要使用测验的全体对象中，选择有代表性的一部分人（称标准化样本），对此样本施测并将所得的分数加以统计整理，得出一个具有代表性的分数分布，此即为该测验的常模。



第六节 常模的制定

- 无论哪种测验，都要参照某个系统对原始分数加以转换，才能作出有意义的解释。
- 常用的心理测验常模有：

- 发展常模
- 百分位常模
- 标准分常模
- 商数
- 等



第七节 测验的标准化

- 标准化，指测验的编制、实施、评分以及解释测验分数的程度的一致性。
- 一、测验内容
- 二、施测过程
- 三、测验评分
- 四、测验分数的解释

第八节 鉴定测验

- 对测验的鉴定，主要就是确定其信度系数与效度系数。
- 一、信度
- 二、效度



第九节 测验说明书

- 测验说明书向测验使用者说明如何使用该测验，以此保证测验的信度和效度。
- 具体应包括以下内容：
 - (1) 本测验的目的与功用
 - (2) 本测验的理论依据
 - (3) 测验内容及实施测验的方法，包括a、何种测验； b、内容几部分； c、每部分有多少测验项目； d、如何作答等。
 - (4) 测验的标准答案以及评分方法。
 - (5) 关于测验的信度、效度资料的说明。
 - (6) 常模表，即如何依据常模解释测验结果。

瑞文标准推理测验 中国城市版

主修者 北京师范大学 张厚粲 王晓平

协作单位与各单位主要负责人

华北区：北京师大 张厚粲 王晓平

天津师大 石德澄 张菁

河北沧州师专 吴长亮

山西大学 章竞思

太原师专 晁连生

东北区：哈尔滨师专 刘庆生

哈尔滨 32 中 秦立群

华东区：上海教科所 顾天楨、高德建

江苏淮阴师范学校 张绍京

福建三明师专 白永莲、黄草木

中南区：河南大学 丁秀峰

河南新乡心理学会 杜希敏、杜兴翼

湖北沙市教科所 陈光浩

西南区：四川教育学院 陈国英

云南教育学院 胡育英、唐忠

西北区：西安交大 宋合义

甘肃教育学院 童长江

谢谢大家！