

第三节 测验的信度

- 第一单元 信度的概念
- 第二单元 信度评估的方法
- 第三单元 信度与测验分数的解释
- 第四单元 影响信度的因素



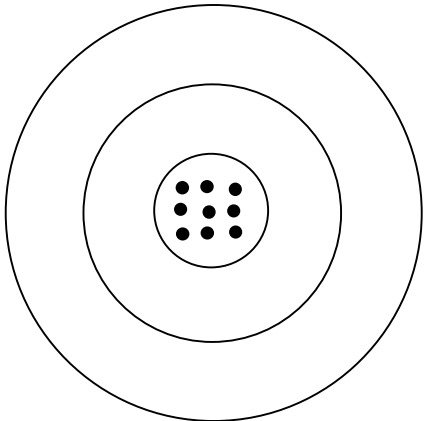
补充内容：

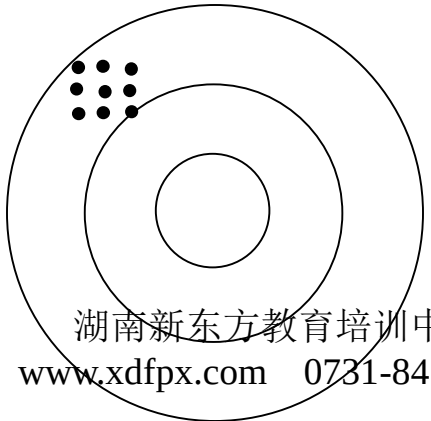
经典测验理论的基本假设

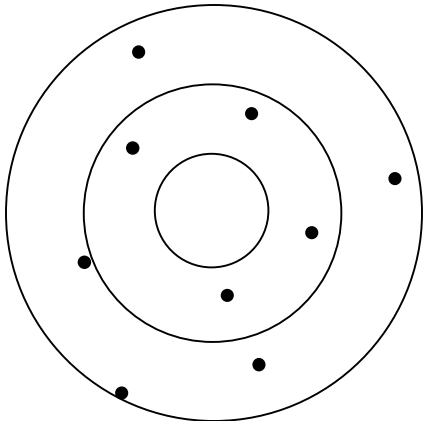


一、误差的含义和种类

- （一）含义
- 测量误差指的是在测量过程中与测量目的无关的变化因素所产生的一种不准确或不一致的测量效应。
- （二）种类　系统误差　随机误差
- 随机误差：由与测量目的无关的、偶然因素引起的而不容易控制的误差。
- 系统误差：由与测量目的无关的变因引起的一种恒定而有规律的误差









下列误差哪些是随机误差?哪些是系统误差？

- 1.水果贩子偷换秤砣来短斤少两
- 2.水果贩子称重时故意快速地耍手法来短斤少两
- 3.在心理测验的过程中，由于被试的身体状况、情绪、动机等因素引起的误差
- 4.不同的老师在作文评分时宽严标准不一，导致评分误差
- 5.智力测验的标准答案出现错误，导致测量误差
- 6.一名老师作文评分时标准过严，导致评分误差



二、误差的来源

（一）测验本身引起的误差

- 取样缺乏代表性，词语模棱两可

（二）施测过程引起的误差

- 1.主试者方面：年龄、性别、外表、动作等
- 2.环境因素
- 3.时间安排
- 4.意外干扰
- 5.记分



二、误差的来源

（三）受测者本身引起的误差

- 应试动机
- 测验焦虑
- 学习和训练
- 测验经验
- 生理状态
- 反应定势



拓展知识（二）

- 1、真分数T：理论值，即没有误差时测得的值，通过无数次抽样（测量），得出平均数；
- 2、任何测验成绩看做是真分数和测量误差的和：X=T+E（只包括随机误差）；
- 3、对于一个团体：实得变异=真分数变异数+误差变异数；$S^2X = S^2T + S^2E$
- 4、真变异数=与测验目的有关的变异（方差）+与测验目的无关的变异（方差）（系统误差）
- $S^2X = S^2V + S^2I + S^2E$



三、真分数

（一）含义

- 测量没有误差时所得到的真值
- 操作性：无数次测量结果得到的平均值

（二）实得分数、真分数与测量误差

（三）性质

- 一组测验分数的变异性是由测量目的有关的变异数、稳定的但出自无关来源的变异数和随机误差变异数所决定的

$$S_x^2 = S_T^2 + S_E^2$$



三、真分数

（四）真分数的数学模型

$$X=T+E$$

（五）三个假设公理

- 若某种心理特质可以用平行测验反复测量足够多次，观察分数就会接近真分数

  $E（X）=T；E（E）=0$

- 真分数与误差分数之间的相关为零

  $\rho（T，E）=0$

- 各平行测验上的误差分数之间的相关为零

  $\rho（E1，E2）=0$

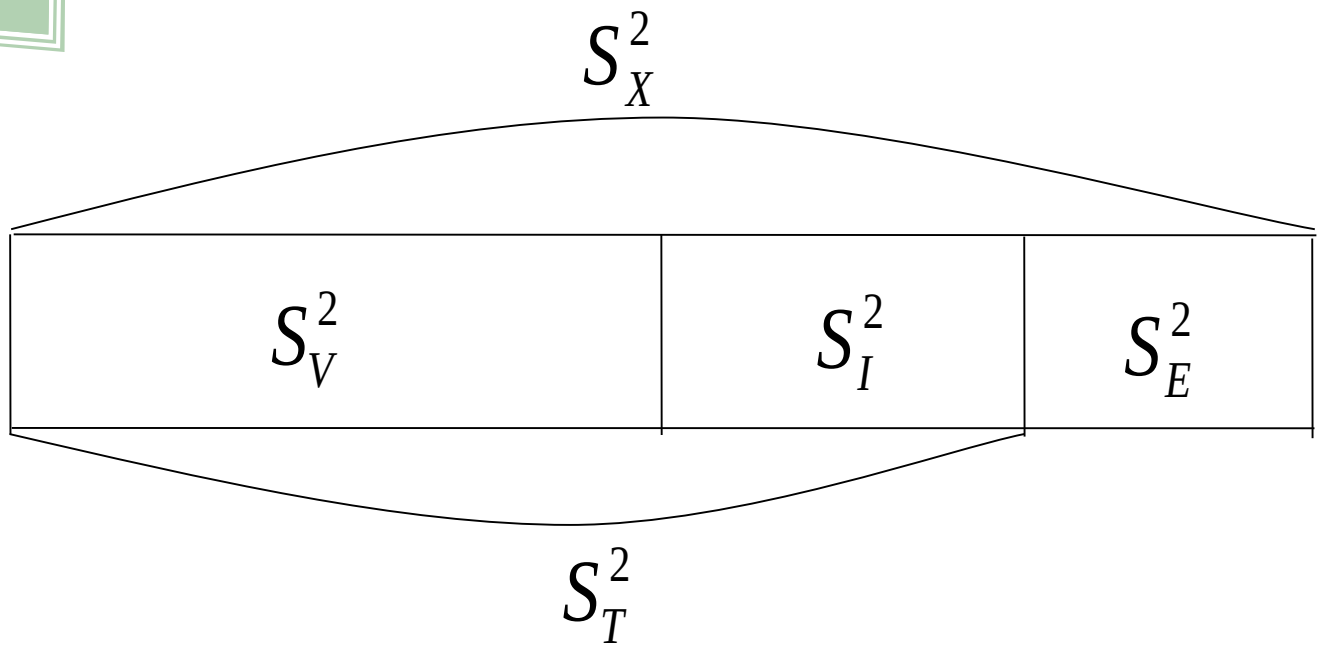


三、真分数

（六）平行测验

- 如果两个题目不同的测验测的是同一种特质，并且题目的形式、数量、难度、区分度以及测查等值团体后所得分数的分布都是一致的，则这两个测验被称为平行测验。





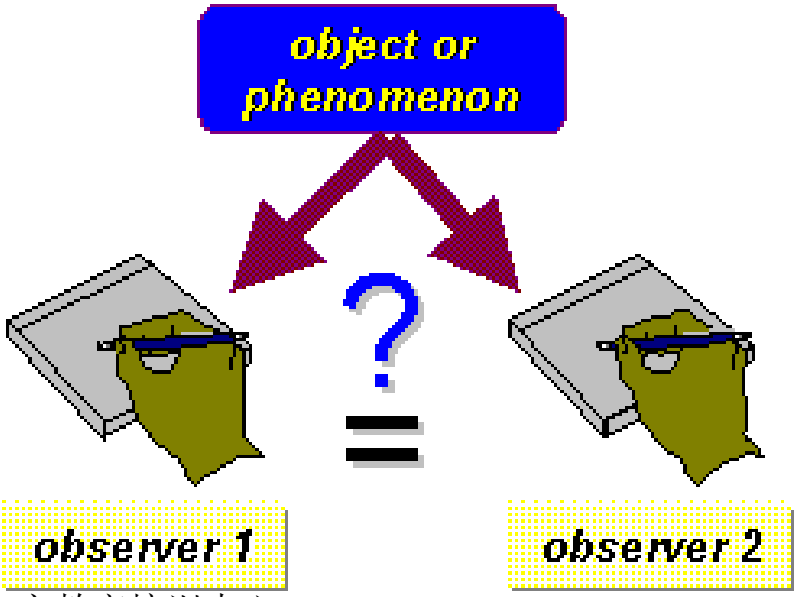
$$S_X^2 = S_V^2 + S_I^2 + S_E^2$$



湖南新东方教育培训中心
www.xdfpx.com 0731-84165117

第一单元 信度的概念

- 一、信度的定义
  - 信度是指同一被试者在不同时间内用同一测验(或用另一套相等的测验)重复测量，所得结果的一致程度。





– 信度只受随机误差的影响。

$$X = T + e$$

Two Components:

- e_r • Random Error
- e_s • Systematic Error

$$X = T + e_r + e_s$$



- 如果我们讨论一组测验分数的特性时，可用方差代表具体分数，得到与公式5-7类似的公式：

$$S_x^2 = S_T^2 + S_E^2$$

- 式中，S_x^2 是测验实得分数的方差，S_T^2 是测验真分数的方差，S_E^2 是测验误差的方差。



- 在测量理论中，信度被定义为:一组测量分数的真分数方差与总方差(实得分数的方差)的比率。即：

- $$r_{XX} = \frac{S_T^{\ 2}}{S_X^{\ 2}}$$ （公式5－9）

- 由于真实分数的方差是无法统计的，因此公式5—9还可以转化为：

$$r_{XX} = 1 - \frac{S_E^{\ 2}}{S_X^{\ 2}}$$

- 因此，信度也可以看作在总的方差中非测量误差的方差所占的比例。



二、信度的指标

- (一)信度系数与信度指数
  - $r_{xx} = r_{XT}^2 = \frac{S_T^2}{S_X^2}$ **(公式5-11)**
  - **信度指数的平方就是信度系数。**
- (二)测量标准误（误差的标准差）

$$SE = S_x \times \sqrt{1 - r_{xx}}$$

  - **信度越高，标准误越小；信度越低，标准误越大。**



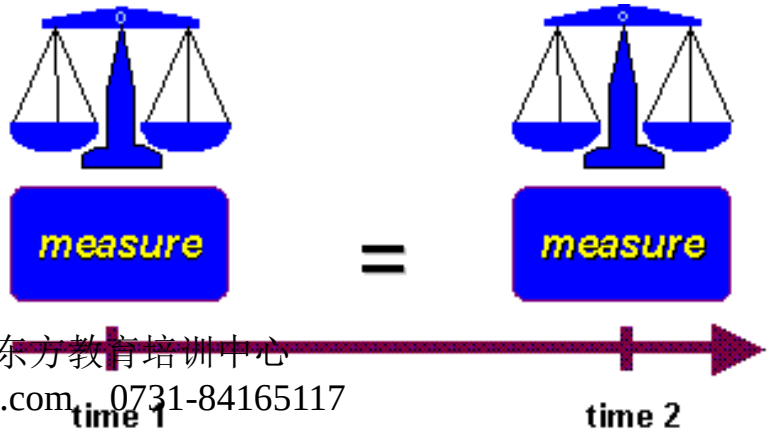
第二单元 信度评估的方法

- 重测信度
- 复本信度
- 内部一致性信度
- 评分者信度



一、重测信度

- 又称稳定性系数。
- 使用同一测验，在同样条件下对同一组被试者前后施测两次测验，求两次得分间的**相关系数**。
- 一般是两周到四周较宜，间隔时间最好不超过六个月。





被试	第一次测验	第二次测验
A	70	77
B	80	72
C	75	83
…	…	…
…	…..	….

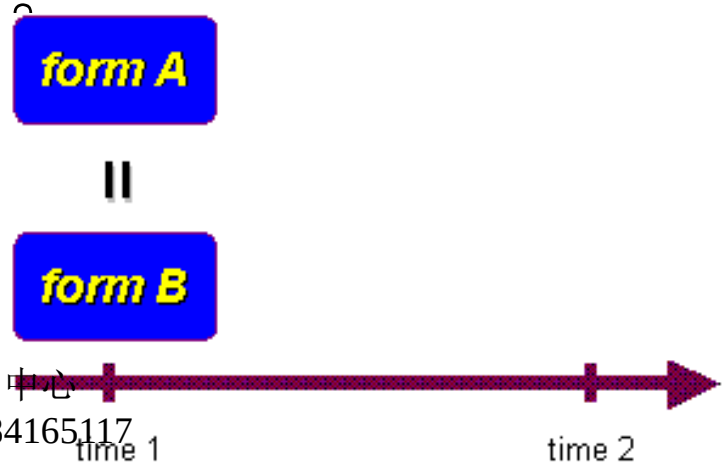


- 优点：在于能提供有关测验是否随时间而变异的资料，可作被试将来的行为依据
- 缺点：易受练习和记忆的影响。



二、复本信度

- 以两个等值但题目不同的测验(复本)来测量同一群体，然后求得被试者在两个测验上得分的相关系数
- 复本信度反映的是测验在内容上的等值性，故又称等值性系数。





- 复本测验要求题目内容、数量、形式、难度、区分度、指导语、时限、例题等相同或相似

被试	测验1	测验 2
A	70	77
B	80	72
C	75	83
…	….	….
…	….	….



A卷：你是一个喜欢热闹的人吗？

1=绝对是

2=是

3=不一定

4=不是

5=绝对不是



1=绝对是

2=是

3=不一定

4=不是

5=绝对不是

B卷：你是一个喜欢安静的人吗？

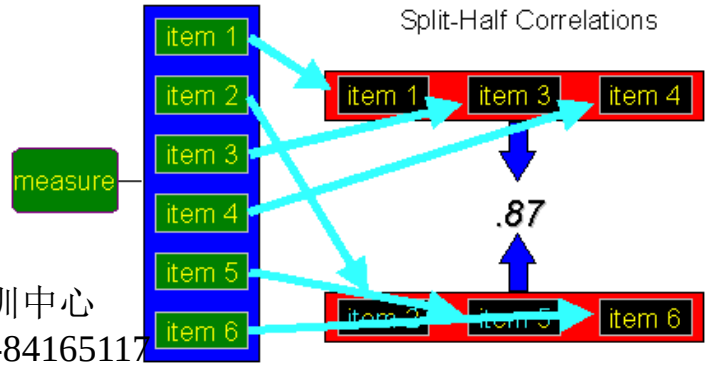
BACK



- 优点：能够避免重测信度的一些问题。
- 缺点：1.如果测量行为受影响，复本信度只能减少而不能完全消除。
- 2.由于第二个测验只改变了题目内容，已经掌握的解题原则，可以很容易地迁移到同类问题。

三、内部一致性信度

- (一)分半信度(split－half reliability)
  - 这个相关系数就代表了两半测验内容取样的一致程度
- $$r_{xx} = \frac{2 \times r_{hh}}{1 + r_{hh}}$$ (公式5-14)
  - 式中为 r_{hh} 一半分数的相关系数，r_{xx} 为测验在原长度时的信度估计值。





How to divide the test?

- Random（随机法）
- odd-even（奇偶法）
- 1st half of items & last half of items（顺序法）

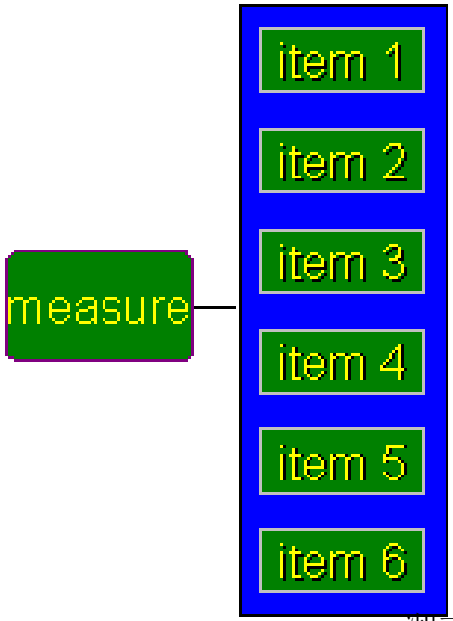


- (二)同质性信度(homogeneity reliability)
  - 代表测验内部所有题目间的一致性。
    - （1）指所有题目都测的是同一种心理特质。
    - （2）指所有题目得分之间都具有较高的正相关。
- 同质性信度就是一个测验所测内容或特质的相同程度。

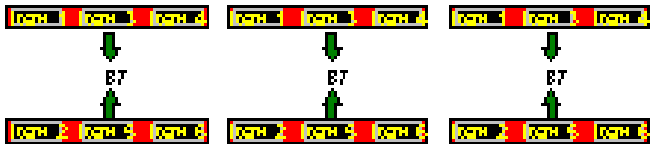


➢ 克伦巴赫α系数

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\Sigma S_i^{\,2}}{S_X^{\,2}}\right)$$



Cronbach's alpha (α)



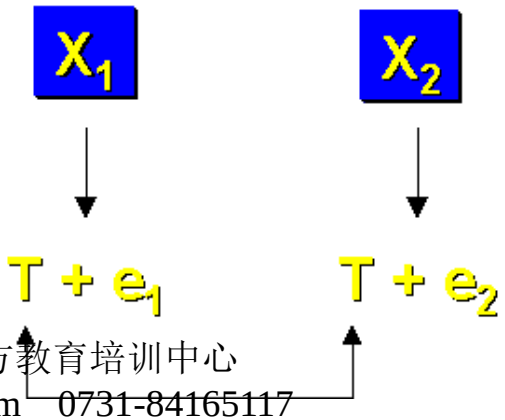
SH_1	.87
SH_2	.85
SH_3	.91
SH_4	.83
SH_5	.86
...	
SH_n	.85

α = .85



四、评分者信度

- 随机抽取若干份测验卷，由两位评分者按评分标准分别给分，然后再根据每份测验卷的两个分数计算相关，即得评分者信度。
  - **一般要求在成对的受过训练的评分者之间平均一致性达0.90以上，才认为评分是客观的。**





斯皮尔曼等级相关

$$r_R = 1 - \frac{6\sum_{i=1}^{N} D_i^2}{N(N^2-1)}$$



评分者信度:肯德尔和谐系数

$$W = \frac{\sum R_i^2 - \frac{(\sum R_i)^2}{N}}{\frac{1}{12}K^2(N^3 - N)}$$

$$W = \frac{\sum_{i=1}^{N} R_i^2 - \frac{(\sum R_i)^2}{N}}{\frac{1}{12}K^2(N^3 - N) - K\sum T}$$



被评者

评定者	一	二	三	四	五	六
A	4	1	3	5	2	6
B	2	3	4	5	1	6
C	3	2	5	4	1	6
R	9	6	12	14	4	18

公式中　K=3（评定者数）

N=6（被评者数）

W＝0.86



第三单元 信度与测验分数的解释

- 解释真实分数与实得分数的相关
- 确定信度可以接受的水平
- 解释个人分数的意义
- 比较不同测验分数的差异



一、解释真分数与实得分数的相关

- **信度系数可以解释为测验的总变异中，真分数造成的变异占百分之几。**
- 在极端的情况下，如有r_{xx}＝1.00，则表示完全没有测量误差，所有的变异均来自真实分数；若有r_{xx}＝0，则所有的变异和差别都反映的是测量误差。



一组测验分数的真实分数方差与实得分数方差的比(真分数变异/实得分数变异)

$$r_{xx} = S_T^2/S_x^2 = 1 - S_E^2/S_x^2$$

$$0 < \text{信度系数} = \frac{\text{真实分数的方差}}{\text{总分数的方差}} < 1$$

测验全是误差

测验没有误差

说明：（1）方差即为标准差的平方



二、确定信度可以接受的水平

- 一般原则是：
  - 当r_{xx}<0.70时，测验不能用于对个人做出评价或预测，而且不能作团体比较；
  - 当0.70≤r_{xx}＜0.85时，可用于团体比较；
  - 当r_{xx}≥0.85时，才能用来鉴别或预测个人成绩或作为。



三、解释个人分数的意义

- 两个作用
  - 其一是估计真实分数的范围；
  - 其二是了解实得分数再测时可能的变化情形。
- 人们一般采用95%的或然率（概率）水平，其置信区间为：
  -

(公式5-16)

$$X - 1.96 \times Se \le X_T \le X + 1.96 \times Se$$

$$SE = S_X \sqrt{1 - r_{xx}}$$



例如：在一次测验中有些学生得80分，这是否反映了他们的真实水平？如果再测一次他们的分数将改变多少？已知该次测验的标准差为5，信度系数为0.84

$$SE = 5 \times \sqrt{1 - 0.84} = 2$$

$$T = 80 \pm 1.96 \times 2 = 80 \pm 3.92 = 76.08 \sim 83.92$$

我们可说这些学生的真正分数有95%的可能性落在76与84分之间。即若再测一次，他们的分数低于76、高于84的可能性不超过5%。



四、比较不同测验分数的差异

- 测量标准误和测验信度在评价两个不同测验的分数是否有明显差异时也非常重要。这种比较包括两个人不同分数的差别和同一被试在两个测验上的差别。
- 差异分数的标准误公式为：

$$SEd = S \times \sqrt{2 - r_{xx} - r_{yy}}$$ （公式5－17）

两个测验务必有相同的标准差



某班期末考试，张生语文、数学的成绩转换成T分数（平均数为50、标准差为10）分别为65和70，由此我们可以知道张生的数学比语文考得稍好些，但二者差异是否有意义，仍不清楚。为了说明个人在两种测验上表现的优劣，我们可用“差异的标准误”来检验其差异的显著性，常用的公式如下：

$$SEd = \sqrt{SE_1^2 + SE_2^2}$$

式中SEd为差异的标准误，SE_1、SE_2分别是两组测验分数的标准误，用$SE_1 = S\sqrt{1 - r_{xx}}$和$SE_2 = S\sqrt{1 - r_{yy}}$代入公式

$$SEd = S\sqrt{2 - r_{xx} - r_{yy}}$$



假定此次语文、数学考试的信度系数分别为0.84和0.91，张生的两个分数转化成T分数后，其差异的标准误为

$$SEd = 10 \times \sqrt{2 - .84 - .91} = 5$$

若采取95%的置信区间（即.05显著水平），则张生在这两门课上T分数的差异必须达到或超过$1.96SEd = 1.96 \times 5 = 9.8$；始能认为二者真有差异。因为数学的$T$分数只比语文高5分，所以差异并不显著。



测验类型	信度		
	低	中	高
成套成就测验	0.66	0.92	0.98
学术能力测验	0.56	0.90	0.97
成套倾向性测验	0.26	0.88	0.96
客观人格测验	0.46	0.85	0.97
兴趣测量	0.42	0.84	0.93
态度测量	0.47	0.79	0.98



已考题次：10　　三级：3
二级：7

第四单元　影响信度的因素

影响因素	影响方式
样本特征	1）样本团体异质性越高，越会高估信度。 2）样本团体平均能力会影响信度。
测验长度	1）测验越长,测题的取样越有代表性,结果越可靠。 2）测验越长,受测者猜测的可能性越小,结果越可靠。 ＊测验太长，也会得不偿失
测验难度	1)　难度和信度之间没有简单的对应关系 2)　故，一个标准的测验应根据不同的能力水平报告的测验难度，作为选择的参考
时间间隔	1）只对复本信度和重测信度有影响。 2）时间间隔越短，信度越高，越长，信度越低。



第四单元 影响信度的因素

- 一、样本特征
- 二、测验长度
- 三、测验难度
- 四、时间间隔



一、样本特征

- 信度常用信度系数来表示，信度系数就是相关系数，相关系数受样本是否异质及样本团体平均能力水平的影响。
  - **(一)样本团体异质性的影响**
  - **(二)样本团体平均能力水平的影响**



㈠样本团体异质性的影响

– 任何相关系数都要受到团体中分数分布的范围所影响，而分数范围与样本团体的异质程度有关。若获得信度的取样团体较为异质的话，往往会高估测验的信度，相反则会低估测验的信度。



㈡样本团体平均能力水平的影响

- 测验的信度还受取样不同团体间平均能力水平的不同而不同。这是因为，对于不同水平的团体，题目具有不同的难度，每个题目在难度上的微小差异累计起来便会影响信度。
- 显而易见，每个信度系数都要求有对建立信度系数的团体的描述。在编制测验时，应把常模团体按年龄、性别、文化程度、职业等分为更同质的亚团体，并分别报告每个亚团体的信度系数，这样测验才能适用于各种团体。



二、测验长度

- 测验长度，亦即测验的数量，也是影响信度系数的一个因素。一般来说，在一个测验中增加同质的题目，可以使信度提高。
  - 第一，测验越长，测验的测题取样或内容取样就越有代表性。
  - 第二，测验越长，被试的猜测因素影响就越小。长有时反而会引起被试的疲劳和反感而降低可靠性。

注意：　增加测验长度的效果应遵循报酬递减率原则，测验过长是得不偿失的，有时会引起受测者疲劳和反感从而降低可靠性。



- $K=r_{kk}(1-r_{xx})/r_{xx}(1-r_{kk})$
  - K为改变后的长度与原长度之比
  - r_{xx}为原测验的信度
  - r_{kk}为测验长度是原来的k倍时的信度估计



三、测验难度

- 难度对信度的影响，只存在于某些测验中，如智力测验、成就测验、能力倾向测验等，对于人格测验、兴趣测验、态度量表等不存在难度问题，因为这些测验的题目没有正确或错误答案之分。

在实际情况中，如果某个测验使用范围很广，难度水平通常只适用于中等能力水平的受测者，而对较高水平和较低水平的受测者可能较易或较难，使得分数分布范围缩小，信度降低。因此，一个标准化的测验，应根据不同能力水平报告测验的难度，以作为选择测验的参考。



- 在实际情况下，如果某个测验适用范围很广，其难度水平通常适合于中等能力水平的被试，而对较高水平和较低水平的被试可能较易或较难，使得分数分布范围缩小，信度水平降低。因此一个标准化的测验，应根据不同能力水平报告测验的难度，以作为选择测验的参考。



四、时间间隔

- 时间间隔只对重测信度和重测复本信度有影响。
- 以再测法或复本法求信度，时间间隔越短，其信度系数越大；间隔时间越久，其他变因介入的可能性越大，受外界影响也越大，信度系数便越低。



真题练习

- **200705-52、一般要求在成对的受过训练的评分者之间平均一致性达到（　）以上，才认为评分是客观的。** C
  - **（A）0.70　（B）0.80　（C）0.90　（D）1.00**
- **200705-53、以再测法或复本法求信度，两次测验相隔时间越短，其信度系数（　）。** A
  - **（A）越大　（B）越低　（C）越小　（D）不变**
- **200705-55、信度只受（　）的影响。** B
  - **A）系统误差B）随机误差C）恒定效应D）概化作用**
- **200711- 54、同质性信度主要代表测验内部（　）间的一致性。** C
  - **A）两半测验B）题目与分测验（C）所有题目（D）分测验**
- **200811-59、信度只受(　)的影响。** B
  - **(A）系统误差　(B)随机误差　(C)恒定效应　(D)概化理论**

- **200905-35、以再测法或复本法求信度，两次测验相隔时间越短，其信度系数越（ ）。** A
  - **（A）大 （B）低 （C）小 （D）不确定**
- **200905-77、重测信度即（ ）。** B
  - **（A）等位性系数B）稳定性系数（C）相关系数 D）X系数**
- **200905- 92、关于心理测验，下列说法正确的是（ ）。** ABD
  - **（A）测验越长，测验的测题取样或内容取样越有代表性**
  - **（B）测验越长，被试者受猜测因素的影响就越小**
  - **（C）测验越长，越遵循报酬递增率原则**
  - **（D）测验太长，有时会引起被试者的疲劳和反感而降低可靠性**
- **200911-76、在不同时间内用同一测验重复测量同一被试者，所得结果的一致程度称为( )。** A
  - **(A)信度 (B)效度 (C)难度 (D)区分度**

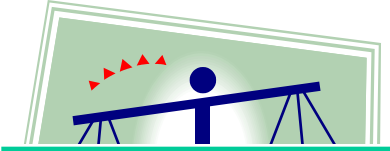


- **20105 -77．同质性信度主要反映测验内部（C ）间的一致性。**
  - **A两半测验 B题目与分测验 C所有题目 D分测验与分测验**
- **201011-61、信度只受（）的影响。P348** B
  - **A系统误差B随机误差C恒定效应D实验误差**



小结：第三节 测验的信度

第一单元 信度的概念
第二单元 信度评估的方法
第三单元 信度与测验分数的解释
第四单元 影响信度的因素



本节考点提要

第三节　信度

第一单元	1、被测在不同时间内用同一测验（或用另一相等的测验）重复测量。所得结果的一致程度。 2、信度只受随机误差的影响。随机误差越大，信度就越低。 3、表示方法：信度系数、信度指数、测量标准误
第二单元	1、信度评估方法：重测又叫稳定性系数、复本又叫等值性系数、内部一性信<分半信度、同质信度>、评分者
第四单元	1、影响信度的因素（样本特征、测验长度、测验难度、时间间隔） 2、样本特征：样本团体异质性影响、样本团体平均能力水平的影响。 3、一个测验中增加同质的题目，可以使信度提高。 4、两次测验的间隔越短，信度系数越大。

湖南新华教育培训中心
www.xdfpx.com　0731-84165117



结果会怎样？

- 用一杆秤来秤一个人的高度
- 用成人智力量表来测儿童的智力
- 妈妈，我的脚现在几点钟？



一个稳定但无效的智力测验

- 1. 你出生在几月份?
- 2. 你母亲姓什么?
- 3. 1+1=?
- 4. 一个星期有几天?
- 5. 下面哪个是三角形?

A. □　　B. ○　　C.▲



一小学生的数学测验

- 想要测验的心理特质：数学能力
- 测验成绩同时受数学能力和语文能力的影响
- 该测验是否有效？



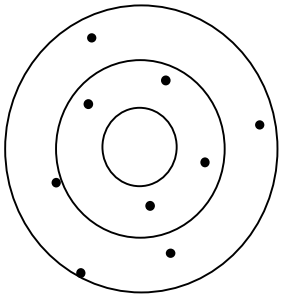
第四节 测验的效度

- 第一单元 效度的概念
- 第二单元 效度评估的方法
- 第三单元 效度的功能
- 第四单元 影响效度的因素

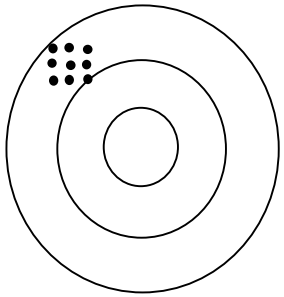


第一单元　效度的概念

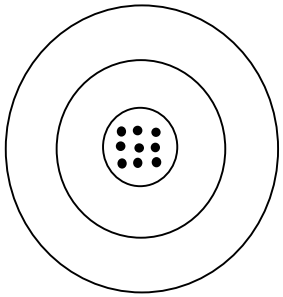
- 效度是指所测量的与所要测量的心理特点之间符合的程度，或者简单地说是指一个心理测验的准确性。
- 测量的效度既受随机误差的影响，同时也受系统误差的影响。
- 可信不一定有效，有效必可信。效度受信度的制约。



信度低
效度也低



信度高
效度低



效度高
信度也高



一、效度的定义

- 效度是指所测量的与所要**测量的心理特点之间符合的程度**，或者简单地说是指一个心理**测验的准确性**。
- 在测验理论中，效度被定义为在一组测量中，**与测量目标有关**的方差(或称有效方差)与总方差的比率，即：

$$r_{xy}^2 = \frac{S_V^2}{S_X^2}$$

(公式5-21)



二、效度的性质

- (一)效度具有相对性
  - 在评鉴测验的效度时，必须考虑其目的与功能。



二、效度的性质

- (二)效度具有连续性
  - 测验效度通常用相关系数表示，它只有程度上的不同，而没有“全有”或“全无”的区别。因此，我们评价一个测验时，不应该说“有效”或“无效”，而应该用效度较高或较低来评价。
  - 另外，效度是针对测验结果的，即测验效度是“测验结果”的有效性程度。

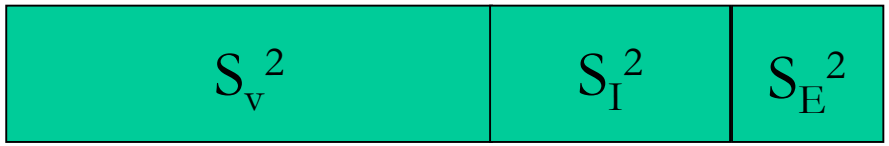


- 3.效度是测量的随机误差和系统误差的综合反映。
- $r^2_{xy}=S^2_v / S^2_x$
- $S_x^2 = S_V^2 + S_I^2 + S_E^2$

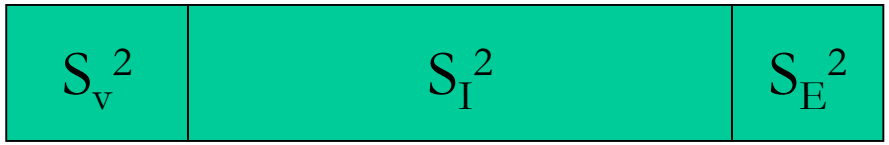


根据公式 $S_X^2 = S_V^2 + S_I^2 + S_E^2$

1. 高信度是高效度的必要而非充分条件



高信度 高效度



高信度 低效度

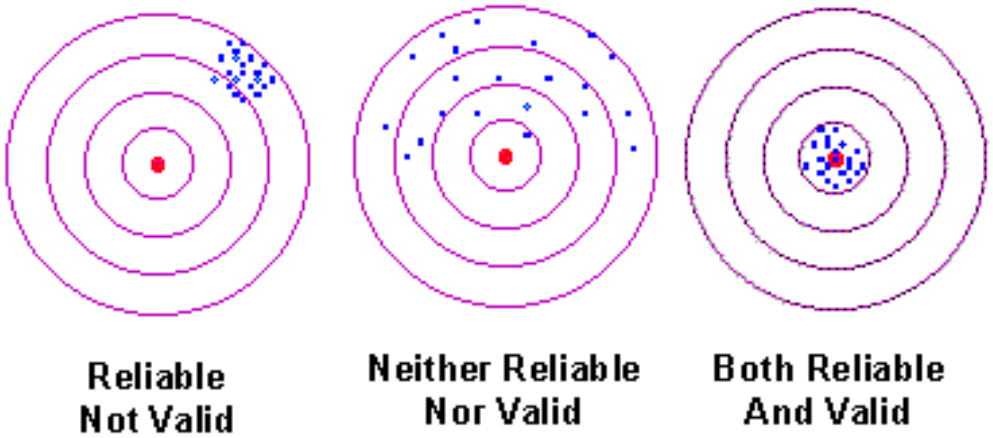


低信度 低效度



信度和效度的关系

- 一、信度是效度的**必要而非充分条件**
- 二、效度是受信度制约的





第二单元 效度评估的方法

- 美国心理学会在1974年所发行的《教育与心理测量之标准》一书中将效度分为三大类：即内容效度、构想效度和效标效度。



- 教师为了了解学生在某一学科或某一课题上对知识的掌握情况，若时间许可，可以进行一个全面的考试，包含所有有关的内容。在实际上这显然行不通。于是，就从这一范围总体中选一样本，也就是从可能的题目中取样来编制测验，根据测验分数推论学生在该范围总体的知识。若测验题目是这个范围的好样本，则推论将有效；若选题有偏差，则推论将无效。
- 由于这种测验的效度主要与测验内容有关，所以叫内容效度。



一、内容效度

- (一)什么是内容效度
  - 内容效度指的是测验题目对有关内容或行为取样的适用性，从而确定测验是否是所欲测量的行为领域的代表性取样。
- 想要编制一个具有较高内容效度的测验需要
  - 对所测心理特征有明确的概念
  - 测验题目应是所界定的内容范围的代表性取样



(二)内容效度的评估方法

- 1. 专家判断法
- 2. 统计分析法
- 3. 经验推测法



1．专家判断法

(1)含义：请有关专家对测验题目与原定内容的符合性做出判断，看测验题目是否代表规定的内容。由于这是一个逻辑分析（定性分析）的过程，也叫逻辑效度。

(2)具体操作步骤：以教育测验为例

测验编制者	专家
①明确欲测内容的范围，包括知识范围和能力要求两个方面 ②将测验目的具体化为不同层次的测验目标 ③确定每一层测验目标在整个测验中的比重； ④编制双向细目表； ⑤确定每一小格中的测题数量； ⑥按比例随机抽取测验题目；	①明确欲测内容的范围，规定好各纲目的比例； ②确定每个题目所测的内容，与双向细目表对照，比较、记录； ③制定评定量表，作总的评价

高职数学科命题细目表

考查内容			了解	理解	掌握	应用	小计	比例
代数	幂函数、指函数、对数函数	题号	①	⑧	⑨	㉙	4 / 15	19 / 60%
		题型	M	M	M	N		
		分值	3	3	3	6		
	三角函数	题号	②		⑩		2 / 6	
		题型	M		M			
		分值	3		3			
	两角和差的三角函数	题号		③	⑪	⑫	3 / 9	
		题型		M	M	M		
		分值		3	3	3		
	不等式	题号	④		㉔		2 / 6	
		题型	H		M			
		分值	3		3			
	数列、极限、数学归纳法	题号	㉕	⑬		㉑	3 / 9	
		题型	H	M		M		
		分值	3	3		3		
	复数	题号		⑭	⑮		2 / 6	
		题型		M	M			
		分值		3	3			
	排列、组合、二项式定理	题号		⑯	㉒	㉓	3 / 9	
		题型		M	M	M		
		分值		3	3	3		
立体几何	直线和平面	题号	⑰		㉚		1.5 / 8	4 / 17%
		题型	M		E			
		分值	3		5			
	多面体和旋转体	题号	⑱	㉖	㉚		2.5 / 9	
		题型	M	H	E			
		分值	3	3	3			
解析几何	直线	题号	⑤	⑲	⑳	㉗	4 / 12	7 / 23%
		题型	M	M	M	H		
		分值	3	3	3	3		
	圆锥曲线	题号		⑥ ⑦	㉘		3 / 11	
		题型		M M	N			
		分值		3 3	5			
合计			21	27	34	18	30 / 100	100

湖南新东方教育培训中心
www.xdfpx.com 0731-84165117

简约式命题细目表

说明：

1. 全卷满分100分，表中分值即为百分比。

2. M为选择题；H为填空题；N为解答题；E为综合题。

3. 试题难度分为容易题（**0.7**以上）、适中（**0.4~0.7**）、难（**0.4**以下）三级，三种试题试题的分值比例约为**5:4:1**。

4. 考试时间**120**分钟，其中第I卷50分钟，第II卷70分钟。

全息式命题细目表

104050 课程命题细目表

考试类别：课程结业考试　　设计人：mao　　设计时间：2002 年 6 月 10 日　　共 1 页　第 1 页

考核内容	目标层次	多项选择题（2分/题） A	B	C	D	单项选择题（1分/题） A	B	C	D	简述题（5分/题） A	B	C	D	设计题（10分/题） A	B	C	D	综合分析题（15分/题） A	B	C	D	小计	合计	备注
考试学基本理论	I					3																3/3	11/20	
	II		4				2															6/10		
	III		1									1										2/7		
	IV																							
考试设计、实施；备考与应考	I	1				1	2	1														5/6	15/40	
	II	2	1	2						1												6/15		
	III			2								1			1							4/19		
	IV																							
考试信息处理	I						1															1/1	6/15	
	II					4																4/4		
	III														1							1/10		
	IV																							
现行考试制度及其改革创新	I	2																				2/4	5/25	
	II						1															1/1		
	III											1										1/5		
	IV																				1	1/15		
小计		5/10	6/12	4/8		8/8	6/6	1/1		1/5		3/15			2/20						1/15	37/100		
合计		15/30				15/15				4/20				2/20				1/15						

难度结构	A	B	C	D
	14/23	14/38	8/24	1/15

目标结构	I	II	III	IV
	11/14	17/30	8/41	1/15

备注：考试时间：I卷 40′；II卷 80′

A.容易　B.适中　C.较难　D.难　　I.识记　II.领会　III.简单应用　IV.综合应用

湖南新东方教育培训中心
www.xdfpx.com　0731-84165117



2．统计分析法

- 重测信度（学习前后测验）:**在学习某一知识前后各做一次测验，进行比较**
- 评分者一致性信度
- 复本信度:**来自同一范围的复本测同一批被试，看其相关**



3．经验推测法

- 即通过实践来检验效度。
  - **不同的被试团体在测验上的得分和对每题的反应有较大差异，一般总分和各题通过率随年级上升**
  - 如儿童发展量表是否有效，经过对不同年龄阶段的儿童进行调查，然后分析其结果。

例如，经验表明测验成绩随年龄提高，因此可分析测验成绩或题目通过率与年龄大小的关系。如果测验分和题目通过率随年龄而提高，则说明测验内容有效。



(三)内容效度的特性

- 内容效度与所有效度的性质一样，不是普遍适用的，而是根据具体情况分析得来。如果测验分析者和测验使用者定义的内容范围相同，则编制者报告的内容效度对使用者而言是有意义的，否则就没有意义。



内容效度与表面效度

- 1. 表面效度是由外行对测验作表面上的检查确定的，它不反映测验实际测量的东西，只是指测验表面上看来好像是测量所要测的东西；
- 我的喉咙里总好像有一块东西堵着似的。（可否用来测个性？）
- 你是否有过随口骂人的时候？你是一个爱交往的人吗？（*可否用来测个性偏好？*）

- 内容效度是由够资格的判断者（专家）详尽地、系统地对测验作评价而建立的。
- 不过，在编制测验时，表面效度是一个必须考虑的特性。
  - 例如，最高行为的测验要求有较高的表面效度



二、构想效度

- (一)什么是构想效度
  - 构思效度，还有人叫结构效度。
  - **一个测验实际测到所要测量的理论结构和特质的程度，或者说分数能够说明心理学理论的某种结构或特质的程度。**
  - 基本特征
    - **结构效度的大小取决于事先假定的心理结构理论；**
    - **无法证实时，有可能理论有问题**



二、构想效度

- **(二)构想效度的估计方法**
  - 1．对测验本身的分析（内容效度、内部一致性信度、被试对题目的反应特点）
  - 2．测验间的相互比较（相容效度、区分效度、因素分析法）
  - 3．效标效度的研究证明
  - 4．实验法和观察法证实



三、效标效度

- (一)什么是效标效度
  - 反映的是测验预测个体在某种情境下行为表现的有效程度。被预测的行为是检验效度的标准，简称效标。
  - 效标就是衡量一个测验是否有效的外在标准，它是独立于测验并可以从实践中直接获得的我们感兴趣的行为。
  - 根据效标资料是否与测验分数同时获得，又可分为同时效度和预测效度两类。



- 常用效标：学业成就、等级评定、临床诊断、特殊训练成绩、实际工作表现等
- 好的效标测量的条件：（1）真实反映观念效标的重要侧面；（2）效标测量必须稳定、可靠、客观；（3）在有效性的前提下，效标测量要尽可能简单、省时、省力



(二)效标效度的评估方法

- 1. 相关法
- 2. 区分法
- 3. 命中率法



1．相关法

- 相关法是评估效标效度最常用的方法，它是求测验分数与效标资料间的相关，这一相关系数称为效度系数。



2．区分法

- 区分法是检验测验分数能否有效地区分由效标所定义的团体的一种方法。
  - 某政府机关通过测验录用了一批公务员，经半年试用期后，根据他们的实际工作成绩，将这批公务员区分为称职和不称职两组，然后回过头来检查这两组人在测验上的分数是否有显著差异，如果差异显著，说明这个测验在录用公务员的选拔性测验中是有效的。



- 由于统计上的显著性受到团体大小的影响．当团体较大时，平均数之间小的差异在统计上也显著，因而变得有意义了。但实际上测验很难将这两个团体的人区分开。为避免这个缺点，还要看这两个团体分布的重迭量，这里有两个指标：
  - 一个是其中一组内被试得分超过另一组平均数的人数与另一组中被试超过其平均数的人数比率，如果比率小，说明两组差异大，测验是有效的。
  - 另一个指标是看两组分数分布共同区域的百分比例，如果百分比例不大，说明两组差异显著比较可信，测验是有效的。



3. 命中率法

- 命中率法是当测验用来做取舍的依据时，用其正确决定的比例作为效度指标的一种方法。
- 使用命中率法，可将测验分数和效标资料分为两类。在测验分数方面是确定一个临界分数（即分数线），高于临界分数者预测其成功，低于临界分数者预测其失败。

测验预测 \ 效标成绩	失败（-）	成功（+）
成功（+）	(A) 失误	(B) 命中
失败（-）	(C) 命中	(D) 失误

1. 总命中率（P_{CT}）

指正确决定数目（命中）对总决定数目（N）的比例：

$$P_{CT}=\frac{\text{命中}}{\text{命中}+\text{失误}}=\frac{B+C}{A+B+C+D}=\frac{\text{命中}}{N}$$

2. 正命中率（P_{CP}）

即所选的人成功的比例，又称选择效率。

$$P_{CP}=\frac{B}{A+B}=\frac{\text{成功人数}}{\text{选择人数}}$$

湖南新东方教育培训中心
www.xdfpx.com 0731-84165117



第三单元 效度的功能

- 一、预测误差
  - 效度系数常常以决定性系数来表示，意旨相关系数的平方，它表示测验正确预测或解释的效标的方差占总方差的比例。
  - $$Sest = S_y \times \sqrt{1 - r_{xy}^2}$$



第三单元 效度的功能

- 二、预测效标分数
  - $\hat{Y} = a + b_{yx}X$ **(公式5-26)**

$b_{yx}=r_{xy}\cdot s_y/s_x$

$a=\hat{Y}-b_{yx}X$

- 三、预测效度指数
  - **E值大小表明使用测验比盲目猜测能减少多少误差**
  - **E=100(1-K)** **k=** $\sqrt{1-r_{xy}^2}$



第四单元 影响效度的因素

一、测验本身的因素

二、测验实施中的干扰因表

三、样本团体的性质

四、效标的性质



一、测验本身的因素

- 1. 测验材料必须对整个**内容具有代表性**；
- 2. 测题设计时应尽量**避免容易引起误差的题型(如是非题)**；
- 3. 测题**难度要适中**，具有**较高的区分度**；
- 4. 测验**长度要恰当**，即要有一定的测题量；
- 5. 测题的排列按**先易后难**顺序。



二、测验实施中的干扰因表

- (一)主试的影响因素
  - 标准化
  - 预测效度的时间间隔
  - 测验情境
- (二)被试的影响因素



三、样本团体的性质

- (一)样本团体的异质性
  - **如果其他条件相同，样本团体越同质，分数分布范围越小，测验效度就越低；**
- (二)干涉变量
  - 年龄、性别、教育水平、智力、动机、兴趣、职业和任何有关的特征不同，同一测验的效度也会有很大不同。



四、效标的性质

- 效标与测验分数之间的关系是否线性
  - 皮尔逊积差相关系数的前题是假设两个变量的关系是线性的分布。
- 效标测量的可靠性即效标测量的信度。



真题练习

- 200705-105、内容效度的评估方法有(　　)。ACD
  - A专家判断法B双向细目表法C统计分析法 D经验推测法
- 200711-102、效度具有相对性，因此在评鉴测验的效度时，必须考虑测验的（　　）。（A）信度（B）目的（C）功能（D）长度 BC
- 200711-103、同200705-105
- 200711-104、测验实施中的干扰因素包括（　　）。ABCD
  - （A）主试者的影响因素（B）测验情景
  - （C）被试者的影响因素（D）其他干扰因素
- 200711-105、构想效度的估计方法包括(　　)。ABCD
  - （A）对测验本身的分析 （B）测验间的相互比较
  - （C）效标效度的研究证明（D）实验法和观察法证实
- 200806-58、效标与测验分数之间的相关，用皮尔逊积差相关系数的前提是假设两个变量的关系是（　　）的分布。B
  - A）非线性　（B）线性　（C）抛物线　（D）正态曲线

- **200806-107．信度与效度的关系可表述为（ ）。** CD
  - **A）信度是效度的必要而充分条件B）信度是效度的充分条件**
  - **C）信度是效度的必要而非充分条件D）效度是受信度制约的**
- **200806-108、根据效标资料是否与测验分数同时获得，可将效标效度分为（ ）。** CD
  - **A）相容效度 （B）区分效度 （C）同时效度 （D）预测效度**
- **200811-60、在心理测验中，效度是指一个心理测验的( )。** B
  - **(A)稳定性 (B)准确性 (C)可信度 (D)区分性**
- **200905-57、检验测验分数能否有效地划分由校标所定义的团体的一种方法是（ ）。** B
  - **（A）相关法 （B）区分法 （D）命中率法 （D）失误法**
- **200905-119、同200705-105**
- **200911-35、测验题目对有关行为取样的适用性是( )。** A
  - **(A)内容效度 (B)构想效度 (C)效标效度 (D)区分效度**
- **201005-107．同200711-102**
- **201011-104、效标效度可分为（ ）。P363** CD
  - **（A）相容效度 （B）区分效度（C）同时效度 （D）预测效度**
- **105、同200711-102**



201105-57、在心理测验中，效度是指所测量的与所要测量的心理特点之间符合的程度，或者简单地说是指一个心理测验的( )。



小结：第四节 测验的效度

第一单元 效度的概念
第二单元 效度评估的方法
第三单元 效度的功能
第四单元 影响效度的因素



编制测验

测验质量

信 度

效 度

测验项目

如何分析
项目的好坏？

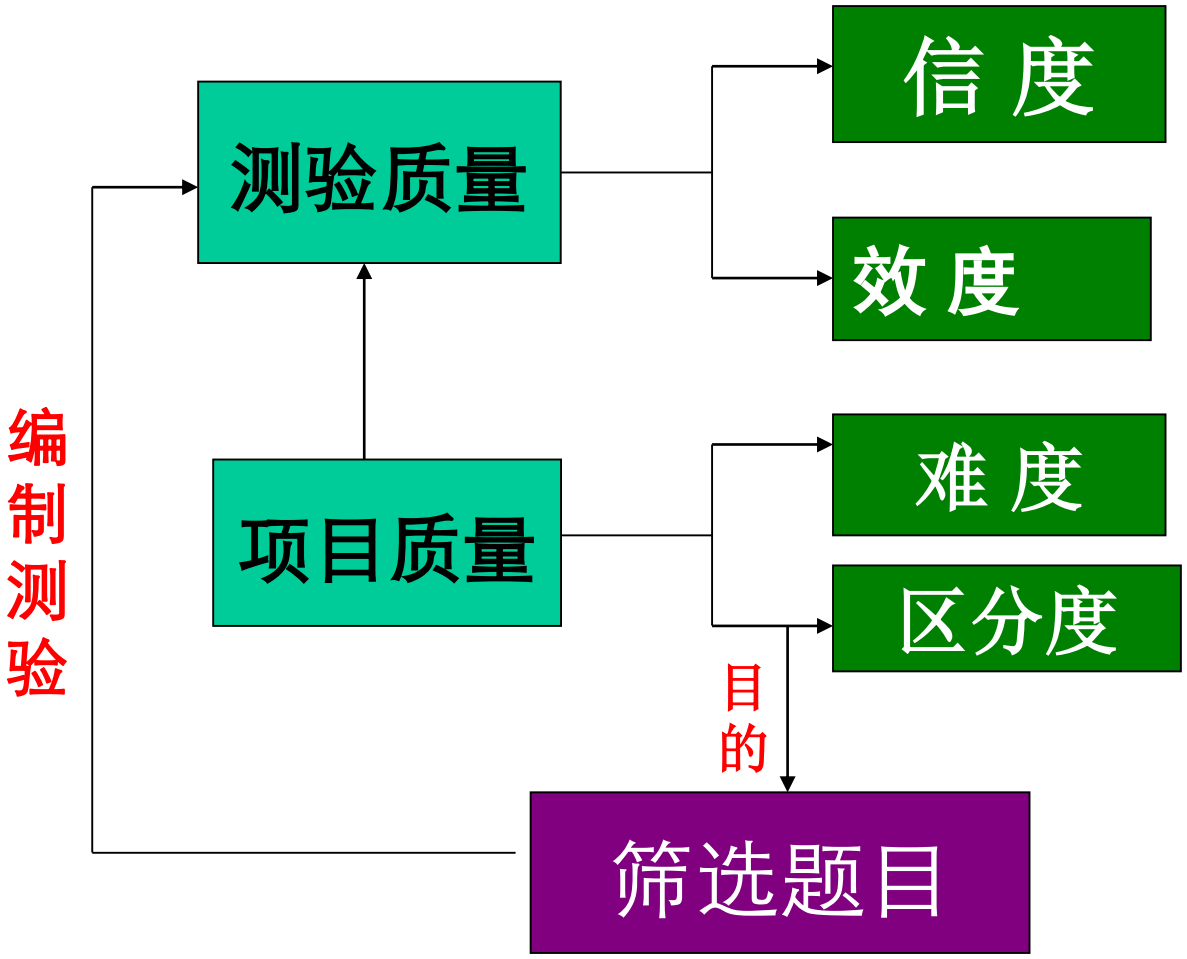
筛选题目



项目分析的内涵

- 项目分析是对测验中每个题目的质量进行的分析，包括质的分析和量的分析。
- 质的分析:主要由专家对项目的内容和形式进行分析，包括考察内容效度、题目编写的恰当性和有效性等。
- 量的分析：是通过计算项目的一些统计量作为项目质量高低的指标，包括项目的难度和区分度。
- 目的：对项目进行筛选和修订，提高测验的信度和效度。
- 如题库的建立




测验质量
信 度
效 度
项目质量
难 度
区分度
编制测验
目的
筛选题目



第五节 项目分析



- 项目分析就是对组成测验的每个测题进行分析。
- 项目分析包括定性分析和定量分析。
  - 定性分析：内容效度、题目编写的恰当性和有效性；
  - 定量分析：题目的难度和区分度。
- 项目分析的应用
  - 对项目进行编制、筛选和修订，可以提高测验的信度和效度
  - 试题库、计算机自适应考试



第一单元 项目的难度



一、定义

- 难度(difficulty)，顾名思义，是指**项目的难易程度**。
- 难度的指标通常以**通过率**表示，即以答对或通过该题的人数**百分比**来表示：
- $$P = \frac{R}{N} \times 100\%$$ (公式5-31)

- 通过人数越多(即P值越大)，难度越低；



二、计算方法

- (一)二分法记分的项目
  - $$P = \frac{P_H + P_L}{2}$$ **(公式5-32)**
- (二)非二分记分的项目
  - $$P = \frac{\overline{X}}{X_{\max}} \times 100\%$$ **(公式5-34)**



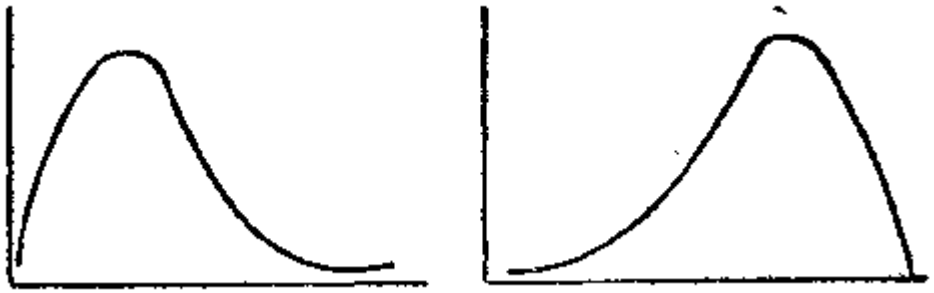
三、难度水平的确定

- (一)项目的难度
  - **选择题目时，最好使试题的平均难度接近0.50，而各题难度在0.50士0.20之间。**
  - **选拔测验一般难度接近录取率**
  - **速度测验、难度测验，一般都应防止被试得满分，因为满分的意义是不明确的**
  - **选择题的难度应大于概率水平**
    - **是非题0.75最为合适，四选一的0.63最为合适**



三、难度水平的确定

- (二)测验的难度
  - 如果被试样本具有代表性，对于中等难度的测验，其测验总分应该接近常态分配。





第二单元 项目的区分度



一、定义

- 项目区分度(item discrimination)，也叫鉴别力，是指测验项目对被试的心理特性的**区分能力**。
- 具有良好区分度的项目，能将不同水平的被试区分开来。
- 区分度（D）的范围在－1.00至1.00之间，有积极区分、消极区分和无区分作用三种类型。



二、计算方法

- (一)鉴别指数
- (二)相关法



(一)鉴别指数

- 根据测验总分取两端的人作为高分组和低分组（一般取上下各占27%的人）

$$D = P_H - P_L$$

- 步骤
  - 按测验总分高低排列
  - 确定高分组和低分组，一般是27%
  - 分别计算P_H 、 P_L
  - 最后计算D



题目鉴别指数与评价价标

鉴别指数	题目评价标准
0.4以上	很好
0.3-0.39	良好、修改后更好
0.2-0.29	尚可，仍需修改
0.19以下	差、必须淘汰



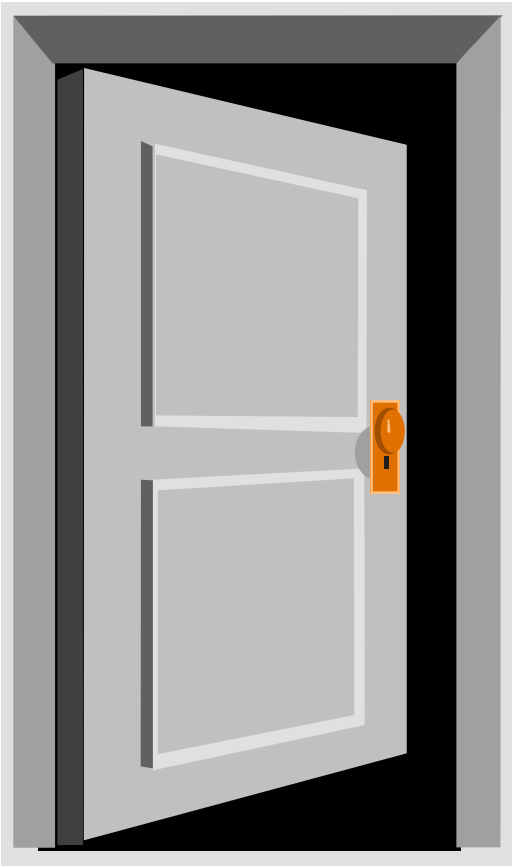
（二）相关法

- 相关法：以项目分数与效标分数或测验总分的相关作为项目区分度指标。
  - 点二列相关
  - 二列相关
  - Φ相关
  - 积差相关



三、区分度与难度的关系

- **关系密切:**
- 难度越接近0.5,项目潜在的区分度越大,难度越接近1.00或0,区分度越小
- 实际编制时,不能要求所有的项目难度都是0.5





D的最大值与项目难度的关系

项目通过率（P）	D的最大值
1.00	0.00
0.90	0.20
0.70	0.60
0.60	0.80
0.50	1.00
0.40	0.80
0.30	0.60
0.10	0.20
0.00	0.00



四、区分度的相对性

- 不同的计算方法，所得区分度不同
- 样本容量大小影响相关法区分度值的大小
- 分组标准影响鉴别指数值（D）
- 被试样本的同质性程度影响区分度质的大小



真题练习

- **200705-54、项目区分度也叫（B），是指测验项目对被试者的心理特征性的区分能力。**
  - **（A）难度　（B）鉴别力　（C）信度　（D）效度**
- **200711-55、项目的难易程度，用P表示，P值越（　　），难度越低。A**
  - **（A）大（B）低（C）小（D）接近于0**
- **201105-58，同上**
- **200806-59．难度与区分度的关系，C一般来说，较难的项目对高水平的受测者区分度（　　）。**
  - **（A）低　（B）一般　（C）高　（D）中等**



小结：第五节 项目分析

- 第一单元 项目的难度
- 第二单元 项目的区分度



第六节 测验编制的一般程序

- 第一单元 测验的目标分析
- 第二单元 测题的编写
- 第三单元 测验的编排和组织



第一单元 测验的目标分析

- 一、测验的对象
- 二、测验的用途
- 三、测验的目标



一、测验的对象

- 只有对受测者的年龄、受教育程度和社会经济以及阅读水平等心中有数，编制测验时才能有的放矢。



二、测验的用途

- 所编出的测验是要对被试做描述，还是做诊断，亦或是选拔和预测，
- (一)显示性测验
  - **显示性测验是指测验题目和所要测量的心理特征相似的测验。表明被试具有什么能力、能完成什么任务。**
- (二)预测性测验
  - 指预测一些没有被测量的行为的测验



三、测验的目标

- 心理测验的目标是指编制的测验是测什么的，即用来测量什么样的心理变量或行为特征。
- (一)工作分析
- (二)对特定概念下定义
- (三)确定测验的具体内容



第二单元 测题的编写

- 一、搜集资料
- 二、命题的一般原则
- 三、测题的编制要领



一、搜集资料

- (一)已出版的标准测验
- (二)理论和专家的经验
- (三)临床观察和记录



二、命题的一般原则

- (一)内容方面
  - 符合测验目的
  - 内容取样具有代表性
  - 题目间内容相互独立
- (二)文字方面
  - 准确、简明扼要
  - 一句话一个概念，尽量不用双重否定



二、命题的一般原则

- (三)理解方面
  - 除创造力和人格测验外，应有明确答案
  - 被试能够理解
  - 题目格式不要被曲解
- (四)社会敏感性方面
  - 假定被试具有某种行为
  - 假定规范不一致
  - 指出该行为是常见的



三、测题的编制要领

- 可以分为两大类：**提供型(supply)和选择型**(selection)题目。



(一)选择题

- 由两部分构成：**题干(stem)和选项**(options，alternatives)。
  - 题干所提问题必须明确，尽量使用简单明晰的词语
  - 选项简明扼要
  - 每题一个正确答案
  - 各选项长度应该相等
  - 避免题干用词选项与选项用词一致
  - 选项最好用同一形式



(二)是非题

- 内容以有意义的概念、事实或基本原则为基础
- 每道题只能包含一个概念
- 尽量避免否定叙述
- 若是表达意见，最好说明来源和根据
- “是”、“非”题的比例适当



(三)简答题

- 宜用文句形式
- 填充形式，空格不能太多
- 有唯一正确答案，评分标准客观



(四)操作题

- 明确目标
- 尽量选择真实程度较高的项目
- 指导语简明扼要
- 制定评分标准，确定记分方法



第三单元 测验的编排和组织

- 一、合成测验
- 二、预测与项目分析
- 三、信度和效度考察
- 四、常模制定
- 五、编写指导手册



一、合成测验

（一）选择与审订试题

（二）测题的编排



(一)选择与审定试题

- **1．选择试题形式**
- **2．审定题目**



(二)测题的编排

- 测验编排的一般原则：

1. 测题的难度排列宜逐步上升。
2. 尽可能将同类型的测题组合在一起。
3. 注意根据各种类型测题本身的特点排列题目。

- 两种常见的排列方式。

1. 并列直进式：

2. 混合螺旋式：



二、预测与项目分析

- (一)预测
- (二)项目分析



(一)预测

- 对象应来自正式测验的总体，要具有代表性
- 实施情境应尽量与正式测验一致
- 预测时间可以宽松一些，让每个被试都能完成题目
- 随时记录被试反应



(二)项目分析

- 质：内容取样的适当性，题目的思想性，表达是否清楚
- 量：预测结果统计分析，确定项目的难度、区分度、备选答案的适宜性等
- 有时候需要从同一总体中抽取两个样本分别进行项目分析（复核）



三、信度和效度考察

- (一)信度
- (二)效度



五、编写指导手册

- 测验指导手册的内容有：
  - **1．测验的目的和功用。**
  - **2．测验编制的理论背景**
  - **3．如何实施测验的说明。**
  - **4．测验的标准答案和记分标准。**
  - **5．常模资料**
  - **6．测验的基本特征**



真题练习

- **200705-102、测验指导语一般由（ ）等内容组成。** ABCD
  - **（A）如何选择反应形式 （B）如何记录这些反应**
  - **（C）时间限制**
  - **（D）给例题，有时告知被试者测验目的**
- **20095-106、测题常见的排列方式有（ ）。** AC
  - **（A）并列直进式 （B）公共汽车式**
  - **（C）混合螺旋式 （D）专列式**
- **201105-109**



小结：第六节 测验编制的一般程序

- 第一单元 测验的目标分析
- 第二单元 测题的编写
- 第三单元 测验的编排和组织



第七节 心理测验的使用

- 第一单元 主试的资格
- 第二单元 测验的选择
- 第三单元 测验前的准备工作
- 第四单元 测验实施的程序及要素
- 第五单元 被试误差及控制方法
- 第六单元 测验的评分
- 第七单元 测验结果的报告



第一单元 主试的资格

- 一、知识结构
- 二、专业理论知识和专业技能
- 三、职业道德



一、知识结构

- 主试者的知识结构是指开展心理测验工作所必须具备的**基础知识和与专业相关的知识**。
- 还应根据自己的工作领域具备相应的本行业的专业知识。



二、专业理论知识和专业技能

- (一)专业理论知识
- (二)专业技能
  - **熟悉有关测验的内容、适用范围、测验程序和记分方法等。**



三、职业道德

- (一)测验的保密和控制使用
- (二)测验中个人隐私的保护



第二单元 测验的选择

- 一、所选测验必须适合测量的目的
- 二、所选测验必须符合心理测量学的要求



第三单元 测验前的准备工作

- 一、测验前的准备工作
- 二、测验中主试的职责
- 三、建立协调关系



一、测验前的准备工作

- (一)预告测验
- (二)准备测验材料
- (三)熟悉测验指导语
- (四)熟悉测验的具体程序



二、测验中主试的职责

- 首先，应按照指导语的要求实施测验
- 其次，测验前不讲太多无关的话。
- 第三，对于被试的反应，主试不应做出点头、皱眉、摇头等暗示性反应
- 第四，对特殊问题要有心理准备



三、建立协调关系

- 这种关系指的是主试和被试之间一种友好的、合作的、能促使被试最大限度地做好测验的一种关系。



第四单元 测验实施的程序及要素

- 一、指导语
- 二、时限
- 三、测验的环境条件



一、指导语

- (一)对被试的指导语
  - 1．如何选择反应形式(划“√”、口答、书写等)；
  - 2．如何记录这些反应(答卷纸、录音、录像等)；
  - 3．时间限制；
  - 4．如果不能确定正确反应时该如何操作(是否允许猜测等)；
  - 5．例题
  - 6．有时告知被试测验目的。
- (二)对主试的指导语



二、时限

- 大多数典型行为测验是不受时间限制的
- 但在**最高作为测验**中，速度是需要考虑的重要因素之一。



三、测验的环境条件

- 首先必须完全遵从测验手册的要求；
- 其次是记录下任何意外的测验环境因素；
- 第三，在解释测验结果时也必须考虑这一因素。



第五单元 被试误差及控制方法

- 一、应试技巧与统习效应
- 二、动机与焦虑因素
- 三、反应定势



一、应试技巧与练习效应

- (一)应试技巧
  - **“测验油子”**
- (二)练习效应



二、动机与焦虑因素

- (一)应试动机
  - 在测量成就、智力和能力倾向等变量时，只有被试动机强烈，才可能尽力回答，取得好成绩。
- (二)测验焦虑



三、反应定势

- 反应定势亦称反应风格(responsesetsorstyles)，是指独立于测验内容的反应倾向，即由于每个人回答问题的习惯不同，而使能力相同的被试得到不同的测验分数。



几种常见的反应定势：

- (一)求“快”与求“精确”的反应定势
  - 除非“反应速度”本身即为重要的研究目标，否则应让被试有充分的时间反应，同时应该注明每题的答题时间，以减少求“快”与求“精确”定势的影响。
- (二)喜好正面叙述的反应定势
  - 在编制是非题时，“是”“否”题大致相等或答“否”题略多，是控制肯定定势约有效方法。



- (三)喜好特殊位置的反应定势
  - **正确答案的位置在整个测验中出现在各位置的概率相等，就可以控制这种位置定势。**
- (四)喜好较长选项的反应定势
- (五)猜测的反应定势



第六单元 测验的评分

- 一、原始分数的获得
- 二、原始分数的转换



一、原始分数的获得

- 为使评分尽可能客观，有三点要求：
  - **1．及时而清楚地记录反应情况。**
  - **2．要有一张标准答案或正确反应的表格，即记分键。**
  - **3．将被试的反应和记分键比较，对反应进行分类。**



二、原始分数的转换



第七单元 测验结果的报告

- 一、测验分数的综合分析
- 二、报告分数的具体建议



一、测验分数的综合分析

- 根据心理测验的特点进行分析
- 不能把分数绝对化，更不能仅仅根据一次测验的结果轻易下结论
  - **一个人在任何一个测验上的分数都是其遗传特征、测验前的学习与经验、以及测验情境的函数**
- 为了对测验做出确切的解释，仅有常模资料是不够的，还需要有信效度资料
- 不同测验的分数不能直接加以比较



二、报告分数的具体建议

- 不把测验分数直接告诉被试本人、家长、学校老师等，应告诉测验分数的解释和建议
- 避免使用专业术语
- 保证当事人指导测验测量或预测什么
- 使当事人知道自己和什么团体进行比较
- 要使当事人知道如何运用他的分数
- 考虑测验分数会给当事人带来什么心理影响
- 让当事人积极参与测验分数的解释



真题

- **201105-110、测验主试者的职业道德包括(　)。**
  - **(A)对测验分数的保密(B)测验的控制使用**
  - **(C)对受测者隐私的保护(D)对测验特别熟悉**



小结：第七节 心理测验的使用

- 第一单元 主试的资格
- 第二单元 测验的选择
- 第三单元 测验前的准备工作
- 第四单元 测验实施的程序及要素
- 第五单元 被试误差及控制方法
- 第六单元 测验的评分
- 第七单元 测验结果的报告



《心理测量学·基础知识》串讲结束

谢谢！