




## 第五章 测验的效度

- ❶ 第一节 效度概述
- ❷ 第二节 效度的种类及其评估方法
- ❸ 第三节 效度的功能
- ❹ 第四节 影响效度的因素

Li, J. School of Psychology, Beijing Normal University




### 第一节 效度概述

- ❶ 1.1 效度的定义
- ❷ 1.2 效度的性质

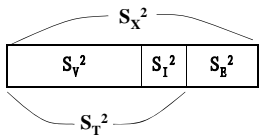
2

Li, J. School of Psychology, Beijing Normal University




### 1.1 效度的概念

- ❶ 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。
- ❷ 在测量学中，效度被定义为与测量目的有关的变异（有效变异）与总变异之比： $r_{XY}^2 = \frac{S_V^2}{S_X^2}$
- ❸ 其中  $r_{XY}$  为效度系数
- ❹  $r_{XY}^2$  为决定系数



3

Li, J. School of Psychology, Beijing Normal University



### 信度与效度的关系


- ❶ 测量的效度除受随机误差影响外，还受系统误差的影响。可信的测验未必有效，而有效的测验必定可信。简言之，信度是效度的必要条件。

$$\because r_{XY}^2 = \frac{S_V^2}{S_X^2} = 1 - \frac{S_E^2}{S_X^2} - \frac{S_I^2}{S_X^2} = r_{XX}^2 - \frac{S_I^2}{S_X^2}$$

$$\because S_I^2 \text{ 最小为 } 0 \quad \therefore r_{XY} \leq \sqrt{r_{XX}}$$

4

Li, J. School of Psychology, Beijing Normal University




### 1.2 效度的性质

- ❶ 效度具有相对性
  - ❶ 任何测验的效度是对一定的目标来说的，或者说测验只有用于与测验目标一致的目的和场合才更有效。
- ❷ 效度具有连续性
  - ❶ 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。
- ❸ 效度是针对测验结果而非测验工具本身

5

Li, J. School of Psychology, Beijing Normal University



### 第二节 效度的种类及其评估方法

- ❶ 美国心理学会1974出版的《教育与心理测验标准》将测验的效度分为三类：
  - ❶ 2.1 内容效度 (content validity)
  - ❷ 2.2 效标效度 (criterion validity)
  - ❸ 2.3 构想效度 (construct validity)

6

Li, J. School of Psychology, Beijing Normal University

## 2.1 内容效度 (content-related validity)



- ① 内容效度指的是测验题目对内容或行为取样的恰当程度，用以确定测验是否是所欲测量的行为领域的代表性取样。
- ① 多用于成就测验、职业资格测验、工作样本测验
- ① 基本要求
  - ↳ 要对所测量的心理特性有明确的概念，并划定出哪些行为与该心理特性有关。
  - ↳ 测验项目应是已界定的内容范围的代表性样本。

7

Li, J. School of Psychology, Beijing Normal University

## 内容效度的评估方法



### ①1) 专家判断法:

- ① 为了确定一个测验的内容效度高低，最常用的方法是请相关专家对测验题目与原定内容的符合性做出判断，看测验的题目是否代表了规定的内容。
- ① 这种估计效度的方法是一个逻辑分析的过程，因此也被称为“逻辑效度” (logical validity)。

8

Li, J. School of Psychology, Beijing Normal University

## 内容效度的评估方法



### ↳ 具体做法

- ① 定义好测验内容的总体范围，描述有关的知识、技能及所用材料的来源；
- ① 编制双向细目表，确定内容和技能各自所占的比例，并由测验编制者确定各题所测的是何种内容与技能；
- ① 制定评定量表来衡量测验的效度，如测验包括的内容、技能、材料的重要程度、题目对内容的适用性等。

9

Li, J. School of Psychology, Beijing Normal University

## 具体量化方法



- ① 例：对一个工作选拔测验的每一个项目进行内容效度分析

$$CVR = \frac{n_e - N/2}{N/2}$$

- ①  $n_e$  = no. of panelists indicating “essential” (out of “essential”, “useful but not essential”, “not necessary”)
- ①  $N$  = total no. of panelists
- ① CVR measures the degree of agreement among raters over the 50% baseline

10

Li, J. School of Psychology, Beijing Normal University

## 内容效度的评估方法



### ①2) 其它方法:

- ↳ 统计分析法：克伦巴赫 -> 复本相关
- ↳ 再测法：前测 -> 教学/训练 -> 后测
- ↳ 经验法：
  - ① 检查不同年级的学生在测验上的得分和在每个题目上的反映情况。例如对于成就测验，如果总分和题目的通过率随着年级而增高，说明测验对教学具有较好的内容效度。

11

Li, J. School of Psychology, Beijing Normal University

## 内容效度的特性



- ① 内容效度与其它效度的性质一样，不是普遍适用的，而是根据具体情况分析得来。
- ↳ 如果测验编制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。

12

Li, J. School of Psychology, Beijing Normal University

## 内容效度 vs. 表面效度

- ① 表面效度 (face validity) 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。
- ② 内容效度是由够资格的判断者 (专家) 详尽地、系统地测验作评价而建立的。

13

Li, J. School of Psychology, Beijing Normal University

## ① 对表面效度的要求

- ② 最高水平测验: 高; 典型行为测验: 低

## ① 提高表面效度可以起到“包装”的作用

14

Li, J. School of Psychology, Beijing Normal University

## 2.2 效标效度 (criterion-related validity)

- ②.2.2.1 什么是效标效度
- ②.2.2.2 效标效度的评估方法

15

Li, J. School of Psychology, Beijing Normal University

## 2.2.1 什么是效标效度

- ① 效标效度反映的是通过测验预测个体在某种情境下行为表现的有效性的程度, 即将测验分数与希望做出推断的行为 (效标/criterion) 联系起来。
- ② 这种效度需在实践中检验, 所以又称实证效度。
- ③ 效标效度是基于数据的 (data based)
  - ④ 一个雇员选拔测验与工作绩效的相关
  - ④ 一个小学数学成就测验与年级的相关
  - ④ 一个诚实测验与偷窃行为的相关

16

Li, J. School of Psychology, Beijing Normal University

## 两种效标关联效度

- ① 同时效度 (concurrent validity): 效标分数与测验分数同时收集。检查测验测量现有的某种能力或特质的有效性, 即描述当前状态时的有效性。
- ② 预测效度 (predictive validity): 先收集测验分数, 过一段时间后再收集效标分数, 表明测验对某种行为的预测的有效性。

17

Li, J. School of Psychology, Beijing Normal University

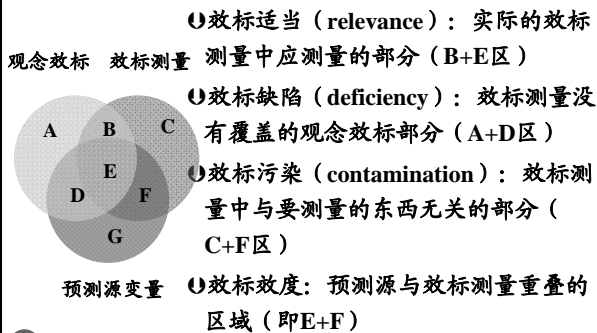
## 观念效标与效标测量

- ① 观念效标是指从概念上确定所要选择的测验效标是什么。
  - ② 如确定工人的工作绩效为特殊能力倾向测验的效标
  - ② 将大学学业成就作为高考的效标
- ② 效标测量是指以具体的测量指标表示效标行为水平的高低。
  - ③ 例如用学生第一学年平均成绩代表大学学业成就。
- ③ 预测源 (predictor) 是指用于预测效标的任何东西 (如高考, 特殊能力倾向测验)

18

Li, J. School of Psychology, Beijing Normal University

## 效标的适当性与效标的污染



19

Li, J. School of Psychology, Beijing Normal University

❹效标的污染: 个人的效标成绩由于评定者知道其预测源分数而使效标“纯度”降低的情况 (Anastasi, 1988)。

20

Li, J. School of Psychology, Beijing Normal University

## 好效标应具备的条件

- ❶效标能最有效地反映测验的目标, 即效标测量本身必须有效;
- ❷效标必须具有较高的信度, 稳定可靠, 不随时间等因素而变化;
- ❸效标可以被客观测量, 可用数据或等级来表示;
- ❹效标测量的方法简单, 省时省力, 经济实用。

21

Li, J. School of Psychology, Beijing Normal University

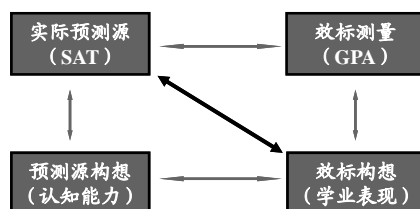
### ❶常见的效标

- ↳ 学业成就: 学科成绩, 学历, 获奖
- ↳ 实际工作表现评定: 产量、治愈率、违规次数等
- ↳ 已知有效的测验
- ↳ 临床诊断
- ↳ 来自他人的等级评定
- ↳ 特殊训练成绩: 例如飞行训练成绩
- ↳ 不同的团体:
  - ❶智力测验-不同年龄团体, 兴趣测验-不同职业团体

22

Li, J. School of Psychology, Beijing Normal University

## 构想、效标与测量的关系



23

Li, J. School of Psychology, Beijing Normal University

## 2.2.2 效标效度的评估方法

### ❶1) 相关法

↳ 求测验分数与效标分数的相关, 这一相关系数称为效度系数。评估效标效度最常用的方法。

❶积差相关/等级相关/点二列相关/二列相关等。

$$r_{pb} = \frac{\bar{X}_p - \bar{X}_q}{S_t} \cdot \sqrt{pq} \quad r_b = \frac{\bar{X}_p - \bar{X}_q}{S_t} \cdot \frac{pq}{Y}$$

24

Li, J. School of Psychology, Beijing Normal University



①效度系数越大越好，虽没有最低要求，但实际使用中，相关系数至少要达到统计显著。

①相关法的优点：

①提供了一个可比较的统计指标

①可用于建立回归方程

①相关法的缺点

①如果预测源与效标的关系非线性则无法使用

①不能提供有关取舍正确性的指标

25

Li, J. School of Psychology, Beijing Normal University



## ②) 区分法

①检验测验分数能否有效地区分由效标所定义的团体（成功、失败）的一种方法。

①可采用Z检验或t检验。

26

Li, J. School of Psychology, Beijing Normal University



## ③) 命中率法

①当测验用来做取舍的依据时，可用其正确决定的比例作为效度的指标。

①概念：总命中率、正命中率

①例：70人参加测验，其中50人通过；一年后评价工作绩效，发现测验通过者中有10人不合格，测验未通过的人中有8人不合格。该测验的总命中率、正命中率分别是多少？

27

Li, J. School of Psychology, Beijing Normal University



效标成绩 测验成绩	效标成绩	
	不合格 (-)	合格 (+)
未通过 (-)	(A) 命中8	(B) 失误12
通过 (+)	(C) 失误10	(D) 命中40

28

Li, J. School of Psychology, Beijing Normal University



## 2.3 构想效度 (construct-related validity)

①构想：一种说明个体之间差异的，理论的、抽象的性质或特质。例如智力、内外向、成就动机、焦虑等。

①构想是抽象滴，但通过对外显行为的观察，也可以将个体拥有该性质的程度数量化。

①对同一构想有不同的测量方法。但无论怎样去测量，对一种特质的测量结果必须与该特质的理论解释相符合。

29

Li, J. School of Psychology, Beijing Normal University



## 2.3 构想效度 (construct-related validity)

①构想效度主要涉及的是心理学的理论概念问题，是指测验能够测量到理论上的构想或特质的程度，即测验的结果是否能证实或解释某一理论的假设、术语或构想，解释的程度如何。也称为构思效度或结构效度。

①内容效度和效标关联效度都体现了构想效度

30

Li, J. School of Psychology, Beijing Normal University



### 检验流程:

- ①建立构想效度，必须先从某一构想的理论出发，提出关于某种心理特质的假设；
- ②然后设计和编制测验并进行施测；
- ③最后对测验的结果采用相关或因素分析等方法进行分析，验证测验结果与理论假设的符合程度。

31

Li, J. School of Psychology, Beijing Normal University



## 确定构想效度的具体方法

- ②.3.1 测验内方法
- ②.3.2 测验间方法
- ②.3.3 效标效度法
- ②.3.4 实验法
- ②.3.5 因素分析法

32

Li, J. School of Psychology, Beijing Normal University



### 2.3.1 测验内方法

①这类方法是通过研究测验内部结构，如测验的内容以及题目间的关系来分析测验的构想效度。

- 1) 测验的内容效度可作为构想效度的证据
- 2) 分析被试对项目作反应的过程
  - ✎哪些因素影响了反应？社会称许性？
- 3) 考察测验的同质性
  - ✎是否测量了单一特质？内部一致性指标

33

Li, J. School of Psychology, Beijing Normal University



### 2.3.2 测验间方法

①通过分析几个测验间的相互关系，找出其共同与不同之处，进而推断这些测验测量的特质是什么，确定其构想效度。

- ✎1) 相容效度 (congruent validity)
  - ①考察一个新测验与测量同一构想的现有测验 (已有效度证据) 的相关系数，以证明新测验的效度。
  - ✎如，从瑞文渐进矩阵推理测验 → 新编数字推理测验
- ✎2) 多质多法矩阵 (MMM)

34

Li, J. School of Psychology, Beijing Normal University



### 多质多法矩阵(multitrait-multimethod matrix)

- ①即以一种以上的方法 (methods) 测量一种以上的特质 (traits)。 (Campbell & Fiske, 1959)
- ②具有良好构想效度的测验，其测验结果不仅应与测量相同构想的变量有高相关，也应与测量不同构想的变量有低相关，前者为会聚效度，后者为区分效度。

35

Li, J. School of Psychology, Beijing Normal University



### MMM实例

- ①假设我们期望测量三种构想：
  - ✎创造力 (Creativity, C)
  - ✎外向性 (Extraversion, E)
  - ✎开放性 (Openness, O)
- ②假设我们可以采用三种不同的测量方法：
  - ✎自我报告 (Self-report, SR)
  - ✎他人评定 (Rating by others, RO)
  - ✎行为测验 (Behavior test, BT)

36

Li, J. School of Psychology, Beijing Normal University

## MMM实例

		Method I 自我报告			Method II 他人评定			Method III 行为测验		
		A1	B1	C1	A2	B2	C2	A3	B3	C3
M	A1 创造力	(.89)								
I	B1 外向性	.43	(.89)							
	C1 开放性	.46	.32	(.76)						
M	A2 创造力	.62	.13	.20	(.93)					
II	B2 外向性	.12	.70	.03	.40	(.94)				
	C2 开放性	.20	.09	.64	.42	.30	(.84)			
M	A3 创造力	.59	.16	.22	.60	.20	.21	(.94)		
III	B3 外向性	.14	.82	.11	.13	.61	.13	.39	(.92)	
	C3 开放性	.21	.10	.72	.19	.09	.52	.40	.36	(.85)

## 会聚效度 (convergent validation)

①会聚效度：用不同方法测量同一种构想所得的相关。在MMM中，这些相关系数值应该较高。

①红色数字是三种方法测量同一种构想的相关。

①这些同一构想的测量称为“效度对角线” (validity diagonal)

38

Li, J. School of Psychology, Beijing Normal University

## 区分效度 (discriminant validation)

①区分效度包括三种：

①不同方法测量不同特质之间的相关应小于效度对角线。

①同一方法测量不同特质之间的相关应小于效度对角线。

①可比三角形中应表现同一相关模式。

39

Li, J. School of Psychology, Beijing Normal University

## MMM的汇总

①Reliability (信度)

①以相同方法测量相同特质 (应有最高相关)

①Convergent validity (会聚效度)

①以不同方法测量相同特质 (应有次高相关)

①Discriminant validity (区分效度)

①以相同或不同方法测不同特质 (应有最低相关)

①Method bias/effects (方法偏误或方法效应)

①以相同方法测不同特质的相关高于以不同方法测不同特质的相关 (但差异应越小越好)

40

Li, J. School of Psychology, Beijing Normal University

## 2.3.3 效标效度法

①一个测验若效标效度理想，那么该测验所预测的效标的性质和种类就可以作为分析测验构想效度的指标。

①1) 对照组与高低分组

①2) 发展变化

①当心理特质具有发展变化特点时方能使用

41

Li, J. School of Psychology, Beijing Normal University

## 2.3.4 实验法

①考察经过某种实验训练后，测验得分是否有所变化。例如测量空间能力。

①出声思考法：在实验前实施测验，实验中要求被试出声思考，考察测验结果与出声思考 (更加真实、客观) 结果的一致性。

①例如测量阅读策略的测验，可先实施测验，之后让受测者完成短文阅读的实验，并出声思考。

42

Li, J. School of Psychology, Beijing Normal University

### 2.3.5 因素分析法 (Factor analysis)

- ① 一种多元统计分析技术，主要目的是降维，即用更少量的因素概括解释许多相互关联的变量。
- ① 通过对一组题目进行因素分析，可以找到影响测验分数的共同因素，这种因素可能就是我们要测量的心理特质（构想）。

43

Li, J. School of Psychology, Beijing Normal University

- ① 因素分析法对构想效度的证明既可以是项目为单位，也可以分测验为单位。

① 例如，韦氏成人智力量表中国修订版（WAIS-RC）的言语分量表包括常识、理解、类比、词汇、算术、记忆广度六个分测验。

- ① 通过因素分析法，我们得到两个因素

44

Li, J. School of Psychology, Beijing Normal University

### 旋转后的韦氏言语分量表因素矩阵

分测验	因素负荷		共同度
	因素I	因素II	
常识	0.8117	0.0921	0.6681
理解	0.8043	0.0000	0.6469
类比	0.8012	0.1499	0.6644
词汇	0.8133	0.0832	0.6684
算术	0.6012	0.4423	0.5570
记忆广度	0.4729	0.5522	0.5285
方差贡献	3.1942	0.5391	3.7333
所占比例	53.23	8.99	62.22

45

Li, J. School of Psychology, Beijing Normal University

- ① 各因素在各分测验（或项目）上的负荷量，即各分测验（或项目）与因素的相关，被称为测验的因素效度（factorial validity）。

- ① 共同度指抽取出的因素能解释测验分数变异的多大比例（可作为测验构想效度的指标）。例如，因素1和2一共可以解释常识测验分数66.81%的变异。

46

Li, J. School of Psychology, Beijing Normal University

## 第三节 效度的功能

- ③.1 预测效标分数
- ③.2 确定预测的误差
- ③.3 人事测评中的实际问题
- ③.4 利用测验做决策时需注意的问题

47

Li, J. School of Psychology, Beijing Normal University

### 3.1 预测效标分数

- ① 如果X与Y两变量呈直线相关，只要确定出二者间的回归方程，就可以从一个变量推估出另一个变量，例如用测验分数来预测效标成绩。

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

其中， $\hat{Y}$  是预测的效标分数

a为纵轴的截距，用于纠正平均数差异

$b_{yx}$ 为斜率，即Y对X的回归系数

X为测验分数

48

Li, J. School of Psychology, Beijing Normal University





### ①预测效标分数的过程:

- ①根据已经收集到的测验分数与对应的效标成绩建立回归方程;
- ②当新的受测者(必须来自于导出回归方程的样本所在的总体)出现,我们可以通过其测验分数,预测其效标的成绩。

49

Li, J. School of Psychology, Beijing Normal University



### ①建立回归方程的本质就是求解两个常数

$$b_{yx} = r_{xy} \times \frac{S_y}{S_x} \quad a = \bar{Y} - b_{yx} \bar{X}$$

其中,  $r_{xy}$  是测验分数与效标成绩的相关, 即效度系数  
 $S_y$  和  $S_x$  分别是效标成绩与测验分数的标准差  
 $\bar{Y}$  和  $\bar{X}$  分别为效标成绩与测验分数的平均数

50

Li, J. School of Psychology, Beijing Normal University

①例: 已知往年20位员工的入职测验得分和工作一年后的绩效考核成绩, 问通过入职测验分数来预测绩效考核成绩的效度是多少, 并据此求绩效考核对入职测验分数的回归方程。若今年新招聘的一位员工在入职测验上得了120分, 请预测其一年后的绩效成绩。

应聘者	01	02	03	04	05	06	07	08	09	10
推理测验X	89	97	126	87	119	101	130	115	108	105
面试Y	55	74	87	60	71	54	90	73	67	70
应聘者	11	12	13	14	15	16	17	18	19	20
推理测验X	84	121	97	101	92	110	128	111	99	120
面试Y	53	82	58	60	67	80	85	73	71	90

### ①解: 经计算得

$$\bar{X} = 107, \bar{Y} = 71, S_x = 13.69, S_y = 11.63, r_{xy} = 0.86$$

52

Li, J. School of Psychology, Beijing Normal University



## 3.2 确定预测的误差

回归方程解决的问题是如何预测效标成绩, 但我们也关心预测的效果到底好不好, 测验分数与效标成绩的关联程度高不高。

- ①效度系数的实际意义常常以决定系数 (coefficient of determination) 来表示, 即相关系数的平方。
- ②表示测验正确预测或解释的效标成绩的方差占效标成绩总方差的比例。
- ③也可看作回归平方和在总平方和中所占比例。

53

Li, J. School of Psychology, Beijing Normal University



①前例中  $r_{xy} = 0.86$ , 则其决定系数  $r^2_{xy} = 0.74$ , 说明效标成绩 (Y) 的变异中有74%是由测验分数 (X) 的变异引起的, 即Y的变异有74%可以被X的变异解释, 剩下的26%无法解释。

X的变异  Y的变异

②再如若复本信度  $r_{xx} = 0.7$ , 则说明两次测验分数的共同变异为0.49, 因此测验的稳定性并不高!

54

Li, J. School of Psychology, Beijing Normal University

## 预测的标准误 $SE = S_x \sqrt{1 - r_{xy}^2}$

①对误差的估计也可通过“回归方程估计效标成绩的标准误”来表示

② $S_{est}$ ，指具有相同测验分数的受测者，其效标成绩（Y）分布的标准差，即预测误差大小的估计值。

$$S_{est} = S_y \times \sqrt{1 - r_{xy}^2}$$

③ $S_y$ 为效标成绩的标准差

55

Li, J. School of Psychology, Beijing Normal University

④当决定系数=1， $S_{est}=0$ ，测验分数可以完美替代效标成绩；

④当决定系数=0， $S_{est}=S_y$ ，纯属随机猜测，预测源测验形同虚设。

①区间估计：置信度分别为95%、99%时

预测效标成绩 -  $1.96S_{est}$  ≤ 真正效标成绩 ≤ 预测效标成绩 +  $1.96S_{est}$

预测效标成绩 -  $2.58S_{est}$  ≤ 真正效标成绩 ≤ 预测效标成绩 +  $2.58S_{est}$

56

Li, J. School of Psychology, Beijing Normal University

③3.1的例子中，

$r_{xy} = 0.86$  ∴ 决定系数为0.74

∴  $S_{est} = S_y \times \sqrt{1 - r_{xy}^2} = 11.63 \times \sqrt{1 - 0.74} = 5.93$

∴ 当 $X$ 为120， $\hat{Y} = 80.5$ ，

∴ 则其95%的置信区间为

$(80.5 - 1.96 \times 5.93, 80.5 + 1.96 \times 5.93)$

即  $(68.88, 92.12)$

57

Li, J. School of Psychology, Beijing Normal University

## 3.3 选拔性测验的实际问题

①当使用测验来选人时，只有高效度（validity coefficient）并不能保证测验一定有用，还需考虑其它因素：

④对成功的界定（definition of success）

④基础率（base rate）

④录取率（selection ratio）

④分数线（cut score）

58

Li, J. School of Psychology, Beijing Normal University

④基础率：在未经选择的总体中某现象或行为的出现率。

①测验要有用，必须比单纯依据基础率能做出更多的正确决定。

①当要预测的行为的基础率特别高或者特别低时，没有必要使用测验。

①效度一定时，基础率越接近50%，测验的贡献越大。

59

Li, J. School of Psychology, Beijing Normal University

④录取率：从申请者中选出的人数比例。

①效度一定时，录取率越低，选出的人成功的可能性越大。

④例如高考

④分数线：合理设置可提高测验使用的价值

①正命中率 vs. 总命中率

①若工作对所测能力的要求不高，可以按照最低能力水平划线

60

Li, J. School of Psychology, Beijing Normal University

### 评价测验效度的一个原则

#### ① 功利率：使用测验的花销 vs. 得到的利益

- ↳ 使用测验的效益 >> 不使用测验的效益
- ↳ 测验的好处 >> 测验所耗时间、精力、经费

61

Li, J. School of Psychology, Beijing Normal University

### 3.4 利用测验做决策时需注意的问题

#### ① 测验只是做决策的辅助工具

#### ① 确定测验效度值的最低标准

- ↳ 相关系数达到统计显著
- ↳ 使用测验必须比依据基础率产生更多的正确决策
- ↳ 使用测验后能达到一定的功利率
- ↳ 测验的有效性大于其它现有工具，或形成补充

#### ① 对效度的评价是基于团体数据，而非某个人的数据

#### ① 实际工作中可以采用多个测验进行连续选择

62

Li, J. School of Psychology, Beijing Normal University

### 概念补充

① A significant trend is a new focus in meta-analytic research on incremental validity. There are many meta-analytic summaries of the validity of individual predictors, and recent work focuses on combining meta-analytic results across predictors in order to estimate the incremental validity of one predictor over another

① Incremental validity: gain in validity resulting from adding one or more new predictors to an existing selection system.

#### ① 增益效度

63

Li, J. School of Psychology, Beijing Normal University

### 材料补充—人事选拔中的效度系数

- ↳ When using cognitive ability tests to select employees, average observed validity coefficients are almost all in the .20s.
- ↳ When using personality sources to predict performance outcomes, observed validity are .04 to .22 across the Big Five.

① Result supports the use of tests of major constructs, certainly cognitive ability and conscientiousness across most occupations

(Neal Schmitt, ITC2010, HK)

64

Li, J. School of Psychology, Beijing Normal University

## 第四节 影响效度的因素

影响效度的因素很多，凡能产生随机误差和系统误差的因素都会降低测验的效度。

- ↳ 4.1 测验本身的因素
- ↳ 4.2 测验施测过程的干扰因素
- ↳ 4.3 样本团体的性质
- ↳ 4.4 测验分数与效标测量的信度
- ↳ 4.5 效标的性质

65

Li, J. School of Psychology, Beijing Normal University

## 4.1 测验本身的因素

① 测验取材的代表性、测验长度、试题类型、难度、区分度以及编排方式等都会影响效度。要保证测验具有较高效度，要做好以下几点：

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

66

Li, J. School of Psychology, Beijing Normal University

## 4.2 测验实施中的干扰因素

### 1) 主试的影响因素

- ① 是否遵从测验使用手册的各项规定进行标准化的实施
- ① 指导语是否统一正确
- ① 测验的时限是否一致
- ① 评分是否合理
- ① 测验与效标数据收集的时间间隔
- ① 施测环境如何

67

Li, J. School of Psychology, Beijing Normal University

### 2) 受测者的影响因素

- ① 受测者在测验时的兴趣、动机、情绪、态度和身心状况、健康状态以及是否充分合作与尽力而为等，都会影响受测者在测验情境中的反应，因而影响测验结果的效度。
- ① 受测者的反应定势也会降低测验的效度。

68

Li, J. School of Psychology, Beijing Normal University

## 4.3 样本团体的性质

- ① 同一个测验，效度团体的性质不同，效度也会有很大差别。

- ↳ 1) 样本团体的异质性
- ↳ 2) 干涉变量

69

Li, J. School of Psychology, Beijing Normal University

### 1) 样本团体的异质性

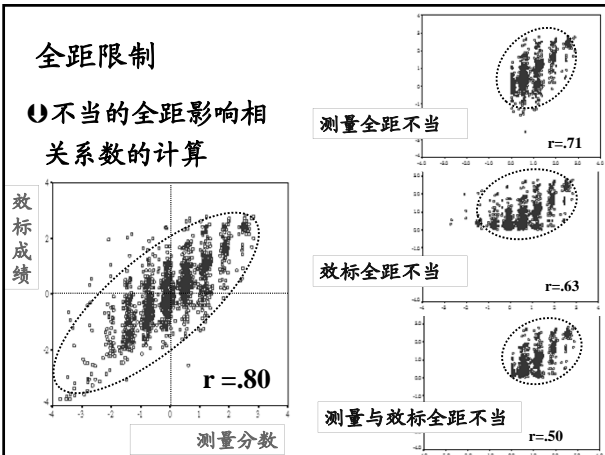
- ① 类似于信度，如果其他条件相同，样本团体越同质，分数分布范围越小，测验效度就越低；样本团体越异质，分数分布范围越大，测验效度就越高。
- ① 样本团体异质性的影响因素：
  - ↳ 只以选拔上（下）的被试为样本研究效度；
  - ↳ 只以绩优（劣）者作为样本研究效度。

70

Li, J. School of Psychology, Beijing Normal University

### 全距限制

- ① 不当的全距影响相关系数的计算



### Estimation for effects of range restriction on correlation

#### ① Guion(1965)

- ↳ 全距限制越多，对相关的低估越严重！

$$r_c = \frac{r(\sigma_\mu / \sigma_{res})}{\sqrt{1 - r^2 + r^2(\sigma_\mu^2 / \sigma_{res}^2)}}$$

$\sigma_x$ before restriction	$\sigma_x$ after restriction	Estimation of r (If sample $r = .3$ )
.8	.7	.33
.8	.6	.38
.8	.5	.45
.8	.4	.53
.8	.3	.64
.8	.2	.78

72

Li, J. School of Psychology, Beijing Normal University



## 2) 干涉变量

对于不同性质的团体，同一测验的效度会有很大的不同。这些性质包括年龄、性别、教育水平、智力、动机、兴趣、职业和任何有关的特征。由于这些特征的影响，使得测验对于不同的团体具有不同的预测能力，故测量学上称这些特征为干涉变量（moderator variable）

73

Li, J. School of Psychology, Beijing Normal University



## ① 吉赛利（E. E. Ghiselli）提出处理干涉变量的方法：

- ① 第一步：用回归方程求得每个人的预测效标分数，将该分数与实际效标分数相比较，获得差异分数D。如果D的绝对值很大，说明测验中可能存在干涉变量。
- ② 第二步：根据样本团体的组成分析，找出对照组，分别计算效度，找出干涉变量。
- ③ 第三步：对于欲测团体，根据某些易见的干涉变量将其区分为预测性高和预测性低的两个亚团体。对于预测性高的团体，获得的测验效度会有所提高。

74

Li, J. School of Psychology, Beijing Normal University



## 4.4 测验分数与效标测量的信度

① 效度受到信度的制约，如果预测源测验和效标测量两者本身信度低，也会导致对效度的低估。

$$r_c = \frac{r_{XY}}{\sqrt{r_{XX} \times r_{YY}}}$$

① 例如原有  $r_{XY}=0.30$ ， $r_{XX}=0.70$ ， $r_{YY}=0.79$

① 则校正后的  $r_c=0.38$

① 即真实的决定系数最大可能值是0.14（非0.09）

75

Li, J. School of Psychology, Beijing Normal University



## 4.5 效标的性质

① 效标效度是以测验分数与效标测量的相关系数来表示的，因此效标的性质如何，在评价测验的效度时是值得考虑的。

② 效标成绩与测验分数的相关是否为线性？

③ 效标测量的有效性如何？

76

Li, J. School of Psychology, Beijing Normal University