

2003年北京师范大学心理学研究生入学考试试题  
心理学研究方法

一、不定项选择题

1. 需要采用恒定法控制的额外变量有
  - a. 实验室物理环境
  - b. 实验过程控制
  - c. 被试的主观态度
  - d. 仪器的性能指标
2. 在一个混合实验设计中，交互作用不显著说明
  - a. 各因素之间不存在内在联系
  - b. 各因素的主效应是相互独立的
  - c. 各因素之间可能存在内在联系
  - d. 各因素的主效应可能会相互影响
3. 在复杂任务的行为中，反应时间常包括
  - a. 基线时间
  - b. 一个或者多个认知加工过程的时间
  - c. 反应键的延迟反应时间
  - d. 其他随机因素导致的反应时间误差
4. 根据信号检测论，我们之所以感受不到信号的存在是因为
  - a. 人自身生理局限
  - b. 有背景噪音的干扰
  - c. a+b
  - d. 被试反应不认真
5. 如果要被试对正确肯定的句子和正确否定的句子做反应，反应速度可能会
  - a. 正确肯定 > 正确否定
  - b. 肯定
  - c. 否定
  - d. 不确定
6. 拉丁方设计的主要作用是
  - a. 平衡实验中可能产生的各种误差
  - b. 降低实验误差
  - c. 收集更多的实验数据
  - d. 合理分配实验材料或者实验处理
7. 练习实验的主要作用是
  - a. 帮助被试学习和掌握反应原则
  - b. 降低实验误差
  - c. 提高被试反应精确性和可靠性
  - d. 练习实验没有什么作用
8. 因素设计的主要特点是
  - a. 可以观察因素之间的相互作用
  - b. 揭示不同因素之间的内在联系
  - c. 获得更生态化的实验结果
  - d. 实验设计与实施相对复杂
9. 下面属于主试效应的有
  - a. 罗森塔尔效应
  - b. 空间误差

c. 被试数量太少

d. 没有考察交互作用

二、简答题（1至11题五分，12至16题七分）

1. 正态分布的标准差有何统计意义，在统计检验中为什么会用到标准差？

2. 正态分布的特征是什么，统计检验中为什么经常要将正态分布转化成标准正态分布？

3. 标准参照测验和常模参照测验的区别是什么，你认为会计师职业资格考试是什么性质的考试。

4. 为什么说测验的信度高是效度高的必要条件而非充分条件？

5. 在调查研究中，如果取样范围比较狭小（如研究中国儿童的问题，只在大中城市选取样本），一般会对相关研究的结果产生什么影响。

6. 学习的智力和学习成绩高度相关，能否以为智力测验分数预测学生的学习成绩，如果不能说明理由，如果能，说明怎样建立具体的预测方程。

7. 某单位在两地招考新生，使用了同一测验工具，但因阅卷者对评分标准理解和掌握不一致，致使两地考生测验的平均分相差很大，重新评分已不可能能否对分数加以统计处理，以便解决这一问题，是录取更公平。

8. 确定常模团体时，应注意哪些问题。

9. 在进行差异的显著性检验时，若将相关样本误作独立样本处理，对差异的显著性有何影响，为什么？

10. 为什么要做去间估计怎样对平均数作区间估计。

11. 为什么调查研究只能确定相关关系，不能确定因果关系？既然如此，调查研究有什么意义？

12. 测验为什么要标准化，怎样保证测验的标准化。

13. 2002年10月29日，《江南日报》发布中华英才网的调查报告，调查结果显示南京职工的人均月薪已达2690元，有人认为这一结果高估了南京人的月收入。你怎么看这个结果，试分析高估的原因。

14. 抽样调查要想得到比较准确的结果，需要控制哪些技术环节？

15. 测验分数为什么要合成，说明合成测验分数的主要方法。

16. T检验 F检验卡方各自适用于什么情况？

三、问答题（每题十分）

1. 在心理学实验中，随机误差和系统误差对实验结果可能产生什么影响。

2. 阐述心理学设计中平衡误差的方法。

3. 阐述减数法的基本思想及其应用。在表象心理旋转实验中，如何运用减数法的原理获得表象心理旋转的速度。

4. 采用信号检测论的方法设计一个实验，考察某一群体被试对图形或文字材料是否存在。

## 参考答案

### 2003年北京师范大学心理学研究生入学考试试题 心理学研究方法

#### 一、不定项选择题

1. 需要采用恒定法控制的额外变量有

a. 实验室物理环境

b. 实验过程控制

c. 被试的主观态度

d. 仪器的性能指标

2. 在一个混合实验设计中，交互作用不显著说明

- a. 各因素之间不存在内在联系
  - b. 各因素的主效应是相互独立的
  - c. 各因素之间可能存在内在联系
  - d. 各因素的主效应可能会相互影响
3. 在复杂任务的行为中，反应时间常包括
- a. 基线时间
  - b. 一个或者多个认知加工过程的时间
  - c. 反应键的延迟反应时间
  - d. 其他随机因素导致的反应时间误差
4. 根据信号检测论，我们之所以感受不到信号的存在是因为
- a. 人自身生理局限
  - b. 有背景噪音的干扰
  - c. a+b
  - d. 被试反应不认真
5. 如果要被试对正确肯定的句子和正确否定的句子做反应，反应速度可能会
- a. 正确肯定 > 正确否定
  - b. 肯定
  - c. 否定
  - d. 不确定
6. 拉丁方设计的主要作用是
- a. 平衡实验中可能产生的各种误差
  - b. 降低实验误差
  - c. 收集更多的实验数据
  - d. 合理分配实验材料或者实验处理
7. 练习实验的主要作用是
- a. 帮助被试学习和掌握反应原则
  - b. 降低实验误差
  - c. 提高被试反应精确性和可靠性
  - d. 练习实验没有什么作用
8. 因素设计的主要特点是
- a. 可以观察因素之间的相互作用
  - b. 揭示不同因素之间的内在联系
  - c. 获得更生态化的实验结果
  - d. 实验设计与实施相对复杂
9. 下面属于主试效应的有
- a. 罗森塔尔效应
  - b. 空间误差
  - c. 被试数量太少
  - d. 没有考察交互作用

答案： 1. a b d    2. b    3. a b c    4. a b c    5. a    6. a    7. a b c  
8. a b c d    9. a

## 二、简答题（1至11题五分，12至16题七分）

### 1. 正态分布的标准差有何统计意义，在统计检验中为什么会用到标准差？

答：指表示数据分散程度的一组统计量，即离中趋势量数的选择与计算。其中最常用的

是标准差或方差。标准差的常用符号为： $SD$ 、 $S$ （统计量）， $\sigma$ （总体参数）；方差的常用符号为： $S^2$ 、 $SD^2$ （统计量）， $\sigma^2$ （总体参数）。

（1）表示数据的分散程度，标准差大表示分散，标准差小表示相对集中。

①若一个班的分数之标准差大，说明该班学习成绩不齐，好的好，差的差。此时标准差小好，说明成绩整齐。

②若一个老师所出的试卷，学生考完后标准大，说明这张试卷出得好，把不同学生的水平区分开了。此时标准差小不好。

③同一测量的标准差大，说明误差较大。

（2）标准差的应用：在正态分布的情况下标准差与平均数之间有一定关系；用于个别数据的取舍；变异系数。

（3）方差和标准差的优点：反应灵敏，随任何一个数据的变化而表示；一组数据的方差和标准差有确定的值；计算简单；适合代数计算，不仅求方差和标准差的过程中可以进行代数运算，而且可以将几个方差和标准差综合成一个总的方差和标准差；用样本数据推断总体差异量时，方差和标准差是最好的估计量。

在正态分布中，用标准差可以表示：不同观测值在原有数据团体中的相对位置。一般三个标准差以外的认为是异常值，作为取舍依据。

## 2. 正态分布的特征是什么，统计检验中为什么经常要将正态分布转化成标准正态分布？

答：（1）如果连续型随机变量  $X$  的概率密度为：

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{其中 } \mu, \sigma \text{ 为常数, 且 } \sigma > 0, \text{ 我们称 } X \text{ 服从正态分布。}$$

正态分布在概率统计中占有重要的位置，现实中如果随机变量的分布具有“中间大，两头小”的特点，都可以认为服从正态分布。

正态分布也叫常态分布或常态分配。是连续随机变量概率分布的一种。正态分布的形式是对称的。它的对称轴是过平均数点的垂线。正态分布中，平均数、中数、众数三者相等，此点  $Y$  最大。正态分布的中央点最高，然后逐渐向两侧下降。为一族对称分布，曲线下面积为 1。

（2）因为  $X$  不是标准正态分布，所以不可直接查表求出结果，必须先将  $X$  化为标准正态分布，然后再用公式计算或查表求出概率值。标准正态分布是  $N(0, 1)$  的正态分布。这时密度函数和面积（概率）计算比较方便，常常转化为标准正态分布。正态分布只是一个理想的分布，所以需要原始分数导出。

## 3. 标准参照测验和常模参照测验的区别是什么，你认为会计师职业资格考试是什么性质的考试。

答：（1）标准是指在编制测验和解释测验分数时所依据的知识和技能领域，而不是指分数的分界标准。标准参照性测验就是在对测验结果进行评价时不是以常模为标准，而是根据特定的操作标准和行为领域，对个体做出是否达标或达到什么程度的判断，而不考虑他人分数的测验。

标准参照测验是将被试的分数与某种标准进行比较来解释。这种测验常常用来检验学习效果，看对指定的内容范围掌握得如何或达到某一标准。衡量测验优劣得主要指标是内容效度。

（2）常模是测验分数的总体分布形态，一般用测验分数的平均数和标准差表示。用常模可以确定一个被试测验分数的相对高低，即他在所属群体的能力或知识连续体上的相对位置。常模参照性测验就是以常模为评价测验分数优劣标准的测验，常模被视为测验分数的参

照，它关心的不是一个人能力或知识的绝对水平，而是他在所属群体的能力或知识连续体上的相对位置。

常模参照测验是将一个人的分数与其他人比较，看其在某一团体中所处的位置。也就是把受测者的成绩与具有某种特征的人所组成的有关团体做比较，根据一个人在团体内的相对位置来报告他的成绩。所谓常模就是指具有某种共同特征的人组成的一个群体。

(3) 会计师职业资格考试属于标准参照测验。这种考试并不要求该考生与同时参加一类考试的其他考生进行比较，通过比较获得考生在团体中的相对位置，比如说排名等。而是以考试大纲为依据，确定试题范围为测试考生是否达到某一级别的要求，从而划定一定的分数作为标准(比如60分)，达到这个标准的考生即认为是合格的，可以颁发相应的资格证书。因此这种考试关键取决于试题的内容效度，即内容是否能衡量考生的真实水平。

#### 4. 为什么说测验的信度高是效度高的必要条件而非充分条件?

**答:**(1) 信度指测验结果的可靠性或可靠程度。所谓可靠性是指对同一组对象进行两次相同测量所得结果的一致性和稳定性程度。可见，信度是实测值和真实值之间差距大小之间的量度，信度值在(0~1)之间，测验信度越高，误差就越小，测验的结果也就越可靠。影响信度的因素很多：被试方面有身心健康状况、注意力、持久性、态度等；主试方面有不按规定程序、制造紧张、评判主观等；研究设计方面有取样不当、题目过少、陈述不清等；实施方面有环境的各种难以控制的变化条件。判定研究信度的方法主要有重复法、相似法和独立评判法。

(2) 效度指测验结果的有效性或正确性。所谓有效性，是指一种测验能够正确的测量出它所要测量的特性或功能的程度。可见，效度要解决两个问题：本测验的测量目标是什么；本测验对于所要测量的目标能测到什么程度。效度的值在[0, 1]之间。常见的效度主要有构思效度、内部效度、统计结论效度和外部效度。

构思效度是指理论构思的合理性及其转换为抽象与操作定义的恰当性程度；研究的内部效度是指在研究的自变量与因变量之间存在一定关系的明确程度；统计结论效度是关于研究的数据分析处理程序的效度检验，或者说，它是检验研究结果的数据分析程序与方法的有效性的指标；研究的外部效度是指研究结果能够一般化和普遍适用到样本来自的总体和到其它的总体中的程度，即研究结果和变量条件、时间和背景的代表性和普遍适用性。

(3) 两者之间的关系：信度是效度的必要条件，但不是充分条件。

效度是测量准确性的量度，信度是测量结果一致性或稳定性的量度，两者之间存在着辩证的关系。信度低，效度不可能高。因为如果测量的数据不准确，也不能有效地说明所研究的对象。比如一个研究使用的数据收集方法本身有问题，所得的数据自然就不具有说服力，更谈不上有效了；信度高，效度未必高。例如，如果我们准确地测量出某人的经济收入，也未必能够说明他的消费水平；效度低，信度很可能高。例如，即使一项研究未能说明社会流动的原因，但它很有可能很精确很可靠地调查各个时期各种类型的人的流动数量；效度高，信度也必然高。

影响信度的因素都会影响效度。测量的效度除受随机误差影响外，还受系统误差影响。可信的测验未必有效，而有效测验必定可信。因此，信度是效度的必要条件而非充分条件。例如：一个量尺对量身高是比较可信的，但用来量身高却是无效的。

#### 5. 在调查研究中，如果取样范围比较狭小（如研究中国儿童的问题，只在大中城市选取样本），一般会对相关研究的结果产生什么影响。

**答:**当取样范围比较狭小（如研究中国儿童的问题，只在大中城市选取样本）时，所取样本不具有代表性，用这些样本的分析结果来研究整个总体是不科学的，存在取样误差。

(1) 取样的基本原则是随机性原则，所谓随机性原则，是指在进行抽样时，总体中每一个体是否被抽选的概率（即可能性）是完全均等的。由于随机抽样使每个个体有同等机会

被抽取,因而有相当大的可能使样本保持和总体有相同的结构,或者说,具有最大的可能使总体的某些特征在样本中得以表现,从而保证由样本推论总体。此外,遵循随机性原则进行抽样,可以对抽样误差的范围进行预算或控制,使研究者能够客观地评价研究结果的精确度和按照所要求的精确度决定样本的容量大小。

(2) 在调查研究中,取样的基本方法有以下几种:

①简单随机取样法

简单随机取样法就是按随机原则直接从总体  $n$  个单位中,抽取几个单位作为样本。常见的有“抽彩法”和“随机表取样法”。这是最简便的取样方法。随机取样又分为置换性随机取样和非置换性随机取样两种。

②系统随机取样法

系统随机取样法是从总体中取一随机起点,从该起点开始选取每  $K$  项元素(个数或分数等),直至取满所需要的样本量。 $K$  值一般以样本量和总体量的比值来确定。系统随机取样法较简单随机取样法更为简便,误差更小,但如果总体中存在周期性的波动或变化,则不宜采用。

③分层随机取样法

分层随机取样法就是先将总体分为若干层次或子总体,然后独立、随机地从每一层次抽取样本。分层随机取样法的代表性和推论的精确性较好,但作出科学的分类则较为困难。

④聚类取样法

聚类取样法是将总体各单位按一定的标准分成许多群,然后按随机原则从这些群中抽取若干群作为样本。该方法的样本单位比较集中,易于组织,适合规模较大的调查研究,但其取样误差常因各群间差异较大而比较大。

⑤多段随机取样法

多段随机取样法是将研究总体各单位按一定标准分为若干群,作为取样的第一级单位,然后再按一定标准将第一级单位分成若干子群,作为取样的第二级单位,如此类推。在各级单位中依照随机原则抽取样本。该方法可综合运用各种取样方法,在研究总体范围大、单位多、情况复杂时十分有用,但取样误差较大。

(3) 取样误差的计算

取样误差是指样本统计值与总体相应参数值之间的差异。取样误差越小,则根据样本的研究结果对总体作出的推论就越为精确,取样就越为成功。取样误差的计算公式为:

$$S_x = \frac{S}{\sqrt{n}}$$

其中,  $S_x$  为取样误差(即样本标准误),  $S$  为样本的标准差,  $n$  为样本容量。当小样本时,可用“ $n-1$ ”代替。一般地,对于特定的研究总体来说,要减小取样误差,就必须增大样本容量。

(4) 样本大小的确定

一般观点认为,只有采用大样本,才能从一定研究数据作出有效的结论;但方法论研究证明,当样本增大到一定程度时,更大的样本对提高研究的准确性的作用就不显著了。样本大小一般取决于以下几方面因素:①研究的类型;②预定分析的精确程度;③允许误差的大小;④总体的同质性;⑤研究者的时间、人力和物力;⑥取样方法等。如果要求的精度高,允许的误差值小,研究总体的异质性大,就要考虑使用较大的样本。

**6. 学习的智力和学习成绩高度相关,能否以为智力测验分数预测学生的学习成绩,如果不能说明理由,如果能,说明怎样建立具体的预测方程。**

答:(1)心理学上把智力定义为,是人们在获得知识和运用知识解决实际问题时所必须

具备的心理条件或特征。智力是一个综合性的概念。

在衡量智力测验时常常涉及一些心理学问题，对这些问题的正确理解，无论对于测验的编制者还是使用者都是十分重要的。

① 智力并不是静止不变的。人的智力不是一成不变的，它不但随着年龄或成熟而发展，还可因教育和训练而改变。智力发展的一般趋势是：智力在小学急速增长，在青春期和成年初期增长缓慢，在成年期达到最大值，保持稳定到中年后期，在老年稍有下降。

② 智力之间存在着诸多的差异。比如：智力的个别差异，研究表明：在一般人群中智力极高（IQ 在 140 以上）与极低（IQ 在 70 以下）者均占少数，智力属于中等或接近中等（IQ 在 80~120）之间者约占全部人口的 80%。智力的团体差异，比如：性别差异、种族差异等。多数学者认为，智力差异都是在遗传与环境两大因素支配下经由成熟与学习交互作用形成的结果。

### ③ 智力测验的公平性问题

智力测验经常受到的批评是：被试在测验上的反应受知识经验的影响，因此对文化背景、教育水平不同的团体是不公平的。

实际上，任何一个测验都不能对所有的团体同样公平，因为每种测验都是在一定文化背景下发展起来的。最好的办法是根据亚文化群的特点为不同的团体编出不同的测验；若使用同一个测验，则要为少数特殊团体制定单独的常模；在从外来文化引进测验时，简单翻译是不够的，还必须对内容加以修订，使之适合本国国情和本民族文化的特点。总的要求是，使每个人都同有类似背景和学习经历的人比较，此种测验叫做特定文化测验。尽管如此，所谓公平仍然是相对的，因为团体有不同的层次，不可能为每一个团体都制定单独的测验或常模。

(2) 智力测验是通过设计量表对智力的某些方面指标进行评价。目前倾向于综合能力的评价。智力测验能从

某种角度上表明儿童在同龄常模群体中的大致位置，可以作为一些参考。但是由于智力测验也具有局限性，因此不能迷信智力测验。学习的智力和学习成绩高度相关，并不能说明二者之间存在因果关系，因此很难以智力测验分数预测学生的学习成绩。

在教学实际中，应该以学生的实际表现等能力发展情况作综合评价。学生是一个发展变化的过程，智力是一个发展变化的过程，所以用智力测验分数预测学生的学习成绩是不科学的。

**7. 某单位在两地招考新生，使用了同一测验工具，但因阅卷者对评分标准理解和掌握不一致，致使两地考生测验的平均分相差很大，重新评分已不可能。能否对分数加以统计处理，以便解决这一问题，是录取更公平。**

**答：**由于使用的是同一测验工具，只是阅卷者对评分标准理解和掌握不一致。所以可以通过将分数合成和解释来达到使分数公平的目的，具体可以通过一些统计方法实现。

(1) 在使用测验时，常常需要将几个分数或几个预测源组合起来以获得一个合成分数或作出总的预测。分数的组合可以在不同层次上进行。

### ① 题目的组合

每个测验都包含许多独立的题目，除非测验使用者对个别题目特别感兴趣，否则总要把各个题目分数组合起来。不同的题目可以组成量表或分测验，而得到量表分或分测验分；所有题目也可以合成一个测验总分。在组合时，对各题目可以等量加权（如选择题），也可以不等量加权（如问答题）。

### ② 分测验或量表的组合

有些测验是由几个分测验或量表组成的，每个分测验或量表都有自己的分数，这些分数可以组合到一起得到一个合成的分数。如韦氏成人智力量表有 11 个分测验，其中 6 个分测验构成言语量表，其合成分数叫做言语智商，另外 5 个分测验构成操作量表，其合成分数叫

做操作智商，此外还将所有 11 个测验分数合成而得到总智商。但有时各量表分也可单独使用而不必合成，如从职业兴趣测验上得到的各科分数就不需要合成。

### ③ 测验或预测源的组合

在做实际决定时，常常将几个测验或预测源同时使用。如美国雇佣服务中心对申请者实施 12 个测验，用来预测在各种职业上的成功。

由于测量目的和所用材料不仅仅是统计的，也可以是推理的或直觉的。包括临床判断、推理方法（包括单位加权和等量加权）、多重分段（包括综合分段和连续栅栏）、多重回归、区分分析、因素分析、多元变量技术及完形计分、轮廓分析等。

采用哪种组合方法取决于使用测验的目的。测验的目的可以二分为预测和描述，前者是用测验分数来预测某种效标行为。后者是用测验分数对人的某种行为做出一般性的描述。在用于预测时，还可进一步分成两类问题：选人问题，分数与安置问题。前者是从申请人中挑选出最佳者，后者是将每个人分派至最适当的位置或类别。

#### (2) 可以将测得的分数转化为标准分数

标准分数是将原始分数与平均数的距离以标准差为单位表示出来的量表。因为它的基本单位是标准差，所以叫标准分数。标准分数可以通过线形转换或者通过非线性转换得到，由此可将标准分数分为两类：

##### ① 线性转换的标准分数

根据标准分数的定义，可通过下式将原始分数直接转换成标准分数：

$$Z = (x - X) / S$$

因为标准分数是从原始分数（ $x$ ）中减去一个恒定值（平均数  $X$ ）再除以一个恒定值（标准差  $S$ ）得到的，所以这是一种线形转换。有时也把线性转换的标准分数简称作标准分数或  $Z$  分数。

##### ② 常态化的标准分数

当原始分数不是常态时，为了使导出的量表分数呈常态分布，可先把原始分数转化为百分等级，然后从正态曲线面积表中便可得到对应的标准分数。由这种方式所得到的分数就叫常态化的标准分数。在将分数常态化时有一个前提：只有所测特质的分数实际上应该是常态分布，只是由于测验本身的缺陷或取样误差而使分布稍有偏斜时，才能计算常态化标准分数。

$T$  分数： $T$  分数以 50 为平均数，以 10 为标准差。

$$T = 50 + 10Z$$

标准九：标准九的全称为标准化九级分制，是一种比  $T$  分数粗糙的量表。它以 5 为平均数，以 2 为标准差。

#### (3) 统计处理方法

① 可以将全部成绩汇总，计算标准分来将全部分数转化为正态分数，根据计算的结果，查正态分布表，计算其在正态分布图中的相对位置。

② 将两地的分数分别考虑，利用各自的分布进行统计，确定考生在各自团体中的相对位置，然后根据录取比例分配两地的录取人数。

### 8. 确定常模团体时，应注意哪些问题。

**答：**常模团体是由具有某种共同特征的人所组成的一组群体，或是该群体的一个样本。确定和选择常模时，要注意常模的要求：

#### (1) 群体必须明确

一个测验可能有许多常模团体，而确定多少常模团体，依赖于对测验将要施测的群体的认识。在告诉测验使用者时，也同样必须有对每个常模团体的性质与特征的简短而明确的叙述。假如常模团体包括某种特征行业或职位时，则职务高低、行业种类、被试资历以及有关材料也应包含在内。

(2) 常模团体必须是所测群体的代表性样本

常模团体缺乏代表性, 会使常模资料产生偏差而影响对测验分数的解释。为了克服取样偏差, 在搜集常模资料时, 一般采用随机取样或分层取样的方法, 有时也可把两种策略结合起来使用。如要把一智力测验应用于全国, 在常模团体的组成上, 城乡比例、男女比例、不同年龄、职业、民族, 以及不同文化程度、社会经济地位等的人数比例, 都应与人普查的资料相符。在保证上述比例的前提下, 再作随机取样。

(3) 样本大小要适当

所谓“大小适当”并没有明确的指标。一般是从经济的或实用的可能性和减少误差这两方面来综合考虑样本大小的。常模大小取决于总体的数目、研究需要和群体性质, 以及试测的结果。

(4) 注意常模的时间性

在考虑常模的合适性时, 不能忽略对常模时间性的要求。由于几年前所编制的常模往往不适于现在的要求, 因此常模必须定期修订。在选择合适常模时, 注意选择较为新近的常模。

(5) 注意一般常模与特殊常模的结合

测验手册上所列的常模通常是典型团体建立的, 比较一般化, 不一定适合使用者的具体情况, 对此问题的一个解决办法是为每一个特定目的的测验建立特殊常模。特殊常模是为典型团体建立的, 一般比子团体建立的常模范围更窄。特殊常模的优点是, 将个人同与其背景相近的人比较, 但这同时也是它的缺点, 不容许分数在较广的范围内作解释。不过, 测验使用者可将特殊常模与一般常模结合起来, 从而获得最大量的信息。

**9. 在进行差异的显著性检验时, 若将相关样本误作独立样本处理, 对差异的显著性有何影响, 为什么?**

**答:** 在进行差异的显著性检验时, 首先需要考虑样本是否正态分布, 如果是正态分布, 还需要考虑总体方差是否已知, 然后看样本是否是独立样本。

根据相关样本与独立样本不同, 会运用不同的计算方法计算显著性。

相关样本与独立样本是根据两个样本是否来自同一个总体来划分的。相关样本之间存在着一一的对应关系。如果是相关样本, 则两组数据的离散程度是相同的。如果是独立样本, 则两组数据的离散程度不同, 需要进行方差的差异性检验。

如果将相关样本误作独立样本处理, 则忽视了样本数据之间的一致性, 导致错误地运用计算公式, 差异的显著性也会受到误估。

**10. 为什么要做区间估计? 怎样对平均数作区间估计。**

**答:** (1) 区间估计是用数轴上的一段距离表示未知参数可能落入的范围, 它不具体指出总体参数等于多少, 但能指出总体的未知参数落入某一区间的概率有多大。进行区间估计主要是想通过样本提供的一些信息, 对总体特征进行估计。同时这也是根据数学的概率原理。

进行区间估计需要考虑置信区间和显著性水平。置信区间或称置信间距, 是指在某一置信度时, 总体参数所在的区域距离或区域长度。置信度又称显著性水平, 意义阶段, 信度系数等, 是指估计总体参数落在某一区间时, 可能犯错误的概率, 用符号  $\alpha$  表示。

(2) 区间估计实质是根据样本分布的理论, 用样本分布的标准误 (SE) 计算区间长度, 解释总体参数落入某置信区间可能的概率。区间估计存在成功估计的概率大小及估计范围大小两个问题。依据小概率事件原理, 0.01 或 0.05 属于小概率事件, 在一次抽样中是不可能出现的。

(3) 对平均数的区间估计, 首先要考虑样本或总体平均数抽样的分布, 如  $t$  分布, 确定查什么统计表。然后求标准差, 根据区间估计的公式求出置信区间。所用的公式有:

$$\bar{X} - Z_{1-\frac{\alpha}{2}}\sigma_{\bar{X}} < u < \bar{X} + Z_{1-\frac{\alpha}{2}}\sigma_{\bar{X}}$$

### 11. 为什么调查研究只能确定相关关系，不能确定因果关系？既然如此，调查研究有什么意义？

答：（1）心理学研究中常用的调查研究方法有以下几种：

①观察法是心理学研究中最基本、最普遍、历史最悠久的方法之一。观察是日常生活和科学研究不可缺少的活动和手段。观察法是研究者通过感官或借助于一定的科学仪器，在一定时间内有目的、有计划地考察和描述人的各种心理活动和行为表现并收集研究资料的一种方法。

观察法的类型，根据不同的划分角度可以有不同的分类，如根据是否借助仪器可分成直接观察与间接观察，根据观察地点可分为实地观察和实验室观察，根据观察结果可分为量的观察与质的观察等等。

观察法具有：能动性、选择性和客观性等特点，但也有一些局限性，如人的生理局限、观察仪器的局限、观察者对所获材料的解释的主观性等。

②访谈法是指调查者依据调查提纲与调查对象直接交谈，收集语言资料的方法，是一种口头交流式的调查方法。访谈法的主要特点是：采用对话、讨论等面对面的交往方式，是双方相互作用、相互影响的过程。访谈法的另一个显著特点是它具有特定的科学目的和一整套设计、编制和实施的原则。它保证了访谈法的科学性、有效性、访谈结论的客观性。

③问卷法是通过书面形式、以严格设计的心理测量项目或问题，向研究对象收集研究资料和数据的一种方法。问卷法是心理学研究中使用最为普遍的方法之一。问卷法有两大特点：一是标准化程度较高，整个研究过程严格按照一定的原则和要求进行，从而保证了研究的准确性和有效性，避免了盲目性和主观性；二是能在短时间内收集到大量的资料并得出结论。

④测验法是指使用特定的量表为工具，对个体的心理特征进行间接了解，并做出量化结论的研究方法。使用测验法，可以研究：了解个体或团体的心理特征，如用智力量表测量儿童的智力水平，用人格量表了解人各不同的心理特征；探讨心理特征与外界因素的关系，如考察智力与学习成绩是否相关，性格内向是否影响社会交往；可以比较不同个体或团体之间的心理差异。

调查研究是相对于实验室研究的研究方法。通过调查研究可以获得大量的实验数据，这些研究具有真实性、系统性和全面性等特点，因此是心理学研究中的重要手段。

（2）调查研究只能确定相关关系，不能确定因果关系，是因为相关性只是表明两个变量之间的变化方向。如果两个变量的变化方向没有什么联系，说明两者没有相关；如果两变量变化方向相反，则呈负相关；如果两变量呈现一致的变化方向，则呈正相关。我们的调查研究很多时候只需要发现变量之间的相关性，以此来确定变量之间的关系。

事物之间的因果关系是相当复杂的，在实际研究中很难完全寻找发现出变量之间的完全因果关系。调查研究的目的是为了发现变量之间的相关程度，为以后的实证研究提供相关的信息资料。因此是特别重要的。

### 12. 测验为什么要标准化，怎样保证测验的标准化。

答：（1）测验标准化是因为：

- ① 可以用等距量表来表示测验分数，使进一步统计成为可能。
- ② 常态化标准分数可参照常态曲线面积表直接转换成百分等级，因而容易解释。
- ③ 允许将几个测验或量表上的分数作直接的比较。

（2）测验的标准化包括测验编制的标准化和测验使用的标准化。

①测验编制的标准化主要是指编制测验的一般程序：

a. 确定测验的目的与对象。验的目的在于指测验作什么用。测验有许多不同的功用，测验的编制程序由于其目的不同而有许多变化。比如，常模参照性测验与标准参照性测验在测验编制过程中就有许多不同的侧重点。在编制测验时，还应该考虑

受测团体的组成和特点。例如被试的年龄、智力水平、受教育程度、社会经济和文化背景以及阅读水平等。不同的变因，对测验的难度、取样范围、形式等要求也不同。

b. 确定并分析测量的目标。测量目标是指所编测验用来测量什么心理变量。在实际工作中，测验的编制者必须把测试目标转换成可操作的术语。这种过程一般可分三种情况：一是编制带有预测功能的心理测验；二是编制针对某种心理特质的测验，首先要给所测量的心理特质下一操作定义，然后计划通过哪些方面加以测量。三是编制标准参照测验，首先要确定测验的内容以及每项内容应掌握的程度（行为目标），然后设计一个双向细目表，作为编制测验的蓝图。

c. 测验题目的收集与编写。测验编制过程中最重要的步骤是测验题目的编写和搜集。一般认为，为了有所筛选，编写出多于所需题目一倍到几倍的题目也不为过。

d. 测验的编排。在测验题目编写完成之后，需请有关专家进行检查，并根据测验目的、性质与功能来选择和编排测题。具体编排形式可以采用并列直进式，也可以采用混合螺旋式。

e. 测验的预试和题目分析。为了获得被试对题目反应的材料，进一步筛选题目和为编排测验提供客观依据，必须将预备测验对一组被试施测。预试的目的一方面是了解受测者对测验的适应情况，借以发现测验编制本身的缺陷；另一方面则主要是对测验及题目进行统计分析，取得必要的数量指标，以便筛选题目、组成正式测验。预试完成后，可以根据预试结果进行题目分析。题目分析主要是确定题目的难度、区分度、备选答案的合适度等数量指标。根据分析结果，再行筛选题目，编制出符合要求的正式测验。

f. 编制测验复本。在实际工作中经常遇到要多次使用某个测验的情况。必须有等值的测验可以替换，这就需要编制测验复本。

②为了使测量结果准确可靠，必须将测验的使用标准化，为此要注意以下几点：

a. 确保测验实施过程的规范性与统一性，包括测验指导语、测验时间、测验情境等。

b. 制定客观的评分标准与解释分数的常模。

c. 务必使测验的效度与信度达到一定的要求。

d. 编写测验指导手册

**13. 2002年10月29日，《江南日报》发布中华英才网的调查报告，调查结果显示南京职工的人均月薪已达2690元，有人认为这一结果高估了南京人的月收入。你怎么看这个结果，试分析高估的原因。**

**答：**如果是高估，那么可能是因为在调查的过程中选择的样本不具有代表性。也就是说选择到的样本只是一部分月薪偏高的，不能代表整个南京人的平均水平。从根本上说是因为没有在抽样调查中运用随机化原则。

抽样调查是根据部分实际调查结果来推断总体标志总量的一种统计调查方法，属于非全面调查的范畴。它是按照科学的原理和计算，从若干单位组成的事物总体中，抽取部分样本单位来进行调查、观察，用所得到的调查标志的数据以代表总体，推断总体。

与其它调查一样，抽样调查也会遇到调查的误差和偏误问题。通常抽样调查的误差有两种：一种是工作误差（也称登记误差或调查误差），一种是代表性误差（也称抽样误差）。但是，抽样调查可以通过抽样设计，通过计算并采用一系列科学的方法，把代表性误差控制在允许的范围之内；另外，由于调查单位少，代表性强，所需调查人员少，工作误差比全面调查要小。特别是在总体包括的调查单位较多的情况下，抽样调查结果的准确性一般高于全面调查。因此，抽样调查的结果是非常可靠的。

抽样调查数据之所以能用来代表和推算总体,主要是因为抽样调查本身具有其它非全面调查所不具备的特点,主要是:

(1) 调查样本是按随机的原则抽取的,在总体中每一个单位被抽取的机会是均等的,因此,能够保证被抽中的单位在总体中的均匀分布,不致出现倾向性误差,代表性强。

(2) 是以抽取的全部样本单位作为一个“代表团”,用整个“代表团”来代表总体。而不是用随意挑选的个别单位代表总体。

(3) 所抽选的调查样本数量,是根据调查误差的要求,经过科学的计算确定的,在调查样本的数量上有可靠的保证。

(4) 抽样调查的误差,是在调查前就可以根据调查样本数量和总体中各单位之间的差异程度进行计算,并控制在允许范围以内,调查结果的准确程度较高。

基于以上特点,抽样调查被公认为是非全面调查方法中用来推算和代表总体的最完善、最有科学根据的调查方法。

但是,要使抽样调查更具有代表性,需要在其中运用随机化原则,即在抽样时,总体中每一个个体是否被抽取,并不由研究者主观决定,而是每一个个体按照概率原理被抽取的可能性相等的。只有这样,才能根据对样本统计量的分析,以样本统计量来估计总体参数。

#### 14. 抽样调查要想得到比较准确的结果,需要控制哪些技术环节?

**答:** (1) 抽样调查是根据部分实际调查结果来推断总体标志总量的一种统计调查方法,属于非全面调查的范畴。它是按照科学的原理和计算,从若干单位组成的事物总体中,抽取部分样本单位来进行调查、观察,用所得到的调查标志的数据以代表总体,推断总体。

(2) 与其它调查一样,抽样调查也会遇到调查的误差和偏误问题。通常抽样调查的误差有两种:一种是工作误差(也称登记误差或调查误差),一种是代表性误差(也称抽样误差)。但是,抽样调查可以通过抽样设计,通过计算并采用一系列科学的方法,把代表性误差控制在允许的范围之内;另外,由于调查单位少,代表性强,所需调查人员少,工作误差比全面调查要小。特别是在总体包括的调查单位较多的情况下,抽样调查结果的准确性一般高于全面调查。因此,抽样调查的结果是非常可靠的。

抽样调查数据之所以能用来代表和推算总体,主要是因为抽样调查本身具有其它非全面调查所不具备的特点,主要是:

① 调查样本是按随机的原则抽取的,在总体中每一个单位被抽取的机会是均等的,因此,能够保证被抽中的单位在总体中的均匀分布,不致出现倾向性误差,代表性强。

② 是以抽取的全部样本单位作为一个“代表团”,用整个“代表团”来代表总体。而不是用随意挑选的个别单位代表总体。

③ 所抽选的调查样本数量,是根据调查误差的要求,经过科学的计算确定的,在调查样本的数量上有可靠的保证。

④ 抽样调查的误差,是在调查前就可以根据调查样本数量和总体中各单位之间的差异程度进行计算,并控制在允许范围以内,调查结果的准确程度较高。

基于以上特点,抽样调查被公认为是非全面调查方法中用来推算和代表总体的最完善、最有科学根据的调查方法。

(3) 抽样调查要想得到比较准确的结果,需要控制的技术环节:

① 随机化原则:在抽样时,总体中每一个个体是否被抽取,并不由研究者主观决定,而是每一个个体按照概率原理被抽取的可能性相等的。只有这样,才能根据对样本统计量的分析,以样本统计量来估计总体参数。

② 根据实际需要选定具体方法,如:简单随机抽样、分层随机抽样或者两阶段随机抽样。

③ 根据实际情况确定样本的容量。样本容量太大太小都不利于研究。确定样本容量要

依据对参数的估计和一定的假设检验进行。

### 15. 测验分数为什么要合成，说明合成测验分数的主要方法。

答：使用测验时，需要将几个分数或几个预测源组合起来获得一个合成分数或作出总的预测。包括项目的组合、分测验或量表的组合、测验或预测源的组合等。

合成测验分数的主要方法有：

(1) 临床判断。在实际工作中，人们常常会根据经验对各变量做直觉的组合。这时，研究者是根据直觉经验，主观地将各种因素组合以得出结论或预测得方法。临床法的优点是能从整体上对各因素加以综合考虑，不但考虑到各个因素的相当重要性，还能考虑到各因素间的交互作用，这种考虑具有高度完形性质；同时，每个判断都是针对特定的个人做出的，能考虑到每个人的具体情况。临床法的缺点是主观加权可能受判断者的偏见的影响，不够客观；没有精确的数量指标；判断者需要受过训练并具有丰富的经验。

(2) 推理方法。不考虑各个变量之间的经验关系，而是根据某种先验的理想程序来作推理性加权。包括单位加权和等量加权。单位加权最简单的方法是将各个变量（题目、分测验或测验）直接相加而得到一个合成分数。等量加权比较麻烦，需要将所有分数转换成标准分数（Z 分数）。

(3) 多重分段。当用测验来决定取舍时，必须确定一个分数线，分数在这条线以上的人接受，在这条线以下的拒绝，这是只有一个预测源的情况。在实际决策中，人们往往不只使用一个预测源。这种方法只是把人分为达到最低标准（接受）与未达到最低标准（拒绝）两类，而不在两组人内部作进一步的区分根据确定分数线的不同情况，多重分段可有两种主要模式：综合分段、连续栅栏等。

(4) 多重回归。多种分段假设预测源间不具有互偿性，但这对许多心理变量来说是不适用的。当同时采用几个预测源来预测一个效标，而这些预测源变量之间又具有互偿性时，多重回归是最常用来组合分数的模式。多重回归方程式的导出相当复杂，一般是借助计算机来进行的。

(5) 合成分数的特殊方法。在某些情况下需要采用一些特殊的方法来组合分数，比如：完形记分、轮廓分析等。所谓完形记分就是将各个变量看做一个整体，不是孤立地看每一个反应结果，而是看总的反应模式。在某些情况下，完形记分可以使效度增加。轮廓分析主要是考虑被试在各个测验或量表上所得分数的轮廓，而不是将各个变量做简单的线性组合。

(5) 合成分数的特殊方法。如：完形记分、轮廓分析等。

### 16. T 检验 F 检验卡方各自适用于什么情况？

答：(1) t 检验运用于总体分布已知的参数检验法中。需要满足总体正态分布，总体方差未知的情况下的显著性、差异性检验。比较适合于小样本 ( $n < 30$ )。这时需要数据符合 t 分布。当样本含量 n 小时，若观察值 x 符合正态分布，则用 t 检验（因此时样本均数符合 t 分布）。

常见的 t 检验形式有：样本均数与总体均数比较的 t 检验；配对设计的 t 检验；成组设计两样本均数比较的 t 检验。

两个小样本均数比较的 t 检验有以下应用条件：

- ①两样本来自的总体均符合正态分布，
- ②两样本来自的总体方差齐。

因此在进行两小样本均数比较的 t 检验之前，要用方差齐性检验来推断两样本代表的总体方差是否相等，方差齐性检验的方法使用 F 检验，其原理是看较大样本方差与较小样本方差的商是否接近“1”。若接近“1”，则可认为两样本代表的总体方差齐。判断两样本来自的总体是否符合正态分布，可用正态性检验的方法。若两样本来自的总体方差不齐，也不符

合正态分布，对符合对数正态分布的资料可用其几何均数进行  $t$  检验，对其他资料可用  $t'$  检验或秩和检验进行分析。

(2)  $F$  检验常常用于方差的显著性检验中。要检验两组数据的离散程度是否有显著不同，需要对两组数据的方差进行差异检验。这时数据符合  $F$  分布。在平均数差异检验时，如果不是相关样本，需要进行方差起行检验。单因方差分析 ( $F$  检验) 常用来检验一个变异因素对试验结果的显著性。作为参数检验法的一种，单因方差分析通常需要假设数据为服从正态分布的随机样本和方差齐性。

方差分析的基本条件是：总体正态分布；变异的可加性；各处理内的方差一致。

(3) 卡方运用于非参数检验。适用于样本是频数分布的情况。其数据是属于点计而来的间断变量；总体分布未知；不是对总体参数的检验，而是对总体分布的假设检验。计数资料的统计检验主要用卡方检验，可以用来同时检验一个因素两项或多项分类的实际观测数据，与某理论次数分布是否相一致的问题，或有无显著差异的问题；还可用于检验两个或两个以上因素各有多项分类之间，是否有关联或是否具有独立性的问题。

卡方检验用于计数资料的分析，对于数据资料本身的分布形态不作任何假设，所以从一定的意义上来讲，又是一种非参数检验的方法。

### 三、问答题（每题十分）

#### 1. 在心理学实验中，随机误差和系统误差对实验结果可能产生什么影响。

答：心理实验中，测量值与真值之差叫误差。它反映了测量值偏离真值的大小和方向。

误差包括随机误差和系统误差。系统误差是指具有确定性的一类误差；随机误差是指不具有确定性的一类误差。系统误差是指在心理物理实验中，由于各种实验条件的影响，造成实验结果经常指向同一方向的误差。随机误差是在对同一被测量的多次测量过程中，绝对值与符号以不可预知的方式变化的测量误差分量。

精密度与准确度分别是对两类不同性质的系统误差和随机误差的描述。只有当系统误差和随机误差都很小时才能说精确度高。精确度是对系统误差和随机误差的综合描述。随机误差与系统误差的合成是没有确定的原则可遵循的，造成对实验结果处理时的差异和混乱。

随机误差和系统误差会直接影响实验的结果。对实验结果产生错误的分析解释。影响实验的效度。

#### 2. 阐述心理学设计中平衡误差的方法。

答：心理学设计中平衡误差可以从以下几个方面考虑：

##### (1) 对自变量方面的控制

##### ① 对自变量规定操作定义

实验中规定操作性定义有很多好处，可以使研究者的思考具体而清晰；可以增进科学研究者之间沟通的正确性；可减少一门学科所用概念或变量的数目。

##### ② 检查点的确定

心理学研究中的自变量，有些有质的区别，是连续的，例如不同的感觉通道，不同的学习方法等，这些不同的质就是不同的检查点，或称为实验处理。一般检查点确定在三到五个为宜。

##### ③ 仪器的校准

实验中对仪器的控制要求是：仪器不仅能对自变量的操作达到严密准确，同时也应具有控制无关变量的能力。

##### ④ 控制呈现刺激方式

包括呈现时间的长短、呈现的顺序、空间位置等，都应根据要求加以控制。

##### (2) 对因变量方面的控制

① 对反应规定操作定义。规定反应变量的操作性定义便于数量化，但是同时要注意指标

的单一性偏差。

②反应变量应具备的条件。实验中的反应指标一般需要具备几个特征：指标要具有有效性；要具有客观性；指标要能够数量化；指标要能够精确地、真实地度量反应的变化。

③反应指标的平衡。

(3) 对无关变量的控制

①对部分被试变量及暂时被试变量的控制方法。如指导语、主试对被试的态度等。

②对环境变量及部分被试变量的控制方法：

a. 设置实验组和控制组。通过实验组和控制组实验结果的对照达到平衡的效果。

b. 随机化法。随机化分配被试，利用数学概率原理，从理论上使各个处理条件下的情况相似。达到平衡误差的目的。

c. 利用 ABBA 或者 AB 法平衡。采取抵消、循环法，利用实验顺序的变化抵消平衡实验误差。

d. 统计方法的运用。在实验数据的统计分析中充分考虑实验误差的影响。通过统计方法观察实验可能存在的误差。

**3. 阐述减数法的基本思想及其应用。在表象心理旋转实验中，如何运用减数法的原理获得表象心理旋转的速度。**

**答：**(1) 减数法就是指唐德斯 ABC 反应时，是一种用减数方法将反应时间分解成各个成分，然后来分析信息加工过程的方法。A 反应就是指简单反应。B 反应是指选择反应。C 反应是指辨别反应。其中 C 减去 A 就得出辨别时间有多长，B 减去 C 就得出选择时间有多长。

①视觉编码和听觉编码实验

波斯纳 (Posner, 1970) 通过应用减法反应时间实验，证明了在短时记忆的短暂时间内，存在着视觉的编码。这说明短时记忆中，先出现一个短暂的视觉编码，然后出现听觉编码，所以随着两个字母相继呈现时间的加大，视觉编码效应逐渐消失，听觉编码效应增大，其反应时间也加大，从而缩小了与 A、a 字母对反应时间的差别。这就是应用减法反应时间的方法，证明了某些短时记忆的信息加工，可有视觉与听觉编码两个连续阶段。

②句子---图画匹配实验

这一实验是由柯拉克 (H. H. Clak) 和蔡斯 (W. G. Chase) 设计的，实验时给被试看一个句子和一个图画，例如“星形在十字之上”要求被试判断二者是否一致并作出反应，记下反应的时间。句子有八种，主语有“星形”和“十字”，谓语有“在之上”和“在之下”、“不在之上”、“不在之下”。

③心理旋转实验

1973 年库伯 (L. A. Cooper) 和谢帕德 (R. N. Shepard) 设计该实验来证明心理旋转的实际存在。实验选取不同的字母和数字 (如 R、J、G、2、5、7 等) 为实验材料，将这些材料取正面或反面以及六种不同的倾斜度，让被试反应后记录反应时间。

(2) 1971 年美国斯坦福大学的心理学家谢帕德 (R. Shepard) 和梅茨勒 (Metzler) 等做了一系列实验。实验的材料是一对对不同方位的立方体的二维形式图。要求被试判断所用的时间。实验中 A 和 B 是两对完全相同的图形，所不同的仅仅是它们的方位，A 中两个物体在平面上相差  $80^\circ$  角，B 中两个物体的深度上相差  $80^\circ$  角。C 中的一对物体是两个方位和结构都不同的物体。谢帕德和梅茨勒制作了一千六百对这类图片，他们请了八位成人被试者进行判断实验。被试者报告了他们判断时使用的方法：首先是把一个物体图形在心理旋转，直到与另一物体的方位相同，然后进行匹配比较，从而作出完全相同或完全不同的判断。以被试者对八百对同物图片判断的结果，可以看出，无论图片所示的物体是在平面上调转 (即通过旋转画纸就可以实现)，还是在三维深度中旋转 (把物体方位旋“进”画纸中去)，判断所用的时间同两物图片上的角度差异成线性关系，即旋转的度数越多，反应所用的时间越长。

这就是著名的心理旋转的实验。后来，谢帕德和库珀（Cooper）合作，以不同倾斜角度的正向和反向的字母 R 为实验材料，对心理旋转作了进一步的研究。

两类心理旋转实验说明了以下重要事实：①表象这一心理现象是客观存在的，是可以用科学的实验方法证明的。心理旋转的实验恰恰证明了表象的存在，并非像行为主义心理学所断言的那样，外界刺激传入大脑，就如同进了黑箱子，人们无法知道头脑是怎样活动的。也不像机能主义心理学那样，单凭被试者的内省报告就来分析人的心理活动。认知心理学用实验进一步揭露信息在大脑中进行加工的过程。②心理旋转实验证明了表象是物体的抽象的类似物的再现。在没有刺激呈现的情况下，头脑中会对视觉信息和空间信息进行加工。表象是真实物体的类似物，它是以观念的形式存在于头脑中的，它具有直观性。大脑对表象的加工操作类似于对真实物体进行知觉时的信息加工。事实上，心理旋转正是真实的物理旋转的一种类似物。只不过表象是这种实物旋转在头脑中的复现而已，并且复现时不受任何感觉通道的束缚。

#### 4. 采用信号检测论的方法设计一个实验，考察某一群体被试对图形或文字材料是否存在

**答：**（1）信号检测论是信息论的一个分支，研究的对象是信息传输系统中信号的接受部分。信号检测理论将被试的感受性和辨别力分离出来，是对传统的心理物理学方法的重大突破。

（2）信号检测理论的基本原理

##### ① 统计学原理

统计决断理论是信号检测理论的数学基础。统计学的主要工具是概率论。心理学的信号检测实验一般是在信号和背景不易分清的情况下进行的。对信号检测起干扰作用的背景叫噪音。

在信号检测实验中，被试者对有无信号出现的判定，可以有四种结果：击中、虚惊、漏报和正确排斥。从统计学的观点，信号检测就是要检验两个统计假设  $H_0$ （无信号）和  $H_1$ （有信号）的真伪。

##### ② 最优决策原则正确判断原则

人类在做出一个判断标准时，一般是按照最优原则进行的。在确定判断信号的标准时，常常受到三个因素的影响：信号和噪音之先验概率的大小；判断结果的奖惩（价值）多少；被试所要达到的目的以及其他一些有关因素等。

（3）实验举例：

被试选择：随机安排大学生进行测试，人数为 60 人。

实验材料：选择一些有相关性的图形和文字材料，输入电脑中，这些材料中图形和文字有对应关系。

实验过程：要求被试通过操作电脑对图形或文字材料进行判断和描述，并进行一些再记忆研究。先让被试记忆一些学习材料。间隔时间后再呈现给被试，让被试作出“是”还是“不是”的判断。对被试完成任务的情况进行分析判断。

统计分析：由于信号检测论的统计基础是概率论。它将噪音从背景中分离出来。将信号检测实验中，被试者对有无信号出现的判定，分为四种结果：击中、虚惊、漏报、正确拒斥，形成两择一判断矩阵，即要检验两个统计假设  $H_0$ （无信号）和  $H_1$ （有信号）的真伪。将其两个独立指标：反应倾向和感觉辨别力区分开来。

通过统计分析，判断被试对图形或文字材料是否存在差异。