



問卷設計與條件

顧志遠 教授

中華民國113年4月16日



好的問卷設計，
要符合研究本身的需求，
更重要的是考量受測者的因素。

課程綱要

第六講 問卷設計與條件

● 問卷設計 (questionnaire)

- ◎問卷設計概論
- ◎問卷的功能
- ◎問卷使用原則
- ◎問卷應用假定
- ◎問卷的優缺點
- ◎優良問卷的特徵
- ◎編製問卷主要過程
- ◎問卷的類型

● 問項的寫作要點

- ◎問項擬定注意事項
- ◎問項的內容
- ◎問項的基本要求

● 問卷前言(卷頭語)內容

● 問卷驗證

- ◎信度
- ◎效度
- ◎信度與效度關係



問卷設計(questionnaire)

■ 問卷設計概論

- 問卷(questionnaire)，意為「為了統計或調查用的問題表格」，直譯為「問題表格」，現在大家習慣稱為「問卷」。
- 最直接得研究方法是現場目擊的分析，但不一定都能做到或者有些調查是不可目擊，故用間接方法，由研究者發問(口語/問卷)，受試者回答(自我陳述)，蒐集研究資料的手段。
- 社會科學與自然科學實證比較：
 - ✓ 自然科學(直接測量)：儀器->度量衡校正->數據分析
 - ✓ 社會科學(間接測量)：問卷->**信度和效度**->數據分析
- 溝通的過程，隔離接觸相對於面對面的訪問。
- 問卷設計是研究的重要工作，其影響研究品質與結果甚鉅。
- 問卷設計原則：有效展現研究問題內涵與假設、精確衡量受測者態度強弱度。

問卷設計(questionnaire)

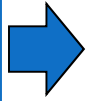
- ❑ 一種測量技術，尤其是對具有多向度問題的量度，一種控制式的測量，用一些變項了解另一些變項，結果可能是因果或相關關係。
- ❑ 建立問卷必須充分瞭解研究的問題/假設/客觀事實、資料的性質/模式/行為、其他有關資料，否則資料難符合研究者的需求。

問 卷

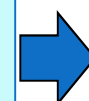
研究主題



- 基本資料(X)
- 量題資料(Y)
 - ❑ 事實問題
 - ❑ 意見問題
 - ❑ 態度問題



- ✓ 資料分析 X_i 、 Y_i
- ✓ 相關分析 (X_i, Y_i) 、 (Y_i, Y_j)
- ✓ 迴歸分析 $Y_i = F(X)$
- ✓ 變異數分析
- ✓ 多變量分析
- ✓ 因徑分析
- ✓ 結構方程式



驗證假設

問卷設計(questionnaire)

■ 問卷的優缺點 ■ 影響與解決方案？

問卷優點：(我門應善用)	問卷缺點：(我門需克服)
經濟省成本 節省時間大量施測 避免訪問者偏差 填答者有充分時間作答 填答者可利用最方便時間作答 具隱密性 題目標準化 實施範圍較廣 適宜用電腦處理 可郵寄亦可訪問 可完全控制變項找出因果關係	缺乏彈性 回收率低 只能了解填答者語文行為 不能控制填答的情境 無法掌握填答者的自發性回答 不能控制回答順序 不能確定係由填答者所填 得到回答可能不完整 問卷設計者有問題無法補救 無法確定填答日期 不易採用複雜的問題深入研究

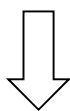
問卷使用基本原則

- 僅在收集由其他方法所無法獲得的資料。
- 問卷要收集的資料與研究目的符合。
- 不超過受測者心力負荷。
- 等值交換受測者的付出。
- 問卷信度與效度的確保。
- 注意問卷倫理尊重研究。

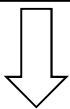
問卷收集資料型式

- 問卷收集的資料分為「屬性/行為/意見/態度」四種資料。

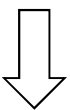
基本資料



行為資料



意見資料



態度資料

- 基本資料**[屬性資料/事實問題]：定義特定人群特徵的數據，譬如「年齡」、「性別」、「婚姻」、「星座」、「收入」等，穩定性較強，比較不會隨著時間有所變動，或是改變較為緩慢。不變者為屬性。通常為自變項。
- 行為資料**[實際行為/事實問題]：特定人群可被記錄的行為數據，譬如「購物行為」、「用餐行為」、「健身頻率」。其為實際的動作，可以直接被衡量或觀察。
- 意見資料**[暫定看法/態度問題]：特定人群對於特定現象展現的現在看法。譬如「喜好」、「意見」、「傾向」等等。意見是一種淺層的態度，會隨著外在環境或現象狀態的改變而改變。
- 態度資料**[深層看法/態度問題]：特定人群對於特定現象展現的態度和價值觀。譬如「風格」、「價值觀」、「個性」等。態度屬於深層或不變的看法，不容易被調查出來的一種數據。

- 問卷收集的資料：行為、態度/信仰/意見、特徵、期望、自我分類、知識。

問卷功用與假定

■ 問卷的功用

- 初步了解研究問題的重要變項與內容，進一步形成研究假設。
- 作為特定研究的主要資料，用來回答問題或驗證假設。
- 用來佐證其他資料的有效程度。

■ 問卷應用三項假定

- 受訪者有足夠的能力看懂題目，且能以書寫回答 ->問卷設計
- 受訪者有充份智識了解題目內容，且能正確回答 ->找對樣本
- 受訪者願據實以答->問卷篩選(反向題、同意副題、間接題)

■ 以受訪者為核心的問卷設計

- 受訪者對問題的認知程度
- 受訪者對於問卷中問題的經驗
- 受訪者對問卷中的資訊能否記得？
- 受訪者願意提供資訊嗎？
- 受訪者是否要費很大的工夫去蒐集資訊？
- 受訪者在哪裡？可以找到與找足夠嗎？

問卷編輯過程

1. 確定所要收集的資料與變項。(由構念依次展開構面)
2. 決定問卷的型式。(由實施方式、研究問題性質、研究樣本性質、資料分析與解釋的方法)
3. 撰擬問卷的題目。(遣詞用字、順序、加入相等副題/反向題)。
4. 修正問卷。(去除無區辨力題目，找較精細專家修正)
5. 預試。(個別訪談或集中測試，可能經幾回預試收斂)
6. 編輯問卷和實施說明。(問題順序，如敏感、難答，特殊、時間近等於後)
7. 考驗問卷性能。(信度與效度)
8. 發展常模。

問題的形式

- ❑ 直接問題與間接問題
- ❑ 特定問題與普遍問題
- ❑ 事實問題與意見(態度)問題
- ❑ 威脅性問題與非威脅性問題
- ❑ 集體問題與個人問題
- ❑ 疑問句與敘述句
- ❑ 獨立問題與關聯問題

問卷的類型

■ 結構型問卷 (structured questionnaire)

□ 適用在1.非發覺新研究構想2.對於問題的變項與答案胸有成竹

□ 根據研究假設與目的之需要，展開設計

□ 問題標準化，受測者方便性高，處理簡單

□ 研究的結構化程度影響問卷的結構化程度

□ 研究越具有結論性，問題越需要結構化

□ 訪談者與受訪者不易犯錯

□ 適用於大樣本研究

□ 等級法：通常用來了解態度的相對程度。

請將下列因素依其對一個人事業成功的重要性,依序排列等級(等級越低,表示重要性越低;等級越高,表示重要性越高)

教育	家世
智力	運氣
人格	經歷

□ 檢核法：通常在初步了解有關行為與知識的事實。

您做什麼戶外活動？(可複選)

☐ 爬山	☐ 打籃球	☐ 打棒球
☐ 游泳	☐ 打網球	☐ 慢跑
☐ 騎馬	☐ 打羽球	☐ 打排球
☐ 騎車	☐ 其他	

□ 評定法：通常用來評量態度的程度。

您認為下列對一個人事業成功的重要性?(請圈選適當的號碼)

重要性	低				高
教育	1	2	3	4	5
智力	1	2	3	4	5
人格	1	2	3	4	5
家世	1	2	3	4	5
運氣	1	2	3	4	5

□ 類別法：通常用在行為或態度或意見可以明確採用。

您目前就讀的學校是：

☐公立學校 ☐私立學校

您贊成中學男女合班嗎？

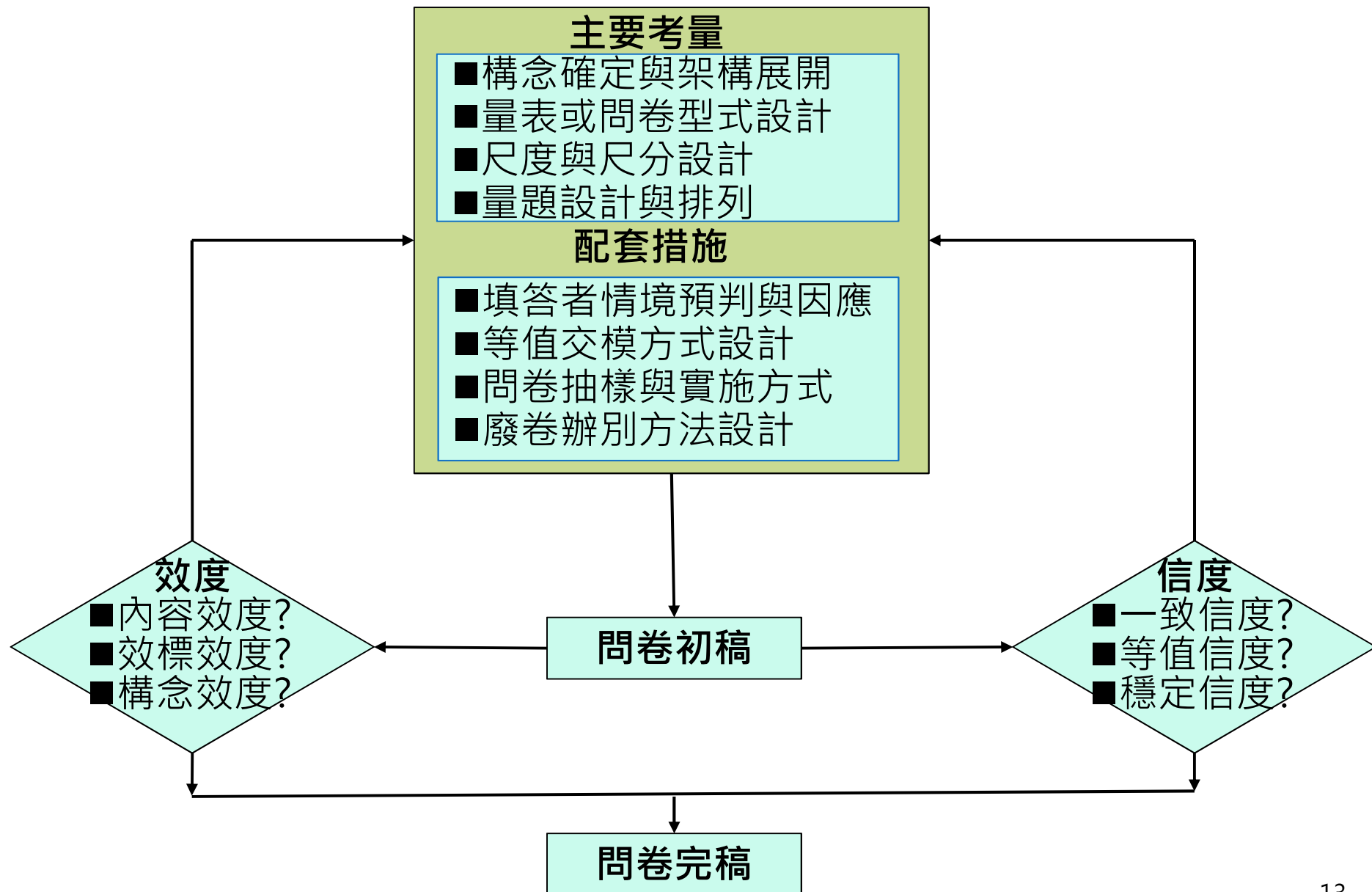
☐贊成 ☐無意見 ☐不贊成

問卷的類型

■ 非結構型問卷 (unstructured questionnaire)

- 無結構問卷均為開放式問卷，這樣研究者和被調查者之間方可自由溝通。並非完全沒結構，只是沒有固定的選答項目。
- 回答者可自由表達意見及看法，訪員在不變更內容與方向前提下，可深入了解問題。
- 適用在1. 發覺新研究構想2.對於問題的變項與答案於形成階段。
- 優點是可以收集廣泛的資料和探詢特殊研究對象的特殊意見和看法。缺點是無法作統計分析之用。
- 研究越具有探索性，問題越需要非結構化。
- 適用在深度訪問場合(適用在少樣本調查)。
- 對大樣本研究具有補助功效。
- 另外訪問過程較困難，往往要求研究者具有較高的研究水平和迅速反應的能力。

問卷設計(questionnaire)



優良問卷的特徵

- 優良問卷的主要特徵：(Best,1977;Mouly,19787)
 - 問卷僅在收集由其他方法所無法獲得的資料。
 - 所有問項與研究目的符合，量題均表達重要變項。
 - 問卷應盡量簡短，太長影響回收率。
 - 問卷顯現研究主題的重要性，並清楚表達，使答題者願意花時間回答。
 - 指導語要清楚，每個問題處理一個概念，問語力求清楚，易於回答。
 - 問卷題目要客觀，沒有引導回答的暗示。
 - 問卷的題目依心理，邏輯次序安排。
 - 問卷收集的資料易於列表說明及解釋。
 - 問卷外觀具有吸引力，且印刷精美。
 - 問卷中應包含研究目的,指導語,個人基本資料與問卷的題目。

問題擬定的注意事項

■ 問題擬定的注意事項(語文問題)

- 避免提空泛性問題(你對某百貨公司印象如何？你認為網路媒體對青少年會產生影響嗎？)
- 避免用不精確的用詞(你經常購買洗髮乳？你每週運動幾次？)
- 避免使用容易誤解的字詞 (你最近是出門旅遊，還是休息？)
- 語文應淺顯易懂不超過受測者領悟範圍
- 問題說明要具體/簡單/清楚/容易理解
- 避免使用專業術語/行話/俚語/雙關語/文言文/倒裝句/縮寫...

■ 問題擬定的注意事項(情緒問題)

- 避免主觀及情緒用語(你認為死刑是野蠻與不文明的處罰嗎？)
- 避免隱含性選擇問題(如果可能的話,你喜歡有工作(家事)嗎？)
- 避免涉及隱私的問題(詢問酗酒/犯罪/離婚/夫妻吵架等事情)
- 問題是否涉及社會禁忌(貴公司有沒有漏稅？)
- 避免引起填答者情緒困擾(敏感問題，如過去犯法違規紀錄)

問題擬定的注意事項

■ 問題擬定的注意事項(理解問題)

- ❑ 避免專業性問題(核能安全？維他命是否有益健康？)
- ❑ 避免沒有區辨力的問題(剎車安全重不重要？)
- ❑ 問題項目是否含混不清(你常看電視嗎？公務員待遇嗎？)
- ❑ 避免使用雙重否定敘述(你認為大多數沒有不贊成的嗎？)
- ❑ 問句要具體且單一問題(你為何不看電影而看電視？)
- ❑ 問題不超過受測人知識與能力(明年經濟成長率，捷運高或低運載)
- ❑ 避免包含兩個問題的概念(下面哪項是成功的關鍵與瓶頸？你喜歡運動與旅遊嗎？品質與價格)
- ❑ 不要引起誤解與爭論(你贊成合理的加稅嗎？)
- ❑ 必須在填答者記憶範圍內(十年來你曾經被開過幾次罰單？)
- ❑ 不要用假設或猜測語句(你是否滿意去年的加薪？)
- ❑ 不要花太多時間填寫(青少年犯罪的原因)
- ❑ 避免使用不當的選目(請問多久逛街？不曾/偶爾/不常/常常)

問題擬定的注意事項

■ 問題擬定的注意事項(理解問題)

- ❑ 避免誘導回答與暗示回答(抽菸有害身體，請問一天抽幾包菸？你是否贊成合理的加稅？是否贊成國父說的話？額外語句引導)
- ❑ 避免有不適當假定或使用假設性的問題(對於台灣高犯罪率，你是否贊同晚上不宜外出?)
- ❑ 避免很難回答的問題(需親身經歷/久遠問題/非科學問題/以後才知)
- ❑ 注意邏輯順序，簡單/認知負擔低前放
- ❑ 避免長句，越短越好
- ❑ 題目避免冗長，複雜問題可分為連貫問題
- ❑ 盡量不要超過20題，題數越多信度越低
- ❑ 題目中需特別強調的觀念，用下加線表示
- ❑ 盡量將同樣性質的問題放在一起
- ❑ 類別項目問，須列舉選目完整
- ❑ 類別項目問題，選目必須互斥
- ❑ 評比項目問題，必須提供參照點
- ❑ 用詞深度適當，不超過填答者領悟能力

問題的順序

■ 問題的順序

□ **時間順序**：先近後遠或先遠後近均可，應依次序排列，不要雜亂，以避免干擾填答者思考時軸為原則。

□ **內容順序**：

- ✓ 一般先 / 特殊後
- ✓ 易答先 / 難答後
- ✓ 封閉先 / 開放後
- ✓ 熟悉先 / 生疏後
- ✓ 有趣先 / 枯燥後
- ✓ 非敏感先 / 敏感後

□ **類別順序**：

- ✓ 基本資料→行為資料→意見資料→態度資料。
- ✓ 若問卷涉及隱私，基本資料置前較容易被拒絕訪問。

問題擬定的注意事項(整理)

- **是否易懂**，問卷的目的是了解用戶的真實想法，所以問卷的問題及其選項應儘量易懂，避免使用專業詞彙，導致用戶在填寫問卷時因職業或經驗背景而有所差異。
- **接受度較高**，一方面是問題表達的意思，用戶能樂於接受，另一方面是問題選項也同樣要設置的能讓用戶接受，避免出現用戶牴觸心理。
- **測試要點要單一**，問題的設定要明確是想了解用戶具體某一個需求或問題點，儘量不要有一個題目出現多個需求點組合的問題，這樣不利於用戶選擇問題選項，也不利於之後的問題分析。
- **問題順序要合理**，問卷整體的問題順序儘可能遵循先易後難原則；先問正面問題，再問反面問題；先問已接受問題，再問涉及敏感的問題等。

問題擬定的注意事項(整理)

- **選項要全面**，選擇性題目的選項要全面，包含所有問題所出現的可能選項，也就是上述“封閉式問題”的問題選項要拿捏好。
- **避免誘導用戶**，問題及其選項的描述儘可能中立，不要誘導用戶進行選擇，這樣無法了解用戶真實想法，使得用戶調研失去意義。
- **獨立選項**，問題選項之間不要出現交集的情況，每個問題選項都是個獨立的個體，除非“承上題”這種問題，但一定要知道往下問是否有意義，不然放上去也只是增加問題。使用戶填寫疲勞。
- **表述清晰**，問題描述要夠清晰，不要出現描述的事情存在模稜兩可的情況，不便於用戶進行問題選項的選擇。
- **問題數量要適宜**，相信大家都有體驗過落落長的問卷調查，俗話說現代人注意力只有3秒，該如何把握住用戶注意力同時又能兼顧設計問卷的同時可以搜集到真正想要的數據。

問卷前言(卷頭語)之內容

- 自我介紹
- 扼要說明研究內容
 - 解釋研究問題與目的
 - 說明進行研究的機構或研究者
 - 強調研究進行的必要性與重要性
- 說明受測者為何務必要回答此問卷
- 強調完成問卷不需花太多時間
- 強調受測者的自願性
- 保證匿名與保密
- 提供聯絡電話
- 提供回郵地址及回覆期限
- 說明如果受測者有興趣，可以索取研究結果簡本
- 感謝對研究的參與

問卷衡量-信度與效度

■ 問卷衡量準則

□ 效度(validity)

考量測量工具是否能測量到欲衡量的特性

□ 信度(reliability)

考量測量工具所衡量結果的穩定性與一致性

□ 敏感度(sensitively)

考量測量工具是否能區辨微小的變化

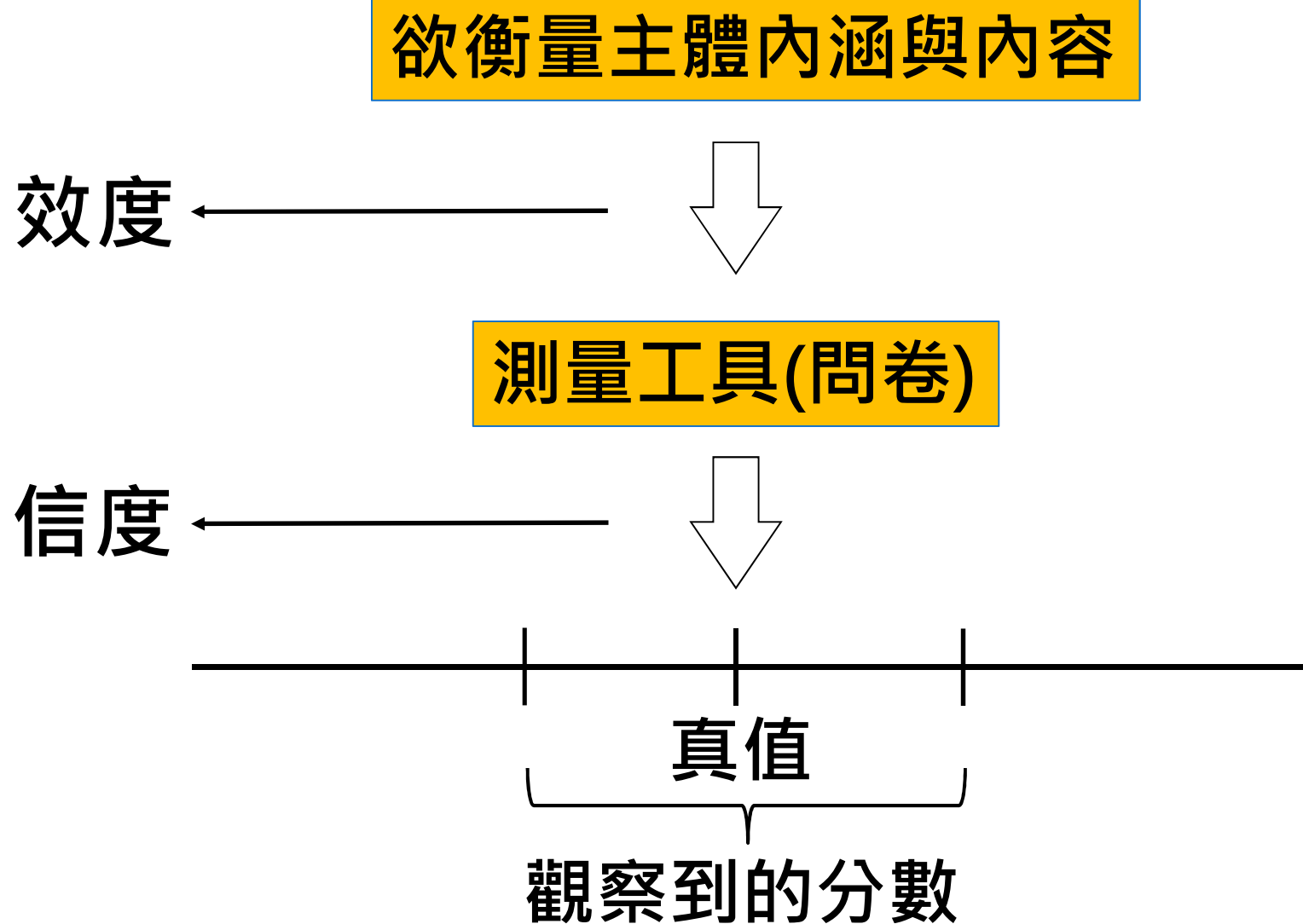
□ 實用性(practicality)

考量測試工具的經濟性, 便利性, 可解釋性等

□ 常模(norm)

提供個人分數在常態樣本中的相對地位
提供比較的量數, 使不同的分數可為比較

問卷衡量-信度與效度



信度

- 美國教育與心理測驗標準定義，信度是分數未受測驗誤差影響的程度。
- 測量結果能反應受測者真實的特徵，不受其他因素(情境、情緒、量題順序、性質等)而影響測驗分數，那麼這個測驗結果是可靠的。
- 「信度」是透過衡量工具所獲得的結果(分數)，能反應出受測者真實(分數)的程度。這種程度可透過對受測者進行不同形式(如不同版本的問卷)或不同時間的測量，並判斷其結果是否一致來獲得。換句話說，一致性(consistency)是信度的基本概念。
- 測量的穩定度(前後測)與一致性(類似測)。
- 信度公式：
假設一個測量工具所測得的值為 X_o ： $X_o = X_t + X_e$
觀察值 X_o ：真實值 X_t ：誤差值 X_e
而假設測量所得的變異量為 V_o ， V_o 可分解為： $V_o = V_t + V_e$
信度 = $V_t / V_o = 1 - (V_e / V_o)$

信度

- 衡量誤差的意義與來源
衡量誤差可分為系統性誤差及隨機性誤差。
- 一般而言，大部份的誤差是系統性的(從偏差而來)。所謂系統性誤差也被視成常數性(constant)誤差。而隨機性誤差(random error)則不是一種常數性誤差，其原因可能來自情境因素，或者被受測者一時的情緒而影響。
- 衡量誤差可能的來源如下：
 - 由回應者(respondent)產生的誤差
 - 由情境因素產生的誤差
 - 由衡量者產生的誤差
 - 由衡量工具產生的誤差
- 以測驗分數變異數理論為基礎，有系統及非系統變異，信度指後者。
 - 同樣測量工具重覆測驗是否能得到相同結果。
 - 測量工具能否減少隨機誤差，而得到某項特質真實數量。
 - 測量的穩定性【前後測】及一致性【類似測】。

信度

衡量工具

		相同	不同
衡量時間	相同	內部一致性度	複本信度
	不同	再測信度	複本再測信度

- 再測信度 (test-retest reliability) : [穩定性]
 - 用同一個測驗但不同時間點測試同一群受試者，又稱穩定係數。
 - 測量間隔時間太短，會有學習與記憶效果
 - 測量間隔時間太長，測量的特質有所成長與改變
 - 用相關分析測量再測信度
 - 一些人類特質，例如：智力、性向、人格等基本上是穩定的，故在測量這類心理特質時，測驗分數具有高度穩定性是必要的。最適宜的相隔時間隨測驗目的和性質而異，少者兩週、多則六個月到1、2年。

信度

■ 複本信度 (inter-method reliability) : [等值性]

- 用內容/形式/難度相當之複本測量，求相關係數，稱為複本信度
- 用不同相似工具來量度同一個項目，例如用邁爾斯-布里格斯性格分類測驗及大五測驗人格。
- 可連續實施或相距一段時間實施，前者稱複本信度，又稱等值係數，後者稱穩定與等值係數
- 說明由於內容與時間變異所造成的誤差情形
- 用相關分析測量複本信度

■ 一致信度(split-half reliability) : [一致性]

- 內部一致性信度越高，代表各個衡量項目能夠測量到相同的構念。
- 內部一致性指量表能否測量單一概念，同時反映組成量表題項之內部一致性程度。
- 通常按奇數題和偶數題或隨機方法將試題分為兩半，將受試者在這兩部份試題的得分求相關，稱為折半信度。

信度

□ 一致信度公式(折半信度)

$$r_x = \frac{2r_{hh}}{1+r_{hh}} \text{ (Spearman-Brown Formula)}$$

r_x ：全測量信度， r_{hh} ：兩半測量相關

□ 一致信度公式(折半信度)

$$r_x = 2\left(1 - \frac{S_a^2 + S_b^2}{S^2}\right) \text{ (Flanagan Formula)}$$

S_a^2 、 S_b^2 ：兩半測驗分數之變異數， S^2 ：整個測驗總分之變異數

□ 一致信度公式(折半信度)

$$r_{xy} = 1 - \frac{S_d^2}{S^2} \text{ (Pulon Formula)}$$

S_d^2 ：兩半測驗分數差之變異數， S^2 ：整個測驗總分之變異數

□ 一致信度公式(Kuder-Richardson)

$$r_{KR20} = \left(\frac{K}{K-1}\right)\left(1 - \frac{\sum pq}{S^2}\right) \text{ (Kuder-Richardson 20 Formula)}$$

K：測驗總題數， $\sum pq$ ：每題答對與答錯比率乘積， S^2 ：測驗總分變量

信度

□ 一致信度公式(Kuder-Richardson)

$$r_{KR21} = \frac{KS^2 - \bar{X}(K - \bar{X})}{(K-1)S^2} \text{ (Kuder-Richardson 21 Formula)}$$

$\bar{X}$ ：測驗總分之平均數

□ 一致信度公式(Kuder-Richardson三)：Cronbach's α 係數：又稱 **alpha信度**，是目前社會科學研究最常使用的信度分析方法。

α 係數是Cronbach由KR20公式發展出測驗信度方法：

$$\alpha = [n / (n-1)][1 - (\sum S_i^2 / S_x^2)]$$

n ：題數。

S_i^2 ：每一題目分數的變異數。

S_x^2 ：測驗總分的變異數。

KR20信度公式只能適用於對錯類型的試題，而 α 係數則可適用於多重計分的測驗，其應用有下列兩方面的限制：(1)不適於速度測驗信度的估計；(2)無法反應時間取樣對於測驗分數的影響，而只能指出內容取樣與內容異質性的誤差。 α 係數和KR20信度時常被稱為估計信度的最低限度。

Cronbach's α 係數	Internal consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

信度(Cronbach's α 係數)

- Cronbach's α 係數：又稱alpha信度，Cronbach在1951年提出，是目前社會科學研究最常使用的信度分析。原理如下：

α 係數是由Cronbach由KR20公式所發展出測驗信度的方法：

$$\alpha = [n / (n-1)][1 - (\sum S_i^2 / S_x^2)]$$

n ：題數。 S_i^2 ：每一題目分數的變異數。 S_x^2 ：測驗總分的變異數。

	1	2	3	4
A	4	3	4	2
B	4	3	4	2
C	4	3	4	2
D	4	3	4	2
S_i^2	0	0	0	0

$$S_x^2 = 11, \sum S_i^2 = 0$$
$$\alpha = [n / (n-1)][1 - (0 / S_x^2)] \rightarrow 1$$

	1	2	3	4
A	4	2	4	1
B	1	4	4	1
C	4	3	2	4
D	3	1	2	4
S_i^2	6	5	4	9

$$S_x^2 = 25, \sum S_i^2 = 24, \sum S_i^2 \rightarrow S_x^2$$
$$\alpha = [n / (n-1)][1 - (S_x^2 / S_x^2)] \rightarrow 0$$

信度(Cronbach's α 係數)

- 某研究設計一份有關“商店服務屬性”之問卷初稿，為慎重起見，選取30位受測者進行前測。

- α 係數 = 0.7984，代表十五道題目測量結果之內部一致性具高可信度。

- 若是想要再提高信度，可由：

- ✓ Corrected Item-Total correlation

- ✓ Cronbach Alpha if item Deleted

- 兩欄數字來判斷刪除那些題目可提高內部一致性。

	種類多寡	品質	價格	取得便利	購物便利	廣告吸引	氣氛服務	整齊乾淨	內外
1	2	3	3	2	1	2	3	2	
2	2	2	3	2	3	3	3	2	
3	2	3	2	2	2	2	2	2	
4	3	2	3	2	2	3	3	3	
5	2	3	3	3	3	2	3	3	
6	2	2	3	3	3	3	3	3	
7	2	2	2	2	2	3	3	2	
8	2	3	3	2	2	3	3	3	
9	3	3	3	2	3	3	3	3	
10	2	2	2	1	1	2	2	2	
11	2	1	2	2	4	3	1	2	
12	1	3	3	3	1	3	3	3	
13	2	3	3	2	2	3	2	3	
14	2	2	3	2	2	3	2	2	
15	2	2	2	2	2	3	3	3	
16	2	3	3	2	2	3	3	3	
17	2	3	2	3	3	3	3	4	
18	2	3	2	2	3	3	2	3	
19	2	3	2	2	2	3	3	3	
20	3	3	2	2	2	3	2	2	
21	1	2	3	2	3	3	3	3	
22	2	2	2	2	2	2	3	3	
23	1	2	3	1	1	1	1	1	

信度(Cronbach's α 係數)

- 價格的「Corrected Item-Total correlation」為-0.0693，表示「價格」與總分間的相關非常的低，如果將之刪除，則可提高 α 係數至0.8156。

```
***** Method 1 (space saver) will be used for this analysis *****
```

RELIABILITY ANALYSIS - SCALE (ALPHA)				
Item-total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Alpha if Item Deleted
種類多寡	35.0333	21.1368	.1999	.8004
品質	34.5333	20.6713	.2650	.7969
價格	34.6000	22.4552	-.0693	.8156
取得便利	34.9000	20.5069	.3512	.7910
購物便利	34.7667	19.9782	.2802	.7990
廣告吸引	34.3000	20.7690	.2818	.7953
氣氛服務	34.4000	18.5931	.6424	.7685
整齊乾淨	34.4667	18.8782	.5713	.7740
內外裝潢	34.2000	20.6483	.3003	.7942
商店印象	34.5000	18.7414	.5954	.7720
服務態度	34.4333	18.4609	.5972	.7708
結帳速度	34.1667	18.0747	.4871	.7813
售後服務	34.7000	19.3897	.4907	.7806
營業時間	34.8333	20.3506	.3469	.7912
信賴度	34.6333	17.9644	.6922	.7624

Reliability Coefficients	
N of Cases =	30.0
Alpha =	.7984
N of Items =	15

信度

- 一致信度公式(評分者間信度)
評分者間信度(Inter-rater Reliability)：評分者間對於某件事情彼此同意程度。其分數顯示評分者之間想法相似(homogeneity) 和共識的程度。
Kendall和諧係數(the Kendall's coefficient of concordance; W)

測量編號評等						
A	3	1	2	5	4	6
B	2	1	3	4	5	6
C	3	2	1	5	4	6
D	4	1	2	6	3	5
E	3	1	2	6	4	5
F	4	2	1	5	3	6
Ri	19	8	11	31	23	34

- (1)二個變項**沒有一致**的等級：

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

$$S = \sum R_i^2 - \frac{(\sum R_i)^2}{N} = \sum (R_i - R)^2$$

R：評分等級；

K：評分者人數；

N：被評分的人數或作品件數。

- (2)二個變項**有一致**的等級：

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

$$S = \sum R_i^2 - \frac{(\sum R_i)^2}{N} = \sum (R_i - R)^2$$

$$\sum T = \frac{t^3 - t}{12} \quad ; (K \geq 3)$$

R：評分等級；K：評分者人數；

N：被評分的人數或作品件數；t：表示得到相同等第的人數。

由 $W=0.87$ 表明6位老師的評定結果有較大的一致性

效度

- 考量測量工具是否能測量到欲衡量的特性。
 - 效度是科學測量的最重要必備條件，若無效度，則其他任何要件一律無法發揮功能。
 - 效度是指一個測量在其使用目的上的有效性。
 - 效度有外部效度(研究成果概化能力)與內部效度(測驗設計能得出其所欲衡量的特質)。
 - 效度是指測量結果推論所欲衡量構念的適切性。
 - 如果把 V_t 分解： $V_o = V_{co} + V_{sp} + V_e$
 - V_{co} (correlated V)：與測量特質相關的共同變異量
 - V_{sp} (specific V)：與測量特質無關的個別變異量
- 效度 = V_{co} / V_o

效度

■ 效度種類

- 內容效度旨在確定驗測內容的適切性。
- 效標效度旨在確定測驗分數與未來成就程度。
- 構念效度旨在確定測驗分數能說明構念是什麼及其程度。

■ 內容效度 - 指測量工具內容的適切性

- 測量內容涵蓋研究所探討架構與內容的程度
- 建立內容效度的過程
 1. 有系統的將所研究構念或問題展開成構面組成
 2. 依每個構面建立諸多量題，行程編製問卷或量表題庫
 3. 邀請專家評鑑架構與題庫量題
 4. 依事先規劃好的量題總數，依適當比例分配與各構面
 5. 從題庫選取各構面最適當之量題
 6. 問卷或量表完成，再請予以評鑑

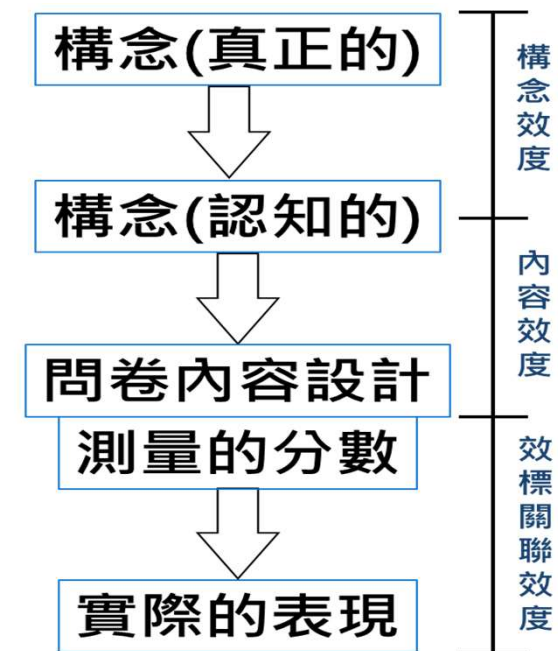
效度

■ 效標關聯效度 - 測驗分數與外在效標間的相關係數

- 測驗分數與效標準間的相關係數。
- 如用偏見表分別那些人有無偏見，GMAT預測學生未來商業能力。
可再分為**預測效度**與**同時效度**。(外在效標分數取得時間)
- 良好效標的特性：攸關係、公平性、可靠性、可獲性。
- 最常見的效標有學業成績，訓練成績，實際工作成績對照團體，評定成績，以往有效的測驗。

■ 構念效度 - 測驗能測量理論的概念或特質之程度

- 構念視一個抽象存在的實體，需證明問卷或量表之測量，可確實測出被測者該「構念」的多少程度
- 構念效度的檢驗方法，並沒有單一方式，它需要從多種不同的資料來源中，逐漸累積證據



構念效度量測法（一）

■ **構念效度量測法（一）** - 依照所欲衡量構念理論，衍出多項假設，並因該發展之問卷或量表測試，若實證結果均與假設符合，則稱此問卷或量表具有高的構念效度。

1. 根據構念的理論分析，發展出一套測量的題目，亦即從構念的有關理論中，預測可能的行為，並據此設計題目。
2. 提出可考驗的構念與其他變項間關係的預測。例如：預測焦慮測驗上的分數與臨床的評定的焦慮程度有關，或焦慮分數和報復水準有關等。
3. 從事實證性的研究，以驗證上述的預測，亦即採用各種方法收集實際的資料，考驗第二步驟所提出的預測（假設）。
4. 淘汰與理論的構念相反的題目（或修正理論）。並從第二和第三個步驟再開始。如果上述的預測成立，測驗的效度就獲得支持。相反的，如預測未成立，不是效度有問題，就是理論有問題，或是兩者。

構念效度量測法（一）

茲舉一個例子，說明新編智力測驗的構念效度的考驗方法。從智力這個構念的有關理論中，我們可以導出下列的預測：

- （1）測驗分數再發展成熟之前隨年齡而增加；
- （2）測驗分數可預測學業成就；
- （3）測驗分數和教師評定的智力等級有正相關；
- （4）測驗分數和其他的智力測驗分數有正相關；
- （5）測驗分數可以鑑別不同能力的團體（如資賦優異與不足者）；
- （6）測驗分數受到直接教學的影響很小。

這六項的預測，可以採用實證性的方法收集證據，逐一與以驗證。若是每項預測均獲得支持，綜合這些正聚集可確信此智力測驗具有良好的構念效度。但是，如果其中多數的預測未獲得支持，不是測驗缺乏效度，就是理論有問題。

構念效度量測法（二）

- 構念效度量測法（二） - 以多特質多項方法(multi-traits multi-methods analysis)找出收斂效度(convergent validity)與區別效度(discriminant validity)。基本原理是一個問卷或量表不僅要和測量相同構念的其他問卷或量表有高度相關，而且要和測量不同構念的其他問卷或量表有低度相關，前者為收斂效度，後者區別效度。(以機械性向量表為例)

對十五位受訪者以三種「方法」測量其三種不同的特質所得數據MTMM相關矩陣

受訪者 變數	自陳量表(1)			投射測驗(2)			同伴評分(3)		
	依賴性 A1	社會性 B1	成就動機 C1	依賴性 A2	社會性 B2	成就動機 C2	依賴性 A3	社會性 B3	成就動機 C3
1	9	5	2	11	5	0	20	9	2
2	8	5	3	9	5	2	20	9	5
3	7	4	3	9	4	3	19	8	8
4	7	5	2	9	6	1	19	11	4
5	8	5	2	10	5	1	21	10	3
6	5	9	5	3	9	3	5	16	9
7	5	8	5	4	8	4	6	18	10
8	4	7	4	4	8	4	6	18	10
9	5	7	5	4	7	4	8	15	12
10	5	8	5	4	8	4	6	18	10
11	2	2	9	2	2	12	1	1	26
12	3	3	8	3	3	10	4	4	27
13	3	3	7	3	3	9	5	5	24
14	2	2	7	3	3	10	4	4	27
15	2	2	8	2	2	11	3	3	26

構念效度量測法（二）

		Method1			Method2			Method3		
Traits		依賴性	社會性	成就動機	依賴性	社會性	成就動機	依賴性	社會性	成就動機
		A1	B1	C1	A2	B2	C2	A3	B3	C3
自陳量表 Method1	A1	(0.98)								
	B1	0.40	(0.98)							
	C1	-.94*	-.47	(0.87)						
投射測驗 Method2	A2	0.95*	0.10	-.89*	(0.96)					
	B2	0.35	0.99*	-.48	0.07	(0.97)				
	C2	-.93	-.67	0.97*	-.80*	-.65*	(0.95)			
同伴評分 Method3	A3	0.94*	0.11	-.90*	0.99*	-.80*	(0.94)			
	B3	0.37	0.97*	-.51	0.11	0.99*	-.67*	0.12	(0.99)	
	C3	-.93*	-.65*	0.96*	.81*	-.81*	0.99*	-.81*	-.64*	(0.96)

註：* 代表 $P < .05$

- 小括弧內數字為重測信度(r 值應最大)；
- 三段有網點的數字為收斂效度(不同方法測相同特質， r 值應第二大)；
- 實線三角形內的數字為區別效度(相同方法測不同特質， r 值應第三大)；
- 虛線三角內的數字為不同方法不同特質(r 值應最小)。

信度與效度 - 關係

■ 信度

假設一個測量工具所測得的值為 X_o : $X_o = X_t + X_e$

X_o (observed X) : 觀察值

X_t (true X) : 真實值

X_e (error X) : 誤差值

而假設測量所得的變異量為 V_o , V_o 可分解為 : $V_o = V_t + V_e$

信度 = $V_t / V_o = 1 - (V_e / V_o)$

■ 效度

如果把 V_t 分解 : $V_o = V_{co} + V_{sp} + V_e$

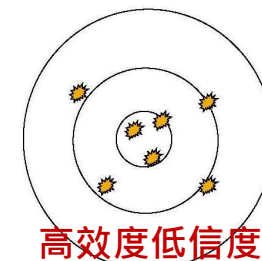
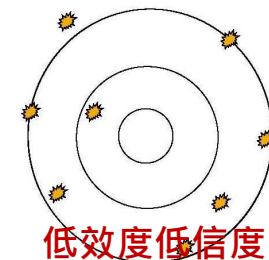
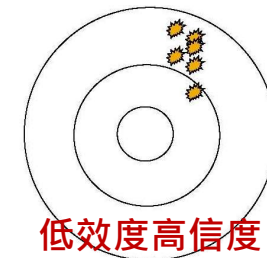
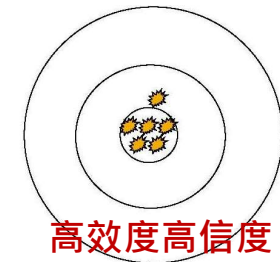
V_{co} (correlated V) : 與測量特質相關的共同變異量

V_{sp} (specific V) : 與測量特質無關的個別變異量

效度 = V_{co} / V_o

■ 效度與信度關係

- ✓ 效度是信度的充分條件—信度高效度低，無意義。
- ✓ 信度是效度的必要條件—效度高信度低，枉然也。
- ✓ 信度過低 $\Rightarrow$ 何來效度可言；效度過低 $\Rightarrow$ 信度沒有意義。



信度與效度

■ 影響信度與效度的因素

產生誤差的原因是多方面，研究者必需面面俱到，才能提高信度與效度。測量中為何會產生誤差，主要可以歸納為五方面：

□ 受訪者的答復

受訪者可能因為個性、情緒、注意力、年齡、性別、疲倦、反應力、知識背景、社會地位以及其他心理、生理因素，語言能力、答題技巧等影響答復的正確性。

□ 測量內容

遣詞用字、問題形式、以及內容是否敏感等。

□ 情境

訪問時間長短、訪問當時的氣氛、有沒有干擾、及開頭的引導說明等。

□ 研究者本身

訪員是否盡責，事前研究設計是否妥善，事後研究分析解釋是否合理。

□ 疏忽

如聽錯、記錯、轉錄錯誤等。



問卷設計與條件

感謝聆聽、敬請指教!

顧志遠 教授

中華民國113年4月16日



好的問卷設計，
要符合研究本身的需求，
更重要的是考量受測者的因素。