

經濟部標準檢驗局臺南分局
問卷設計與服務滿意度調查研習座談

盧昆宏 博士

國立高雄大學
亞太工商管理學系 教授

簡 歷

1. 國立高雄大學亞太工商管理學系教授
2. 國立高雄大學教務長
3. 國立高雄大學圖書資訊館館長暨
4. 國立高雄應用科技大學管理學院院長
5. 國立高雄應用科技大學圖書館館長
6. 國立高雄應用科技大學人力資源發展系系主任、所長
7. 國立高雄應用科技大學工業工程與管理學系副教授、教授

8. 國科會工程處工業工程學門規劃委員
9. 中華民國品質管理學會理事長
10. 研考會提昇為民服務品質考核委員
11. 公共工程委員會公共建設甄審委員
12. 勞委會職訓局人力資源發展服務團團長(南區)
13. 經濟部工業局「全國團結圈活動競賽」評審委員

大 綱

- 問卷內容設計要點與設計步驟
 - 問卷設計的原則
 - 語句之設計要點
 - 影響問卷效度的因素
 - 問卷設計常犯的錯誤
 - 抽樣程序
 - 樣本大小的考量準則

問卷內容設計要點

- 問題的必要性
- 把握問題範圍適切性
- 受測者要能回答

*

問卷設計的步驟(1/2)

1. 決定所需要的資訊
2. 根據所需資料，發展個別的問題
3. 決定問卷的順序(ordering)
4. 擬定問卷初稿

問卷設計的步驟(1/2)

5. 問卷實體製作

6. 檢討和修正

7. 前測(pretest)

8. 定稿

問卷設計的原則(1/2)

- 問題要讓受訪者充分了解，內容不可超出受訪者之知識及能力之範圍。
- 問題是否切合研究主題與內容。
- 要能引發受訪者真實的反應，而非敷衍了事。
- 問項是否含混不清，易引起受訪者的誤解。

問卷設計的原則(2/2)

- 問題是否涉及社會禁忌、偏好。
- 問題是否產生暗示作用。
- 便於忠實的記錄。
- 便於資料處理及資料分析：包括編碼、問卷資料鍵入到電腦檔…。

語句之設計要點(1/2)

1. 使用通用的辭彙要淺顯易懂
2. 問題描寫要簡單明瞭
3. 語句意義要清楚不能模糊
4. 不能假設受測者都懂
5. 不能用有偏差誤導的字句

語句之設計要點(2/2)

6. 不要有暗示的作用

7. 不要隱藏其它的方案

8. 間接問題的利用

9. 句子要短而集中，且一個問句只問一個事物、概念或事件

*

影響問卷效度的因素(1/2)

1. 問題的性質
2. 資料蒐集的媒體
3. 施測的樣本特性
4. 調查訪問員的特質及技巧

影響問卷效度的因素(2/2)

5. 主辦單位

6. 訪問場合

7. 所使用分析方法

8. 問卷印刷

信度分析的原理

- 評估一個系統的信度是以組成此系統的各個項目 (Item) 或量測之間的相關係數為基礎。
- 所謂信度也稱為可靠度，指的是一份量表所測得的分數之一致性 (consistency) 與穩定性 (stability)。

信度分析的意義

- 假設在衡量潛在變數時，量測後所得到的**衡量分數**(measurement score, x)與**真實分數**(true score, T)不相等，其差異為**誤差分數**(error score, E)， $x=T+E$ ，誤差愈小信度愈高。信度可定義如下：
 1. 衡量分數與真實分數 T 相關係數平方，即 $\rho_{x,T}^2$
 2. 真實分數的變異數占衡量分數的變異數之比例，亦即： $R_x^2 = \rho_{x,T}^2 = \sigma_T^2 / \sigma_x^2$

信度係數的種類與衡量方法

主要分成再測信度、複本信度、折半信度與 Cronbach α 信度。

再測信度(不同時間，相同測驗)

- 使用同一份衡量量表，對同一群受訪者，在不同的時間前後測試兩次，這種相關係數又稱**穩定係數**。
- 如兩次衡量相隔太短，受測時間者記憶猶新；
- 但兩次衡量時間如相隔太久，則受測者受心智成長的影響，穩定係數也可能會降低。
- 再測信度在實務上只適用在**固定樣本研究 (Panel Study)**。

複本信度(不同時間，不同測驗)

- 當某一套量表有**兩種以上**版本時可以交替使用，可避免再測信度的缺點。
- **複本**是指內容相似、難易度相當的兩份量表。
- 複本信度的變異來源是受測者在兩種不同版本量表的**得分差異**，差異的存在是因為量表的內容造成，而非時間所造成。
- **複本法**是信度量測的一種好方法，但要編製複本量表有時相當困難。

折半信度(相同時間，不同測驗)

- 如果沒有複本且只能施測一次的話，可將題目分成兩半，分別計算這兩半的總分及相關係數
- 最好這兩半的題目內容性質與難易度能夠盡可能一致。
- 在折半法的優點：只需舉行一次衡量。
- 但折半法的缺點：折半法需將一份量表分成兩半，分法有很多種，不同的分法所得信度可能不一樣。

Cronbach α 係數(1/4)

- Cronbach於1951年提出計算一份問卷或測驗的衡量工具之信度稱為Cronbach α 係數（簡稱 α 係數，或稱為alpha係數），是目前社會科會研究最常使用的信度。
- 設一份量表包括 n 個項目， x_1, \dots, x_n ，假設這 n 個 x_i 分數都與真實分數 T 有關，即每一個項目除與 T 有關外，也有獨立的誤差項 E_i 。

Cronbach α 係數(2/4)

- 令量表總和為： $H = x_1 + x_2 + \cdots + x_n$

$$\alpha = \frac{n}{n-1} \left[1 - \frac{\sum S_i^2}{S_H^2} \right]$$

Cronbach α 係數(3/4)

- 當題目間的**相關係數愈大**時， α 係數也會愈大。當題目數**n愈多**時， α 係數也會愈接近於1。
- 若只有一個題目，則信度值為1。
- 要作信度分析須先檢查每個問項是否都是同方向的，即都是**正面問法**，如果有一題與其他題的相關係數都是負的，應考慮將此題先「**變號或刪除**」後再進行計算 α 係數。
- α 係數的優點是只需衡量一次。

Cronbach α 係數(4/4)

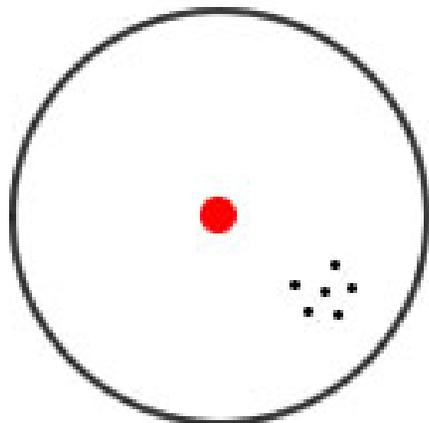
- 可以根據原始資料或標準化後的資料分別加以計算。
- Cronbach α 係數，在基礎研究中，信度至少應達到.80才可接受；在探索性研究中，信度只要達到.70就可接受。
- 實務上， α 值只要.60，即可宣稱該衡量工作信度是可接受

效度分析的原理

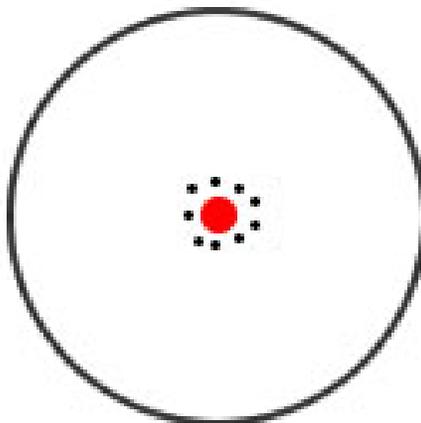
- 效度也稱為**正確性**，是表示一份量表能真正衡量到該量表所要衡量的能力或功能之程度。
- 如果在智力測驗中以英文命題，結果學生不會填答，有可能是他的英文能力不好的緣故，那就測不出學生真正的智力了。
- 因 $\text{效度} \leq \sqrt{\text{信度}}$ ，信度的平方根是效度的上限。
- 故**信度低則效度亦低**，有效度保證有信度，反之，**有信度並不保證有效度**。

信度與效度的差別

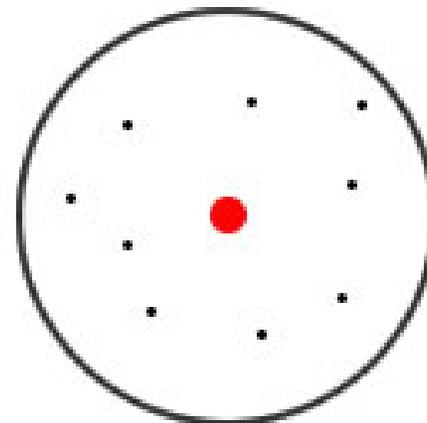
有信度無效度



有信度有效度



無信度無效度



內容效度(content validity)

- 乃指問卷的內容是具有相當高的代表性，即包括被衡量構念所有層面的項目。
- 若問卷內容係以理論為基礎，並參考以往學者類似研究之問卷內容加以修訂，且進行過預試（pretest），即可認為具有相當的內容效度。有時也稱為表面效度（face validity）。

效標關聯效度

(criteria-related validity)(1/2)

- 效標關聯效度乃指衡量結果和效標間的相關程度，其測量的重要關鍵在於效標（criteria）的選用，效標必須能反應衡量分數內涵與特質的獨立衡量，又分為以下兩種：

效標關聯效度(2/2)

- **預測效度 (predictive validity)**：如果效標資料需要衡量完成之後才能蒐集者，則預測效度乃用來衡量測量結果與未來效標的相關程度。例如：編製「成就動機測驗」量表，即可用來「預測」未來該生的成就，以未來成就當作效標。
- **同時效度 (concurrent)**：如果效標資料可在衡量的同時一起蒐集，可用來衡量測量結果與目前效標的相關程度。

建構效度

(construct validity) (1/3)

- 指衡量工具所能衡量到理論構念的程度，所謂建構即構念。
- 它是社會科學研究中所涉及之抽象的概念或變數，如成就動機、忠誠度、焦慮、滿意度、績效等。

建構效度(2/3)

- 建構效度可分為收斂和區別效度 (Convergent & Discriminant Validity)。來自於相同構念的項目，彼此之間的相關要愈高愈好，稱為收斂效度；但屬於不同構念的項目，彼此之間的相關要愈低愈好，稱為區別效度。
- 通常須使用 Multitrait & Multimethod 法來檢定，故需以不同的方法 (method) 來衡量同一個異質的特質 (hetero-trait)。

建構效度 (3/3)

- Kerlinger指出，可用：(1)因素分析；(2)總分和項目分數之相關 (part-whole correlation) 來驗證問卷之建構效度。
- 在因素分析法中須比較該研究抽取的因素是否和以前學者所抽取的因素相同，及每個因素所包含的項目是否相同。
- 由於part-whole correlation即為SPSS信度分析中的Item-Total分析，故目前大部分學者認為，part-whole相關並不是效度而是信度的衡量。

效度值可以計算嗎？

- 每一種效度皆很難衡量，亦即很難驗證量表能夠衡量到真正要衡量的構念。
- 因為沒有方法可真正保證衡量到要衡量的構念。
- 故學術論文中，通常只能透過強調量表設計程序經過嚴謹的過程，或經過專家修正及經過預測，來強化其具有效度，但這也只是內容效度而已。

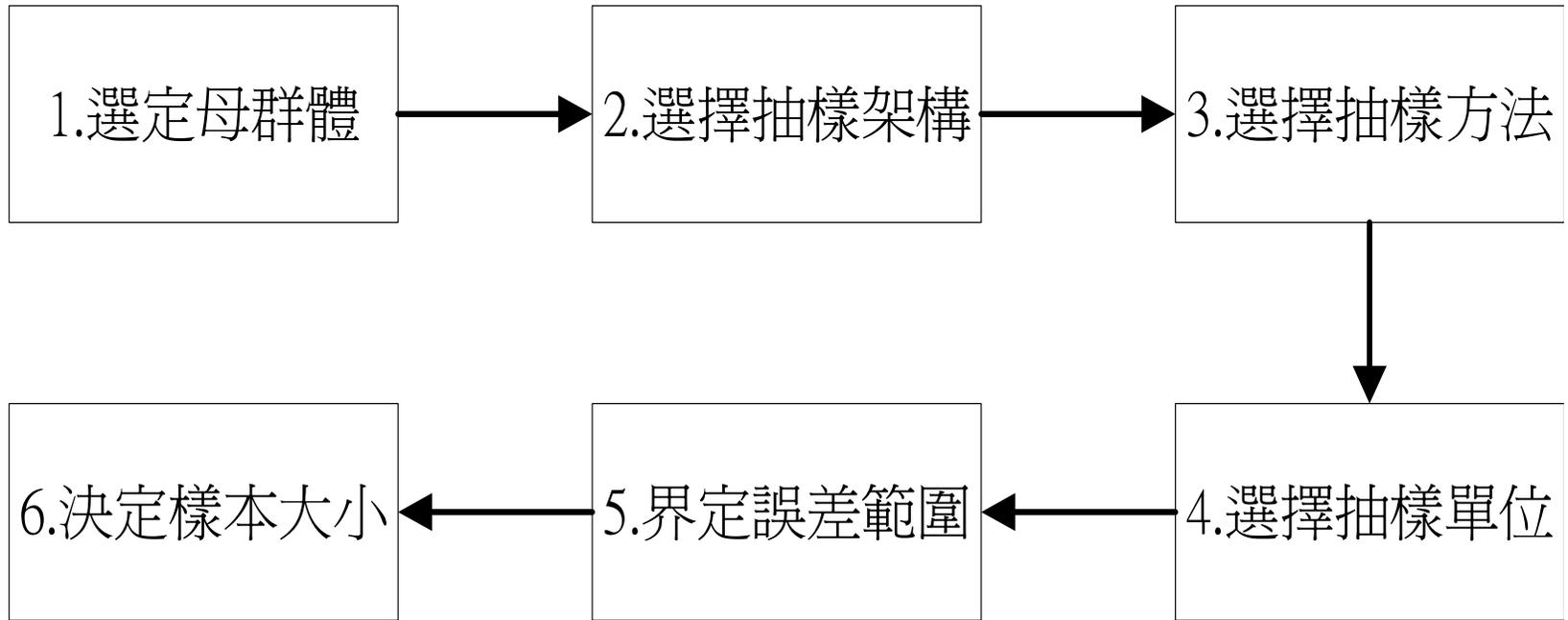
問卷設計常犯的錯誤(1/2)

1. 將問卷與測驗(test)編製混為一談
2. 問卷調查名稱未能明確
3. 遺詞用字過於籠統含混
4. 問卷題目的設計，未能秉持客觀公正
5. 問卷設計中，內容包含備受爭議之部份
6. 問卷印刷編排的錯誤

問卷設計常犯的錯誤(2/2)

7. 文獻探討應先於問卷設計之前，並且注意題目編排順序
8. 問卷題目儘量避免問一些與主題無關或很難回答問題
9. 題目之回答選項不夠互斥性及周延性
10. 問卷設計要考量未來變數如何量化
11. 問卷設計要儘量避開敏感性問題

抽樣程序



樣本大小的考量準則(1/2)

1. 研究的特殊性。
2. 研究的類型：試探性研究、預測、前測所需的樣本就比驗證性、正式研究來得少。
3. 研究假設：當我們預期的實驗處理差異要愈小時，則樣本就要愈大。
4. 經費來源、可用人力的限制。
5. 研究結果愈具重要，則樣本就要愈大。

樣本大小的考量準則(2/2)

6. 研究變數的個數愈多時，則所需的樣本就要愈大。
7. 資料收集的樣本異質性愈高、或不一致性愈大，則所需的樣本就要愈大。
8. 要求的研究結果之正確性/精確度愈高，則所需的樣本就要愈大。
9. 母群體的大小：母群體愈大，則所需的樣本比例就要愈小。

單一樣本大小的公式

$$n=(Z^2 * \sigma^2)/e^2$$

- 母群體變異數 σ^2 愈大，則研究者所需樣本數 n 就愈大。
- 可忍容的誤差 e 愈小，則研究者所需樣本數 n 就愈大。
- 欲使研究推論達到的信賴水準愈大(95%→99%)，則 Z 值就愈高，所需樣本數 n 就愈大。通常我們 Z 值是取1.96。



感謝聆聽！

敬請指教！