政治科學論叢／第二十七期／民國 95 年 3 月／頁 185～210

政治態度與行為量表的信度分析：再測信度與內部一致性檢定*

張佑宗**

摘　要

過去二十多年來台灣有關政治性議題的大型調查研究，往往忽視對其測量指標作有系統地檢驗。然而，唯有提出高品質的信度與效度的量表，將來才能建構可靠與正確的經驗性政治理論。本文主要以二〇〇三年與二〇〇四年「台灣選舉與民主化調查」（TEDS）兩次大型調查爲例，統計檢定在這兩次調查中，幾組重要的政治態度與行爲量表的信度問題。信度問題的檢定方法，大致可分爲再測信度檢定與內部一致性檢定，針對再測信度的檢定，本文主要透過卡帕（Kappa）係數，以及以平方加權法爲計算基礎的加權卡帕（weighted Kappa）係數來作檢驗分析；至於內部一致性的檢定，最常被使用的是克朗巴阿爾發的係數（*Cronbach's* α），以及折半信度中的蓋特門（Guttman）的計算公式。透過這四種方法檢驗 TEDS (2003) 與 TEDS (2004P) 兩次大型調查後發現，整體而言這兩次調查信度的檢驗結果都不錯，過去被認爲非常敏感的政治性題目，例如選民的投票支持對象、族群認同、統獨立場與其政黨認同等問題，都具有穩定性。其他一些政治態度量表，也具有高度的內部一致性。當然，目前仍有少數幾個題目信度的檢驗結果不佳，這些測量題目將來應在問卷討論中進一步修正。

關鍵詞：政治態度、行為量表、再測信度檢定、一致性檢定、台灣選舉與民主化調查

---

* 本文初稿曾在 2005 年 3 月「台灣選舉與民主化調查」計畫（TEDS）所舉辦的研究工作坊中發表，作者非常感謝評論人黃紀教授，以及兩位匿名審查人所提供的修正意見。本文分析的資料來自國科會補助的 TEDS 2003 與 2004P 兩次調查計畫，計畫主持人分別是中研院政治所籌備處特聘研究員朱雲漢教授，以及東吳大學政治系黃秀端教授。此外，感謝研究助理彭士芬、劉美姿、蔡宗漢與李佳怡在資料分析上的協助。

** 台灣大學政治學系，e-mail: yutzung@ntu.edu.tw。

收稿日期：94 年 7 月 28 日；通過日期：94 年 12 月 12 日

臺灣大學學術期刊資料庫

壹、導　論

從古典測量理論（classical test theory）的角度，評斷一個「好」或「正確」的測量方法，至少包括信度（reliability）與效度（validity）這兩個判斷指標。信度是指測量工具正確性（accuracy）與精確性（precision）的程度，亦即測量結果不受測量誤差（errors of measurement）影響的程度。因此，信度具有兩種意義：第一是穩定性（stability），亦即在不同的時間點上，針對同一個對象（不論是人或物）進行施測，然後比較兩次結果相關的程度；第二是一致性（consistency），在所有態度量表中各項目之間必須具有內部一致的同質性（邱政皓，2002：3-7）。信度檢定的目的在瞭解各種調查所使用的測量工具之穩定程度，並確保研究的品質。目前信度檢定主要的評估方法有再測信度檢定（test-retest reliability），折半信度檢定（split-half reliability），複本信度檢定（alternate form reliability），克朗巴阿爾發的係數（*Cronbach's* α）分析等。至於效度，指的是測量結果真正能測量到所要的目的之程度，所涉及的是測量結果是否正確。至於信度與效度兩者的關係，信度與測量結果的正確與否並無關係，信度所關心的只是測量本身是否穩定，結果是否可靠，以及測量是否都一致的問題。因此，信度是效度的必要條件但不是充分條件，如果測驗本身缺少信度，也就無效度可言了（陳順宇，2000：8-7）。至於影響信度與效度的因素，主要受到受訪者的基本背景、心理與情緒、測量內容、訪問情境、訪員因素等所影響（吳統雄，1985；杜素豪，2004）。

過去幾次 TEDS（台灣選舉與民主化調查，英文為 Taiwan's Election and Democratization Study，故英文簡稱為 TEDS）計畫，或是台灣過去二十多年來歷次政治類型的大型調查計畫，在執行過程中都只有注意樣本代表性檢定與樣本加權的問題，而沒有系統地針對問卷中重要的態度量表，例如民主價值量表、統獨立場量表進行信度與效度的檢查。測量工具的穩定性、一致性與正確性，對從事量化研究的人而言，其重要性不可言喻。經驗性理論的建構，不論是從理論開始著手，或是從資料中開始歸納，測量指標的信度與效度是成功與否重要的關鍵之一（Thye, 2000）。

臺灣大學學術期刊資料庫

本文主要針對 TEDS 二〇〇三與二〇〇四年兩次大型調查中幾個重要的量表，例如投票抉擇、政治信任、政黨認同、民主價值取向與統獨立場等指標，運用各種統計模型的分析，進行政治態度與行爲量表的信度檢定與分析。本文以下第一部分將簡單介紹目前各種運用在信度檢定的理論與方法，第二部分將使用卡帕（Kappa）與加權卡帕（weighted Kappa）這兩種方法檢定 TEDS 量表的再測信度，第三部分將使用蓋特門（Guttman）公式計算折半係數（split-half reliability），以及最被廣泛使用於估算量表內部一致性的克朗巴阿爾發的係數（*Cronbach's* α）。透過這四種方法檢驗 TEDS 2003 與 TEDS 2004P 兩次大型調查後發現，整體而言這兩次調查信度的檢驗結果都不錯，過去被認爲非常敏感的政治性題目，例如選民的投票支持對象、族群認同、統獨立場與其政黨認同等問題，都有不錯的穩定性。此外，其他一些政治態度量表，同時也具有高度的內部一致性。當然，目前仍有少數幾個題目的檢驗結果不佳，這些測量題目可在以後的調查中進一步被修正。

貳、信度檢定的理論與方法

古典測量主要的原理，建立在測驗分數的變異理論上，對於某個變項的測量分數，必定是落在一定範圍之內，因此可以得到該變項的離散程度，此爲變異數。這些分數的變異主要由兩個因素造成，一是真實分數的變異，以及因爲隨機誤差（random error）產生的變異。誤差 (E) 是指觀察分數 (O) 與真實分數 (T) 之間的差距，誤差的來源可分爲系統性誤差（systematic）以及非系統性誤差或隨機性誤差（random errors）（Crocker and Algina, 1986）。系統性誤差又可分爲恆定誤差（constant）與偏誤（bias）兩種。對所有受測者皆有影響的誤差爲恆定誤差，只對某些人會產生影響的誤差爲偏誤，至於隨機誤差則是隨機產生的。

最常被用來說明測量分數信度的是心理學家所發展的「真實分數理論信度模式」（林師模、陳苑欽，2003：368）。假設在衡量變數時，得到一觀察值 Xo，而受測者的真實分數爲 Xt，若兩者不相等，則有一誤差爲 Xe。三者的關係如下：

臺灣大學學術期刊資料庫

$$Xo=Xt+Xe$$

其中 Xe 滿足以下假設：

(1)　E(Xe)=0

(2)　Xt 與 Xe 相互獨立，其相關係數爲 0。

由以上假設得

$$E(Xo)=E(Xt)$$

$$\sigma^2_{Xo}=\sigma^2_{Xt}+\sigma^2_{Xe}$$

則信度可表示爲

$$r=\frac{\sigma^2_{Xt}}{\sigma^2_{Xo}}=\frac{\sigma^2_{Xo}-\sigma^2_{Xe}}{\sigma^2_{Xo}}=1-\frac{\sigma^2_{Xe}}{\sigma^2_{Xo}}$$

一般而言，我們可以將信度指標（index）與係數（coefficient）歸類爲三大類型：穩定性（stability）、等值性（equivalence）與內部一致性（internal consistency）。穩定性是指同一測量工具對相同受測者重複施測結果相同；等值性（equivalence）是指不同調查者將同一方法應用在不同情境，得到等值結果；內部一致性（internal consistency）是指許多問項（或測量題目）測量的結果，指向同一現象或面向。

一、穩定性

再測信度（test-retest reliability）主要評估量表的穩定性，又稱爲外在信度（external reliability），主要是利用兩次不同時間、相同受測者、相同題目所測得的結果計算相關係數，此相關係數稱爲穩定係數（the coefficient of stability）。係數越高，表示兩次結果越穩定，通常時間相隔約在兩週至數個月之間。再測信度可以根據資料的性質以不同的方法進行信度檢定。第一種處理方式是，如果資料屬於名目變項（nominal）（如投票取向）或

臺灣大學學術期刊資料庫

二分類，則以兩次調查得到同樣結果的百分比之平方根爲正確率。例如兩次調查所得同樣的比率爲 81%，正確率就等於 90%。如果調查變項爲等距（interval）或等比尺度（ratio），則以兩次調查的的相關係數 (r) 來評估再測信度（Carmines and Zeller, 1979；Prost and Zeifang, 1987；黃毅志，1997）。

第二種方法是 Kappa 係數，如果變項的選項呈現的是兩個方向，則使用 Kappa 係數來評估再測信度；如果變項是同一方向則可以使用成對樣本 t 檢定（paired t-test）。舉例而言，若題目類型是詢問受訪者是同意或不同意，則屬於前者；若是詢問受訪者對我國民主現況的評價，以 0 到 10 給一個分數，則是屬於後者。使用 t 檢定可以觀察其相關性與 t 值的檢定水準。

Kappa 係數：最早由 J. Cohen 所提出，在不同時間，對相同受測者，實施相同測量尺度的測量，所得到的兩個觀察值之間是否一致，也就是此變項是否穩定（Everitt, 1997: 146-148l；林秀娟、張紹勳，1995）。Kappa 係數的計算可利用列聯表（contingency）進行思考，以二維列聯表爲例，假如受測者的回答是穩定的，則在 (O_{11}, O_{21})、(O_{12}, O_{22}) 的次數總和會等於總次數，也就是說 (O_{11}, O_{21})、(O_{12}, O_{22}) 的和佔總次數的百分比會等於 1。其算式如下：

$$P_0 = \sum_{i}^{r} n_{ii} / N$$

但必須考慮到按照邊際（marginal）的次數採取隨機配置的比例，也就是各個格子（cell）的期望值佔總次數的百分比。其計算方式如下：

$$P_c = \frac{1}{N} \sum_{i=1}^{r} n_i n_i / N$$

所以實際次數一致的百分比 (P_0) 與理論上一致且經過機誤（chance）校正之後的最大可能次數百分比 (P_c) 的比率，即爲 Kappa 係數，其算式如下：

臺灣大學學術期刊資料庫

$$\kappa = \frac{P_o - P_c}{1 - P_c}$$

Landis 和 Koch 提出以下 Kappa 值的判斷標準（Everitt, 1997: 148）：

0：完全不一致（poor）
0.01～0.2：稍微一致（slight）
0.21～0.4：普通一致（fair）
0.41～0.6：中度一致（moderate）
0.61～0.8：相當一致（substantial）
0.81～1：幾乎完全一致（almost perfect）

第三種評估穩定性的分法爲 weighted Kappa。weighted Kappa 是由 Cicchetti 所提出的，它適用於等級順序的變數（ranking ordered variables）且必須是 2×2 以上的列聯表才可使用，若列聯表爲 2×2 的，則計算 Kappa 與 weighted Kappa 值是一樣的。與 Kappa 不同的是它考量了尺度間的相關性問題，而 Kappa 值並未考量這個因素。但是，政治態度量表都是屬於 Likert Scale（例如非常同意、同意、不同意、非常不同意），理論上受訪者在第一次回答「同意」，第二次回答「非常同意」的差距，要比第一次回答「同意」，第二次回答「不同意」的差異距離要小，加權卡帕就是利用公式計算因兩次回答的相同或不同型而給予適當的加權值，這要比傳統的 Kappa 將受訪者第一次回答「同意」，第二次回答『非常同意』，與第一次回答「同意」，第二次回答「不同意」一樣，都視爲「不一致」，而沒有考慮其實兩者是有差異的。因此，加權卡帕要比卡帕更精細一些。[1]

weighted Kappa 係數的公式如下所示[2]：

---

[1] 此觀點源自作者參加 2005 年 5 月 TEDS 舉辦的研討會中，由中研院調查研究工作室杜素豪教授針對另一篇有關信度檢定所提出的評論意見。

[2] 有關 weighted Kappa 的公式、證明與特性的介紹，請參閱：Schuster (2004); Martikainen et al. (1999); Murphy et al. (2000)。或參閱 SAS 使用手冊（http://www.id.unizh.ch/software/unix/statmath/sas/sasdoc/stat/chap28/sect26.htm）；stata7.0 使用手冊 pp.149-160。

臺灣小學學術期刊資料庫

$$0 <= w_{ij} < 1 \text{ for all } i \neq j,\ w_{ii} = 1 \text{ for all i, and } w_{ij} = w_{ji}.$$

$$\hat{\kappa}_w = \frac{P_{0(w)} - P_{e(w)}}{1 - P_{e(w)}}$$

$$\text{Where} \quad P_{0(w)} = \sum_i \sum_j w_{ij} p_i p_j \quad \text{and} \quad P_{e(w)} = \sum_i \sum_j w_{ij} p_{i\bullet} p_{\bullet j}$$

目前 weighted Kappa 係數 (w_{ii}) 有兩種計算的方式，一個爲 Cicchetti-Allision 加權的類型，另一個爲 Fleiss-Cohen 加權的類型：

㈠**線性加權法**（Kappa with Linear Weighting）

源自 Cicchetti 和 Allison 兩人，加權的公式爲：

$$w_{ij} = 1 - \frac{|i - j|}{(k - 1)}\ ,$$

（i=列值，j=行值，k=類別數，i, j=1, ….., k）

舉例來說，兩次調查中，如果受訪者對同一個問題第一次時選擇第一個選項（即 j = 1），第二次回答同一個選項（即 i = 1）則權值爲 1，加權值依據兩次回答差距的加大而呈現均等的遞減（每次遞減 0.25）。如果受訪者對同一個問題第一次時選擇第一個選項（即 j=1），第二次回答選擇第五個選項（即 i=5）則權值爲 0。

表 1　線性加權法的權值計算

i \ j	1	2	3	4	5
1	1.00	0.75	0.50	0.25	0.00
2	0.75	1.00	0.75	0.50	0.25
3	0.50	0.75	1.00	0.75	0.50
4	0.25	0.50	0.75	1.00	0.75
5	0.00	0.25	0.50	0.75	1.00

臺灣大學學術期刊資料庫

㈡**平方加權法**（Kappa with Quadratic Weighting）

源自 Fleiss 和 Cohen 兩人，加權的公式爲：

$$w_{ij} = 1 - \left(\frac{i-j}{k-1}\right)^2$$

（i=列值，j=行值，k=類別數，i, j=1, ….., k）

舉例來說，兩次調查中，如果受訪者對同一個問題第一次時選擇第一個選項（即 j=1），第二次回答同一個選項（即 i=1）則權值爲 1，加權值依據兩次回答差距的加大而遞減，但不是平均地遞減。如果受訪者對同一個問題第一次時選擇第一個選項（即 j=1），第二次回答選擇第二個選項（即 i=2）則權值尚有爲 0.9375，而不是 0.75。

表 2　平方加權法的權值計算

i \ j	1	2	3	4	5
1	1.0000	0.9375	0.7500	0.4375	0.0000
2	0.9375	1.0000	0.9375	0.7500	0.4375
3	0.7500	0.9375	0.1111	0.9375	0.7500
4	0.4375	0.7500	0.9375	1.0000	0.9375
5	0.0000	0.4375	0.7500	0.9375	1.0000

二、等值性

複本信度（alternative-form reliability）主要評估量表的等值性，亦即一套測驗採用兩種以上性質相同的複本。複本信度所使用的複本必須與原測驗在題數、題型、內容、難度、鑑別度等皆屬一致才可以。複本測驗若對一群受試者同時接受測驗的得分，其相關係數稱爲等值係數（the coefficient of equivalence）。若在不同時間施測所得係數稱爲穩定與等值係數（the coefficient of stability and equivalence）。複本信度的優點是，可說明內容和時間變異所造成的誤差情形，避免再測信度出現的缺點（例如受訪者記憶、

臺灣大學學術期刊資料庫

受測環境的改變、兩次間隔的時間等）（Carmines and Zeller, 1979；林師模，陳苑欽，2003：370）。

三、內部一致性

評估量表內部一致性的方法很多，本文主要討論以下幾種方法：折半信度（split-half reliability）、庫李信度（*KR* reliability）、克朗巴阿爾發的係數（*Cronbach's* α）與 *Hoyt* 係數。[3]

㈠折半信度（split-half reliability）

將受試者的測驗結果，按題目的單雙數或是其他分法分成兩半，再根據各二半的測驗上的結果，計算其一致性。折半法因爲只需舉行一次測驗，較節省人力及時間，所以是最常用的方法之一，而且也不像再測法是分兩個時間進行，所以不會有記憶效應問題。缺點是折半法必須是平行測驗，也就是假設兩半的誤差項的變異數會相等，另外若分成兩半的分法不同，所計算出的信度也會不同（陳順宇，2000：8-15）。

計算折半信度通常有兩種算法：

1. Guttman 公式：

$$r_{xx'} = 2\left[1 - \frac{S_a^2 + S_b^2}{S_x^2}\right]$$

（S_a^2 與 S_b^2 爲兩部分測驗的變異量，S_x^2 爲整個測驗總分的變異量）

2. 計算兩部分相關係數後，再用斯布公式（Spearman-Brown coefficient）加以校正：

$$r_{xx} = \frac{2r_{hh}}{1 + r_{hh}}$$

（r_{xx} 爲測驗真正的信度係數，r_{hh} 爲原測驗的信度係數）。

[3] 以下有關內部一致性的信度檢定之討論，請參考台灣大學心理系網站上的相關資料：http://www.psy.ntu.edu.tw/pm.method。

臺灣大學學術期刊資料庫

但是，斯布公式的基本假設是在兩半測驗的「變異數相等」時才可使用，如違反此假設，會導致信度被高估。

㈡庫李信度（KR reliability）

庫李信度適用在答案只有對錯兩者、或是選項只有兩個時。有庫李 20 號及 21 號信度，如果假設測驗題目的每題的困難度都相同，亦即測驗中每題答對與答錯人數百分比的乘積都相等，則可以用庫李 21 號公式求得信度係數。庫李 20 與 21 號公式的差別就是對題目困難度的假設不同，當所有題目的困難度都相等，則兩個公式所計算出來的信度係數相同，如果題目困難度不一，則庫李 21 號公式所計算的信度係數會比 20 號公式要來的低。

$$KR\text{-}20 = \frac{K}{K-1}\left[1 - \frac{\sum pq}{S_T^2}\right]$$

（K：測驗的題數。p：答對某一題的人數百分比。q：答錯某一題的人數百分比。$(q = 1 - p)p$ ：答對某一題的人數百分比。S_T^2 ：測驗總分的變異量。）

$$KR\text{-}21 = \frac{k}{k-1}\left[1 - \frac{\overline{X}(k - \overline{X})}{kS_T^2}\right]$$

（K：測驗的題數。$\overline{X}$：測驗總分的平均數。S_T^2：測驗總分的變異量。）

㈢ Cronbach's α的係數

Cronbach's α的係數則是檢視題目間的一致性來決定信度高低的一種信度指標，若同一個分量表中的題目領域愈接近則其一致性便愈高，原則上 α 係數的值愈接近 1 則其信度愈高。Nunnally (1978) 建議，此係數 0.5 或 0.6 以上就可以使用，但若低於 0.35 就過低。Cronbach's α 的係數有兩個缺點，一是對同類測量模式，它低估了信度，一是它對只有單一項目的測量（即 n=1 的情況）沒有辦法處理（陳順宇，2000：8-17）。其公式如下：

臺灣大學學術期刊資料庫

$$\alpha_k = \frac{k}{k-1}\left[1 - \frac{\sum S_i^2}{S_T^2}\right]$$

（K：測驗的題數。S_i^2：每一題分數的變異量。S_T^2：測驗總分的變異量。）

用 *Cronbach'* α的係數計算出內部一致性後，可以採用刪除某一題（If item deleted）後來看是否增加原來的信度。藉由每一個項目中，項目分數與總分的相關性，來刪減可能不具信度的題項。

㈣ Hoyt 係數

Hoyt 信度係數的基本假設是：若問卷的題目是在測量同一種行爲層面時，受測者在所有題目上的表現應該會呈現一致的狀況，因此在每一個分量表題目間的變異情形應該不會有太大的差異，也就是說受測者在題目間的變異量（SSa）與測驗誤差項（Sssa）二者之間的比值應該不會有顯著差異，因此在進行單因子變異數分析時，若 F 值未達到顯著水準，便表示問卷具有不錯的信度。Hoyt 係數和 *Cronbach's* α的係數會相近，他們的代數式被統計學家證明是相等的。（林師模、陳苑欽，2003：374）其公式如下：

$$\text{Hoyt ANOVA} = 1 - \frac{MS_{errors}}{MS_{individuals}}$$

（MS_{errors} 代表殘差的均方，$MS_{individuals}$ 代表受測者間的均方）

參、缺失值的處理方式

本文是檢定 TEDS 量表的效度問題，也就是分析政治態度量表的穩定性與一致性，對於資料中出現缺失值的問題，針對再測信度 Kappa 係數的計算，將不摒除任何缺失值，所有樣本都納入分析，也不合併任何缺失值，唯獨有一、兩個個案因訪員漏問該題而不得不放棄。

另一方面，在計算 weighted Kappa 與 *Cronbach's* α的係數時，因這兩

臺灣大學學術期刊資料庫

種計算方法都假設變項是數值（quantitative）的資料。在此情況下，目前在各種問卷調查資料的登錄方法，有兩種情形需要特別討論，第一種情況是當受訪者對測量題目回答「不知道」。依照目前國際上大型調查研究計畫的處理方法，回答「不知道」並不難以處理，「不知道」通常是合理且有實質意義的答案，因此不該被摒除。如果回答「不知道」被當缺失的個案（missing cases），使其在統計分析中隱形，此種作法的缺點是使得有效案例減少，將造成樣本代表性的偏差。一個比較好的替代方案是，將回答「不知道」者，在態度的反應量表（response scale）上，重新定義爲一個在理論上「可防禦」的數值（defensible spots）。以政治參與的問題爲例，我們假設那些不知道是否曾經參與此類活動的人，定義爲從未參加類似的活動。以政治態度的量表爲例（以 4 個等級代表強烈同意一直至強烈不同意），本文將「不知道」者重新編碼爲中間或中立（middle or neutral）的數值。在這種情況中，我們假設「不知道」等同於沒有影響力（zero effect）（Bratton et al., 2005；Chu et al., 2003）。

第二種有問題的是答案完全缺失的資料（complete absence of data）。例如在少數情況下，受訪者拒絕提供答案（這與回答不知道者不同）。此外，在極少的情況下，可能因訪員的錯誤，造成答案漏填的情形。當嘗試評估資料缺失的個案，是否與觀察到的資料有所不同時，研究者需要界定可忽略與不可忽略（ignorable and nonignorable）的遺漏值。如果爲不可忽略的缺失值，則表示缺失值出自於系統性的偏差（systematic biases），例如一個逃稅者比較不可能回答有關所得的問題；而低薪階級者，可能較無法在職業欄中提供答案。在這些例子中，我們需要再知道更多有關這筆資料的資訊，再來填補缺失值。當缺失值是可忽略時，表示缺失值與該變項 (Y) 真實的數值無關，缺失值是隨機分佈的。同時，我們可以利用其它變項的資料 (X)，預測該變項 (Y) 缺失值可能的數據。例如，一個人回不回答是否支持民主體制的機率，與某人是否對總統的施政表現是否滿意，應該沒有相關（Pharr and Putnam, 2000）。在這樣的條件下，稱之爲隨機遺漏（missing at random）。針對隨機遺漏問題的處理，King 等人（2001）指出，如果使用整筆刪除的方法（listwise deletion），將造成低效率與偏差估計的結果（inefficient and biased estimates）。因此，他們建議使用多重差補（multiple

臺灣博碩士學術期刊資料庫

imputation）的方法，使用已知其他變數的資訊，尋找數個數值替代遺漏值（missing values）。但這是有疑問的，因爲必須假定缺失值是隨機分佈的。另一種方法是利用 Heckman，Dubin 與 Rivers 等人所發展出來兩階段的選樣偏誤模型（selection bias model），但除非得到母體某項特質，例如候選人實際的得票率，我們才能比對修正後的模型是否比沒有修正過的模型更貼近母體。可惜的是，我們並不知道全台灣地區民眾真實的政治態度爲何，例如政黨認同、政治信任度，可能只有「神」或「上帝」知道。因此，由於目前缺失值所佔總樣本數不是很多，大都在 1% 以下，本文認爲這對信度檢定最後結果的影響程度不是很大，除非有更好的方法，本文在此擬不再處理。

肆、再測信度的檢定

TEDS (2004P) 調查計畫，在二〇〇四年的七月開始進行，約在九月初完成調查工作，隨即在二個禮拜後，進行再測信度訪問，總共完成 365 份再測問卷。表 3 顯示，TEDS (2004P) 透過 Kappa 係數的檢定，其中在選民的投票參與部分，問其是否有去投票，共有四種答案：有去投票、沒有去投票、不知道與拒答，再測信度檢定的結果是相當一致，Kappa 係數爲 .617。此外，再問選民投票的對象，共有四種答案：陳水扁與呂秀蓮、連戰與宋楚瑜、不知道與拒答，選民兩次回答的結果幾乎完全一致，Kappa 係數高達 .820。在選舉評估方面，共有六種答案：非常公平、公平、不公平、非常不公平、不知道與拒答，再測信度檢定的結果是中度一致，Kappa 係數爲 .413。如果將選項進一步歸類四分類（回答非常同意與同意者爲一類，非常不同意與不同意者爲一類，另加不知道與拒答），Kappa 係數就升高爲 .584，快要達到相當一致的標準。

在候選人形象部分，如果將選項六分：選民回答非常有魄力、非常清廉、非常誠實、非常瞭解民眾需要、非常有親和力爲一類；有魄力、有清廉、有誠實、有瞭解民眾需要、有親和力爲一類；不太有魄力、不太有清

臺灣大學學術期刊資料庫

表 3　再測信度檢定表：Kappa 係數

	Kappa 係數		樣 本 數
	選項六分	選項四分	
㈠投票參與			
H01 在這一次（三月二十日）舉行的總統大選中，有很多人去投票，也有很多人因各種原因沒有去投票，請問您有沒有去投票？	-	0.6168	359*
H1B 請問您投票給哪一組候選人？	-	0.8203	324（扣除未投票者）
R04 整體而言，請問您認爲這次總統大選的過程公不公平？	0.4132	0.5843	358（扣除 1 位訪員漏問）
㈡候選人形象			
J01 首先，在魄力方面：請問您認爲陳水扁的魄力如何？	0.2779	0.3924	359
J1A 連戰呢？您認爲連戰的魄力如何？	0.3160	0.4130	359
J02 在清廉不貪污方面：請問您認爲陳水扁是否清廉不貪污？	0.4184	0.4929	359
J2A 連戰呢，您認爲連戰是否清廉不貪污？	0.3497	0.3955	359
J03 再來，在誠信（有信用）方面：請問您認爲陳水扁是否誠實有信用？	0.4375	0.5454	359
J3A 連戰呢?您認爲連戰是否誠實有信用？	0.3325	0.4086	359
J04 在瞭解民眾需要方面：請問您認爲陳水扁是否瞭解民眾的需要？	0.3357	0.4238	359
J4A 連戰呢？您認爲連戰是否瞭解民眾的需要？	0.2559	0.3023	358（扣除 1 位訪員漏問）
J05 在具有親和力（平易近人）方面：請問您認爲陳水扁是否具有親和力（平易近人）？	0.3471	0.5009	358（扣除 1 位訪員漏問）
J5A 連戰呢？您認爲連戰是否具有親和力（平易近人）？	0.3527	0.4726	358（扣除 1 位訪員漏問）
㈢候選人能力			
K01 請問您認爲陳水扁這一組候選人促進省籍與族群和諧的能力如何？	0.4245	0.5215	358（扣除 1 位訪員漏問）

臺灣人文學術期刊資料庫

	Kappa 係數		樣 本 數
	選項六分	選項四分	
K1A 連戰這一組呢？	0.3829	0.4459	357（扣除 2 位訪員漏問）
K03 請問您認爲陳水扁這一組候選人促進經濟發展的能力如何？	0.3789	0.4929	359
K3A 連戰這一組呢？	0.3642	0.4326	358（扣除 1 位訪員漏問）
K06 請問您認爲陳水扁這一組候選人推動民主改革的能力如何？	0.2956	0.3958	357（扣除 2 位訪員漏問）
K6A 連戰這一組呢？	0.3749	0.3996	357（扣除 2 位訪員漏問）
㈣族群認同			
N01 在我們（台語：咱）社會上，有人說自己是「台灣人」，也有人說自己是「中國人」，也有人說都是。請問您認爲自己是「台灣人」、「中國人」，或者都是？	-	0.5580**	359
㈤政黨認同			
P01 目前國內有幾個主要政黨，包括國民黨、民進黨、親民黨、新黨、建國黨，以及台灣團結聯盟，請問您是否（台：敢有）偏向哪一個政黨？	-	0.3728	359
P1B 請問是哪一個政黨？	0.8118***	-	206（扣除無政黨傾向者）
P03 這一次（今年三月）的總統選舉，有人認爲是「泛綠」與「泛藍」兩大陣營的競爭；而人們有時候會認爲自己是屬於「泛綠」陣營的，也有些人認爲自己是屬於「泛藍」陣營的。請問您覺得自己是偏向「泛綠」陣營，還是「泛藍」陣營？	-	0.5720	359

註：*再測信度問卷總共完成 359 份，除訪員漏問，以及跳答題外，其餘皆進入分析的個案中。

**此題選項爲五分類。

***此題選項爲九分類。

資料來源：TEDS（台灣選舉與民主化調查），2004P。

臺灣大學學術期刊資料庫

廉不貪污、不太有誠實、不太瞭解民眾需要、不太有親和力爲一類；一點也沒有魄力、一點也不清廉、一點也不誠實、一點也不瞭解民眾需要、一點也沒有親和力爲一類；不知道與拒答各爲一類。再測信度檢定的結果不是很好，介於普通一致到中度一致之間，Kappa 係數介於 .256 到 .438 之間。如果再將選民對兩組候選人形象正面評價的歸爲一類，負面評價的歸爲一類，再加上不知道與拒答者共爲四類，Kappa 係數就提高一些，介於 .302 到 .545 之間。然而，不論是選項六分或四分，都顯示選民對候選人形象評價的穩定程度不是很好，不如上面的投票參與。

在候選人能力部分，如果將選項六分，選民回答非常有能力促進省籍與族群的和諧、發展經濟、推動改革能力歸爲一類；回答有能力促進省籍與族群的和諧、發展經濟、推動改革能力歸爲一類；回答不太有能力促進省籍與族群的和諧、發展經濟、推動改革能力歸爲一類；回答一點也沒有能力促進省籍與族群的和諧、發展經濟、推動改革能力歸爲一類；不知道與拒答各爲一類。再測信度檢定的結果介於普通一致與中度一致之間，Kappa 係數介於 .296 到 .425 之間。如果再將選民對兩組候選人能力正面評價的歸爲一類，負面評價的歸爲一類，再加上不知道與拒答者共爲四類，Kappa 係數就提高一些，介於 .396 到 .522 之間。然而，不論是選項六分或四分，都顯示選民對候選人能力評價的穩定程度雖然比上面候選人形象的評價要好一些，但不如上面的投票參與。

在族群認同上，如果將選項歸爲五類：認同中國人、認同台灣人、兩者都是、不知道與拒答，兩次訪問再測信度檢定的結果中度一致，Kappa 係數達 .558。在選民的政黨認同部分，如果將選項歸爲四類：有政黨認同、無政黨認同、不知道與拒答，Kappa 係數是 .373，檢定結果還算一致。如果再將有政黨認同者繼續追問其支持對象，選項共分九類：國民黨、民進黨、親民黨、台聯、新黨、建國黨、其它政黨、不知道與拒答，Kappa 係數達 .812，兩次訪問檢定結果幾乎完全一致。如果將選民的政黨認同分爲「泛藍」與「泛綠」，檢定結果爲相當一致，Kappa 係數達 .572。

本文上面提過 weighted Kappa 適用於等級順序的變數，它考量了尺度間的相關性問題，且必須是 2×2 以上的列聯表才可使用，若列聯表爲 2×2 的，則計算 Kappa 與 weighted Kappa 值是一樣的。本文將受訪者的回答全

臺灣博碩士論文知識加值系統

部保留，並將回答看情形、無意見、不知道者視爲有意義的回答，計入分析的個案中，以中間值代入。至於因訪員漏問或拒答而出現的遺漏值，本文上面提過，因佔本文分析個案數的比率非常少，通常不到百分之一，故分析時將不處理這些個案。在權值計算的部分，並以平方加權法的方式計算加權值。表 4 中的比較結果顯示如果把題目的選項三分，weighted Kappa 係數只比 Kappa 係數稍微好一點，其中有幾個題目的再測信度檢定甚至更差。所以當選項比較少時，weighted Kappa 與 Kappa 係數差不多。但如果將題目的選項加多時，則 weighted Kappa 係數比 Kappa 係數好很多，幾乎都比沒有使用加權值代入的情況更好。

伍、內部一致性檢定

本文主要透過四個指標來評斷 TEDS 量表的內部一致性的檢定，分別是該題與全組題目的相關程度（Item-total correlation）、刪掉某題後所剩的 Alpha 值（Alpha if item deleted）、Chronbach's α 的係數與 Guttman 折半係數。

表 5 顯示，在民主治理部分，題目與題組相關程度最低，如果因題目過多，將來可以刪除並可以提高內部一致性的前五題，依序是：C09（我們的政治人物與政黨，從不公開究竟從金主財團募集多少政治獻金）、C08（現在所有重要的事情政府官員通常都會隱瞞，而不會主動讓老百姓知道）、C03（司法機關經常在沒有充分證據的情況下扣押調查的對象）、C10（現在的公務人員都用能拖就拖的態度，來處理公家的事情）與 C02（一般而言，我們中央政府官員都能按法律規定辦事，不會做違法的事）。整個題組的 Chronbach's α 值爲 .7108 還算不錯，Guttman 折半係數爲 .6516。有效個案數有 1127，而因漏問或拒答被刪除的個案數只有 37 筆。

表 6 顯示，在政治信任部分，題目與題組相關程度最低，如果因題目過多，將來可以刪除並可以提高內部一致性前三題，依序是： E1J（報紙）、E1H（軍隊）與 E1K（電視）。整個題組的 Chronbach's α 的係數爲 .7334 還算不錯，Guttman 折半係數爲 .6330。有效個案數有 1135，而因漏問與拒答被刪除的個案數只有 29 筆。

臺灣博碩士論文知識加值系統 期刊資料庫

表 4　Kappa 係數與 Weighted Kappa 係數的比較

題號	N	3×3 groups			Weight 3×3 groups			5×5 groups			Weight 5×5 groups		
		agreement	κ	95% CI	agreement	K	95% CI	agreement	κ	95% CI	agreement	κ	95% CI
H01	355	93.49	0.6289	0.462-0.796	95.49	0.6289	0.462-0.796	-	-	-	-	-	-
H1b	271	96.31	0.9281	0.891-0.956	97.69	0.9528	0.920-0.973	-	-	-	-	-	-
R04	347	76.66	0.6123	0.562-0.639	87.03	0.7044	0.693-0.739	58.21	0.4297	0.381-0.449	94.34	0.6957	0.678-0.715
J01	354	66.38	0.3990	0.388-0.428	76.77	0.4109	0.385-0.429	51.69	0.2818	0.230-0.294	92.06	0.4563	0.411-0.511
J1a	356	67.98	0.4180	0.364-0480	77.67	0.3583	0.332-0.397	53.37	0.3195	0.307-0.349	92.63	0.4337	0.364-0.492
J02	350	68.86	0.5117	0.479-0.523	83.64	0.5848	0.580-0.593	58.00	0.4325	0.409-0.502	93.87	0.6165	0.596-0.642
J2a	355	60.85	0.4018	0.376-0.426	80.28	0.4676	0.405-0.516	53.24	0.3551	0.259-0.386	93.47	0.5043	0.468-0.520
J03	354	73.73	0.5586	0.512-0.580	83.90	0.6272	0.571-0.684	59.60	0.4450	0.412-0.490	94.58	0.6923	0.682-0.729
J3a	352	63.07	0.4210	0.339-0.469	78.84	0.4694	0.459-0.508	53.98	0.3422	0.303-0.379	93.54	0.5056	0.451-0.546
J04	355	65.63	0.4320	0.418-0.472	76.20	0.4420	0.418-0.445	53.80	0.3406	0.330-0.361	92.25	0.5051	0.504-0.572
J4a	354	59.89	0.3068	0.215-0.340	71.33	0.2603	0.207-0.330	49.44	0.2597	0.236-0.284	90.98	0.3425	0.304-0.419
J05	355	77.18	0.5018	0.488-0.589	85.00	0.5152	0.508-0.626	57.75	0.3460	0.319-0.284	93.77	0.5549	0.536-0.577
J5a	356	68.82	0.4711	0.421-0.488	78.30	0.4783	0.438-0.517	54.78	0.3507	0.288-0.401	92.98	0.5376	0.505-0.554
K01	355	71.55	0.5288	0.445-0.571	82.54	0.5774	0.532-0.599	57.18	0.4294	0.372-0.461	93.49	0.6439	0.575-0.673
K1a	354	65.54	0.4518	0.427-0.518	75.49	0.4071	0.387-0.454	56.78	0.3876	0.370-0.452	92.36	0.4673	0.444-0.491
K03	357	70.59	0.4964	0.471-0.525	82.14	0.5582	0.456-0.583	56.02	0.3812	0.322-0.411	93.80	0.5949	0.572-0.621
K3a	355	64.79	0.4374	0.433-0.456	78.31	0.4299	0.406-0.501	56.62	0.3681	0.329-0.448	92.94	0.4458	0.388-0.479
K06	353	66.57	0.4018	0.336-0.466	78.47	0.3799	0.205-0.457	52.12	0.2998	0.251-0.324	92.19	0.4712	0.446-0.539
K6a	354	61.63	0.4043	0.372-0.437	75.92	0.3762	0.312-0.481	56.50	0.3792	0.315-0.389	93.01	0.4301	0.388-0.462
M1a	352	66.48	0.4387	0.350-0.472	89.28	0.5235	0.431-0.597	50.85	0.2751	0.223-0.305	94.87	0.4986	0.495-0.533
M1b	353	74.50	0.4160	0.302-0.490	90.23	0.3464	0.248-0.412	39.09	0.2570	0.211-0.284	94.37	0.3726	0.300-0.519
M1c	354	69.49	0.3754	0.369-0.403	89.19	0.2877	0.242-0.374	38.14	0.2222	0.201-0.263	94.34	0.2684	0.167-0.314
N01	350	76.57	0.5769	0.534-0.648	92.86	0.6005	0.575-0.709	-	-	-	-	-	-
P01	348	68.39	0.3957	0.371-0.434	71.41	0.4157	0.396-0.459	-	-	-	-	-	-
P1b	203	89.16	0.8307	0.791-0.846	97.02	0.6256	0.471-0.760	-	-	-	-	-	-
P03	352	73.01	0.5895	0.574-0.653	92.19	0.7342	0.717-0.775	60.51	0.4514	0.427-0.519	96.73	0.7378	0.722-0.743

註：1. 加權方式採用 Fleiss-Cohen Kappa 加權係數（即 Kappa with Quadratic Weighting）。

2. 將回答『不知道』這四類重新歸成中間值；而訪員漏問、拒答以及跳答歸類成遺漏值。

3. P1b 題目含不知道者共有七個選項。

資料來源：TEDS（台灣選舉與民主化調查），2004P。

臺灣大學學術期刊資料庫

表 5　TEDS (2003) 內部一致性信度檢定：民主治理

	Item-total correlation	Alpha if item deleted	Chronbach's α	Guttman 折半係數
民　主　治　理			.7108	.6516
C01 C1．現在去政府機關辦事情，如果不找關係完全按規定來，很難辦得通。	.2962	.6983		
C02 C2．一般而言，我們中央政府官員都能按法律規定辦事，不會做違法的事。	.2618	.7018		
C03 C3．司法機關經常在沒有充分證據的情況下扣押調查的對象。	.2299	.7048		
C04 C4．因爲有定期的選舉，我們的政府首長在做決策時，不敢違背多數的民意。	.3752	.6900		
C05 C5．大部分立法委員選舉的時候口口聲聲都是爲民服務，當選以後都是爲自己利益。	.2746	.7007		
C06 C6．現在任何人都可以隨便批評政府，不必擔心會被找麻煩。	.4256	.6833		
C07 C7．現在任何人都可以自由參與示威抗議活動，不必擔心會被找麻煩。	.4462	.6814		
C08 C8．現在所有重要的事情政府官員通常都會試圖隱瞞，而不會主動讓老百姓知道。	.1704	.7115		
C09 C9．我們的政治人物與政黨，從不公開究竟從金主財團募集多少政治獻金。	.1241	.7169		
C10 C10·現在的公務人員都用能拖就拖的態度，來處理公家的事情。	.2412	.7043		
C11 C11．一般來說，政府在推動重要政策都很有效率。	.3775	.6901		
C12 C12·一般老百姓如果有事情需要跟政府機關打交道（台：辦代誌），您認爲會不會得到公平的對待？（訪員請出示 5 號卡片，並追問強弱度）	.4622	.6777		
C13 C13．一般老百姓如果上法院打官司，您認爲會不會得到公平的審判？（訪員請出示 5 號卡片，並追問強弱度）	.3286	.6946		

臺灣大學學術期刊資料庫

	Item-total correlation	Alpha if item deleted	Chronbach's α	Guttman 折半係數
C14 C14·您認爲現在高層政府官員與財團利益掛勾的情況普遍不普遍？（訪員請出示 6 號卡片，並追問強弱度）	.3862	.6870		
C15 C15·您認爲現在基層公務員貪污、拿紅包的情況普遍不普遍？（訪員請出示 6 號卡片，並追問強弱度）	.3218	.6981		
N=1127				

註：1. C1、C3、C5、C8、C9、C10 回答：「非常同意」爲 1 分、「同意」爲 2 分、「不知道」爲 3 分、「不同意」爲 4 分、「非常不同意」爲 5 分。C2、C4、C6、C7、C11 回答：「非常同意」爲 5 分，「同意」爲 4 分、「不知道」爲 3 分、「不同意」爲 2 分、「非常不同意」爲 1 分。C12、C13 回答：「一定會」爲 5 分、「會」爲 4 分、「不知道」爲 3 分、「不會」爲 2 分、「一定不會」爲 1 分。C14、C15 回答：「極少見」爲 5 分、「少見」爲 4 分、「不知道」爲 3 分、「有些」爲 2 分、「非常普遍」爲 1 分。

2. 資料中漏問與拒答者列爲遺漏值（共 37 筆），不作分析。

資料來源：TEDS（台灣選舉與民主化調查），2003。

表 6　TEDS (2003) 內部一致性信度檢定：政治信任

	Item-total correlation	Alpha if item deleted	Chronbach's α	Guttman 折半係數
政　治　信　任			.7334	.6330
E1A E1a · 下面列出一些機構名稱，請問您對這些機構信任的程度怎樣？您是非常信任，還算信任，不太信任，還是非常不信任？ E1B E1a　法院	.3955	.7360		
E1B E1b · 中央政府	.4726	.7248		
E1C E1c · 地方政府	.5164	.7189		
E1D E1d · 政黨	.3970	.7358		
E1E E1e · 立法院	.4656	.7256		
E1F E1f · 公務人員	.3894	.7370		
E1G E1g · 軍隊	.3561	.7411		
E1H E1h · 警察	.4250	.7319		
E1J E1j · 報紙	.3534	.7420		
E1K E1k · 電視	.3778	.7386		
N=1135				

註：1. 回答「非常信任」爲 5 分，「還算信任」爲 4 分、「不知道」爲 3 分、「不信任」爲 2 分、「非常不信任」爲 1 分。

2. 資料中漏問與拒答列爲遺漏值（共 29 筆），不作分析。

資料來源：TEDS（台灣選舉與民主化調查），2003。

臺灣大學學術期刊資料庫

表 7 顯示，在民主價值部分，題目與題組相關程度最低，如果因題目過多，將來可以刪除並可以提高內部一致性的前三題，依序是：F8A（教育程度很低的人，應該和教育程度高的人一樣，在政治上享有同等的發言

表 7　TEDS (2003) 內部一致性信度檢定：民主價值

	Item-total correlation	Alpha if item deleted	Chronbach's α	Guttman 折半係數
民　主　價　值			.6920	.6608
F8A F8a · 教育程度很低的人，應該和教育程度高的人一樣，在政治上享有同等的發言權。	.0702	.7070		
F8B F8b · 女性不應該像男性一樣參加政治活動。	.2766	.6804		
F8C F8c · 政府自然會爲人民解決困難，謀求福利，我們不必多作主張。	.4094	.6607		
F8D F8d · 政府首長等於是大家庭的家長，一切大小國事，皆應聽從他的決定。	.3792	.6650		
F8E F8e · 強而有力的領袖比優良的法律更（台：卡）重要。	.3650	.6669		
F8F F8f · 大家想法若不一致，社會就會混亂。	.3089	.6757		
F8G F8g · 一種觀念能否在社會上流傳，應由政府來決定。	.4142	.6601		
F8H F8h · 對付殘暴的罪犯應立即處罰，不必經過法院複雜的審理程序。	.2707	.6835		
F8J F8j · 在一個地方（台：社區）上，如果東（台：這）一個團體，西（台：那）一個團體，就會影響到地方的安定與和諧。	.4199	.6583		
F8K F8k · 一個國家如果有很多政黨，就會影響政治安定。	.4367	.6548		
F8M F8m · 法官在審理重大案件時，應該接受行政機關的意見。	.3364	.6716		
F8N F8n · 政府如果時常受到議會的牽制，就不可能有大作爲了。	.2322	.6871		
N=1146				

註：1. F8A F8a 回答：「非常同意」爲 5 分，「同意」爲 4 分、「不知道」爲 3 分、「不同意」爲 2 分、「非常不同意」爲 1 分。其餘各題回答：「非常不同意」爲 5 分，「不同意」爲 4 分、「不知道」爲 3 分、「同意」爲 2 分、「非常同意」爲 1 分。

2. 資料中漏答、跳答與拒答者列爲遺漏值（共 18 筆），不作分析。

資料來源：TEDS（台灣選舉與民主化調查），2003。

臺灣大學學術期刊資料庫

權）、F8N（政府如果時常受到議會的牽制，就不可能有大作爲了）、F8H（對付殘暴的罪犯應立即處罰，不必經過法院複雜的審理程序）。整個題組的 Chronbach's α 的係數爲 .6920 還算不錯，Guttman 折半係數爲 .6608。有效個案數有 1146，而因漏問與拒答被刪除的個案數只有 18 筆。

表 8 顯示，在政治效能感部分，題目與題組相關程度最低，如果因題目過多，將來可以刪除並可以提高內部一致性的前三題，依序是：C04（您覺得您對國家大事的了解，並不比一般人差）、C08（現在的政治人物爲了爭取權位往往不擇手段）與 C05（台灣不管由哪個政黨來執政，結果都

表 8　TEDS (2004P) 內部一致性信度檢定：政治效能感

	Item-total correlation	Alpha if item deleted	Chronbach's α	Guttman 折半係數
政 治 效 能 感			.6015	.6275
C01 我們一般老百姓對政府所作所爲，沒有任何影響力。	.3532	.5524		
C02 政府官員不會在乎（台：不會管）我們一般老百姓的想法。	.4339	.5257		
C03 政治有時候太複雜了，所以我們一般老百姓實在搞不懂（台：不清楚）。	.3264	.5622		
C04 您覺得您對國家大事的了解，並不比一般人差。	.0504	.6359		
C05 台灣不管由哪個政黨來執政，結果都是一樣的。	.3059	.5679		
C06 我們的中央政府官員，都不了解地方基層的需要。	.4049	.5371		
C07 政治是少數政治人物玩的遊戲，跟我們多數人無關，不值得我們關心。	.4119	.5334		
C08 現在的政治人物爲了爭取權位往往不擇手段。	.1180	.6147		
N=1800				

註：1. C04：「非常不同意」爲 1 分，「不同意」爲 2 分、「中間值」3 分、「同意」爲 4 分、「非常同意」爲 5 分。其餘各題：「非常不同意」爲 5 分，「不同意」爲 4 分、「不知道」3 分、「同意」爲 2 分、「非常同意」爲 1 分。

2. 資料中回答跳答、漏問與拒答者列爲遺漏值（共 23 筆），不作分析。

資料來源：TEDS（台灣選舉與民主化調查），2004P。

臺灣大學學術期刊資料庫

是一樣的）。整個題組的 Chronbach's α 值爲 .6015 稍微偏低，Guttman 折半係數爲 .6275。有效個案數有 1800，而因漏問與拒答被刪除的個案數只有 23 筆。

陸、結　論

任何從事經驗性理論的建構，或是從事經驗性研究，都必須要收集可靠的與正確的經驗資料，亦即極大化測量指標的效度與信度（King et al., 1994）。從以上初步的分析可知，整體而言 TEDS 量表的信度，不論就再測信度與內部一致性檢定，檢測的結果都還不錯，但其中有幾題檢測的結果並不如預期，或許以後可以加以修正或刪除。過去常常聽到許多人質疑在台灣執行政治性大型面訪調查的可信度，他們認爲台灣民眾面對政治性的敏感問題會不太願意回答，或是實問虛答，或是不會說實話等問題。但我們從某些特別敏感的題目，例如投票取向、政黨認同、族群認同與統獨立場上，得到的是不錯的信度檢定結果。過去台灣在威權時期，民眾對政治事務抱持觀望或不敢討論的態度，在經歷快速民主化過程後，我們對台灣民眾不談政治事務的刻板印象，似乎應該要拋棄了。

當然，除了使用古典的測量理論來檢視 TEDS 量表的信度問題外，最近幾年在測量理論裡，興起了項目分析理論（Item Response Theory）與概推度理論（Generalization Theory）二種看法。前者部分修正古典測量理論的假設（黃旻華，2004），後者則考慮測試者（訪員）、受試者（受訪者）與訪問情境等因素造成測量誤差。這兩種對古典測量理論的建議與修正，未來在檢定 TEDS 政治態度量表時，也應可一併考慮進來。

臺灣大學學術期刊資料庫

參考書目

一、中文部分

吳統雄，1985，〈態度與行爲研究的信度與效度：理論、應用、反省〉，《民意學術專刊》，民 74 年（夏季號）：29-53。

杜素豪，2004，〈投票意向問題不同類型項目無反應之分析：以 2000 年總統大選爲例〉，《選舉研究》，11(2): 111-132。

林秀娟、張紹勳，1995，《SPSS For Windows 統計分析－初等統計與高等統計》，下冊，台北：松崗。

林師模、陳苑欽，2003，《多變量分析－管理上的應用》，台北：雙葉。

邱政皓，2002，《量化研究與統計分析－SPSS 中文視窗版資料分析範例解析》，台北：五南。

陳順宇，2000，《多變量分析》，台北：華泰。

黃旻華，2004，〈統獨量表的心理計量學分析〉，發表於 2004 年《台灣政治學年會》學術論文發表會，高雄：義守大學。

黃毅志，1997，〈問卷調查資料之再測信度：以台灣地區社會變遷調查爲例〉，《調查研究》，3: 205-242。

二、英文部分

Bratton, Michael, Robert Mattes and E. Gyimah-Boadi. 2005. *Public Opinion, Democracy and Market Reform in Africa*. New York : Cambridge Univesity Press.

Carmines, G. Edward & Richard A. Zeller. 1979. *Reliability and Validity Assessment*. London: Sage.

Chu, Yun-han, Yu-tzung Chang and Fu Hu. 2003. “Regime Performance, Value Change and Authoritarian Detachment in East Asia.” Paper presented at the International Conference of *How East Asians View Democracy: The Region in Global Perspective*, Taipei, December 8-

臺灣大學學術期刊資料庫

9, Taipei.

Crocker, L. and J. Algina. 1986. *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart, and Winstton.

Everitt, B. S. 1992. *The Analysis of Contingency Tables* (2nd ed). London: Chapman & Hall.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation," *American Political Science Review* 95(1): 49-69.

King. Gary, Robert O. Keohane and Sidney Verba. 1994. *Design Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Martikainen, Pekka et al. 1999. "Reliability of Perceived Health by Sex and Age." *Social Science & Medicine* 48: 1117-1122.

Murphy, Debra A. et al. 2000. "Test-Retest Reliability of Dominic Anxiety and Depression Items Among Young Children." *Journal of Psychopathology and Behavioral Assessment* 22: 257-270.

Nunnally, J. C.1978. *Psychometric Theory*. New York: McGraw-Hill.

Pharr, Susan J. and Robert D. Putnam. 2000. *Disaffected Democracies*. Princeton: Princeton University Press.

Porst, R., and K. Zeifang. 1987. "A description of the German General Social Survey test-retest study and a report on the stability of the sociodemographic variables." *Sociology Methods and Research* 15(3): 177-218.

Schuster, Christof. 2004. "A note on the interpretation of weighted Kappa and ITS relations to other rater agreement statistics for metric scales," *Educational and Psychological Measurement* 64: 243-253.

Thye, Shane, R. 2000. "Reliability in Experimental Sociology." *Social Forces* 78: 1277-1309.

臺灣大學學術期刊資料庫

Reliability of Political Attitudes and Behavioral Scales in the Survey of TEDS:

Test-Retest Reliability and Internal Consistency Analysis

*Yu-Tzung Chang**

Abstract

In the last few years, empirical researches in political science have often neglected to systematically examine their measurement indices and methods. Concerning such a general neglect, this paper uses the large-scale surveys of TEDS conducted in 2003 and 2004 as the bases for analyzing a number of possible problems related to the reliability of the measurements of political attitudes and behavioral scales. The degree of reliability can be measured by performing a test-retest and then examining the result for the extent of internal consistency. Test-retest is carried out with the Kappa coefficient and the weighted Kappa coefficient with quadratic weighting. Internal consistency is tested with the widely used Cronbach's α coefficient and the Guttman formula of split-half reliability. Using these four measures to examine the two TEDS surveys in 2003 and 2004, we find that, on the whole, the two surveys pass our tests for reliability. It is found that a number of previously considered very sensitive political topics, such as voter support, ethnic identity, Taiwan independence, and party identification, all relatively reveal stable conditions of results. We also find that the most important measures of political attitude display a high degree of internal consistency. In addition to the stability found in the acquired results, our tests also reveal some unexpected results on a small number of items, which may require further revision in our future research.

Key words: political attitudes, behavioral scales, test-retest reliability, internal consistency, TEDS

* Assistant Professor, Department of Political Science, National Taiwan University. E-mail: yutzung@ntu.edu.tw.

臺灣大學學術期刊資料庫