

國立臺灣師範大學教育心理與輔導學系  
教育心理學報，民 97，39 卷，測驗與評量專刊，21—42 頁

# 以試題反應理論分析 CES-D 量表不同調查方式的差異效果\*

李仁豪

中山醫學大學  
心理學系

謝進昌

國立政治大學  
教育學系

余民寧

國立政治大學  
教育學系

本研究主要是探討網路問卷與紙本問卷這兩種調查方式是否有所差異，問卷內容是採用國際間廣泛使用的 CES-D 憂鬱量表。2400 位樣本分層隨機取自台灣中小學教師，其中共有 1141 位教師在 20 題 CES-D 量表上作答，包括 535 位網路組及 606 位紙本組教師。本研究藉由單參數試題反應理論軟體 ConQuest 開發的多向度潛在回歸模組來比較網路問卷組與紙本問卷組在問卷內容上的差異情形。研究結果顯示，儘管網路問卷回收率稍低於紙本問卷，但並未導致兩組受試者在背景變項上有所差異。在問卷內容的比較上，雖然網路問卷組在問卷內容的潛在平均數顯著地低於紙本問卷組，但是並無實質上的重要差異，亦即問卷組別變項對 CES-D 量表各因素的解釋力趨近於 0。本研究認為仍有必要持續進行調查方式的比較研究，以確認網路問卷是否可取代紙本問卷，以減少研究及社會成本。

**關鍵詞：**多向度潛在回歸、流行病學研究中心憂鬱量表、評定量尺模式、試題反應理論

傳統問卷調查方式不外乎以郵寄紙本問卷調查、電話訪談、面對面訪談方式為主，其中以面對面訪談及電話訪談的回收率較高，但是所耗費的成本太高，較不適合一般有經費限制的學術研究，因此，郵寄紙本問卷成為最節省研究成本及人力的資料收集方式，故最常被採用來作為收集研究資料之用。然而，自從美國國防部在上世紀九 0 年代，將網際網路的技術對民間開放以來，網路的普及率不斷上升，使用網際網路問卷來進行調查研究亦逐漸興盛，其中最主要的理由即是其研究成本更低於郵寄 (Tourangeau, 2004)；除了單獨使用外，也時常與其他問卷調查方式搭配來提高問卷的回收率 (Goho, 2002)。不過，不同調查方式之間所收集到的資料是否具有調查方式效果 (mode effects) 差異的研究，卻一直未停止過 (如 Aquilino, 1994；Carini, Hayek, Kuh, Kennedy, & Ouimet, 2003；Chan, Orlando, Ghosh-Dastidar, Duan, & Sherbourne, 2004；Greenfield, Midanik, & Rogers, 2000；McCabe, Boyd, Young, Crawford, & Pope, 2005；McCabe, Couper, Cranford, & Boyd, 2006；Wright, Aquilino, & Supple, 1998)。所謂問卷調查方式效果的差異，是指受試者作答反應在不同的問卷調查方式上有所不同，受試者可能會因為對不同調查方式所持的態度、感受或認知方式的不同，包括信

\* 感謝國科會補助本研究案全部經費，補助編號為：NSC-93-2413-H-004-008

任度 (confidentiality) 及隱私匿名性 (anonymity) 程度等, 而有社會期許效應 (social desirability distortion) 產生 (Aquilino, 1994; Lautenschlager & Flaherty, 1990; Schwarz, Strack, Hippler, & Bishop, 1991; Wright, Aquilino, & Supple, 1998)。調查方式效果差異的研究之所以重要, 乃因為調查研究結果的推論不應該受到不同調查方式的影響, 否則研究本身便無信效度可言。因此, 調查方式效果差異程度關係著許多利用調查方式進行的研究, 其結果甚至會影響國家政策的擬定。本研究的目的即在探討網路問卷與傳統郵寄紙本問卷這兩種調查方式是否在某種程度上是等價的, 以網路問卷取代紙本問卷是否可行, 特別是針對個體心理取向的問卷 (如 Center for Epidemiological Studies Depression 量表, 簡稱 CES-D) 在特定母群上的調查方式之比較。

對於網路 (或電腦) 問卷與紙本問卷是否具有調查方式效果差異的研究而言, 學術界目前雖然尚未有定論, 但根據筆者評閱相關文獻所獲得的心得指出, 有較高比例的研究結論認為兩種調查方式在問卷內容的比較上並無實質的差異 (Carini, Hayek, Kuh, Kennedy, & Ouimet, 2003; Layne, DeCristoforo, & McGinty, 1999; McCabe, 2004; Pealer, Weiler, Pigg, Miller, & Dorman, 2001; Tomsic, Hendel, & Matross, 2000), 但從樣本回收率來看, 儘管回收率並非很高, 有些研究發現網路問卷的回收率高於紙本問卷 (McCabe, 2004; McCabe, Boyd, Couper, Crawford, & D'Arcy, 2002; Pealer et al., 2001), 但也有些研究則獲得相反的結果 (Goho, 2002; Kwak & Radler, 2002)。以下將針對影響調查效果的幾個因素進行探討, 包括社會期許效應及匿名性、母群體背景變項對網路 (或電腦) 問卷與紙本問卷調查方式的影響, 並針對本研究所界定的母群加以論述。此外, 將進一步針對本研究問卷使用的 CES-D 量表進行因素結構的探討, 以利後續研究的進行。

### 一、社會期許效應及匿名性的影響

當調查研究的受試者認為隱私不會暴露時, 則問卷的匿名性較高, 此時受試者較不會往社會期許的方向上填答問卷, 反之, 當某種調查方式讓受試者感受到隱私被威脅時, 問卷的匿名效果較差, 受試者傾向往對自己有利的方向填答, 通常是出現較多社會期許的態度或行為表現。而 Richman, Kiesler, Weisband 和 Drasgow (1999) 的後設分析結果發現, 調節社會期許效應與問卷調查方式效果的主要因素之一, 就是受試者知覺到的問卷匿名性 (anonymity)。Richman 等人針對電腦問卷、紙本問卷及面對面訪談進行後設分析, 結果指出, 以社會期許效應為依變項時, 電腦化問卷與紙本問卷的社會期許性效應差異在 581 個效果量分析中的平均效果量只有 0.05, 即電腦問卷的社會期許效應較紙本問卷稍高; 而電腦問卷與面對面訪談的社會期許性效應差異在 92 個效果量分析中的平均效果量則有 -0.19。整體而言, 受試者的社會期許性效應在電腦問卷與紙本問卷上並無差異, 但在電腦問卷與面對面訪談上則有較大的差異存在, 且電腦問卷的社會期許效應低於面對面訪談, 儘管如此, 就效果量的標準而言, 上述的效果量都屬於低效果 (Cohen, 1988, 1992), 但電腦問卷與紙本問卷之間的差異似乎是比電腦問卷與面對面訪談之間的差異來得更低。這結果說明了社會期許性效應基本上對電腦與紙本問卷兩種調查方式的影響力相似。在匿名性上, 雖然在電腦問卷與紙本問卷的比較或是電腦問卷與面對面訪談的比較上, 可能是因為效果量間的變異量太小, Richman 等人發現, 匿名與否對社會期許性的效果量變異並無解釋力, 但當以各種量尺測量工具 (除認知測驗外的其它測驗, 包括人格、行為、態度測驗) 的得分為依變項時, 匿名與否對電腦問卷與紙本問卷在這些量尺測量工具上的差異效果量之變異則具有顯著的解釋力。這說明了對心理類量尺問卷而言, 匿名與否對電腦問卷與紙本問卷兩種調查方式仍有不等的影響力, 因此, 在一定程度上控制這兩種問卷調查方式的匿名性相似, 會使兩種調查方式在心理類量尺的研究結果更為可信。此外, 這裡必須特別注意的是, 雖然本文並不刻意區分電腦問卷與網路問卷的不同, 但 Richman 等人後設分析所蒐集的研究報告年代幾乎都在 1995 年網際網路開始興起前, 因此, 匿名性對網路問卷與紙本問卷影響力的差異將可能較電腦問卷與紙本

問卷的差異來得更小，理由誠如 Richman 等人所說的，當電腦問卷的受試者處在單獨、匿名的情境時，其社會期許效應在各種測量量尺上將更小。而本研究將進行的網路問卷與紙本問卷的差異研究中，網路問卷與紙本問卷都是以郵寄方式通知（研究程序詳見研究方法一節），而非一般電腦問卷大都必須很清楚地確認個體在某個電腦教室或實驗室中接受問卷調查。故本研究中的網路問卷與紙本問卷在匿名性及社會期許效應上將更為一致。

此外，匿名性因素可能會透過問卷議題與調查方式之間的交互作用，而導致社會期許效應，造成不同調查方式下所獲得的問卷調查結果產生差異。過去一些研究指出，在敏感議題或問卷工具的使用上確實會使不同的調查方式之間產生差異，例如，非法藥物的使用、具風險的性行為、及潛在令人尷尬的主題等，對於這些議題的調查而言，電話訪談相對於面對面訪談的匿名性較高，因此受試者可能更願意吐露真相（Czaja, 1987 ; Greenfield et al., 2000 ; Schwarz et al., 1991）；然而相反的情形也可能發生，面對面訪談較電話訪談更易使受試者與訪員建立起融洽關係（*rapport*），因而使受試者更願意表達他們真實的困難，此時，受試者視面對面訪談是一種獲取同情或社會支持的機會（Henson, Cannell, & Roth, 1978 ; Holbrook, Green, & Krosnick, 2003）。而 Newman, Des Jarlais et al., (2002) 使用勝算比（odds ratios）的方式分析 1417 位藥物注射使用者參與電腦問卷（688 人）與面對面訪談（729 人）在三種議題上的反應情形，結果發現，對於羞恥性的問題（愛滋病風險行為）而言，電腦問卷較面對面訪談獲得更高頻率的報告，其勝算比達顯著水準，暗示匿名性相較於尋求社會支持而言有較高的重要性；對於中性問題而言，電腦問卷與面對面訪談兩者之間報告的頻率並無差異；而對有關心理困擾或憂鬱的情感問題而言，面對面訪談則較電腦問卷獲得較高的頻率報告，其勝算比亦達顯著水準，此時，個體尋求社會支持的意願似乎高於個人隱私的要求。這個研究不僅指出調查研究收集資料的方式與調查的主題有交互作用，更進一步指出涉及隱私的匿名性效應與尋求社會支持的動機有著複雜的心理需求競爭。另外，Chan 等人（2004）使用試題反應理論（item response theory, IRT）在電話訪談與郵寄紙本問卷上針對 278 位憂鬱患者進行 CES-D 量表的試題差異功能（differential item functioning, DIF）檢定，結果發現紙本問卷的匿名性較電話訪談高，接受電話訪談的受試者會傾向社會期許性的方向回答。

綜上述，整體而言雖然電腦問卷與紙本問卷調查方式的差異不明顯，但也很可能因問卷調查主題而有調查方式效果存在，以筆者將進行的問卷調查主題而言，是屬於憂鬱傾向的問卷調查（CES-D 量尺），儘管尚未有研究比較網路問卷與紙本問卷調查方式在憂鬱量表上的差異效果，但由本節的文獻探討可推論得知網路問卷與紙本問卷的匿名性將較電腦問卷與紙本問卷更為一致，因此，網路問卷與紙本問卷即使有調查方式效果差異，也應該不會有實質意義，其理由亦如同 Tourangeau（2004）對不同調查方式的比較所提供的分析及推論。他提出三個影響調查模式效果的關鍵變項，分別是調查方式所提供的匿名性、合法性、及調查方式對受試者加諸的認知負荷量。就匿名性而言，郵寄紙本問卷與網路問卷所提供的匿名性最高也最為接近，而面對面訪談必須近距離與受訪者接觸，電話訪談則有聲音上的接觸，這兩者都不免令受訪者提高警覺或與受訪者產生交互作用而可能影響問卷的真實性；就合法性來說，面對面訪談有較多的機會讓受訪者獲悉資料蒐集合法性的證據，例如提供照片、背章或身份等可供查詢的線索，另外幾種調查方式則在此點上相對較為吃虧；而就認知負荷量來說，電話訪談的認知負荷最高，受訪者必須記住每一題問卷選項才能適當表達意見，這對受訪者而言是一種折磨，可能剛開始幾題還會認真配合，到後來可能就會分心而隨意應付。就上述推論而言，相較於其他調查方式的配對比較，本研究欲進行的郵寄紙本問卷與網路問卷在匿名性、合法性及認知負荷量上最為接近。因此，在匿名性相似的網路問卷與紙本問卷調查方式下，儘管是憂鬱傾向的調查，也可能較不會出現實質的調查方式效果差異。

## 二、調查母群背景的影響

調查方式與所調查的母群之間也可能有交互作用的效果產生，不同年齡層或不同社經地位背景的母群對網路問卷與紙本問卷的使用經驗或信任程度會影響調查結果。然而，儘管青少年對電腦問卷的正向態度及熟悉度較年長者高，導致青少年（12-18 歲）在電腦問卷上的作答比例較紙本問卷高；但就整體樣本而言（12-34 歲），電腦問卷則與紙本問卷調查的反應情形仍然是相似的（Wright, et al., 1998）。Couper 和 Rowe（1996）及 Wright 等人（1998）認為，隨著時代的進步，對電腦的熟悉及信任度的提高，不同年齡世代（cohort）所產生的調查方式效果會減低。

在其他背景變項上，由於受試者的低教育水準及社經地位會影響其對電腦或網路的熟悉度，相較下，紙本問卷反而較易受到信任，導致使用電腦或網路問卷調查的結果之效度受到質疑（Aquilino, 1994；Couper & Rowe, 1996；Pridemore, Damphousse, & Moore, 2005）。Shaw 和 Davis（1996）研究發現，擁有個人電腦、較常使用電子郵件的受試者有較多的電子科技經驗，導致電子問卷與紙本問卷有不同的調查結果。Bandilla, Bosnjak 和 Altdorfer（2003）使用兩種不同的母群進行網路問卷與紙本問卷的比較，儘管有調查方式效果存在，但在控制兩種母群的社經背景後，受試者在網路問卷與紙本問卷的作答情形則沒有差異。Hayslett 和 Wildemuth（2004）審慎地選擇 300 位背景相似的圖書館員為受試者，比較郵寄紙本問卷、紙本郵寄上網填答通知、電子郵寄上網填答通知三種調查方式後發現，三者問卷內容上幾乎沒有任何差異存在。因此，有效地控制樣本的背景變項會使調查方式的比較具有可信的結果，即使調查的主題具有風險或令人難以啓齒，控制為社經背景一致的受試者不同調查方式上的比較結果也較為可信。McCabe（2004）以自陳報告方式調查 7,000 名大學生非法藥物使用的經驗，研究中以隨機分派的方式控制學生的背景變項，結果發現網路問卷與紙本問卷的作答有相似的藥物使用經驗。Huang（2006）以背景變項作為配對方式，比較網路問卷與紙本問卷在適應性問題、敏感性問題及開放性問題上的差異，研究結果也顯示，兩組在這些問題上幾乎沒有差異。Denscombe（2006）刻意選用同質性高的 15 歲美國中學生為樣本，比較網路及紙本問卷調查結果在青少年健康行為上的作答情形，結果在 23 題中也只有 1 題出現明顯差異。綜上所述，本研究擬選用社經背景相當一致的中小學教師作為研究的母群，調查其憂鬱傾向，並比較網路問卷與紙本問卷在憂鬱傾向的問卷內容上是否有不同的調查方式效果存在。

就本研究關心的中小學教師母群體來說，教師的脆弱性（vulnerability）是值得關心的主題（Bullough, 2005），教師必須面對的壓力來源包括來自外在的教育當局政策與學校或家長要求相衝突所產生的矛盾壓力，內在的壓力源包括個人能力的限制以及面對學生的拒絕等不舒服的感覺等等。長期處於這些壓力狀態，將產生 Brouwers 和 Tomic（2000）所謂的情緒耗弱（emotional exhaustion）現象，導致教師自我效能的降低以及教室管理能力的下降。長期的壓力不可避免地將導致憂鬱症狀的產生，過去已有一些研究指出教師族群與憂鬱症狀（depressive symptoms）的關聯性，研究發現教師的低工作滿意度及高工作壓力與憂鬱症狀有關，女性教師在 CES-D 上的得分高於臨界點 16 分的人數是男性的 2 倍，而 41 ~ 50 歲這個年齡層的教師在 CES-D 得分高於 16 分的人數也多於其他年齡層，另外，有高達 53.2% 教師在 CES-D 得分高於 16 分者希望轉行（Jurado et al., 2005）；另外，工作環境與憂鬱症狀有高度的相關存在，偶發事件壓力源（episodic stressors）以及現況壓力源（ongoing stressors）都與 CES-D 量表分數有顯著的負相關（Schonfeld, 2000）。以台灣來說，台北市教師會曾於民國 90 年針對台北市五十所國中小學教師進行抽樣調查，有效問卷回收九百一十七份，問卷由長庚醫院精神科設計，研究結果指出台北市國中小學教師有近二成八患有潛在憂鬱症，其中約有一成三老師有憂鬱症傾向，一成七必須尋求專業醫師的協助解決壓力（彭駕駢，2001）。故本研究使用中小學教師作為取樣的來源，除了研究上的目的外，也有進一步想瞭解全台灣中小學教師憂鬱傾向之興趣。

由前兩節的探討可知，本研究的目的即在控制研究母群的社經背景地位所產生的干擾，並推論網路問卷與紙本問卷兩種調查方式具有相似的匿名性及其它條件下，進行兩種問卷調查方式效果之比較。

### 三、CES-D 量表的因素結構

自從 Radloff (1977) 在團隊組織的合作下，發展出提供十分完整信效度訊息的 CES-D 量表用來測量憂鬱傾向後，學術界每年不斷地有新研究報告針對 CES-D 量表的應用提出實徵性的研究結果，其中討論較多的是 CES-D 量表的因素結構問題。Radloff (1977) 在美國樣本上獲得四因素的結構：即「憂鬱情緒」(depressed affect)、「正向情感」(positive affect)、「身體症狀」(somatic and retarded activity)、「人際問題」(interpersonal)。但在國際間使用多年來，不同的族群中所獲得的因素分析結果並非十分一致，從二個因素到七個因素結構的情形，都曾出現過 (Boisvert, McCreary, Wright, & Asmundson, 2003)。

Spijker 等人 (2004) 以移民至荷蘭的第一代勞工移民 (330 位來自土耳其及 299 位來自摩洛哥) 及 330 位當地荷蘭人為研究對象，年齡皆在 55-74 歲之間，進行 CES-D 量尺因素結構的探索比較，結果發現，當地荷蘭人在 CES-D 量表上出現五個因素、來自土耳其的移民則出現四個因素，而來自摩洛哥的移民則出現三因素的結構；特別是其中「正向情感」及「人際問題」兩個因素在三種族群樣本中都被穩定的被分離出來。此外，Miller, Markides 和 Black (1997) 以年老的墨西哥裔美國人為樣本，使用探索性因素分析並根據特徵值大於 1 的標準來決定因素個數，在不同次群體下獲得四個因素或三個因素的結構，然而他們發現有太多的題目具有跨因素的情形出現，顯然地，這暗示著有過多的因素需要被合併，因此，進一步使用 MIMIC 模式 (multiple indicators, multiple causes CFA models) 進行驗證性因素分析，在考慮了其他人口變項的影響後，驗證了二因素結構具有較佳的適配度。此二因素結構清楚地將「正向情感」因素與其他三個因素明顯區分開來，就一階因素分析而言，這似乎暗示了 CES-D 量表事實上所包含的是一正一負兩個主要的情緒構念。Edman 等人 (1999) 使用 243 位夏威夷的菲律賓裔美國青少年為樣本，使用探索性因素分析亦獲得一正一負的二因素結構，進一步使用驗證性因素分析發現，此二因素結構的適配度優於 Radloff (1977) 提出的四因素結構。然而，Cheung 和 Bagley (1998) 以香港地區 138 對夫妻為研究對象的結果指出，香港地區華人的 CES-D 量表以「憂鬱症狀」(depressive symptoms) 及「人際問題」(interpersonal problems) 二因素結構的適配度較佳，也就是將人際問題外的其它題目都歸屬在同一個因素下，因此，他們雖獲得與一般研究二因素結構不同的結果，但相同的是憂鬱情緒與身體症狀的題目在 CES-D 上都可被歸為同一個因素。以筆者的研究為例，也發現台灣中小學教師的憂鬱情緒與身體症狀因素無法區分的現象，其 CES-D 量表各試題得分經由主軸法萃取共同性，並以斜交轉軸後的因素分析顯示，屬於「身體症狀」的題目均與「憂鬱情緒」題目被歸屬在同一因素之下。比較特別的是，在本研究中，「人際問題」亦與「身體症狀」及「憂鬱情緒」被歸屬在同一因素下，因素負荷量在 0.31 ~ 0.81 之間，而另外 4 題「正向情感」試題則分屬另外兩個因素，其中第 8、12、16 題屬於第二個因素，因素負荷量在 0.69 ~ 0.80 之間，第四題單獨成爲一個因素，但因素負荷量不到 0.13，幾乎可以忽略，因此，本研究結果也十分接近二因素的結構。

由於 CES-D 量表的因素結構有所歧異，因此，本研究在使用試題反應理論進行網路問卷與紙本問卷的比較時，考慮採用兩種二因素結構，包括「憂鬱症狀」(depressive symptoms) 及「人際問題」(interpersonal problems) 二因素結構、以及一正一負的二因素結構、另外也將 Radloff 的四因素結構、前文暗示的三因素結構 (即正向情感、人際問題、及憂鬱情緒+身體症狀)、以及一因素結構納入研究中處理。之所以特別將一因素結構列入考量，主要是因爲 Radloff (1977) 編製 CES-D 量表

的目的是著眼於憂鬱傾向的測量，它的計分方式仍採用 20 題分數相加而得的總分概念，因此，不論該量表在各族群初始因素個數為何，其背後必定可以合理假設有一個潛在的共同因素（即憂鬱）存在，實徵的驗證性因素分析研究也指出 CES-D 具有一個高階的憂鬱因素，且在年老男女樣本中皆存在（O'Rourke, 2005），縱貫性研究的資料也顯示出此二階的憂鬱因素穩定地存在於青少年樣本中（Motl, Dishman, Birnbaum, & Lytle, 2005）；而游森期、余民寧（2006）已使用結構方程式模型（structural equation model, SEM）對 CES-D 量表進行驗證性二階因素分析，而獲得一個共同的潛在憂鬱因素存在的適配結果。因此，將一因素結構的情形列入兩種調查方式差異效果的比較仍有其必要性。

## 方 法

### 一、研究樣本

本研究使用的樣本取自台灣地區的中小學教師母群體，採用分層隨機抽樣方式，首先由教育部公布的網頁中獲得台灣地區高中、高職、國中、國小學校的名冊及網址，將上述各級學校依台灣地區北、中、南、東四大區域學校數的比例，進行分層隨機抽取學校的抽樣工作，國小共抽得 205 所，其中北部 60 所、中部 64 所、南部 61 所、東部 20 所；國中共抽得 57 所，其中北部 20 所、中部 15 所、南部 16 所、東部 6 所；高中共抽 25 所，其中北部 11 所、中部 6 所、南部 7 所、東部 1 所；高職共抽得 13 所，其中北部 4 所、中部 3 所、南部 5 所、東部 1 所。全部共計 300 所學校，每所學校隨機抽取 8 位教師，包括 4 男 4 女，因此，全台灣地區共抽得教師 2400 人，然後，隨機分派成網路問卷組與紙本問卷組，每組各 1200 人。

本研究在 2004 年 12 月 13 日第一次寄出紙本邀請函 1,200 份給網路組，以及 1,200 份紙本問卷給紙本組，並且在三週之後進行第一次催收，六週之後再進行第二次催收。兩組樣本的回收狀況如表 1 所示：

表 1 問卷回收摘要表

	問卷回收數量				
	初次回收	第一次催收	第二次催收	累積次數	有效問卷
網路組	311	195	35	541	535
紙本組	401	173	57	631	606

由於網路組回收樣本有 6 份遺漏值，紙本組回收樣本有 25 份有遺漏值，因此，本研究將之刪除，網路組有效問卷回收率為 44.58%，紙本組有效問卷回收率為 50.50%，有效的總回收樣本 1141 份（總回收率為 47.54%）。

### 二、研究工具

本研究採用 Radloff（1977）所編製的「流行病學研究中心憂鬱量表」（Center for Epidemiologic Studies Depression Scale, CES-D），包括屬於「憂鬱情緒」（depressed affect）的第 3,6,9,10,14,17,18 題、屬於「正向情感」（positive affect）的第 4,8,12,16 題、屬於「身體症狀」（somatic and retarded activity）的第 1,2,5,7,11,13,20 題、屬於「人際問題」（interpersonal）的第 15,19 題，共 20 題。本量表

雖然是根據臨床個案所表現出來的症狀編製而成，但並卻不是爲了臨床診斷的目的而編製的，反而比較適用於一般性樣本來測量其憂鬱傾向之用（Radloff, 1977），因此也適合本研究的教師樣本。本量表採 Likert 式的四點計分法，要求受試者在 20 題的 CES-D 量表中，針對每題所描述的情況勾選其每週出現頻率的選項。每題選項皆分爲「很少如此（每週少於 1 天）」、「偶爾如此（每週有 1-2 天）」、「經常如此（每週有 3-4 天）」、「總是如此（每週有 5 天以上）」。

根據出現憂鬱症狀的頻率，分別給予 0 分（很少如此）到 3 分（總是如此）不等；但是，其中的第 4、8、12、16 題等四題是屬於反向計分題。將受試者在這 20 題上的作答分數加總起來，即構成受試者在該量表上的總分，總分的值域界於 0 分到 60 分，分數愈高表示憂鬱傾向愈高；得分若超過 16 分以上，即被列爲憂鬱症高危險群的人口。

本量表的  $\alpha$  內部一致性信度在 0.85 ~ 0.90 之間，憂鬱病人組較一般正常人組稍高。一般正常人組在 2 ~ 8 週的再測信度（以郵寄問卷方式實施再測）約爲 0.51 ~ 0.67 之間，3 ~ 12 月的再測信度（以面談方式實施再測）約在 0.32 ~ 0.54 之間；而病人組在經四週的療程後之再測信度爲 0.53（Radloff, 1977）。

本量表在聚斂效度方面，憂鬱病人組的樣本顯示本量表和臨床醫師及護士評定的 Hamilton Clinician's Rating 量表及 Raskin Rating 量表之相關爲 0.44~0.56，經四週的療程後，相關更高達 0.69 ~ 0.75；若以自陳報告作爲效標，一般正常人組之 CES-D 量表與其他憂鬱有關的量表之相關在 0.51 ~ 0.61，病人組則在 0.55 ~ 0.72 之間。在區辨效度方面，本量表對於精神病患和普通人具有良好的區分能力，病人組 16 分以上的比例爲 70%，而一般正常人組 16 分以上的只有 21%（Radloff, 1977）。

### 三、研究程序

本研究紙本組樣本以傳統的郵寄印刷問卷方式通知受試者。問卷含指導語、5 題背景資料、及 20 題的 CES-D 量表題目，並且附上回郵，請受試者填答後寄回。網路組樣本所收到的郵件之中，則只有邀請函，邀請受試者上網填答。網路問卷則暫時建置在國立政治大學的教育測驗與評量網站（<http://edutest.nccu.edu.tw>），受試者依據邀請函所載明的網址以及代碼，上網填答。該網站在完成資料收集目的後隨即拆除。

本研究爲了避免網路問卷可能的重複填答問題，每份網路組的邀請函中給予一組身份代碼，必須輸入正確的代碼，才能夠進行填答，而且只能填寫一次。此外，爲了與紙本組一致，網路問卷並未設計強制填答所有問卷題目後傳回。

在問卷的催收方面，不管是網路組或紙本組仍然依據傳統郵寄的方式，寄至受試者所任教的學校進行催收，網路組仍然是只寄邀請函，紙本組也仍然是寄出紙本問卷並附上回郵信封。本研究不採用更節省成本的電子郵件催收方式，其原因乃是考量到本研究無法獲得每位教師常用的電子信箱號碼，即使能夠獲得電子信箱的訊息，所寄出的催收通知也可能被伺服器視爲垃圾郵件而被攔截下來，這種情形將使影響回收率的因素無法獲得控制，因此，本研究仍然採用傳統郵寄問卷的方式來通知受試者與進行催收。

### 四、資料分析

本研究主要是藉由 ConQuest 軟體（Wu, Adams, & Wilson, 1998）中的評定量尺模式（rating scale model）來適配所蒐集到的資料，該模式一般最常用來處理採用李克特氏計分的量表。評定量尺模式所估計的參數包括個體能力參數、位置參數（location parameter）、及類別交叉參數（category intersection parameter）；位置參數反映出特定題目的相對難度，每一題有一位置參數；但所有試題共

用一組相同的類別交叉參數值，其數目為試題的類別數減 1 個 (Embretson & Reise, 2000)。值得注意的是，類別交叉參數在這裡並不是指兩兩相鄰類別反應曲線的交叉點之橫座標值，而是指交叉點與位置參數的差值，因此，位置參數加上類別交叉參數才是相鄰類別反應曲線的交叉點之值，故試題中的類別交叉參數之和為 0。此外，每一題的位置參數可以由該題所有兩兩相鄰的類別反應曲線交叉點的平均求得。

ConQuest 是應用單參數試題反應理論發展出來的統計套裝軟體，除了可以將次序量尺轉換為具有等距量尺的邏輯斯 (logits) 單位之功能外，更具有處理多向度或多因素資料的功能，而不侷限於單向度的假設，因此，有助於時常使用多因素量表之教育或心理學研究。本研究藉由其提供的「多向度潛在回歸」(multidimensional latent regression) 模組進行網路問卷組與紙本問卷組在各因素上的平均數差異檢定，即以網路問卷組為參照組之二元問卷組別預測變項，分別對各因素進行回歸分析。在 ConQuest 中，將「潛在回歸」定義為由試題反應資料直接進行回歸模式的估計工作，而不需要先估計出個體的潛在能力後再利用一般統計軟體進行回歸分析等多道繁雜的手續。對於兩組母群平均數差異的檢定而言，使用潛在回歸在功能上與 t 檢定一致，但比 t 檢定優越的是「潛在回歸」是針對潛在變項或因素進行組別間平均數的比較而非一般 t 檢定針對觀察變項來進行比較。

## 結果與討論

本節將針對受試者背景資料進行描述以及網路組與紙本組在 CES-D 量表上的差異進行分析，除了提供相關憂鬱研究的討論外，進一步比較網路問卷與紙本問卷在調查研究上相互取代的可能性。

### 一、背景資料描述統計

由表 2 可知，網路組問卷與紙本組問卷的回收樣本數有差異存在，紙本問卷回收量較網路問卷高，顯示郵寄問卷方式仍然是一種較受歡迎的調查方式。在其他背景變項的分析上顯示，性別、學歷、年齡在問卷調查方式之間並無顯著差異存在，表示這兩種調查方式不受該等背景變項的影響。此外，值得注意的是，問卷調查方式與憂鬱傾向（以受試者在 CES-D 總分 16 分以上者為判定標準）之間也無顯著關聯，這也初步暗示這兩種問卷並無顯著的調查方式效果差異存在。以上結果顯示，雖然紙本問卷回收率較高，但基本上其樣本組成特性與網路問卷回收的樣本組成特性並無差異，網路問卷似乎可以視為紙本問卷調查方式外的另一種選擇，甚至是可以合併使用。

另外，在 CES-D 量表分數的描述統計上，由於 CES-D 量表在全球的使用情形十分頻繁，為方便日後相關研究對台灣地區特定母群研究的參考，本研究也列出台灣中小學教師在其他各種背景變項下潛在憂鬱症的盛行率。如表 3 所示，男女性教師潛在憂鬱症的盛行率約在二成五左右，其中男性教師略高於女性教師；就學歷而言，潛在憂鬱症的盛行率在二成到二成五之間，其中以大學學歷者較高；就年齡層而言，潛在憂鬱症的盛行率分佈較廣，約在一成四到三成之間，以 26 ~ 35 歲這個年齡層的教師最高，佔 30.0%，其次則是 25 歲以下的 21.8%，最低的則是 46 ~ 55 歲這個年齡層的 14.0%；另外，可看出 45 歲以下的潛在憂鬱症盛行率遠高於 46 歲以上的中小學教師。

Cheung 和 Bagley (1998) 調查香港華人在 CES-D 平均數為 12.31，標準差約為 6.74，而本研究中的台灣中小學教師紙本組樣本的 CES-D 之平均數為 12.15，標準差為 8.01；此外，香港華人樣本在 CES-D 量表的總分高於 16 分者佔 25.0%，而本研究的中小學教師樣本在 CES-D 量表上的總分高於 16 分者佔 24.5%，雖然使用的樣本不同，但兩個研究的結果十分接近，差異較大的是 Cheung 和 Bagley 使用的樣本人數是 138 對已婚夫妻。而 Schonfeld (1990) 也曾使用 CES-D 量表調查紐約市 67



表 2 網路與紙本問卷回收樣本之背景分析

	問卷調查方式		卡方檢定
	網路問卷 (n=535)	紙本問卷 (n=606)	
			$\chi^2 = 4.42, df=1, p = 0.036$
性別 (%)			
男	48.2	47.0	$\chi^2 = 0.16, df=1, p = 0.687$
女	51.8	53.0	
學歷 (%)			
專科	2.2	3.7	$\chi^2 = 3.31, df=2, p = 0.192$
大學	75.3	76.8	
研究所以上	22.4	19.5	
年齡 (%)			
25 歲以下	5.0	3.9	$\chi^2 = 9.40, df=4, p = 0.052$
26 ~ 35 歲	43.7	50.1	
36 ~ 45 歲	35.3	28.1	
46 ~ 55 歲	14.6	15.7	
56 歲以上	1.3	2.2	
CES-D 總分 (%)			
15 分以下	77.8	73.4	$\chi^2 = 2.87, df=1, p = 0.090$
16 分以上	22.2	26.6	

表 3 不同背景變項下的中小學教師潛在憂鬱症盛行率

背景變項	CES-D16 分以上	有效人數
性別 (%)		
男	26.7	543
女	22.6	598
學歷 (%)		
專科	20.6	34
大學	25.3	857
研究所以上	22.6	235
年齡 (%)		
25 歲以下	24.0	50
26 ~ 35 歲	30.0	533
36 ~ 45 歲	21.8	357
46 ~ 55 歲	14.0	172
56 歲以上	15.0	20

位中小學及幼兒教師的憂鬱傾向，CES-D 的平均分數是 13.03，標準差為 11.03，有 35% 的教師得分超過 16 分。就樣本相似性而言，Schonfeld 的研究顯示，紐約市中小學及幼兒教師的憂鬱傾向似乎較台灣中小學教師嚴重。此外，本研究與彭駕驛（2001）所引述的台北市中小學教師約 28% 具有潛在憂鬱症傾向的結果相比較，以台灣中小學教師為母群的本研究樣本結果，其憂鬱傾向似乎較台北市稍低，或許因為台北市是台灣教育政策的首衝之區，因此教師所背負的壓力亦較全台灣其他地區教師為

高所致。

## 二、問卷內容的差異檢定

由文獻探討一節可知，CES-D 量表在國際間研究及應用極為廣泛，其因素結構在實務應用上雖非重點，但一直被廣為討論。因此，本研究將針對較常出現的因素結構進行網路與紙本問卷調查方式在各因素平均數上的差異檢定，以呈現網路與紙本問卷在心理類量尺的差異情形。

在進行網路與紙本組別對各因素的潛在回歸分析前，必須先討論 CES-D 各題與多向度（因素）評定量尺模式的適配情形，以四向度為例的語法如附錄一所示，如此後續的研究結果較為可信，之後才繼續進行正式的回歸分析，其語法如附錄二所示。以四向度為例，各題都有一個位置參數估計值及其估計標準誤，但其中會有四題沒有估計標準誤，原因在於一般試題反應模式都假設每一向度下所屬試題之難度（位置參數）之和為 0，因此，就四向度的資料而言，會有四題被限制以使所屬試題的位置參數之和為 0，也就是不需要進行估計，故這四題沒有提供估計標準誤，也無適配指標。另外，在類別交叉參數的情形也是大同小異，以 CES-D 四點量表而言，應該會有 3 個所有試題共同享有的類別交叉參數值；在 ConQuest 軟體的設計上，會限制最後一個類別交叉參數值不估計，使 3 個類別交叉參數值的和為 0。

在適配指標上，ConQuest 軟體提供未加權（unweighted）及加權（weighted）兩種適配指標 MNSQ（mean squares）值，並藉由 Wilson-Hilferty 轉換法將 MNSQ 轉換為近似常態化  $t$  分配值以利檢視；MNSQ 值是模式預期值與觀察值（即受試者在量表上圈選的點數）相減後之標準化殘差平方和之平均值，遵循卡方分配的形式，其範圍從 0 到正無窮，期望值是 1（Bond & Fox, 2001）。根據 Linacre（2006）對 MNSQ 合理範圍的建議，MNSQ > 2 表示該題將扭曲或破壞測量系統；MNSQ 在 1.5 ~ 2 之間表示該題對測量的建構雖不具生產性，但也不具破壞性。MNSQ 在 0.5 ~ 1.5 之間表示該題對測量具生產性；MNSQ < 0.5 表示該題對測量有較少生產性。

### （一）四因素 CES-D 結構

由表 4 可知，以 Radloff（1977）所提的四因素 CES-D 結構與模式適配情形的指標來看，MNSQ 值都大於 0.5 以上，並無超過 2 的值出現，因此適配情形良好。不過由  $T$  值來看，出現較多絕對值過高的  $T$  值，一般而言， $T$  值的絕對值在 2 或 3 以內是較佳的，然而根據 Linacre（2006）的建議，如果 MNSQ 值是可接受的，則  $T$  值可以被忽略。

由表 5 的潛在回歸分析可知，以網路與紙本為問卷組別預測變項分別對四因素進行簡單回歸分析，以  $p = 0.01$  的顯著性為標準，值得注意的是，由於 ConQuest 軟體對於各項參數估計值均未提供確實的顯著性  $p$  值，本研究僅以一個星號（ $p < 0.01$ ）的方式註明該參數估計值的顯著性。結果可看出對「憂鬱情緒」及「正向情感」兩個因素而言，組別變項的回歸係數並不顯著，表示網路組與紙本組受試者在 CES-D 量表上對有關這兩個因素的試題反應並無差別；但對「身體症狀」及「人際問題」這兩個因素而言，組別變項的回歸係數達 0.01 顯著水準，表示網路組與紙本組受試者在 CES-D 量表上對有關這兩個因素的試題反應有統計考驗上的不同。然而，從實用的角度而言，可以發現組別變項在這四個因素上的解釋力（ $R$  平方）都非常小，因此，可以推論組別變項回歸係數顯著的原因很可能是大樣本造成的結果。此外，就此四因素的變異量來看，正向情感的變異情形較小，顯示受試者間的正向情感特質變化不大，此外，這可能也暗示該因素所屬題目在區辨受試者正向情感的能力上較差，屬於較不好的題項，可供以後重新命題的參考；而人際問題及憂鬱情緒兩因素上的變異情形較大，表示受試者間對人際問題及憂鬱情緒上的感受有較大的差異。

表4 四因素CES-D結構之各題與評定量尺模式適配情形

位置參數估計值 (標準誤)	unweighted fit		weighted fit		
	MNSQ	T	MNSQ	T	
$\lambda_1$	-0.510 (0.037)	0.95	-1.3	0.96	-0.9
$\lambda_2$	1.168 (0.043)	0.85	-3.8	0.99	-0.1
$\lambda_3$	-0.228 (0.047)	0.70	-8.0	0.82	-3.3
$\lambda_4$	-0.547 (0.028)	1.76	14.9	1.72	14.7
$\lambda_5$	0.092 (0.039)	0.84	-4.2	0.84	-3.5
$\lambda_6$	-0.893 (0.044)	0.72	-7.3	0.72	-5.8
$\lambda_7$	-0.415 (0.037)	0.71	-7.8	0.74	-6.1
$\lambda_8$	0.182 (0.029)	0.72	-7.5	0.70	-8.0
$\lambda_9$	0.672 (0.052)	0.73	-7.0	0.82	-3.0
$\lambda_{10}$	-0.094 (0.048)	0.78	-5.6	0.78	-4.1
$\lambda_{11}$	0.071 (0.039)	0.95	-1.3	1.01	0.2
$\lambda_{12}$	0.319 (0.029)	0.57	-12.1	0.56	-12.7
$\lambda_{13}$	-0.402 (0.037)	1.06	1.4	1.05	1.0
$\lambda_{14}$	-0.822 (0.044)	0.73	-7.1	0.76	-4.8
$\lambda_{15}$	0.086 (0.051)	0.70	-7.9	0.64	-6.7
$\lambda_{16}$	0.046	---	---	---	---
$\lambda_{17}$	1.242 (0.055)	0.63	-10.3	0.72	-4.9
$\lambda_{18}$	0.123	---	---	---	---
$\lambda_{19}$	-0.086	---	---	---	---
$\lambda_{20}$	-0.005	---	---	---	---
類別交叉參數估計值 (標準誤)					
	-1.712 (0.017)	1.83	15.9	1.83	15.4
	0.392 (0.025)	1.28	6.1	1.60	9.6
	1.320	---	---	---	---

註：--- 表示該題限制，使同屬某一因素的各題之參數估計值平均為0，故無法提供標準誤及適配檢定。此外，位置參數估計值的測量單位為邏輯斯 (logit)。

表5 以四因素為依變項之潛在回歸分析

回歸變數	四因素CES-D結構			
	憂鬱情緒	正向情感	身體症狀	人際問題
常數項	-3.835* (0.192)	-0.542* (0.083)	-2.762* (0.117)	-4.512* (0.198)
問卷組別變項	0.234 (0.120)	0.092 (0.051)	0.301* (0.073)	0.429* (0.123)
因素之變異量	4.028	0.749	1.538	4.413
R平方 (%)	0.00%	0.26%	2.54%	2.24%

註：小括號內為估計標準誤。此外，常數項及問卷組別變項之回歸係數的測量單位皆為邏輯斯 (logit)。

\* $p < 0.01$

## (二) 兩種二因素 CES-D 結構

一正一負二因素結構及「憂鬱症狀」與「人際問題」二因素結構曾在文獻探討中提及，因此，這裡亦針對兩種二因素結構與評定量尺模式的適配情形加以描述。由表 6 及表 7 可知，除了表 7 第 4 題及憂鬱與人際二因素結構下的第一個類別交叉參數估計值之 MNSQ 值稍微高於 2 外，所有 MNSQ 值都在 0.5 ~ 2 之間。整體而言，各題在一正一負二因素結構及在憂鬱與人際二因素結構的適配情形相去不遠。

表 6 一正一負二因素 CES-D 結構之各題與評定量尺模式適配情形

位置參數估計值 (標準誤)	unweighted fit		weighted fit		
	MNSQ	T	MNSQ	T	
$\lambda_1$	-0.927 (0.037)	1.09	2.0	0.99	-0.2
$\lambda_2$	0.808 (0.044)	1.02	0.5	1.02	0.4
$\lambda_3$	0.050 (0.041)	0.83	-4.3	0.87	-2.7
$\lambda_4$	-0.538 (0.028)	1.74	14.6	1.70	14.5
$\lambda_5$	-0.307 (0.039)	0.90	-2.4	0.87	-2.8
$\lambda_6$	-0.552 (0.039)	0.62	-10.6	0.64	-8.4
$\lambda_7$	-0.830 (0.038)	0.82	-4.6	0.76	-5.3
$\lambda_8$	0.179 (0.029)	0.71	-7.8	0.69	-8.3
$\lambda_9$	0.863 (0.044)	0.80	-5.2	0.85	-3.0
$\lambda_{10}$	0.167 (0.041)	0.78	-5.8	0.80	-4.3
$\lambda_{11}$	-0.333 (0.039)	1.05	1.3	1.04	0.8
$\lambda_{12}$	0.313 (0.029)	0.56	-12.5	0.55	-13.0
$\lambda_{13}$	-0.820 (0.038)	1.08	2.0	1.08	1.6
$\lambda_{14}$	-0.492 (0.039)	0.77	-6.0	0.77	-5.0
$\lambda_{15}$	0.590 (0.043)	0.83	-4.4	0.92	-1.6
$\lambda_{16}$	0.045	---	---	---	---
$\lambda_{17}$	1.383 (0.047)	0.88	-2.9	0.95	-0.9
$\lambda_{18}$	0.365 (0.042)	0.69	-8.4	0.74	-5.7
$\lambda_{19}$	0.441 (0.043)	0.78	-5.8	0.81	-3.9
$\lambda_{20}$	-0.408	---	---	---	---
類別交叉參數估計值 (標準誤)					
$\delta_1$	-1.667 (0.016)	1.78	15.1	1.77	14.6
$\delta_2$	0.374 (0.025)	1.26	5.8	1.60	9.6
$\delta_3$	1.293	---	---	---	---

由表 8 潛在回歸分析的結果可知，在問卷組別變項的回歸係數上，除了一正一負二因素結構下的「正向情感」因素未達顯著外，其餘三個因素都達 0.01 的顯著水準。而由問卷組別變項的解釋力來看，在憂鬱與人際二因素結構下的「人際問題」因素上，R 平方的值異常的大，達到 17.8%，而其餘三個因素的解釋力都不具實用價值。至於，造成「人際問題」因素的異常高解釋力可能是因為憂鬱與人際二因素結構的適配情形不佳，使結果較不穩定，導致犯第一類型錯誤的機率上升，因此有可能偶然出現異常偏高之值。據此，本研究試著在憂鬱與人際二因素結構下將適配不佳的第四題刪除後，問

表 7 憂鬱症狀與人際問題二因素 CES-D 結構之各題與評定量尺模式適配情形

位置參數估計值 (標準誤)		unweighted fit		weighted fit	
		MNSQ	T	MNSQ	T
$\lambda_1$	-0.328 (0.036)	0.90	-2.5	0.91	-2.1
$\lambda_2$	1.256 (0.042)	0.86	-3.5	0.95	-1.0
$\lambda_3$	0.562 (0.039)	0.73	-7.1	0.80	-4.4
$\lambda_4$	-2.139 (0.032)	2.19	21.4	2.05	19.6
$\lambda_5$	0.235 (0.038)	0.80	-5.0	0.80	-4.7
$\lambda_6$	0.012 (0.037)	0.57	-12.4	0.59	-10.4
$\lambda_7$	-0.240 (0.036)	0.70	-7.9	0.70	-7.4
$\lambda_8$	-1.398 (0.033)	0.93	-1.7	0.91	-2.1
$\lambda_9$	1.310 (0.043)	0.71	-7.6	0.79	-4.6
$\lambda_{10}$	0.672 (0.040)	0.66	-9.1	0.73	-6.1
$\lambda_{11}$	0.216 (0.038)	0.91	-2.2	0.96	-0.9
$\lambda_{12}$	-1.261 (0.033)	0.76	-6.2	0.74	-6.7
$\lambda_{13}$	-0.228 (0.036)	0.98	-0.5	0.99	-0.3
$\lambda_{14}$	0.071 (0.037)	0.66	-9.3	0.71	-6.9
$\lambda_{15}$	0.083 (0.050)	0.80	-5.0	0.66	-6.3
$\lambda_{16}$	-1.535 (0.032)	1.03	0.8	1.01	0.3
$\lambda_{17}$	1.795 (0.045)	0.85	-3.9	0.89	-2.3
$\lambda_{18}$	0.854 (0.040)	0.62	-10.5	0.67	-7.6
$\lambda_{19}$	-0.083	---	---	---	---
$\lambda_{20}$	0.146	---	---	---	---
類別交叉參數估計值 (標準誤)					
$\delta_1$	-1.588 (0.016)	2.01	18.8	2.10	19.9
$\delta_2$	0.255 (0.025)	1.38	8.1	1.76	11.9
$\delta_3$	1.333	---	---	---	---

卷組別變項的 R 平方在二因素上的解釋力則都在 1.10% 以下。此外，憂鬱與人際二因素結構的討論是根據 Cheung 和 Bagley (1998) 調查香港華人的研究而來，就筆者評閱其他相關研究的結果顯示，大部分二因素 CES-D 結構還是以一正一負二因素結構佔絕大多數。

由一般簡單回歸分析的概念可知，二元的問卷組別變項之回歸係數恰巧是網路組與紙本組在依變項平均數上的差異大小，以一正一負二因素結構為例，網路組在「正向情感」與「負向情感」兩個因素的平均數都明顯低於紙本組問卷，儘管我們仍可用非常低的 R 平方解釋力概念來理解此結果的實用價值不大，也即是該結果很可能是大樣本對細微差異的過度檢測所造成的；然而，在仔細檢視本研究的流程後，筆者發現很可能是研究團隊一個不經意的小作法造成此細微差異被大樣本統計偵測出；亦即在我們的研究過程中，雖然網路問卷的邀請函及紙本問卷在寄出的信封上都有打上收件人姓名，但紙本問卷內容上不再出現受試者姓名，可是網路問卷邀請函的接受人在依規定點入網路問卷的網址後，卻可以發現其姓名就出現在問候語中，或許就是因為這個原本想讓網路問卷受試者感到親切友善的作法反而造成網路問卷受試者感到被監控，而導致網路問卷回收率較低以及對問卷內容的作答有所保留。

表 8 以兩種二因素為依變項之潛在回歸分析

回歸變數	一正一負二因素 CES-D 結構		憂鬱與人際二因素 CES-D 結構	
	正向情感	負向情感	憂鬱症狀	人際問題
常數項	-0.550* (0.081)	-3.226* (0.140)	-2.290* (0.102)	-4.106* (0.166)
問卷組別變項	0.098 (0.050)	0.274* (0.087)	0.175* (0.063)	0.370* (0.103)
因素之變異量	0.721	2.188	1.144	3.692
R 平方 (%)	0.14%	1.23%	0.00%	17.8%

註：小括號內為估計標準誤。\* $p < 0.01$ 。

### (三) 三因素與一因素 CES-D 結構

表 9 及表 10 分別是三因素結構及一因素 CES-D 結構與模式的適配情形，除了表 10 中一因素結構下第 4 題的 MNSQ 值以及第一個類別交叉參數 MNSQ 值稍微高於 2 外，其餘 MNSQ 值也都在 0.5 ~ 2 之間。整體而言，各試題在三因素 CES-D 結構下的適配程度較一因素 CES-D 結構時為理想，但相去不遠。

表 9 三因素 CES-D 結構之各題與評定量尺模式適配情形

位置參數估計值 (標準誤)	unweighted fit		weighted fit		
	MNSQ	T	MNSQ	T	
$\lambda_1$	-0.864 (0.037)	1.07	1.7	1.00	0.0
$\lambda_2$	0.889 (0.044)	1.00	0.1	1.03	0.6
$\lambda_3$	0.121 (0.041)	0.82	-4.5	0.89	-2.4
$\lambda_4$	-0.540 (0.028)	1.74	14.5	1.70	14.5
$\lambda_5$	-0.238 (0.040)	0.92	-2.1	0.88	-2.6
$\lambda_6$	-0.485 (0.039)	0.65	-9.6	0.65	-8.3
$\lambda_7$	-0.765 (0.038)	0.78	-5.5	0.77	-5.2
$\lambda_8$	0.180 (0.029)	0.70	-8.0	0.70	-8.2
$\lambda_9$	0.950 (0.045)	0.80	-5.2	0.87	-2.7
$\lambda_{10}$	0.244 (0.042)	0.75	-6.5	0.81	-4.1
$\lambda_{11}$	-0.261 (0.040)	0.98	-0.5	1.05	1.0
$\lambda_{12}$	0.314 (0.029)	0.56	-12.5	0.56	-12.9
$\lambda_{13}$	-0.752 (0.038)	1.11	2.6	1.09	1.8
$\lambda_{14}$	-0.420 (0.039)	0.75	-6.5	0.78	-4.8
$\lambda_{15}$	0.081 (0.050)	0.63	-10.2	0.62	-7.5
$\lambda_{16}$	0.046	---	---	---	---
$\lambda_{17}$	1.478 (0.047)	0.88	-2.9	0.96	-0.8
$\lambda_{18}$	0.443 (0.042)	0.68	-8.6	0.75	-5.4
$\lambda_{19}$	-0.081	---	---	---	---
$\lambda_{20}$	-0.340	---	---	---	---
類別交叉參數估計值 (標準誤)					
$\delta_1$	-1.696 (0.017)	1.85	16.4	1.85	15.8
$\delta_2$	0.388 (0.025)	1.29	6.4	1.63	10.1
$\delta_3$	1.308	---	---	---	---

表 10 一因素 CES-D 結構之各題與評定量尺模式適配情形

位置參數估計值 (標準誤)		unweighted fit		weighted fit	
		MNSQ	T	MNSQ	T
$\lambda_1$	-0.429 (0.036)	0.92	-1.9	0.90	-2.3
$\lambda_2$	1.161 (0.042)	0.86	-3.6	0.95	-1.1
$\lambda_3$	0.464 (0.039)	0.72	-7.5	0.80	-4.5
$\lambda_4$	-2.244 (0.032)	2.16	21.0	2.05	19.7
$\lambda_5$	0.137 (0.038)	0.78	-5.7	0.79	-4.8
$\lambda_6$	-0.087 (0.037)	0.54	-13.4	0.59	-10.5
$\lambda_7$	-0.340 (0.036)	0.68	-8.7	0.69	-7.5
$\lambda_8$	-1.501 (0.033)	0.92	-2.0	0.91	-2.1
$\lambda_9$	1.215 (0.043)	0.70	-8.0	0.79	-4.6
$\lambda_{10}$	0.574 (0.040)	0.64	-9.8	0.73	-6.3
$\lambda_{11}$	0.116 (0.038)	0.87	-3.2	0.95	-1.1
$\lambda_{12}$	-1.363 (0.033)	0.75	-6.5	0.74	-6.6
$\lambda_{13}$	-0.329 (0.036)	0.96	-1.0	0.98	-0.4
$\lambda_{14}$	-0.029 (0.037)	0.67	-8.9	0.71	-7.0
$\lambda_{15}$	0.964 (0.041)	0.75	-6.4	0.84	-3.3
$\lambda_{16}$	-1.638 (0.033)	1.03	0.8	1.02	0.4
$\lambda_{17}$	1.701 (0.045)	0.78	-5.7	0.89	-2.3
$\lambda_{18}$	0.756 (0.040)	0.60	-11.3	0.67	-7.7
$\lambda_{19}$	0.826 (0.041)	0.69	-8.3	0.75	-5.6
$\lambda_{20}$	0.047	---	---	---	---
類別交叉參數估計值 (標準誤)					
$\delta_1$	-1.578 (0.016)	2.11	20.3	2.19	21.2
$\delta_2$	0.231 (0.025)	1.43	9.1	1.71	11.2
$\delta_3$	1.347	---	---	---	---

由表 11 可知，除了三因素結構下的「正向情感」因素未達顯著水準外，其餘皆達 0.01 顯著水準。但以 R 平方來看，問卷組別變項對所有因素的解釋力都非常小，不具實用性。

表 11 分別以三因素及一因素為依變項之潛在回歸分析

回歸變數	三因素 CES-D 結構			一因素 CES-D 結構
	憂鬱情緒 + 身體症狀	正向情感	人際問題	憂鬱
常數項	-3.182* (0.143)	-0.530* (0.082)	-4.082* (0.167)	-2.402* (0.104)
問卷組別變項	0.270* (0.089)	0.089 (0.051)	0.353* (0.104)	0.181* (0.065)
因素之變異量	2.252	0.703	3.161	1.199
R 平方 (%)	0.00%	0.00%	2.90%	1.00%

註：小括號內為估計標準誤。\*  $p < 0.01$ 。

從上述五種 CES-D 量表的因素結構適配度情形來看，以 MNSQ 的適配標準而言，四因素結構、一正一負二因素結構及三因素結構的適配情形較佳，而憂鬱與人際二因素結構及一因素結構的適配情

形稍微不佳，其關鍵都在於測量正向情感的第 4 題及第一個類別交叉參數之  $MNSQ > 2$ 。不過要注意的是，本研究的目的並不在決定哪一種 CES-D 量表的因素結構較佳，而是在各種因素結構所屬的題項皆處於可接受的適配情形下，去證明網路問卷與紙本問卷的調查結果在實質上是無差異的。

## 結論與建議

本研究爲了促進網路問卷在未來調查研究中的使用，以減少相關的研究及社會成本，因此，在經過文獻探討後，合理推論網路問卷與紙本問卷之間在匿名性、社會期許效應、合法性、認知負荷量上的相似性，並融入研究程序中小心地控制兩種調查方式流程的一致性，此外，亦對可能影響兩種調查方式差異效果的母群社經背景變項加以控制，然後進行網路問卷與常用的紙本問卷這兩種調查方式的比較，而獲得如下結論與建議。

### 一、結論

1. 由樣本回收的情形來看，網路問卷與紙本問卷的回收率都達到 5 成左右，雖然網路問卷的回收率較紙本問卷稍低，但其差異並非不可接受。特別在一些基本背景變項的比較上，回收率的些許差異並未導致網路與紙本問卷的受試者在背景變項的檢查上出現差異，表示所回收的樣本組成基本上是符合原先預期的同質性，亦即接受網路或紙本問卷的中小學教師在背景變項上並無不同。

2. 在問卷內容的比較上，由於 CES-D 量表在國際間廣泛的使用性，因此，本研究不厭其煩地針對感興趣的因素結構，進行網路問卷與紙本問卷在不同因素上的平均數差異比較，亦即以潛在回歸的方式探討問卷組別變項在各因素上的影響力。雖然統計考驗的結果大都顯示網路問卷與紙本問卷在平均數上達到顯著差異，但可以由決定係數（R 平方）看出，問卷組別變項的解釋力大都趨近 0，也就是在實用性上，這些差異沒有實質的意義。

### 二、建議

1. 儘管本研究團隊對匿名性的重要性作了文獻探討，但仍在匿名性細節的控制上出現了一個小疏忽，導致網路問卷組不管在回收率或在問卷內容的回答上，其統計檢定的結果都明顯低於紙本問卷組，特別是問卷調查內容涉及個人憂鬱傾向的隱私性問題，使得兩種調查方式之人爲疏忽的細小匿名差異被放大。因此，建議未來研究者不可在網路問卷上讓受試者姓名出現，不可因爲要使受試者感到網路問卷介面的友善性而與紙本組作法有差異。

2. 本研究採用 CES-D 量表作爲問卷主要內容的優點在於其廣泛的使用性，使得跨文化、母群或國際的調查方式效果之比較更具有研究比較上的意義。此外，儘管目前使用潛在變項作爲主要差異檢定的研究仍然少之又少，但使用潛在變項能有效減少一般觀察變項不考慮測量誤差所帶來的研究可信度問題，因此，建議未來的研究仍應以潛在變項爲主，使研究結果更具說服力。

3. 關於問卷調查的議題可能對網路問卷與紙本問卷造成的差異影響，憂鬱傾向以外的其他類型議題仍待更多的研究。余民寧、李仁豪（2006）曾經使用較爲中性的問卷議題來比較網路與紙本問卷的差異，獲得無差異的結論。另外，余民寧、李仁豪（2008 出版中）也曾經使用時事性價值評判的問卷議題來比較網路與紙本問卷的差異，同樣獲得兩種調查方式無差異的結果。儘管單憑本研究及上述文獻的結果，仍無法十分肯定地告訴未來研究者可以完全使用較少成本的網路問卷來取代紙本問卷，但藉由相關研究及本研究的研究過程，未來的研究仍可以繼續改進相關缺失，並持續進行相關調查方式比較的研究，相信對未來學術或公私立機構在調查作業上有一定的幫助。



## 參 考 文 獻

- 彭駕駢 (民 90)：從教師憂鬱症談心理適應。台灣教育，608 卷，26-31 頁。
- 游森期、余民寧 (民 95)：網路問卷與傳統問卷之比較：多樣本均等性方法學之應用。測驗學刊，53 卷，1 期，105-130 頁。
- 余民寧、李仁豪 (民 95)：調查方式與問卷長短對回收率與調查內容影響之研究。當代教育研究，14 卷，3 期，127-168 頁。
- 余民寧、李仁豪 (民 97，已接受)：調查方式與議題熟悉度對問卷回收率與內容的影響。教育學刊。
- Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly*, 58, 210-240.
- Bandilla, W., Bosnjak, M., & Altdorfer, P. (2003). A comparison of web-based and traditional written self-administered surveys using the ISSP environment module. *Social Science Computer Review*, 21, 235-243.
- Boisvert, J. A., McCreary, D. R., Wright, K. D., & Asmundson, G. J. G. (2003). Factorial validity of the Center for Epidemiologic Studies-Depression (CES-D) scale in military peacekeepers. *Depression and Anxiety*, 17, 19-25.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching & Teacher Education*, 16(2), 239-253.
- Bullough, R. V., Jr. (2005). Teacher vulnerability and teachability: A case study of a mentor and two interns. *Teacher Education Quarterly*, 32(2), 23-39.
- Carini, R. M., Hayek, J. H., Kuh, G. D., Kennedy, J. M., & Ouimet, J. A. (2003). College student responses to web and paper surveys: Does mode matter? *Research in Higher Education*, 44, 1-19.
- Chan, K. S., Orlando, M., Ghosh-Dastidar, B., Duan, N., & Sherbourne, C. D. (2004). The interview mode effect on the Center for Epidemiological Studies Depression (CES-D) scale: An item response theory analysis. *Medical Care*, 42(3), 281-289.
- Cheung, C-K., & Bagley, C. (1998). Validating an American scale in Hong Kong: The Center for Epidemiological Studies Depression Scale (CES-D). *The Journal of Psychology*, 132(2), 169-186.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Couper, M. P., & Rowe, B. (1996). Evaluation of a computer-assisted self-interview component in a computer-assisted personal interview survey. *Public Opinion Quarterly*, 60, 89-105.
- Czaja, R. (1987). Asking sensitive behavioral questions in telephone interviews. *International Quarterly of Community Health Education*, 8, 23-32.
- Denscombe, M. (2006). Web-based questionnaires and the mode effect: An evaluation based on completion rates and data contents of near-identical questionnaires delivered in different modes. *Social Science Computer Review*, 24(2), 246-254.
- Edman, J. L., Danko, G. P., Andrade, N., McArdle, J. J., Foster, J., & Glipa, J. (1999). Factor structure of the CES-D (Center for Epidemiologic Studies Depression Scale) among Filipino-American adolescents.

- Social Psychiatry and Psychiatric Epidemiology*, 34(4), 211-215.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.
- Goho, J. (2002). *Mixed mode effects in a community college graduate survey*. Paper presented at the 42nd Annual Forum of the Association for Institutional Research, 2-5 June, Toronto, Canada.
- Greenfield, T. K., Midanik, L. T., & Rogers, J. D. (2000). Effects of telephone versus face-to-face interview modes on reports of alcohol consumption. *Addiction*, 95(2), 277-284.
- Hayslett, M. M., & Wildemuth, B. M. (2004). Pixels or pencils? The relative effectiveness of Web-based versus paper surveys. *Library & Information Science Research*, 26, 73-93.
- Henson, R., Cannell, C. F., & Roth, A. (1978). Effect of interview mode on reporting of moods, symptoms, and need for social approval. *Journal of Social Psychology*, 105, 123-129.
- Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, 67, 79-125.
- Huang, H-M. (2006). Do print and Web surveys provide the same results? *Computers in Human Behavior*, 22, 334-350.
- Jurado, D., Gurpegui, M., Moreno, O., Fernandez, M. C., Luna, J. D., & Galvez, R. (2005). Association of personality and work conditions with depressive symptoms. *European Psychiatry*, 20, 213-222.
- Kwak, N., & Radler, B. (2002). A comparison between mail and web surveys: Response pattern, respondent profile, and data quality. *Journal of Official Statistics*, 18(2), 257-273.
- Lautenschlager, G. J., & Flaherty, V. L. (1990). Computer administration of questions: More desirable or more social desirability? *Journal of Applied Psychology*, 75(3), 310-314.
- Layne, B. H., DeCristoforo, J. R., & McGinty, D. (1999). Electronic versus traditional student ratings of instruction. *Research in Higher Education*, 40, 221-232.
- Linacre, J. M. (2006). *A user's guide to FACETS*. Retrieved August, 15, 2006 from <http://www.winsteps.com>
- McCabe, S. E. (2004). Comparison of web and mail surveys in collecting illicit drug use data: A randomized experiment. *Journal of Drug Education*, 34(1), 61-72.
- McCabe, S. E., Boyd, C. J., Couper, M. P., Crawford, S., & D'Arcy, H. (2002). Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. *Journal of Studies on Alcohol*, 63, 755-761.
- McCabe, S. E., Boyd, C. J., Young, A., Crawford, S., & Pope, D. (2005). Mode effects for collecting alcohol and tobacco data among 3rd and 4th grade students: A randomized pilot study of web-form versus paper-form surveys. *Addictive Behaviors*, 30, 663-671.
- McCabe, S. E., Couper, M. P., Cranford, J. A., & Boyd, C. J. (2006). Comparison of web and mail surveys for studying secondary consequences associated with substance use: Evidence for minimal mode effects. *Addictive Behaviors*, 31, 162-168.
- Miller, T. Q., Markides, K. S., & Black, S. A. (1997). The factor structure of the CES-D in two surveys of elderly Mexican Americans. *The Journals of Gerontology*, 52B(5), S259-S269.
- Motl, R. W., Dishman, R. K., Birnbaum, A. S., & Lytle, L. A. (2005). Longitudinal invariance of the Center for Epidemiologic Studies Depression Scale among girls and boys in middle school. *Educational and Psychological Measurement*, 65(1), 90-108.
- Newman, J. C., Des Jarlais, D. C., Turner, C. F., Gribble, J., Cooley, P., & Paone, D. (2002). The differential

- effects of face-to-face and computer interview modes. *American Journal of Public Health*, 92, 294-297.
- O'Rourke, N. (2005). Factor structure of the Center for Epidemiologic Studies Depression Scale (CES-D) among older men and women who provide care to persons with dementia. *International Journal of Testing*, 5(3), 265-277.
- Pealer, L., Weiler, R. M., Pigg, R. M., Miller, D., & Dorman, S. M. (2001). The feasibility of a web-based surveillance system to collect health risk behavior data from college students. *Health Education Behavior*, 28, 547-559.
- Pridemore, W. A., Damphousse, K. R., & Moore, R. K. (2005). Obtaining sensitive information from a wary population: A comparison of telephone and face-to-face surveys of welfare recipients in the United States. *Social Science & Medicine*, 61, 976-984.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401.
- Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, 84, 754-775.
- Schonfeld, I. S. (1990). Psychological distress in a sample of teachers. *The Journal of Psychology*, 124(3), 321-338.
- Schonfeld, I. S. (2000). An updated look at depressive symptoms and job satisfaction in first-year women teachers. *Journal of Occupational and Organizational Psychology*, 73(3), 363-372.
- Schwarz, N., Strack, F., Hippler, H. J., & Bishop, G. (1991). The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology*, 5, 193-212.
- Shaw, D., & Davis, C. H. (1996). The Modern Language Association: Electronic and paper surveys of computer-based tool use. *Journal of the American Society for Information Science*, 47, 932-940.
- Spijker, J., van der Wurff, F. B., Poort, E. C., Smits, C. H. M., Verhoeff, A. P., & Beekman, A. T. F. (2004). Depression in first generation labour migrants in Western Europe: The utility of the Center for Epidemiologic Studies Depression scale (CES-D). *International Journal of Geriatric Psychiatry*, 19, 538-544.
- Tomsic, M. L., Hendel, D. D., & Matross, R. P. (2000). *A World Wide Web response to student satisfaction surveys: Comparisons using paper and Internet formats*. Paper presented at the Annual Meeting of the Association for Instructional Research, Cincinnati, OH.
- Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology*, 55(1), 755-801.
- Wright, D. L., Aquilino, W. S., & Supple, A. J. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. *Public Opinion Quarterly*, 62, 331-353.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1998). *ACER ConQuest: Generalised item response modeling software manual*. Camberwell, Melbourne: The Australian Council for Educational Research Ltd.

收稿日期：2007年03月09日

一稿修訂日期：2007年07月04日

接受刊登日期：2007年07月09日

Bulletin of Educational Psychology, 2008, 39, Special Issue on Test and Measurement, 21-42  
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

## **Survey Mode Effect Analyzed by Item Response Theory on the Center for Epidemiological Studies Depression (CES-D) Scale**

**Ren-Hau Li**

Department of Psychology  
Chung Shan Medical University

**Jin-Chang Hsieh**

Department of Education  
National Chengchi University

**Min-Ning Yu**

Department of Education  
National Chengchi University

This study investigated whether differences exist between Internet and paper questionnaires in survey mode, where the internationally- and widely-used “Center for Epidemiological Studies Depression” (CES-D) scale was adopted as content of the questionnaire. Of 2400 primary and secondary school teachers stratified and sampled randomly from Taiwan, 1141 teachers responded to the 20\_item CES-D Scale, including 535 teachers in the Internet group, and 606 teachers in the paper group. This research used multidimensional latent regression module developed by ConQuest software. Using one-parameter item response theory, we compared the differences between the Internet and paper questionnaire groups. The results of this study showed no significant difference between the two groups in background variables; nonetheless, response rate in the Internet group was lower than that in the paper group. As for comparisons between content of questionnaires, although the latent means of the Internet group were significantly lower than those of the paper group, these differences had no substantial meanings. The group variable of survey mode explained close to zero variance in factors of the CES-D Scale. To examine whether Internet questionnaire survey can be used to substitute for paper survey in order to reduce research and administration cost, we encourage more comparisons between the two survey modes.

**KEY WORDS: Center for Epidemiological Studies Depression (CES-D) scale, item response theory, multidimensional latent regression, rating scale model**

## 附錄一

---

```

datafile CESD2007.dat;
format group 6 responses 7-26 !tasks(20);
model tasks+step;
score (0,1,2,3) (0,1,2,3) ( ) ( ) ( ) ! tasks (3,6,9,10,14,17,18);
score (0,1,2,3) ( ) (0,1,2,3) ( ) ( ) ! tasks (4,8,12,16);
score (0,1,2,3) ( ) ( ) (0,1,2,3) ( ) ! tasks (1,2,5,7,11,13,20);
score (0,1,2,3) ( ) ( ) ( ) (0,1,2,3) ! tasks (15,19);
export reg_coefficient >> CESd4.reg;
export covariance >> CESd4.cov;
export parameters >> CESd4.prm;
set update=yes,warnings=no;
estimate!fit=no,method=montecarlo,nodes=400,conv=.01;
reset;
datafile CESD2007.dat;
format group 6 responses 7-26 !tasks(20);
model tasks+step;
score (0,1,2,3) (0,1,2,3) ( ) ( ) ( ) ! tasks (3,6,9,10,14,17,18);
score (0,1,2,3) ( ) (0,1,2,3) ( ) ( ) ! tasks (4,8,12,16);
score (0,1,2,3) ( ) ( ) (0,1,2,3) ( ) ! tasks (1,2,5,7,11,13,20);
score (0,1,2,3) ( ) ( ) ( ) (0,1,2,3) ! tasks (15,19);
export parameters >> CESd4.prm;
import init_reg_coefficient << CESd4.reg;
import init_covariance << CESd4.cov;
import init_parameters << CESd4.prm;
set update=yes,warnings=no;
estimate!method=montecarlo,nodes=1000,conv=.002;
show !tables=1:2:3:4:5 >> CESd4.shw;
quit;

```

---

附錄二

---

```
datafile CESD2007.dat;
format group 6 responses 7-26 !tasks(20);
regression group;
model tasks+step;
score (0,1,2,3) (0,1,2,3) ( ) ( ) ! tasks (3,6,9,10,14,17,18);
score (0,1,2,3) ( ) (0,1,2,3) ( ) ! tasks (4,8,12,16);
score (0,1,2,3) ( ) ( ) (0,1,2,3) ( ) ! tasks (1,2,5,7,11,13,20);
score (0,1,2,3) ( ) ( ) ( ) (0,1,2,3) ! tasks (15,19);
import anchor_parameters << CESd4.prm;
import init_covariance << CESd4.cov;
estimate!method=montecarlo,nodes=1000,conv=.003;
show !tables=1:2:3:4:5 >> CESdreg4.shw;
quit;
```

---