

臺灣憂鬱症量表的 實用決斷分數編製報告

余民寧

國立政治大學教育學系教授

劉育如

經國管理暨健康學院講師

李仁豪

中山醫學大學心理學系助理教授

摘 要

本研究根據全人治療取向與精神科醫生的臨床診斷，企圖編製一份自評式臺灣憂鬱症量表，供一般社會大眾作為簡易自評診斷的使用。本量表共包含22題，每題皆以李克特氏四點量表方式計分，分別是0、1、2、3分。本量表內含四個因素，分別是（1）認知向度（6題）；（2）情緒向度（6題）；（3）身體向度（6題）；（4）人際向度（4題）。研究樣本來自大台北地區的213名憂鬱症患者與200一般大學生。本量表經結構方程式模型驗證後獲得合理的信效度證據，並與CES-D憂鬱量表間具有高達0.92的效標關連效度。此外，以四向度分數作為預測因子的區別分析，也獲得93.2%的正確分類百分比。最後，為了實務應用上的目的，最好能界定一個決斷值來區分出潛在的憂鬱症患者與一般正常人的不同。本研究使用試題反應理論中的最大訊息量法，經資料分析後獲得決斷值為37分，並建議以此決斷值供作臨床工作者或一般社會大眾作為篩選潛在憂鬱症患者的依據。

關鍵詞：臺灣憂鬱症量表、試題反應理論、決斷值、最大訊息量法

The Practical Usage of Cutoff Score in the Taiwanese Depression Scale

Min-Ning Yu

Professor, Department of Education, National Chengchi University

Yu-Ju Liu

Lecturer, Ching Kuo Institute of Management and Health

Ren-Hau Li

Assistant Professor, Department of Psychology, Chung Shan Medical University

Abstract

A Taiwanese Depression Scale (TDS) with four factors is newly developed based on the holistic approach to healing theory and empirical psychiatrists' diagnosis. 22 items are included in this scale and each is scored by four-point Likert type scale, namely, 0,1,2,3 points. Four factors included in this scale are (1) cognitive dimension with six items, (2) emotional dimension with six items, (3) physical dimension with six items, and (4) interpersonal dimension with four items. The sample includes 213 depression patients and 200 normal undergraduates. The scale tested by structural equation model reveals reliable and valid factor structures, and the criterion-related validity with CES-D scale is about 0.92. Besides, 93.2% of original grouped cases are correctly classified by discriminant analysis where the four subscale factors are used as predictors. Finally, for practical usage, it is better to decide a cutoff score to discriminate potential depression patients from general normal people. We use the item response theory with its maximum information method to determine the cutoff score (i.e., 37) which making more objective judgment than ever before. We would like to recommend that our newly-developed Taiwanese Depression Scale (TDS) be a good tool for practical usage in screening out potential depression patients, and as a means to replace some out-of-date depression scales.

Keywords: Taiwanese Depression Scale, item response theory, cutoff score, maximum information method

壹、緒論

憂鬱症是現代文明社會的一大隱憂，根據美國國家心理健康機構（National Institute of Mental Health, NIMH）的統計，大約有將近一千八百萬位18歲以上美國成人患有憂鬱症，約佔美國人口數的9.5%，其中女性（12.0%）大約是男性（6.6%）的兩倍（NIMH, 2001）。世界健康組織（World Health Organization, WHO）亦估計，到了廿一世紀的20年代左右，憂鬱症所造成的社會負擔將是所有疾病的第二名，僅次於第一名的心臟病（WHO, 2001）。若排除身體疾病，無疑地，憂鬱症將是影響人類社會最大的心理疾病，嚴重的患者甚至會出現自殺的意圖及行動，導致極高的死亡率。一如國際自殺防治協會（International Association for Suicide Prevention, IASP）2006年在世界自殺防治日的文宣中指出，進入廿一世紀短短幾年，全球已累積超過五百萬人自殺，而在自殺死亡的人數中，約有三分之二的人患有憂鬱症（IASP, 2006），顯見憂鬱症是自殺的潛在風險因子，對國家社會所造成的傷害難以估計。

關於憂鬱症的診斷，精神科醫生最常依據的標準是美國精神醫學協會（American Psychiatric Association, 2000）所頒訂的心理疾病診斷與統計手冊第四版（the Diagnostic and Statistical Manual of Mental Disorders, DSM-IV-TR）。在該手冊中，將一系列可能罹患憂鬱症的症狀條列供醫生判斷，症狀持續二週以上將被診斷為憂鬱症患者。然而，此種由醫生根據DSM-IV標準來進行的診斷是屬於半結構的訪談方式，過程可能較為費時並且僅侷限於有權從事醫療行為的醫生等專業人員。此外，大部分可能有憂鬱傾向的患者並不會主動就醫，即使是這些潛在病患的家人，亦不會主動帶病人到醫院精神科尋求協助。因此，早在二十世紀60年代起，即有許多自評式的憂鬱症量表被開發出來，用來簡易評估潛在患者的憂鬱程度，例如：有名的貝克憂鬱症量表（Beck Depression Inventory, BDI）（Beck, Ward, Mendelsohn, Mock, & Erbaugh, 1961; Beck, Steer, & Brown, 1996）、流行病學研究中心憂鬱症量表（Center for Epidemiologic Studies Depression Scale, CES-D）（Radloff, 1977）等。

本研究的目的，即是在編製一套適合台灣本土的憂鬱症量表，暫時取名為「臺灣憂鬱症量表」（Taiwanese Depression Scale, TDS）。儘管台灣地區已有一些學者或醫生自行發展出憂鬱症量表，例如，董氏基金會網站提供的台灣人憂鬱症量表(李昱、楊明仁、賴德仁、邱念睦、周騰達，2000)，但本研究所採用的編製方法，除了在樣本上擴大及不同統計方法的採用外，乃是根據美國全人醫療協會（American Holistic Medical Association, AHMA）所提出的「全人取向的醫治方式」理念，來組織過去研究的發現，以作為建構一份自評式憂鬱症量表的架構依據。該理念強調從更人性化的角度，來提供全人取向的治療方針；針對憂鬱症患者而言，全人取向的醫治方式需要檢視患者在認知、情緒、身體、與人際關係等四個面向，從這些層面因素來瞭解憂鬱症發生的症狀，並針對個人獨特的情況設計專屬的醫療方式（Jantz & McMurray, 2003）。因此，本研究擬從憂鬱症的認知、情緒、身體、與人際關係等四個面向，來歸納相關的憂鬱症文獻，並徵詢精神科醫師臨床問診的心得，企圖編製與發展一份自評式的「臺灣憂鬱症量表」。底下，茲評閱與簡述本研究自編憂鬱症量表所持四個面向（即認知、情緒、身體、與人際關係）的相關文獻。

一、憂鬱的認知向度

根據憂鬱症者對自我、世界、及未來等方面看法所使用的形容詞可知，其用語多半偏向負面的、消極的特徵描述，諸如：（1）在自我方面：使用「不適當、缺乏、軟弱」等用語，（2）在世界方面：使用「陰沈、孤獨、不公平」等用語，（3）在未來方面：使用「無趣、晦暗、無望」等用語（Greenberg & Beck, 1989）。而Beck, Perkins, Holder, Robbins, Gray,及Allison（2001）在歸納有關區分憂鬱與焦慮之認知內容特徵的研究後指出，憂鬱症患者通常會出現悲觀與自我批判的認知。MacLeod及Salaminiou（2001）的研究則指出，憂鬱症患者對事件正向經驗無法預期的原因，是患者對於正向經驗之心理表徵的觸接有所困難。此外，有報告也指出，憂鬱症與負向的自我評價有很大的關係，包括低自尊與低自我效能（Brosse, Sheets, Lett, &

Blumenthal, 2002)，這些評述都涉及憂鬱症認知層面的探討。

Leary (1995) 的研究指出，根據自尊的社會量尺模式 (socialmeter model)，低自尊者常伴隨著真實、預期或想像的拒絕，研究也發現，憂鬱與低自尊有顯著關係存在。而Roberts及Gotlib (1997) 更進一步以長達六週的追蹤研究發現，自尊與特定的自我評價可以有效地預期生活壓力事件所產生的憂鬱症狀。此外，Tillema, Cervone,及Scott (2001) 的研究也發現，憂鬱症患者在缺乏情緒相關的線索下，所持的表現標準常超過本身自我效能的知覺，因此，他們會展現出自我挫敗 (self-defeating) 的認知組型。另外，Muris, Schmidt, Lambrichs,及Meesters (2001) 以結構方程式模型進行憂鬱症風險及保護因子的研究，結果發現被動的處事風格 (passive coping style) 及負向的歸因風格 (negative attribution style) 可以有效的預測憂鬱程度。最近的研究則指出，憂鬱症狀對病人在魏氏智力測驗上的表現有強烈的影響效果，而在經過適當的藥物治療後，認知功能及症狀都有一定程度的改善(Mandelli, Serretti, Colombo, Florita, Santoro, Rossini, Zanardi, & Smeraldi, 2006)。

綜合上述，我們可以發現，憂鬱症患者在臨床症狀上常表現出認知上的缺陷，包括自我否定及負向思考，嚴重者將導致自殺行為的出現，誠如Oei, Bullbeck, 及Campbell (2006) 所說的，憂鬱症與無望 (hopelessness) 及自殺之間有特別的連結關係。

二、憂鬱的情緒向度

從憂鬱的情緒面向來看，憂鬱症本身在DSM-IV的診斷標準上，即被歸類為一種情感疾病 (affective disorder)。憂鬱情緒是指長期的憂傷、焦慮、罪惡感、羞辱感和沒有希望、憤怒、沮喪等感覺 (Jantz & McMurray, 2003)。Carey, Carey, 及Kelley (1997) 歸納了過去許多有關情緒理論後，暗示生活事件對情緒產生的衝擊，他們也使用階層迴歸進行分析後發現，情緒對憂鬱症狀有很高的解釋力。而Tillema, Cervone,及Scott (2001) 的研究發現，憂鬱症患者與非憂鬱症患者在認知上的自我評判 (self-judgment) 背後可能有不同的

情緒過程，當前的情緒狀態在表現標準的建構與自我效能間扮演一個重要的角色。這些研究發現，意謂著負向的情緒反應不僅是憂鬱症的症狀，同時在憂鬱認知的形成過程中，也展現出一定的重要性。

Beevers及Miller（2004）的研究中即發現，憂鬱有關的負向認知不僅具有情緒的特質屬性，同時也具有隨情緒狀態波動的依賴性。此外，在憂鬱症的神經心理研究上，憂鬱症患者對情緒刺激的編碼與辨識過程所產生的事件關連腦電位（event-related brain potential, ERP）組型與正常人不同（Deldin, Keller, Gergen, & Miller, 2001）。另外，值得注意的是Beck, Perkins, Holder, Robbins, Gray,及Allison,（2001）歸納過去的研究發現，憂鬱情緒與焦慮情緒之間有些重疊特徵，包括非特定的苦惱（distress）與不快樂的情緒，而憂鬱與焦慮不同之處在於憂鬱展現出低度的正向情感，而焦慮則是出現生理上的過度喚起（hyperarousal）。然而，廣義而言，對於躁鬱型的憂鬱症患者來說，仍然會出現情緒或生理上過度喚起的現象，因此，發脾氣或生氣等反應，亦有可能是憂鬱症患者的症狀之一。

三、憂鬱的身體向度

Jantz及Mcmurray（2003）根據醫學上的診斷，認為憂鬱症患者可以從身體上的症狀尋得，例如：他們常會產生睡眠行為的改變，睡太多或太少；在胃口上，會覺得沒胃口，或覺得吃不飽；體重改變、疲倦、無精打采、記憶力減退；或無法專注、容易焦慮；或容易生氣、缺乏成就動機等身體症狀。上述身體上的臨床症狀，通常都不是產生憂鬱症的主要原因，但是在臨床診斷上必須依賴這些身體症狀來判斷患者的憂鬱嚴重程度。過去，主要的治療理論大都強調在心理層面的認知或情感上的介入處理，除了少數十分嚴重憂鬱症患者需要藥物治療外，較少研究強調從身體的層面著手。不過，Brosse, Sheets, Lett,及Blumenthal（2002）從生理、心理社會機制的觀點收集相關研究後，提出運動有助於改善憂鬱症狀的論點，包括運動對生理上神經激素的調整、心理上自我評價的提升、身體意像的改善等等。雖然，目前仍不是很清楚運動對改善憂鬱症狀的詳細機制，但他們確實提出一個治

療憂鬱症方法的未來值得努力研究的新方向。

最近，Zarrinpar, Deldin,及Kosslyn（2006）以視覺心理影像（visual mental imagery）的產生與旋轉作業為工具，對憂鬱症患者與一般正常人進行研究，結果發現憂鬱症患者的感覺及運動處理程序較一般正常人遲緩，造成遲緩的原因可能在於訊息編碼或運動輸出上。而Christopher及MacDonald（2005）的研究則發現，憂鬱症會影響注意力的分配，並傷害工作記憶（working memory）的所有元素，包括聽覺及視覺空間成分。Gallassi, Sarro, Morreale,及Amore（2006）的研究亦發現，憂鬱症對記憶的負向效果雖然在藥物治療後獲得顯著改善，但是並不能完全恢復，他們懷疑憂鬱症可能有狀態（state）及特質（trait）成分在影響記憶功能。上述這些症狀與受影響功能，均會在患者的身體症狀上反映出來。

四、憂鬱的人際關係向度

憂鬱症患者通常會報告出有較多的生活壓力事件，特別是那些因為個人特質或行為所引起的人際互動而產生的生活壓力事件（Hammen, 2006），而在人際關係上的互動困難，可能會阻礙患者的治療效果（Puschner, Kraft, & Bauer, 2004）。另外，有部分學者暗示憂鬱症狀的出現與患者人格特質有所關連與區隔（Mcdermut, Zimmerman, & Chelminski, 2003），例如，依賴傾向強的憂鬱症患者較常出現寂寞、失落及孤離，而自我批評或完美主義傾向者的憂鬱症患者則較常出現挫敗、自我責難及退縮等症狀（Barrett & Barber, 2007；Blatt & Zuroff, 1992）。Barrett及Barber（2007）的研究發現，具有憂鬱人格的憂鬱症患者比沒有憂鬱人格的憂鬱症患者報告更多的人際關係苦惱，並有支配、控制的人際關係問題出現。此外，與人際關係困難有關的憂鬱症在兒童時期出現並不少見，Brendgen, Vitaro, Turgeon,及Poulin（2002）以國小4～6年級的學童為研究對象，結果發現有憂鬱症的學童報告其與同儕的相處有困難，但其同儕的報告則無此現象。Rockhill, Fan, Katon, McCauley, Crick,及Pleck（2007）研究兒童期的憂鬱症與人際功能缺陷的關係，將高低憂鬱兒童分別與其自陳最要好的朋友來配對進行電玩遊戲，結果

發現雖然高低憂鬱兒童在正向或負向情緒上並無不同，但他們的配對同伴之感受則十分不同；與低憂鬱兒童的配對同伴相較下，高憂鬱兒童的配對同伴在競爭遊戲中表現出更多的負向情緒，而在合作遊戲中則表現較少的正向情緒。此外，高憂鬱兒童與配對同伴間的互動，也較低憂鬱兒童與配對同伴間的互動較不愉快。令人吃驚的是，憂鬱症患者可能在很小的時候便已有人際關係上的困難，而後人際關係互動的困難逐漸形成其人格特質的一部份，導致青少年或成年後一再復發的憂鬱症（Rockhill, et al., 2007）。雖然，人際關係困難可能加劇憂鬱症狀的產生，甚至形成人格特質的一部份，影響個人的生命全程，但是人際社會網絡也有其良善的一面，例如：已有研究指出社會支持有助於改善憂鬱症狀，是防止憂鬱症的保護因子（Panzarella, Alloy, & Whitehouse, 2006）。

綜合上述文獻評閱，本研究的目的即在編製一份具有臨床實用意義的自評式憂鬱症量表，並試圖提供決斷分數，以供臨床工作者或一般社會大眾作為初步診斷憂鬱症的篩選工具，從而減少精神科醫生的負荷，並提供臨床診斷的參考依據。因此，本研究在量表編製過程中，較為倚重精神科醫生的臨床診斷經驗，並輔以量化統計分析數據，作為提供相關信效度的證據訊息。

貳、研究方法

一、研究樣本

本研究樣本有兩種來源，其一來自台北市某兩家大型醫院的憂鬱症患者213人，這些患者皆已經過精神科醫生的診斷確認為罹患憂鬱症者，其中男性佔34.0%，女性佔66.0%，婚姻狀態為50.2%已婚，38.4%未婚，11.4%離婚，學歷為高中以下佔23.6%，大專佔56.6%，研究所以上佔12.7%；其二為一般正常人200人，取樣自北部某大學之大學生，其中男性佔41.0%，女性佔59.0%，皆為未婚。本研究以此合併樣本，作為進行信效度分析的對象。本研究樣本除了提供憂鬱症量表編製所需的人數外，還有另一個目的：就是想要利用憂鬱症患者與正常人的憂鬱程度不同，來獲得本研究自編量表的區

別效度 (discriminant validity) 證據，以及作為實用上分類正確性 (hit ratio) 的訊息。這些樣本經確定後，即分別接受本研究自編憂鬱症量表和外在效標測量工具 (即CES-D憂鬱症量表) 的同時施測。

二、研究工具

本研究除了自行編製憂鬱症量表外，為了驗證本自編量表是否具有效標關聯效度 (criterion related validity)，同時使用美國流行病學中心的憂鬱症量表 (CES-D; Radloff, 1977) 作為本研究自編量表的外在效標測量工具，以供作驗證效標關連效度證據之用。

本研究自編量表的原始題目共37題，係由研究者根據相關文獻評閱心得，以及在台北市某兩家大型醫院以一年的時間實際觀察精神科醫生問診所得的資訊編製而成，採用李克特式四點評定量表 (Likert type four-point rating scale) 的作答形式，填答「總是如此」者得3分，填答「經常如此」者得2分，填答「偶爾如此」者1分，而填答「很少如此」者得0分。由於作為迅速便利篩選用途之量表不宜過長，原始題目乃經由三位精神科醫生以專家角度進行內容效度篩選後，取出較具代表性的22題，作為本自編量表的正式施測用量表試題。

而外在效標測量工具方面，美國流行病學研究中心 (the Centre for Clinical Trials and Epidemiological Research, CCTER) 所發展出來的20題CES-D憂鬱症量表，包含憂鬱情緒、正向情感、身體症狀、人際問題四因素，以李特克氏四點量尺方式計分。該量表之中譯版取自楊中芳 (1997) 所譯的「性格與社會心理測量總覽 (上)」一書。其量表內部一致性 α 信度係數與Spearman等級相關係數均為0.90，再測信度為0.32 (間隔12個月) 到0.67 (間隔4週) 之間；CES-D之聚斂效度 (convergent validity) 顯示與臨床醫師及護士評定的Hamilton Clinician's Rating量表及Raskin Rating量表之相關為0.44~0.56，而區辨效度 (divergent validity) 則是與社會期許量表之間有一低度負相關 ($r = -0.18$)，可說是一份兼具信效度特色並適合用來測量憂鬱心理特質的工具 (Radloff, 1977)。CES-D雖然在因素結構上與本自編量表的四個面

向接近，但仍有所區別，首先，其正向情感因素是以正向情緒的方式來撰寫題意，所測得的構念並不完全與憂鬱情緒相等；再者，其人際問題因素只有短短2題，並非十分理想。因此，本研究自編憂鬱症量表基於構念相似而採用CES-D量表作為效標，但也試圖改進其缺失，希望本自編量表能在台灣樣本的條件下，獲得理想的信效度結果。

三、資料分析方法

本研究資料分析有兩個重點，其一是確認所編製量表的信效度係數的大小，其二是由量表原始總分決定一個決斷值，以利判斷得分在幾分以上時可被歸類為潛在的憂鬱症患者。本研究茲使用LISREL 8.52版、WINSTEPS 3.58版、與SPSS 13.0版等統計軟體，來進行有關的資料統計分析。

（一）信效度分析

本研究根據相關文獻評閱所得憂鬱症四個向度為架構依據，假定四個向度背後有一個二階因素（second-order factor）的存在—即「憂鬱症」（depression），並以此架構為基礎，應用結構方程式模型（structural equation modeling, SEM）方法學，進行驗證性因素分析（confirmatory factor analysis, CFA）工作，以檢定本研究所收集資料與模式間是否適配，並獲得相關的信效度訊息。此外，再利用本自編量表四個向度的分數，進行一次區別分析（discriminant analysis），以區別函數（discriminant functions）的分類結果與已知的樣本類別（即憂鬱症患者與正常人），進行分類正確性的計算。最後，本研究亦同時提供本自編量表與CES-D憂鬱量表之間的相關係數，作為本自編量表的效標關連效度之證據。

（二）界定決斷值

在臨床應用上，通常都會根據量表原始總分決定出一個臨界點分數（或稱作「決斷值」（cutoff score）），以利實務工作者未來使用的方便。本研究係應用試題反應理論（item response theory, IRT）中的評定量尺模式（rating scale model, RSM），以最大訊息量法（maximum information method）來判

斷臨界點分數的所在。未來，在實務應用上，凡受試者的得分在此臨界分數以上者，即將被視為憂鬱症的潛在患者，需要再進一步經精神科醫師的確認，以決定是否接受後續的醫療程序；當然，在此臨界分數以下的人，並不表示即可忽略，仍然值得相關醫護人員適當的注意與關心。為使此決斷值有實務應用上的價值與意義，本研究將依據此決斷值分數進行樣本的重新歸類分析，再與原本已知的樣本類別，進行分類正確性的計算，以提供實用上可靠的區辨效度證據。

從理論觀點而言，試題反應理論下的最大訊息量概念，可以簡單地說是最能反應出該量表最高測量精確性之分數所對應的能力值，亦即，最大訊息量所對應的能力值之測量誤差（measurement error），是所有能力值所具有之測量誤差中最小的一者，也就是說，在最大訊息量對應下的能力估計值是最精確的一者。因此，從鑑別度的觀點來看，最大訊息量所對應的受試者能力量尺分數，即代表該分數最能區分出樣本憂鬱程度的分類指標；若以此臨界點分數作為決斷值來判定個體是否罹患憂鬱症，是一項很不錯的鑑別方式。而在較具代表性的樣本下，以最大訊息量方式來決定本量表的臨界點分數，將比使用傳統主觀或隨意界定的臨界點分數來得客觀。

參、結果與討論

本節將呈現三個部分，一是專家對內容效度的審查與選題工作，二是量化信效度分析的結果，三是決斷值及分類正確性的計算。

一、專家內容效度

本研究著眼於本自編量表能在未來臨床與自我診斷上的實務應用價值，特別商請三位具有不同經驗的精神科醫師，針對原始37題試題進行審題及措辭修改工作。根據這些專家修改措辭及刪題的結果，最終獲得22題，包括「認知向度」六題、「情緒向度」六題、「身體向度」六題、與「人際關係向度」四題，茲將題號重新整理後，陳列於表1裡。

表1 本研究自編憂鬱症量表最後保留的正式施測用題目

保留的題目與原題號	重新排序的題號及向度歸屬	
3. 有自殺的念頭	1	認知向度
6. 對什麼事都失去興趣	2	
9. 凡事往壞的方向想	3	
10. 有罪惡感	4	
14. 感覺自己很沒用	5	
25. 無力感	6	
31. 有壓力	7	情緒向度
7. 發脾氣、生氣	8	
12. 擔心、煩惱	9	
16. 害怕、恐懼	10	
19. 想哭	11	
20. 心情低落	12	
1. 胃口不好（或暴飲暴食）	13	身體向度
2. 睡眠狀況不佳	14	
11. 身體疲憊	15	
13. 無法專心做事	16	
17. 身體不舒服	17	
18. 記憶力不好	18	
5. 不想與他人往來	19	向度人際關係
21. 少說話（或不太愛說話）	20	
22. 不想出門	21	
36. 生活圈小	22	

二、信效度分析

本小節將描述驗證性因素分析結果、區別效度及效標關聯效度證據等訊息。

(一) 驗證性因素分析結果

關於所蒐集的資料與驗證性因素分析理論模式的適配情形，可以從基本適配度指標、整體模式適配度指標、與模式內在品質三個方面來評估（Bagozzi & Yi, 1988）。但值得注意的是，資料與模型的適配好壞需以整體的角度來審視，不宜僅根據少數指標不佳便否定整體評估的價值，而仍需考慮量表發展過程中所意涵的理論建構。

1. 基本適配度檢定

本研究自編憂鬱量表之驗證性因素分析路徑關係模式如圖1所示，參數估計後發現，所有誤差變異數皆為正值，且皆達0.05的顯著水準；所有估計參數的標準誤並未有特別不尋常之處；各單題變項與一階的因素間因素負荷量都在0.5以上、0.95以下，但是二階的因素負荷量則在0.92~0.99之間，這意謂著雖然在理論構念上憂鬱症可以從四個向度來判斷，但這四個向度的關聯性也相當高。因此，除了四個一階的因素彼此間的相關過高外，其它參數估計值間的相關也都非常地小，幾乎都為0。

2. 整體適配度檢定

本模式的 χ^2 值為597.60，自由度為205， p 值為0.00，未能獲得不顯著結果，這可能與本樣本較大有關。因此，進一步計算卡方值與自由度的比例結果，約為2.80，顯示在可接受的範圍。而RMSEA = 0.068、RMR與SRMR皆為0.032，也都在理想的範圍內。另外，雖然GFI = 0.88、AGFI = 0.86未能達到0.9以上的標準，但其他適配指標如NFI = 0.98、NNFI = 0.99、CFI = 0.99、IFI = 0.99、RFI = 0.98都達到0.9以上的標準。整體而言，整體適配度指標顯示資料與模式間的適配情形是可以接受的。

3. 模式內在品質檢定

個別題目（如x1~x22）的信度，除了x1 = 0.49、x4 = 0.49、x8 =

0.45、 $x_{13}=0.45$ 、 $x_{18}=0.45$ 與 $x_{22}=0.40$ ，低於0.5之外，其他各題皆介於0.53~0.79之間；而認知因素、情緒因素、身體因素、與人際關係因素的潛在變項平均變異抽取量約在0.56~0.62；所有潛在變項的成份信度，則在0.87~0.91之間；而四個向度的 α 信度係數分別為認知因素的 $\alpha=0.91$ ，情緒因素的 $\alpha=0.90$ ，身體因素的 $\alpha=0.88$ ，人際關係因素的 $\alpha=0.86$ 。此外，在所有231個標準化殘差值中，有29個標準化殘差超過2.57，其中最大是6.97，最小的是-3.37；出現大於修正指標（MI值）6.64以上者（即 $\alpha=0.01$ ， χ^2 自由度為1時的臨界標準）的比例並不高。

由以上三個面向的適配情形來看，除了整體適配性指標GFI、AGFI未達0.9的標準、一階的四個向度因素相關太高（介於0.89~0.98之間，但只有兩個相關係數高於0.95）及少數單題的信度小於0.5外，其它的指標都顯示合理的適配結果。因此，本研究將以此結構方程式模型的適配結果，作為支持本量表因素結構的基本信效度證據之一。

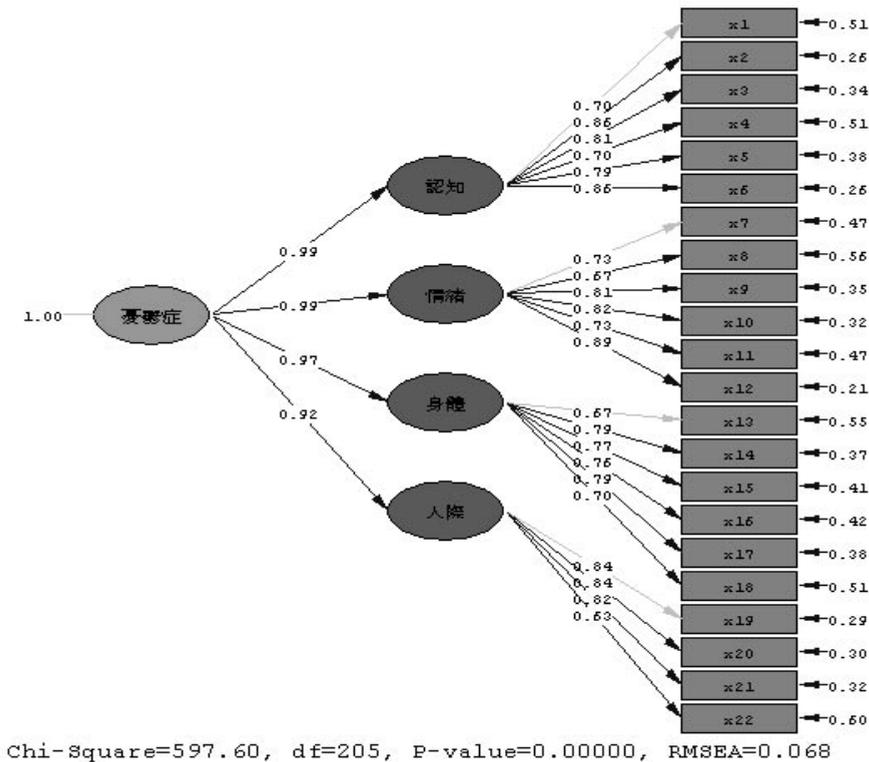


圖1 本自編量表的驗證性因素分析之路徑關係圖

（二）區別效度與效標關連效度

為了提供更多本研究量表信效度的證據，本研究利用區別分析的技術，以認知、情緒、身體、人際關係四向度的分數作為預測變項，來針對依變項進行區別分析，依變項即是憂鬱症患者與一般正常人所形成的組別變項。分析結果顯示，無論是使用逐步方式或同時進入法，四個預測變項都具有顯著的效果，所獲得的區別函數將樣本重新歸類為兩群的結果如表2所示，正確分類的百分比為 93.2%（即 $(189+196) / 413 = .93$ ），已高於80%的建議標準（Hair, Black, Babin, Anderson, & Tatham, 2006）。而使用Kappa分類一致性檢定的結果，其值為0.87，非常接近1，並達 $\alpha = 0.001$ 的顯著水準。這兩種結果都表示使用四個向度分數所建構的區別函數具有良好的區別效果，能使分類後的樣本歸屬結果與已知的原本樣本組別達到相當一致的程度；同時，也證明本研究量表中的四個一階因素之構念並非是完全重疊的，仍然具有不錯的區別性。另外，本研究自編憂鬱症量表與國際知名的CES-D憂鬱症量表間之效標關聯效度亦達0.92，顯見本量表所測量的憂鬱症構念與CES-D憂鬱症量表所測得的構念間，具有高度的相似性。

表2 本研究自編量表的區別分析結果摘要表

		預測組別		總人數
		患者組	正常組	
原來 組別	患者組	189	24	213
	正常組	4	196	200
總人數		193	220	413

三、決斷值與分類正確性

在實務應用上，本研究自編憂鬱症量表的應用功能頗類似於標準參照測驗（criterion-referenced test），希望能有效界定個體在量表得分多少以上可被判定為「潛在的憂鬱症患者」（potential depression patient）。因此，本研究擬採用較為客觀的實徵分析方法，即利用試題反應理論中的評定量尺模式，將這22題所產生的最大訊息量所對應到受試者能力量尺上的分數作為

決斷值。值得注意的是，本節所使用的Winsteps統計軟體主要是依據單參數Rasch模式而設計的，資料必須符合單向度潛在能力的基本假設，而本研究由之前結構方程式模型的檢定可知，本量表具有一個二階的「憂鬱症」因素，因此，整體來說，本樣本資料符合評定量尺模式應具備單向度潛在能力特質的基本假設要求。

在使用試題反應理論模式進行有關的參數估計之前，需先確認作答反應資料與模式適配的情形是否合理。Winsteps軟體提供兩類適配指標可供參考使用，分別是INFIT及OUTFIT指標，每類指標皆計算MNSQ及ZSTD兩種數值，MNSQ係指所有受試者在該題項上的實際觀察值與模式期望值之間差異的標準化殘差均方和，即標準化殘差變異的平均值，以1.0為標準，大於1.0即表示資料填答情形超出模式的預期，小於1.0即表示優於模式的預期；而ZSTD則是將MNSQ進行Wilson-Hilferty轉換後，以方便對照常態Z分佈機率表來判定MNSQ是否達顯著水準（Bond & Fox, 2001）。根據Linacre及Wright（1994）的看法，對於評定量尺而言，MNSQ在0.6~1.4之間是合理的，代表受試者在題項上的反應符合模式預期的範圍。而Linacre（2006）對MNSQ合理範圍的建議則為MNSQ>2.0表示該題將扭曲或破壞測量系統；MNSQ在1.5~2.0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。此外，當MNSQ值可以被接受時，ZSTD值是可以被忽略的（Linacre, 2006）。

表3所示，即是22題作答資料與評定量尺模式適配的情形。不論是從INFIT或OUTFIT來看，MNSQ值大約在0.57~1.63之間，儘管部分的ZSTD值有些過大，仍表示資料與模式的適配情形可以被接受。表3中所示的測量值，係指各試題的閾限參數值（threshold parameter），其範圍值在-1.08~1.60之間，連同其估計標準誤一併在此提供參考。

表3 本研究自編量表中各試題與評定量尺模式適配情形

題號	測量值	標準誤	INFIT		OUTFIT	
			MNSQ	ZSTD	MNSQ	ZSTD
1	1.60	.09	1.33	3.7	1.17	1.5
2	.40	.08	.73	-4.2	.67	-4.7
3	-.28	.08	.86	-2.2	.86	-1.9
4	.74	.08	1.21	2.7	1.22	2.5
5	.35	.08	.93	-1.0	.91	-1.2
6	-.25	.08	.63	-6.2	.64	-5.4
7	-1.08	.08	.97	-.5	1.04	.5
8	.17	.08	1.27	3.7	1.23	2.8
9	-1.02	.08	.71	-4.7	.70	-3.9
10	.42	.08	.84	-2.4	.75	-3.4
11	.37	.08	1.19	2.6	1.11	1.4
12	-.46	.08	.58	-7.4	.57	-6.7
13	.58	.08	1.56	6.8	1.63	6.5
14	-.71	.07	1.15	2.2	1.33	3.8
15	-1.03	.08	.89	-1.6	.89	-1.4
16	-.32	.08	.87	-2.1	.86	-1.9
17	-.04	.08	.96	-.5	.98	-.3
18	-.15	.08	1.15	2.1	1.28	3.4
19	.40	.08	.92	-1.1	.93	-.9
20	.33	.08	.93	-1.0	.94	-.8
21	.02	.08	.98	-.2	.95	-.7
22	-.05	.08	1.53	6.8	1.63	7.1

表4是經評定量尺模式處理後的能力參數值（ability parameter）、訊息量與原始分數的對照情形。由此表可知，在原始分數37時，其最高的訊息量為12.32，其估計標準誤最小，為0.28；因此，研究者決定以此原始分數37分作為決斷值，凡得分在37分以上（含37分）者，即可判定個體是個「潛在的憂鬱症患者」，個體應該進一步尋求精神科醫生的專業診斷，以確認是否確實罹患憂鬱症。

表4 本研究自編量表的原始分數、能力參數值與訊息量對照表

原始分數	能力參數值	標準誤	訊息量	次數
0	-6.23	1.84	.30	4
1	-5.00	1.02	.96	1
2	-4.26	.74	1.83	1
3	-3.81	.62	2.63	6
4	-3.47	.55	3.36	11
5	-3.20	.50	4.04	6
6	-2.97	.46	4.66	10
7	-2.77	.44	5.24	9
8	-2.59	.42	5.78	14
9	-2.42	.40	6.27	13
10	-2.27	.39	6.74	16
11	-2.12	.37	7.17	10
12	-1.99	.36	7.57	17
13	-1.86	.35	7.95	8
14	-1.74	.35	8.30	17
15	-1.62	.34	8.63	14
16	-1.50	.33	8.95	7
17	-1.39	.33	9.24	12
18	-1.29	.32	9.52	7
19	-1.18	.32	9.79	7

表4 本研究自編量表的原始分數、能力參數值與訊息量對照表 (續)

原始分數	能力參數值	標準誤	訊息量	次數
20	-1.08	.32	10.04	5
21	-.98	.31	10.28	10
22	-.89	.31	10.50	2
23	-.79	.31	10.72	4
24	-.70	.30	10.92	4
25	-.61	.30	11.11	5
26	-.52	.30	11.29	8
27	-.43	.30	11.46	2
28	-.35	.29	11.61	5
29	-.26	.29	11.75	6
30	-.18	.29	11.88	6
31	-.09	.29	11.99	7
32	-.01	.29	12.09	4
33	.07	.29	12.17	7
34	.15	.29	12.24	2
35	.24	.29	12.28	5
36	.32	.29	12.31	6
37	.40	.28	12.32	5
38	.48	.29	12.30	6
39	.56	.29	12.27	5
40	.64	.29	12.21	13
41	.73	.29	12.13	8
42	.81	.29	12.03	9
43	.89	.29	11.90	14
44	.98	.29	11.74	6
45	1.06	.29	11.56	7

表4 本研究自編量表的原始分數、能力參數值與訊息量對照表 (續)

原始分數	能力參數值	標準誤	訊息量	次數
46	1.15	.30	11.35	5
47	1.24	.30	11.11	8
48	1.33	.30	10.84	4
49	1.42	.31	10.54	3
50	1.52	.31	10.21	1
51	1.62	.32	9.85	3
52	1.72	.33	9.46	8
53	1.83	.33	9.03	3
54	1.94	.34	8.57	1
55	2.06	.35	8.08	5
56	2.19	.36	7.54	2
57	2.33	.38	6.98	4
58	2.48	.40	6.37	3
59	2.65	.42	5.73	4
60	2.83	.45	5.05	2
61	3.04	.48	4.32	0
62	3.30	.53	3.56	0
63	3.62	.60	2.74	1
64	4.05	.73	1.88	2
65	4.78	1.02	.97	3
66	6.01	1.83	.30	0

為了檢視此決斷值的可信度，研究者將樣本依此臨界點分數重新歸類後，如表5所示。原來的憂鬱症患者有78人被重新歸類為正常人組，只有135人被歸類為「潛在的憂鬱症患者」；而原來的正常人組200人，則仍全部被歸類為正常人。因此，正確分類的百分比為81.1%（即（135+200）/ 413

=.811)；而Kappa分類一致性係數為0.63，顯著性p值亦小於0.001。另外，也可以由原來憂鬱症患者組在量表上的原始分數介於8~65的分佈（平均數為39.84），以及正常人組在量表上原始分數介於0~33的分佈（平均數為11.72），知道37分的決斷標準會使原來的正常人組全部仍被歸類為正常人組，而原來的憂鬱症患者將有部分人被歸類為正常人。

儘管依此決斷值分類的正確性，不像之前使用四個向度分數所建立的區別函數歸類來得高，但對於臨床工作者或一般社會大眾的使用而言，以此臨界點分數作為判定類別歸屬的依據，將具有更大的簡易性與方便性。所以，本研究將以此37分（含37分），作為建議使用臺灣憂鬱症量表來判定潛在的憂鬱症患者的決斷值標準。

此外，為方便醫護領域人員使用本量表，這裡也提供醫學上常用的專有名詞之統計量。根據表5計算，敏感性（sensitivity）是指經決斷分數判斷有病的個案佔確實有病個案的比例，或稱真陽性比率，其值為0.63；特殊性（specificity）是指經決斷分數判斷無病的個案佔確實無病個案的比例，或稱真陰性比率，其值為1.00。而陽性預測值（positive predictive value）是指依決斷分數判斷有病的個案中，確實有病個案所佔的比例，其值為1.00；陰性預測值（negative predictive value）是指依決斷分數判斷無病的個案中，確實無病個案所佔的比例，其值為0.72。ROC（Receiver Operating Characteristic）曲線以下之面積佔高達0.99的比率，達到0.001以下的顯著水準。

表5 依據臺灣憂鬱症量表之建議決斷值的分類正確性結果摘要表

		依決斷值歸類組別		總人數
		患者組	正常人組	
原來 組別	患者組	135	78	213
	正常人組	0	200	200
總人數		135	278	413

若將表3所示各試題的閾限參數值（及其測量內容）及表4所示各原始分數的能力參數值，重新排列並合併呈現的話（如表6所示），則可以進一

步幫助我們解釋本研究自編憂鬱症量表的實質內涵。由表6所示可知，如果我們採酌本研究的建議標準，當個體的原始分數在37分（或相當於能力參數值在 .40）以上，且在人際關係向度（如：19.不想與他人往來）、情緒向度（如：10.害怕、恐懼）、身體向度（如：13.胃口不好或暴飲暴食）、及認知向度（如：2.對什麼事都失去興趣、4.有罪惡感、1.有自殺的念頭）等典型試題上的作答，都出現勾選「總是如此」的嚴重症狀（依據試題閾限參數值大小來判斷憂鬱的嚴重程度）時，即很有可能表示該個體是個「潛在的憂鬱症患者」，該個體更應該進一步尋求精神科醫生的專業診斷，以確定是否確實罹患憂鬱症。

表6 本自編量表試題閾限參數值與決斷值診斷對照表

根據決斷值進行診斷歸類	建議的決斷值	量表試題	
		閾限參數值	題號及所測量內容
憂鬱患者組	.40 (含) 以上 (或原始分數37 (含) 分以上)	1.60	(1.有自殺的念頭)
		.74	(4.有罪惡感)
		.58	(13.胃口不好或暴飲暴食)
		.42	(10.害怕、恐懼)
		.40	(2.對什麼事都失去興趣)
		.40	(19.不想與他人往來)
正常人組	.39 (含) 以下 (或原始分數36 (含) 分以下)	.37	(11.想哭)
		.35	(5.感覺自己很沒用)
		.33	(20.少說話或不太愛說話)
		.17	(8.發脾氣、生氣)
		.02	(21.不想出門)
		-.04	(17.身體不舒服)
		-.05	(22.生活圈小)
		-.15	(18.記憶力不好)
		-.25	(6.無力感)
		-.28	(3.凡事往壞的方向想)
		-.32	(16.無法專心做事)
		-.46	(12.心情低落)
		-.71	(14.睡眠狀況不佳)
		-1.02	(9.擔心、煩惱)
		-1.03	(15.身體疲憊)
-1.08	(7.有壓力)		

肆、結論與建議

本研究的目的，即在編製一份自評式憂鬱症量表，以提供臨床工作者及一般社會大眾一個篩選「潛在的憂鬱症患者」的實用工具。在量表的編製過程中，本研究特別倚重精神科醫生對量表試題的內容審查，再輔以統計的實徵分析證據，來判定此量表的實用性與可行性。由於編製過程中所使用的樣本皆來自台灣地區，再加上負責審題的精神科醫生也具有多年的看診經驗，因此，本自評式憂鬱症量表應當是一份非常適合台灣民眾使用的工具。

本量表內容除了具有專家內容效度外，亦根據憂鬱症四向度，使用結構方程式模型來進行模式適配度的檢定工作，以提供基礎的信效度證據。根據資料與模式適配度檢定的結果顯示，除了少數適配度指標未達標準外，大多數的指標都呈現指出本量表具有一階的四個向度因素與二階的一個向度因素模型。此外，本研究亦提供其它統計數據來呈現額外的信效度證據，包括根據區別分析技術，研究發現，不論是使用強迫進入法或逐步分析法，四個面向的因素均在所獲得的區別函數中，以之來預測樣本的新組別，獲得93%的正確分類百分比，顯示四個向度的因素架構儘管其背後有一「憂鬱症」的二階因素，且認知、情緒、身體、人際關係等四個向度之間，仍具有一定的區別性。最後，本研究自編量表與國際知名的CES-D憂鬱症量表的效標關聯效度為0.92，亦顯示本量表所依據的憂鬱症四向度架構，確實能掌握住憂鬱症的概念精髓。

本研究為方便一般臨床工作者或一般社會大眾在使用本量表時，能夠快速篩選出需要進行後續醫療流程的「潛在的憂鬱症患者」，特根據試題反應理論中的最大訊息量方法，找出最適切作為決斷值分數的界定標準。根據此最大訊息量在其所對應的能力量尺分數位置，即最能有效地區別出憂鬱程度的差異來，因此，最大訊息量所對應的能力量尺分數（即本量表中的原始分數37分），適合作為本研究量表在篩選憂鬱症傾向時的決斷值。而此決斷值之界定，在樣本具有代表性的條件下，預期將會十分的穩定。

本量表的編製過程不免有些不足及限制之處，未來可以繼續努力研究

的方向，包括：使用更多地區的樣本來建立交叉驗證效度、使用配對方式尋找更具代表性的正常人作為對照、服藥前與服藥後憂鬱症患者在症狀上的輕重對量表信效度的影響、有利憂鬱症篩選工作之決斷值的界定在不同樣本下的穩定性檢定等。詳細言之，儘管本研究認為一般大型醫院的憂鬱症患者已具有一定的樣本代表性，但為了使本量表的應用面更廣泛，仍有必要繼續提供更多的信效度證據，當然在時間及經費許可下，實施更大規模的控制研究，可以減少混淆變項對量表信效度的影響。此外，憂鬱症傾向的嚴重程度，一般都會隨著用藥情形而有所減緩或降低，如果時間與經費允許的話，未來也可以控制此一混淆變項對量表編製的影響。最後，使用最大訊息量法作為判定「潛在的憂鬱症患者」是一項創新，過去有許多類似量表的決斷值都是由人為的主觀判斷來決定，但本研究率先應用試題反應理論所提供的最大訊息量概念，因為其具有客觀的數理基礎與統計數據，因此會比主觀的決定來得有說服力。未來，仍可在其它代表性樣本的使用條件下，繼續檢視本研究所獲得的37分決斷值之跨樣本的穩定性。

參考文獻

- 李昱、楊明仁、賴德仁、邱念睦、周騰達（2000）。台灣人憂鬱問卷之發展。*長庚醫學*，23（11），688-694。
- 楊中芳總校訂譯（1997）。*性格與社會心理測量總覽（上）*（John P. Robinson原著）。台北：遠流出版。
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text revision). Washington, DC: American Psychiatric Association.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Academic of Marketing Science*, 16, 76-94.
- Barrett, M. S., & Barber, J. P. (2007). Interpersonal profiles in major depressive disorder. *Journal of Clinical Psychology*, 63(3), 247-266.
- Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory manual* (2nd ed.). San Antonio, TX: The Psychological Corporation.

- Beck, R., Perkins, T. S., Holder, R., Robbins, M., Gray, M., & Allison, S. H. (2001). The cognitive and emotional phenomenology of depression and anxiety: Are worry and hopelessness the cognitive correlates of NA and PA. *Cognitive Therapy and Research, 25*(6), 829-838.
- Beevers, C. G., & Miller, I. W. (2004). Depression-related negative cognition: Mood-state and trait dependent properties. *Cognitive Therapy and Research, 28*(3), 293-307.
- Blatt, S. J., & Zuroff, D. C. (1992). Interpersonal relatedness and self-definition: Two prototypes for depression. *Clinical Psychology Review, 12*, 527-562.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Brendgen, M., Vitaro, F., Turgeon, L., & Poulin, F. (2002). Assessing aggressive and depressed children's social relations with classmates and friends: A matter of perspective. *Journal of Abnormal Child Psychology, 30*(6), 609-624.
- Brosse, A. L., Sheets, E. S., Lett, H. S., & Blumenthal, J. A. (2002). Exercise and the treatment of clinical depression in adults. *Sports Medicine, 32*(12), 741-760.
- Carey, T. C., Carey, M. P., & Kelley, M. L. (1997). Differential emotions theory: Relative contribution of emotion, cognition, and behavior to the prediction of depressive symptomatology in non-referred adolescents. *Journal of Clinical Psychology, 53*(1), 25-34.
- Christopher, G., & MacDonald, J. (2005). The impact of clinical depression on working memory. *Cognitive Neuropsychiatry, 10*(5), 379-399.
- Deldin, P. J., Keller, J., Gergen, J. A., & Miller, G. A. (2001). Cognitive bias and emotion in neuropsychological models of depression. *Cognition and Emotion, 15*(6), 787-802.
- Gallassi, R., Sarro, R. D., Morreale, A., & Amore, M. (2006). Memory impairment in patients with late-onset major depression: The effect of antidepressant therapy. *Journal of Affective Disorders, 91*, 243-250.
- Greenberg, M. S., & Beck, A. T. (1989). Depression versus anxiety: A test of the content-specificity hypothesis. *Journal of Abnormal Psychology, 98*(1), 9-13.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Pearson Education.
- Hammen, C. (2006). Stress generation in depression: Reflections on origins, research, and future directions. *Journal of Clinical Psychology, 62*(9), 1065-1082.
- International Association for Suicide Prevention [IASP] (2006). *Information Flyer of 2006 World Suicide Prevention Day: With understanding, new hope*. Available: http://www.med.uio.no/iasp/english/wspd/2006/wspd_menu2006.html. Accessed: Dec 12, 2007.
- Jantz, G. L., & McMurray, A. (2003). *Moving beyond depression: A whole-person approach to healing*. Colorado: Waterbrook Press.
- Leary, M. R. (1995). The role of low self-esteem in emotional and behavioral problems:

- Why is low self-esteem dysfunctional? *Journal of Social Clinical Psychology*, 14(3), 297-314.
- Linacre, J. M. (2006). *A user's guide to FACETS*. Available: <http://www.winsteps.com/facets.htm>. Accessed: Dec 12, 2007.
- Linacre, J. M., & Wright, B. D. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- MacLeod, A. K., & Salaminiou, E. (2001). Reduced positive future-thinking in depression: Cognitive and affective factors. *Cognition and Emotion*, 15(1), 99-107.
- Mandelli, L., Serretti, A., Colombo, C., Florita, M., Santoro, A., Rossini, D., Zanardi, R., & Smeraldi, E. (2006). Improvement of cognitive functioning in mood disorder patients with depressive symptomatic recovery during treatment: An exploratory analysis. *Psychiatry and Clinical Neurosciences*, 60, 598-604.
- Mcdermut, W., Zimmerman, M., & Chelminski, I. (2003). The construct validity of depressive personality disorder. *Journal of Abnormal Psychology*, 112, 49-60.
- Muris, P., Schmidt, H., Lambrichs, R., & Meesters, C. (2001). Protective and vulnerability factors of depression in normal adolescents. *Behavior Research Therapy*, 39(5), 555-565.
- National Institute of Mental Health [NIMH] (2001). *The numbers count: Mental disorders in America 2001*. Available: <http://www.nimh.gov/publicat/numbers.cfm>. Accessed: Dec 12, 2007.
- Oei, T. P. S., Bullbeck, K., & Campbell, J. M. (2006). Cognitive change process during group cognitive behavior therapy for depression. *Journal of Affective Disorders*, 92, 231-241.
- Panzarella, C., Alloy, L. B., & Whitehouse, W. G. (2006). Expanded hopelessness theory of depression: On the mechanisms by which social support protects against depression. *Cognitive Therapy and Research*, 30(3), 307-333.
- Puschner, B., Kraft, S., & Bauer, S. (2004). Interpersonal problems and outcome in outpatient psychotherapy: Findings from a long-term longitudinal study in Germany. *Journal of Personality Assessment*, 83(3), 223-234.
- 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.
- Roberts, J. E., & Gotlib, I. H. (1997). Temporal variability in global self-esteem and specific self-evaluation as prospective predictors of emotional distress: Specificity in predictors and outcome. *Journal of Abnormal Psychology*, 106(4), 521-529.
- Rockhill, C. M., Fan, M.-Y., Katon, W. J., McCauley, E., Crick, N. R., & Pleck, J. H. (2007). Friendship interactions in children with and without depressive symptoms: Observation of emotion during game-playing interactions and post-game evaluations. *Journal of Abnormal Child Psychology*, 35, 429-441.

- Tillema, J. L., Cervone, D., & Scott, W. D. (2001). Negative mood, perceived self-efficacy, and personal standards in dysphobia: The effects of contextual cues on self-defeating patterns of cognition. *Cognitive Therapy and Research, 25*(5), 535-549.
- World Health Organization [WHO] (2001). *The world health report: Chapter 2: Burden of mental and behavioral disorders*. Available: <http://www.who.int/whr/2001/chapter2/en/index4.html>. Accessed: Dec 12, 2007.
- Zarrinpar, A., Deldin, P., & Kosslyn, S. M. (2006). Effects of depression on sensory/motor vs. central processing in visual mental imagery. *Cognition and Emotion, 20*(6), 737-758.

銘謝：本文作者非常感激台北市立萬芳醫院沈武典與盧孟良、私立長青醫院吳佳憬等三位精神科醫師，針對本量表試題內容，提供逐題審查與臨床判斷的專家效度證據。本文初稿曾於今（2008）年1月22-24日在澳洲伯斯市的西澳洲大學舉行之「第三屆Rasch 測量模式的發展國際研討會」上口頭發表，作者擬感謝與會學者專家提供寶貴的修改意見。

