

「成人批判思考技巧測驗」之發展

葉玉珠
國立中山大學

陳月梅
國立中山大學
教育碩士

謝佳蓁
國立中山大學
教育碩士

葉碧玲
國立中山大學
教育碩士

摘要

本研究的主要目的在發展一份適用於成人（含大學生）的本土化批判思考能力測驗——成人批判思考技巧測驗(The Test of Critical-thinking Skills for Adults, TCTS-A)，並建立常模。預試的對象包含大學生 88 人，研究生 99 人，共 187 人。經預試及項目分析後，TCTS-A 共包含五個分測驗——辨認假設、推論、演繹、解釋及評鑑，每個分測驗包含五題，共計二十五題；填答時間為 25 分鐘。測驗中各題之鑑別指數為.21 ~ .53，平均值為.35；難度指數為.24 ~ .84，平均值為.58。信度及效度分析的結果顯示：TCTS-A 的試題難易適中，並具有良好的信度與效度。本研究除了發展「成人批判思考技巧測驗」之外，也進一步以大學生 191 人、研究生 135 人，共計 326 人為研究對象，考驗性別及學歷的差異。研究結果發現：大學部男女學生的批判思考能力無顯著差異，而研究所男女學生在演繹及評鑑兩個分測驗上有顯著差異；在學歷的效果方面，則發現大學生與研究生的批判思考能力無顯著差異。

關鍵字：大學生、批判思考技巧、研究生、性別、常模

批判思考為一複雜的認知歷程；此一思考歷程涉及思考者的知識、意向與技巧與所在情境的互動。一位良好的批判思考者除了必須具備足夠的知識、意向及技巧之外，尚需視問題發生的脈絡，建立一套有效及合理的判斷規準，對陳述或問題加以澄清與評估，以做成決策並解決問題(葉玉珠, 民 89)。批判思考不僅是教育的理想,也是民主社會的基石(Paul, 1990; Taube, 1997) 許多學者認為批判思考為一有效的學習策略(Halpern, 1998; Klein, Olson, & Stanovich, 1997; Lawson, 1999)。最近, Elder(1997)更指出：批判思考是增進情緒智商的關鍵。因此，批判思考無論對學生本身的學習、個人的生活品質或社會的進步均有正面的功能。

近年來，各國對於培養具有批判思考能力的 21 世紀公民及專業人員有更為重視的趨勢。如美國國會明述：促進大學生「批判思考、有效溝通、及問題解決」的能力為西元 2000 年的國家教育目標 (Facione, Sanchez, Facionem, & Gainen, 1995, p. 2); 美國八大事務所白皮書(Big Eight White Paper, 1989)及美國會計教育改革委員會(AECC, 1990)則強調會計人員必須擁有新的技術與能力，而批判思考即其中一項必備的能力。IBM 的主要執行長官(Chief

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

Executive Officer)也曾在 1996 年提出批判思考能力是現今商業領袖必備的條件之呼籲(Harris & Eleser, 1997)。有鑑於此，最近教育部呼籲各大專院校加強通識教育，並重點補助包含提升大學生批判思考等能力的通識課程。許多研究發現有效的教學能提升學生的批判思考能力(Browne & Meuti, 1999; Gadzella & Masten, 1998; Hittner, 1999; McCarthy-Tucker, 2000)。在大專院校進行批判思考教學，如何評估其教學效果，伴隨而來的問題便是如何評量大學生的批判思考能力。因此，發展適用於大學生的批判思考測驗是必要的。

國外的批判思考運動發展至今，不僅將批判思考列入教育目標中，也已編製出許多適用於大學生的測驗工具（如 Ennis, Millman, & Tomko, 1985; Norris & Ennis, 1989; Watson & Glaser, 1980）。反觀國內，除了對國外批判思考測驗的修訂以及鄭英耀、王文中、吳靜吉、黃正鵠（民 85）針對高中生以上為對象所編之「批判思考量表」之外，尚無針對國內大學生及成人所編製的本土化批判思考測驗。因此，本研究嘗試綜合學者對批判思考技巧的定義，發展一份適用於大學生及成人的批判思考測驗，希冀能為我國大學生及成人批判思考能力之發展，提供評估、教學及研究之參考。

壹、批判思考的技巧

一、批判思考的定義

綜合學者們對於批判思考的定義，大致可將其歸納為下列三種看法或取向：批判思考即心理歷程(critical thinking as processes)、批判思考即程序(critical thinking as procedures)以及批判思考即技巧(critical thinking as skills)。

持心理歷程取向者認為批判思考基本上是心理歷程的有效運用。例如 Facione 等人(1995)認為批判思考為一具有目的性及自我調節判斷的認知過程；藉由此一複雜的認知過程，我們在特定的情境下做成何者應為與何者應信之判斷。Chaffee(1990)認為批判思考是在仔細檢驗思考過程以澄清和改善我們對事物的理解，進而使這個世界所發生的現象對我們而言是有意義的。

持程序取向者認為批判思考的展現必須有一些必要的程序或步驟。例如 Daniel Eckber 認為批判思考包含下列六個步驟：定義兩難情境(define dilemma)、檢驗選擇方案(examine electives)、考慮後果(consider consequences)、調查重要性(investigate importance)、決定方向(decide direction)、評鑑最後結果(evaluate ends)(Bailin et al., 1999)。Marzano 等人於 1988 年認為批判思考包含下列八大步驟：概念的形成(concept formation)、原則的形成(principle formation)、理解(comprehension)、問題解決(problem solving)、決策(decision making)、研究(research)、統整(composition)、口頭講述(oral discourse)(Bailin, Case, Coombs, & Daniels, 1999)。

持技巧取向者認為批判思考涉及運用認知及後設認知的技巧以澄清、評鑑或解決問題。例如 Halpern(1997)認為批判思考乃使用許多認知技巧與策略以增進獲致所希望結果的機率，它是目的導向及理性導向的；此一思考涉及在特定的情境中(context)解決問題、形成推論、計算可能發生的機率、和提出決策等。Paul(1990)認為：批判思考即運用思考去分辨

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

或分析事物；我們應用語言、邏輯等技巧去解釋事象，並設定判斷的標準，據此給予適當的評價。

上述三種取向的定義各有其強調的重點，對於有效的批判思考教學均有非常重要的啟示。這三種取向的觀點與 Costa(1985)所強調之「思考教學三原則」——思考技巧的教學(teaching of thinking)、為思考而教學(teaching for thinking)以及關於思考的教學(teaching about thinking)——有異曲同工之妙；即技巧取向適用於強調「批判思考技巧的教學」，心理歷程取向適用於強調「為批判思考而教學」，程序取向則適用於強調「關於批判思考的教學」。

二、批判思考的技巧

目前大部分的訓練課程都以技巧的訓練為主；而一般的批判思考測驗也均以技巧的測驗為主。以下即對學者們的看法作一簡介。

Ennis 是較早從事批判思考相關研究及測驗發展的學者之一。Ennis 與 Norris 將批判思考的技巧歸類為如下四大領域的技巧(Norris & Ennis, 1989)：

1. 澄清的技巧 (clarification)：包括能注意問題的重點所在、能對具有爭議性的問題加以分析、能提出有助於澄清或改變問題的疑問或解答、能從形式、定義和策略三個層面對一名詞下適當的定義、能辨認假設。
2. 建立完善推論基礎的技巧 (basic support)：包括能判斷訊息的可信度、能觀察或判斷他人的觀察報告的可信程度。
3. 推論的技巧 (inference)：包括能演繹、能歸納、能做價值判斷。
4. 運用策略的技巧：包括能決定採取行動的適當時機、能與他人產生互動（即和他人討論、議論及表達自己看法）。

Michelli Pines 與 Oxman-Michelli(1990)以及 Norris 與 Ennis(1989)認為批判思考涉及許多認知及後設認知的技巧，諸如分析、詮釋、推論、解釋、評鑑、有創意地運用訊息、做結論、與他人互動及自我調節。在 1990 年，由跨學科學者所完成為時二年的 Delphi project 中，學者們達成一個共識：即批判思考者必須具有下列認知技巧：分析、詮釋、推論、解釋、評鑑與自我調節 (Facione et al., 1995)。Chaffee(1990)認為批判思考者應具備下列能力：能仔細周全地探索情境、能自我省思、能從不同觀點分析情境、能主動思考、能有組織地與他人討論自己的觀點。Dick(1991)則認為批判思考者應具備 15 種能力，他並把這些能力分為五大類：辨認論點(identifying arguments)、分析論點(analyzing arguments)、外在資源(external resources)、科學分析推理(scientific analytical reasoning)、推理與邏輯(reasoning and logic)。

Halpern(1997)認為批判思考者應具備下列能力：

1. 在解決問題時能確認、控制、或評估多重影響因素的重要性；
2. 當無法瞭解教材時，能區辨究竟是不能定義一個名詞或是沒有足夠的訊息；
3. 當一個概念未被明確地定義時，能瞭解提出明確定義的重要性並能產生新的定義；
4. 能思考所提出的支持理由是否能有效支持所做成的結論；

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

5. 能從實徵資料和證據(包括相關的推理)中做成推論；
6. 能使用理性的判斷標準權衡可能解決方案的利弊，並作成決定；
7. 當解決問題時能清楚覺知有助於問題解決的策略並能有系統地加以應用；
8. 在閱讀時遇到困難的內容，能監控自我的理解過程與方法；
9. 當解決問題時能預期困難所在；
10. 當說和寫時能清楚且有效地溝通；
11. 能辨識設計來引發某種情緒反應的宣傳技巧。

國內學者張玉成(民81)認為批判思考者具有七類(廿二項)行為特質：把握重點、條理貫達、態度嚴謹、事證舉例、正確推理、掌握變項以及價值判斷。陳膺宇(民83)則認為批判思考技巧可以下列五類的能力予以涵蓋：釐清問題的能力、鑑定資訊的能力、正確推論的能力、審辨假設的能力以及合理行動的能力。

目前較被廣為使用的批判思考測驗應為 Ennis、Millman 與 Tomko(1985)所編製的康乃爾批判思考測驗 X 級與 Z 級(Cornell Critical Thinking Test, Level X 與 Cornell Critical Thinking Test, Level Z)以及 Watson 與 Glaser(1980)的華格批判思考測驗(Watson-Glaser Critical Thinking Appraisal)。「康乃爾批判思考測驗 X 級」所測量的技巧包括歸納、觀察與決定可信度、演繹及辨認假設四種技巧；「康乃爾批判思考測驗 Z 級」所測量的技巧包括歸納、決定可信度、預測及實驗計畫、演繹、定義及辨認假設七種技巧；「華格批判思考測驗」所測量的技巧則包括歸納、演繹、辨認假設、判斷結論及評鑑五種技巧。Ennis 等人(1985)並認為批判思考能力為許多複雜認知技巧的整體表現，在測驗分數的解釋上應以總分為主要考量。

參酌上述學者們對批判思考技巧的看法及過去批判思考測驗的內涵(Ennis et al., 1985; Watson & Glaser, 1980; 毛連塏、陳麗華、劉燦樑, 民80; 吳靜吉、鄭英耀、王文中, 民81; 葉玉珠、吳靜吉, 民81), 本研究認為一份適用於大學生及成人的批判思考測驗應包含辨認假設(recognition of assumptions)、推論(inferences)、演繹(deductions)、解釋(interpretations)及評鑑(evaluation of arguments)五種技巧, 研究者因而以此為架構, 編製批判思考技巧的測驗題目。茲將這五種技巧的意義說明如下:

1. 辨認假設: 是指能夠辨認出陳述或宣稱中所被視為理所當然的隱含前提或預設立場; 亦即當我們在敘述一件事情或表達一個看法的時候, 都會先有一些自己認為「理所當然」的看法, 辨認假設就是指能夠辨認這些「理所當然」看法的能力。
2. 推論: 是指能夠由已知資料中推論出最有可能的結果; 亦即把每個題目中的敘述當作是真的, 並且決定由這個敘述最有可能導致的結論是什麼的能力。
3. 演繹: 是指能夠從已知的陳述或前提中, 辨認論述之間所隱含的關係, 並從已知前提中, 確定該推論是否確實為隱含的或必然的結果; 亦即將兩個敘述當作是真的, 且決定由這兩個敘述一定會導致什麼結果的能力。
4. 解釋: 是指能夠自陳述中, 尋找證據並評估推論的可能性; 亦即把每個題目中的敘述當作是真的, 然後根據這個敘述去判斷最合理而最有可能成立的論述之能力。

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

5. 評鑑：是指能夠評估問題中論證的支持程度；亦即當我們在面臨重要的問題時，通常會提出一些論點來支持我們的看法。有些論點比較強，有些比較弱。所謂「強」的論證，必須是直接與問題有關而且是重要的；所謂「弱」的論證，是與問題沒有直接關係或是不重要的論點，評鑑即能夠判斷每個論點，對於所提出問題的支持程度是「強」或「弱」。

貳、研究方法

一、研究對象

本研究分兩階段進行：預試及正式施測。預試的目的在使所編測驗具有良好的信度與效度。正式施測的目的在於建立常模並進行性別與學歷的差異分析。預試的對象包含兩所國立大學及一所國立科技大學的大學生 88 人，研究生 99 人；其中男生共 102 人，女生共 85 人，有效樣本共為 187 人。正式施測的對象為兩所國立大學及一所國立科技大學的學生，共計有大學生 191 人，研究生 135 人；其中男生有 160 人、女生有 166 人、共計 326 人。

二、研究工具

本研究在預試階段所採用的工具為自編的基本資料題目(包含學校名稱、學號、性別、年齡、系所、學歷、父母的教育程度)、初編的「成人批判思考技巧測驗(TCTS-A)」。在正式施測階段，本研究所使用的工具為修訂後之 TCTS-A、「中小學批判思考技巧測驗(TCTS-PS)」及「語文智慧量表」。

(一) 成人批判思考技巧測驗

初編之 TCTS-A 共分為五個部分，第一部分為辨認假設，第二部分為推論，第三部分為演繹，第四部分為解釋，第五部分為評鑑。初編的 TCTS-A 每一部分有 7 小題，共 35 題，答題時間為 30 分鐘。在計分上，每題在答對得 1 分、答錯得 0 分、答錯不倒扣的情況下，本測驗的總分為 35 分。總分愈高，表示批判思考能力愈強。預試後之 TCTS-A 仍包含五個部分，但每一部分各有 5 題，共 25 題。因此，TCTS-A 的總分為 25 分，答題時間為 25 分鐘。

在試題型態方面，TCTS-A 每個部分在測驗題目之前均有一段描述該批判思考技巧的說明及作答方式，之後並有一舉例及說明。以下即就五個分測驗，分別各舉第一題說明其題型：

1.辨認假設：

「選擇住在同一縣市的對象結婚，婚後不但可以就近照顧雙方父母，而且每當有緊急狀況發生時都可以立即處理。」

假設一：雙方及其父母均住在同一縣市。

假設二：至少有一方有照顧雙方父母的意願。

(甲) 假設一必然存在。

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

- (乙) 假設二必然存在。
- (丙) 假設一和假設二都必然存在。
- (丁) 假設一和假設二都不必然存在。

2.推論：

政府為了落實環保及資源回收，開始強制實施垃圾分類及清潔費隨袋徵收，許多民眾因此抱怨連連。

推論一：短期之內民眾不會改變態度，配合政府措施。

推論二：抱怨的民眾不會按照規定將垃圾分類。

- (甲) 推論一必然成立。
- (乙) 推論二必然成立。
- (丙) 推論一和推論二都必然成立。
- (丁) 推論一和推論二都不必然成立。

3.演繹：

很多人喜歡熱氣球；熱氣球可以做為旅行的工具。

結論一：很多人喜歡坐熱氣球旅行。

結論二：坐熱氣球旅行的人都喜歡刺激。

- (甲) 結論一必然成立。
- (乙) 結論二必然成立。
- (丙) 結論一和結論二都必然成立。
- (丁) 結論一和結論二都不必然成立。

4.解釋：

醫學報導指出：在懷孕期間吸煙的孕婦比不吸煙者，容易流產或產下畸形兒；報導中也指出吸二手煙的孕婦流產或產下畸形兒的機率比自己抽煙的孕婦高。

解釋一：孕婦吸二手煙對嬰兒的危害程度更甚於自己抽煙。

解釋二：部份流產或產下畸形兒的孕婦，在懷孕期間有抽煙的習慣。

- (甲) 解釋一必然成立。
- (乙) 解釋二必然成立。
- (丙) 解釋一和解釋二都必然成立。
- (丁) 解釋一和解釋二都不必然成立。

5.評鑑：

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

e 時代的企業經營講求團隊合作與創新，但在組織團隊中，人與人之間的衝突在所難免。當我們的看法與他人發生衝突時，應不應該勇於立即表達自己的意見？

論點一：應該。因為團隊中個人意見的充分表達與良好溝通是促成企業改革與創新的最重要因素。

論點二：不應該。因為當發生衝突時，雙方較易堅持己見且缺乏理性思考，應該等彼此心情平緩之後，再做適當的表達與溝通，才能真正達到合作與創新的目的。

(甲) 論點一是強的論證。

(乙) 論點二是強的論證。

(丙) 論點一和論點二均是強的論證。

(丁) 論點一和論點二均是弱的論證。

(二) 中小學批判思考技巧測驗

中小學生批判思考技巧測驗(TCTS-PS)為葉玉珠、葉碧玲、謝佳蓁(民 89)所編製。TCTS-PS 包含五個分測驗：辨認假設、推論、演繹、解釋及評鑑，共計 24 題。測驗中各題之難度指數為.37 ~ .87，平均值為.66；鑑別指數為.25 ~ .71，平均值為.46。總量表係數為.80。TCTS-PS 總分與分測驗分數之間均有高度的相關，其相關係數為.63 ~ .79, $ps < .001$ ；各分測驗之間有中度的相關，其相關係數為.31 ~ .54, $ps < .001$ 。在計分上，每題答對為 1 分、答錯為 0 分，在答錯不倒扣的情況下，本測驗的總分為 24 分。

(三) 語文智慧量表

本研究所使用的「語文智慧量表」是採自 Armstrong 所著的《經營多元智慧》一書中所附的成人多元智慧核對表中的語文部分(李平譯，民 86)。此量表共 10 題，每一題的計分為：勾選得 1 分，未勾選則為 0 分。

三、研究過程

(一)編擬「成人批判思考技巧測驗(TCTS-A)」的試題

經文獻探討後，研究者開始著手收集適合成人的議題，並於八十九年九月至十一月間，每週約花四小時，由四位研究者共同參與試題的討論與編擬。針對所欲測量之五種批判思考技巧(辨認假設、推論、演繹、解釋及評鑑)，研究者首先決定試題題型，並以適用對象(成人)的生活經驗為主要考量。主題的選擇以配合時事及受試者較熟悉的題材為篩選的標準；題目的陳述則以配合成人的語文及理解能力為原則編擬。完成編擬試題後，研究者請 3 位就讀於中山大學教育研究所的學生進行測驗的試做與提出改進意見。經修改後，進而編製成正式的預試量表。在預試問卷中，除了問候語和指導語外，分成二大部分。第一部分為填答者的基本資料，第二部分為選擇題，共分成「辨認假設」、「推論」、「演繹」、「解釋」、「評鑑」5 個部分，每個部分有 7 題，總共有 35 題，每題有四個選項，有一個正確的標準答案。

(二)預試

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

於八十九年十二月間，研究者以 187 人為對象進行預試。在施測 TCTS-A 時，研究者先施以指導語說明，再請學生開始作答，作答時間為 30 分鐘。

(三)預試結果的項目分析及刪題

預試資料收回後，研究者隨即進行測驗题目的難度及鑑別度分析，並根據項目分析的結果，將不良的题目予以刪除。

(四)刪題後之試題分析

為確認刪題後的 TCTS-A 具有良好的信度與效度，本研究再一次進行項目分析，並進行分測驗之間的相關分析以及 TCTS-A 與效標變項的相關分析。

(五)正式施測

確定 TCTS-A 為一份具有良好的信度與效度的測驗後，研究者於九十年二月至三月間，以 326 人為樣本進行正式施測。

(六)進行性別與學歷的差異分析

資料回收後，研究者隨即以大學生及碩士班以上學生為樣本，分析性別及學歷與批判思考能力的關係。

四、資料分析

在預試階段，本研究採用的資料分析方法包括項目分析、內部一致性信度分析、皮爾森積差相關分析。在正式施測階段，本研究採用的資料分析方法包括描述統計及多變量及單變量變異數分析。

參、研究結果

一、預試

(一)項目分析與刪題

經以 ITEMAN 3.0 進行項目分析後，TCTS-A 的鑑別指數為-.02 ~ .47，平均值為.25；難度指數(答對百分比)為.11 ~ .87，平均值為.55。

Ebel 認為鑑別指數(D)為.19 以下之題目應淘汰，.20 至.29 之間則尚可接受，但須修改；而難度指數最佳為.40 至.80 之間(郭生玉，民 85)。本預試量表經項目分析後，以鑑別指數為主要的篩選標準，鑑別指數小於.20 者，即予以刪除。根據此一標準，研究者先將「辨認假設」中的第 3 題(D = .14)、第 4 題(D = .03)；「推論」中的第 11 題(D = .10)、第 12 題(D = .09)；「演繹」中的第 15(D = .23)題；「解釋」中的第 26 題(D = -.02)；「評鑑」中的第 29 題(D = .07)、第 33 題(D = .08)刪除。在「演繹」及「解釋」部分，雖然每題的鑑別指數均達.20 以上，但由於「辨認假設」、「推論」及「評鑑」在刪題後均為 5 題，為使試題平均分佈，故將「演繹」中鑑別指數最低的第 18 題(D = .24)及「解釋」中鑑別指數最低的第 28 題(D = .23)予以刪除。總計將第 3、4、11、12、15、18、26、28、29、33 題刪除，共刪除

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊, 48(2), 35-50。

10 題。

(二)刪題後之試題分析

1. 試題分佈：

刪題後之 TCTS-A 共 25 題，刪題後之試題分佈與題號如表 1。刪題後的計分方式與預試時相同，但總分為 25 分。

表 1 刪題後之 TCTS-A 的試題分佈

量表名稱	辨認假設	推論	演繹	解釋	評鑑	總量表
部分	—	二	三	四	五	共五部分
題號	1-5	6-10	11-15	16-20	21-25	共 25 題

2. 難度及鑑別度分析：

表 2 刪題後之 TCTS-A 的鑑別指數與難度指數(N = 187)

分測驗	題號	難度指數	低分組	高分組	鑑別指數	t 值
			答對百分比	答對百分比		
辨認假設	1	.39	.24	.53	.30	3.39***
	2	.24	.10	.37	.27	3.65***
	3	.84	.61	.95	.34	4.63***
	4	.47	.37	.68	.30	3.37***
	5	.42	.35	.56	.21	2.27*
推論	6	.67	.37	.89	.51	6.45***
	7	.46	.29	.63	.33	3.73***
	8	.83	.61	.94	.33	4.32***
	9	.84	.67	.98	.32	4.63***
	10	.65	.37	.90	.53	6.79***
演繹	11	.55	.33	.76	.42	4.92***
	12	.63	.39	.81	.41	4.84***
	13	.72	.45	.85	.40	4.83***
	14	.68	.43	.84	.41	4.83***
	15	.57	.29	.71	.42	4.79***
解釋	16	.42	.25	.61	.36	4.08***
	17	.50	.25	.74	.49	5.85***
	18	.48	.33	.65	.31	3.44***
	19	.33	.16	.48	.33	3.98***
	20	.84	.67	.92	.25	3.36***
評鑑	21	.39	.24	.55	.31	3.58***
	22	.69	.49	.84	.35	4.10***
	23	.80	.63	.90	.28	3.53***
	24	.71	.53	.84	.31	3.65***
	25	.31	.22	.44	.22	2.55*

* p<.05。*** p<.001。

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

刪題後的 TCTS-A 之鑑別指數為 .21 ~ .53，平均值為 .35；難度指數（答對百分比）為 .24 ~ .84，平均值為 .58；高、低分組在每一題得分的 t 考驗也均達顯著水準(詳見表 2)。可見，TCTS-A 的題目難度適中且具有不錯的鑑別力。

3. 信度分析：

本研究的內部一致性係數採 Iteman 3.0 分析所得之 係數。此 係數乃 Crocker 和 Algina 在 1984 年所提出，專門用於評分值為 1 和 0 的測驗之內部一致性分析。此 數值通常被認為是測驗中信度的下限 (lower-bound) (User's manual for the ITEMAN conventional item analysis program, 1989-1996, pp.3-17)。

刪題後之 TCTS-A 的 係數為 .67，較預試時 (.61) 為高。除了分測驗「辨認假設」之外，其他各分測驗的 係數均較預試時為高。此外，TCTS-A 總分與分測驗分數之間均有中度的相關，其相關係數為 .59 ~ .69， $ps < .01$ ；各分測驗之間亦有顯著相關，其相關係數為 .17 ~ .36， $ps < .05$ (見表 3)。

表 3 刪題後之 TCTS-A 分測驗與總分之間的相關 (N = 187)

量表	辨認假設	推論	演繹	解釋	評鑑	總分
辨認假設	1.00					
推論	.31**	1.00				
演繹	.25**	.28**	1.00			
解釋	.17*	.26**	.36**	1.00		
評鑑	.21**	.35**	.17*	.20**	1.00	
總分	.59**	.69**	.68**	.63**	.58**	1.00

* $p < .05$ 。 ** $p < .01$ 。

4. 效標關聯效度分析：

根據 Ennis 等(1985)建議，批判思考能力為多種認知技巧的綜合運用，應以測驗總分為詮釋的依據；因此以下的效標關聯效度分析均以 TCTS-A 的測驗總分進行。本研究以中小學批判思考技巧測驗(TCTS-PS)及語文智慧量表為效標，探討 TCTS-A 的效標關聯效度。刪題後 TCTS-A 與效標變項的相關如表 4。由表 4 得知：TCTS-A 與 TCTS-PS 及語文智慧均有正相關， $rs(108) = .47$ 及 $.20$ ， $ps < .01$ 。

從 Gardner 對七種智能的定義中可知：語文智慧高者能敏銳地辨識語意及語言的功能(李平譯，民 86)；批判思考通常是在特定的領域或情境之下，為澄清概念或解決問題所產生的，因此語文智慧對於批判思考技巧的有效應用而言是必要的。本研究發現 TCTS-A 與語文智慧有正相關；此結果 Gardner 的智能理論相符。而 TCTS-A 與 TCTS-PS 的顯著正向關係，更顯示了 TCTS-A 具有良好的效度。

表 4 刪題後之「成人批判思考技巧測驗」與效標變項的相關

變 項	樣本來源	N	TCTS-A
TCTS-PS	教育學程學生	110	.47**
語文智慧	教育學程學生	110	.20*

* $p < .05$ 。 ** $p < .01$ 。

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

綜合上述的項目分析、內部一致性分析、分測驗之間的相關分析、分測驗與總測驗之間的相關分析以及效標關聯效度分析的結果，成人批判思考測驗為一難度適中且具有不錯的信度與效度。

二、性別及學歷在 TCTS-A 上的差異分析

本研究以大學生 191 人、研究生 135 人，共 326 人為研究對象，進行性別及學歷在 TCTS-A 上的差異分析。本研究認為批判思考能力為辨認假設、推論、演繹、解釋及評鑑五種能力的綜合運用，在比較批判思考能力的表現時應以學生在 TCTS-A 的整體表現為詮釋的依據，但在教學時也應瞭解學生在 TCTS-A 中五種批判思考技巧的表現上是否有差異，以便在教學時能兼顧個別差異，針對學生較弱的批判思考能力予以加強。因此本研究以多變量變異數分析(MANOVA)考驗不同性別及學歷的學生在 TCTS-A 五個分測驗的表現差異情形。大學生、研究生及全體學生在 TCTS-A 得分的平均數及標準差見表 5。男生、女生及全體學生在 TCTS-A 五個分測驗的得分情形如圖 1。

(一) 得分概況分析

表 5 大學生、研究生及全體在 TCTS-A 得分之平均數與標準差

	男生 ($n = 86$)		女生 ($n = 105$)		全體 ($N = 191$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
大學生						
辨認假設	2.42	1.00	2.56	1.18	2.50	1.10
推論	3.57	0.94	3.64	1.01	3.61	0.98
演繹	3.37	1.26	3.10	1.12	3.22	1.19
解釋	2.24	1.31	2.08	1.22	2.15	1.26
評鑑	2.84	1.10	2.93	1.07	2.89	1.08
總量表	14.44	3.05	14.30	3.38	14.37	3.23
研究生						
辨認假設	2.22	1.06	2.23	1.04	2.22	1.05
推論	3.38	1.22	3.38	1.21	3.38	1.21
演繹	3.32	1.34	2.84	1.39	3.10	1.38
解釋	2.30	1.24	2.11	1.10	2.21	1.17
評鑑	2.54	1.02	3.05	0.97	2.77	1.03
總量表	13.76	3.77	13.61	3.25	13.69	3.53
全體						
辨認假設	2.33	1.03	2.44	1.14	2.38	1.09
推論	3.48	1.08	3.54	1.09	3.51	1.09
演繹	3.35	1.29	3.00	1.23	3.17	1.27
解釋	2.27	1.27	2.09	1.17	2.18	1.22
評鑑	2.70	1.07	2.98	1.03	2.84	1.06
總量表	14.13	3.41	14.05	3.34	14.09	3.37

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

由表 5 及圖 1 得知：大學生及研究生在 TCTS-A 五個分測驗上的得分趨勢大約一致；全體學生在推論及演繹兩個分測驗的得分較高，而在辨認假設及解釋兩個分測驗的得分偏低。就總分而言，全體學生的得分偏低；在總分 25 分的測驗中，平均得分僅為 14.09 分。

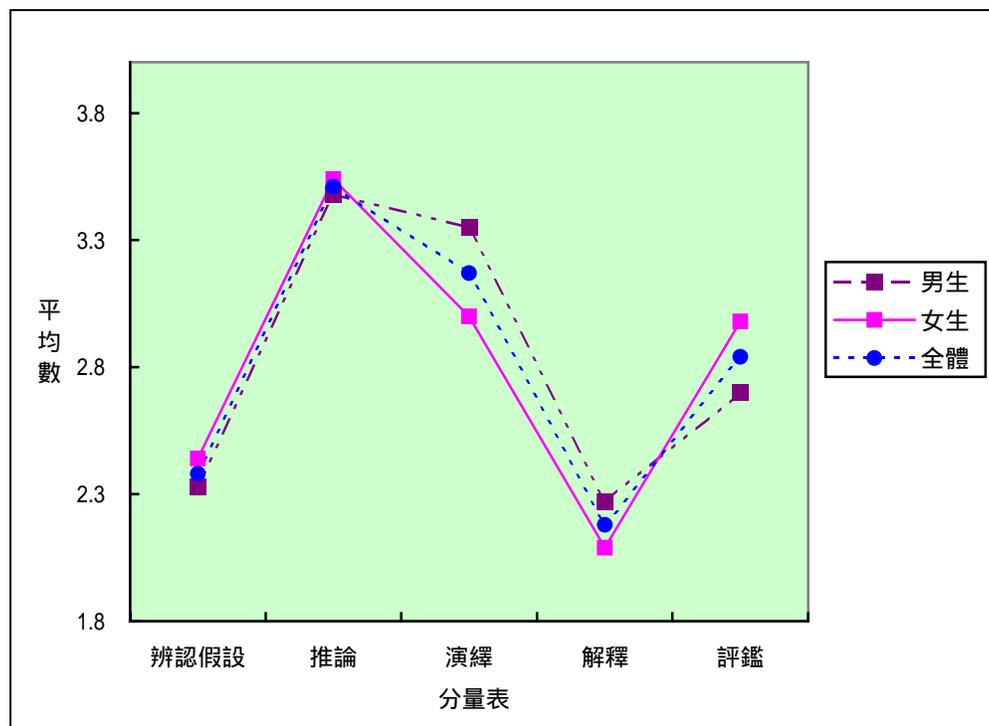


圖 1：男女及全體參與者在 TCTS-A 的得分平均數

(二) 性別在 TCTS-A 上的差異分析

在分析性別差異時，本研究分別就大學生及研究生進行分析。MANOVA 的結果顯示性別對大學生五種批判思考能力的表現沒有顯著效果 ($Wilk's \lambda = .97, ns.$) (見表 6)，即大學男女學生的批判思考能力無顯著差異。然而性別對研究生五種批判思考能力的表現則有顯著效果 ($Wilk's \lambda = .90, p < .05$)；單變量變異數分析的結果發現性別對研究生在演繹及評鑑兩個分測驗的得分上有顯著效果 ($F(1, 189) = 4.30, p < .05$ 以及 $F(1, 134) = 8.64, p < .01$) (見表 7)。由表 5 的平均數得知：研究所男生的演繹能力優於女生，而女生的評鑑能力則優於男生。

表 6 性別對大學生批判思考能力的效果之多變量變異數分析摘要表 ($N = 191$)

變異來源	df	Wilk's λ	F (1, 189)				
			辨認假設	推論	演繹	解釋	評鑑
組間	1	.97	.80	.23	2.58	.83	.37
組內	189						
總和	190						

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

表 7 性別對研究生批判思考能力的效果之多變量變異數分析摘要表(N = 136)

變異來源	df	Wilk's λ	F (1, 134)				
			辨認假設	推論	演繹	解釋	評鑑
組間	1	.90*	.01	.00	4.30*	.81	8.64**
組內	134						
總和	135						

* $p < .05$ 。 ** $p < .01$ 。

(二)學歷在 TCTS-A 上的差異分析

由表 8 的 MANOVA 結果得知：學歷對學生五種批判思考能力的表現沒有顯著效果 (Wilk's $\lambda = .97$, *ns.*) (見表 8)，即整體而言，大學生與研究生的批判思考能力沒有顯著差異。但單變量變異數分析的結果發現學歷對研究生在辨認假設一分測驗的得分上有顯著效果 ($F(1, 234) = 5.15$, $p < .05$)。由表 5 的平均數得知：大學生的辨認假設能力優於研究生。

表 8 學歷對批判思考能力的效果之多變量變異數分析摘要表(N = 236)

變異來源	df	Wilk's λ	F (1, 234)				
			辨認假設	推論	演繹	解釋	評鑑
組間	1	.97	5.15*	3.56	.66	.21	1.01
組內	234						
總和	235						

* $p < .05$ 。

肆、結論

本測驗綜合國內外批判思考的相關文獻，認為辨認假設、推論、演繹、解釋及評鑑為批判思考的重要技巧，並以此五種技巧為中心編製成「成人批判思考技巧測驗」，以測量大學生及研究生的批判思考能力。本研究為更深入了解所編測驗的信度，同時分析了「成人批判思考技巧測驗」分測驗與總測驗的信度，但在批判思考能力表現的詮釋上，研究者仍建議以總量表的分數為主；Ennis 等人(1985)對於其所編製的批判思考測驗也有此建議。唯在教學上，為針對學生較弱的批判思考技巧，施予有效的教學，可分析學生在「成人批判思考技巧測驗」中五種批判思考技巧的表現差異情形，以供教學時的參考。

就測驗的長度而言，本測驗具有良好的信度與效度。通常，當題目間有高一致性時，測驗題目越多，信度會越高。本研究因考慮測驗的實用性，故將題目減少為 25 題，以使學生能在 25 分鐘內完成。經項目分析及其他的信度與效度考驗後發現，「成人批判思考技巧測驗」具有良好的信度與效度。

本研究除了發展「成人批判思考技巧測驗」，也進一步以大學生及研究生 326 人為研究對象，考驗性別及學歷的差異。在得分概況方面，大學生在「成人批判思考技巧測驗」總

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊，48(2)，35-50。

測驗的平均得分為 14.37 ($SD = 3.23$)、研究生的平均得分為 13.69 ($SD = 3.53$)，均屬偏低的狀況(總分為 25 分)；此一結果顯示大學生及研究生的批判思考能力有待加強。在性別方面，本研究發現大學男女學生的批判思考能力均沒有顯著差異；此發現與過去的研究發現男女學生的批判思考能力無顯著差異的結果一致(Ennis et al., 1985；葉玉珠，民 89)。然而研究所男生的演繹能力優於女生，而女生的評鑑能力則優於男生；由於本研究研究生的樣本並不大，此一性別差異的結果有待進一步驗證。在學歷的效果方面，本研究發現大學生與研究生的整體批判思考能力並無顯著差異。礙於時間與人力，本研究未能建立一完整常模，未來的研究可嘗試以科系或學院為單位並包含大學生及研究生以外的成人，進一步建常模。

當今社會科技與資訊發展一日千里，21 世紀的公民亟需批判思考的能力，以提昇民主素養並達終身學習的教育理想。批判思考包含辨認假設、推論、演繹、解釋、評鑑等技巧的綜合應用，值此大力提倡全面提升大學通識教育的品質之時，高等教育相關人員(尤其是教師)，應重新思索如何將學生批判思考能力的提升融入各專業領域及通識教育課程之中。

參考書目

- 毛連塏、陳麗華、劉燦樑(民 80)：康乃爾批判思考測驗之修定。測驗年刊，38，109-123。
- 李平譯(民 86)：經營多元智慧。台北：遠流出版公司。
- 吳靜吉、鄭英耀、王文中(民 81)：華格批判思考量表之修訂。教育心理與研究，15，39-78。
- 張玉成(民 81)：國小語文科實施批判思考教學之實驗研究。台北師院學報，5，1-66。
- 張玉成(民 86)。資優兒童思考特質及其培育之研究。國科會專題研究計劃 (NSC86-2511-S-152-009)。
- 陳膺宇(民 83)：批判思考運動初探。國立政治大學學報，69，141-171。
- 葉玉珠(民 89)：智能與批判思考。國立中山大學社會科學季刊，2(1)，1-28。
- 葉玉珠、吳靜吉(民 81)：中小學生批判思考與學業成績之相關研究。教育心理與研究，15，79-100。
- 葉玉珠、葉碧玲、謝佳蓁(民 89)：「中小學批判思考技巧測驗」之發展。測驗年刊，47(1)，27-46。
- 鄭英耀、王文中、吳靜吉、黃正鵠(民 85)：批判思考量表之編製初步報告。中國測驗學會測驗年刊，43，213-224。
- Accounting Education Change Commission (AECC). (1990). Objective of education for accountants: Position statement number one. *Issues in Accounting Education*, 5(2), 307-312.
- Bailin, S., Case, R., Coombs, J. R., Daniels, L. B. (1999). Common misconceptions of critical

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊 , 48(2) , 35-50。

thinking. *Journal of Curriculum Studies*, 31(3), 269-283.

Big Eight White Paper (1989). *Perspectives on education: Capabilities for success in the accounting profession*. New York: Arthur Anderson & Co.

Browne, M. N., & Meuti, M. D. (1999). Teaching how to teach critical thinking. *College Student Journal*, 33(2), 162-170.

Chaffee, J. (1990). *Thinking critically* (3rd ed.). Boston: Houghton Mifflin.

Costa, A. L. (Ed.)(1985). *Developing minds: A resource book for teaching thinking*. Alexandria, VA: Association for Supervision and Curriculum Development.

Dick, R. D. (1991). An empirical taxonomy of critical thinking. *Journal of Instructional Psychology*, 18, 79-92.

Elder, L. (1997). Critical thinking: The key to emotional intelligence. *Journal of Developmental Education*, 21(1), 40-41.

Ennis, R. H., Millman, J. & Tomko, T. N. (1985). *Cornell critical thinking tests, Level X & Level Z- Manual*. Calif.: Midwest Publications.

Facione, P. A., Sanchez, C. A., Facione, N. C., & Gainen, J. (1995). The dispositions toward critical thinking. *The Journal of General Education*, 44(1), 1-25.

Gadzella, B. M., & Masten, W. G. (1998). Critical Thinking and Learning Processes for Students in Two major Fields. *Journal of Instructional Psychology*, 25(4), 256-261

Halpern, D. F. (1997). *Critical thinking across the curriculum: A brief edition of thought and knowledge*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist*, 53(4), 449-455.

Harris, J. C., & Eleser, C. (1997). Developmental critical thinking: Melding two imperatives. *Journal of Developmental Education*, 21(1), 12-19.

Hittner, James B. (1999). Fostering critical thinking in personality psychology: The trait paper assignment. *Journal of Instructional Psychology*, 26(2), 92-97.

Klein, P., Olson, D. R., & Stanovich, K. (1997). Structuring reflection: Teaching argument concepts and strategies enhances critical thinking. *Canadian Journal of School Psychology*, 13(1), 38-47.

Larson, B. E. (2000). Classroom discussion: a method of instruction and a curriculum outcome. *Teaching and Teacher Education*, 16(5-6), 661-677.

McCarthy-Tucker, S. N. (2000). Teaching style, philosophical orientation and the transmission of critical thinking skills in the U.S. public schools. *Korean Journal of Thinking & Problem Solving*, 10(1), 69-77.

Michelli, N. M., Pines, R., & Oxman-Michelli, W. (1990). *Collaboration for critical thinking in teacher education : The Montclair State College Model* (Series 3, no. 3). NJ: Institute for Critical Thinking.

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking*. CA: Midwest Publications.

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊 , 48(2) , 35-50。

O'Tuel, F. S., & Bullard, R. K. (1993). *Developing higher order thinking in the content areas K-12*. Pacific Grove, CA: Critical Thinking Press and Software.

Paul, R. W. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. Robnert Park. CA: The Center for Critical Thinking & Moral Critique, Sonoma State University.

Taube, k. T. (1997). Critical thinking ability and disposition as factors of performance on a written critical thinking test. *JGE: The Journal of General Education*, 46(2), 129-164.

Assessment Systems Corporation (1989-1996). *User's manual for the ITEMAN conventional item analysis program*. St. Paul, Minnesota: Assessment Systems Corporation

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser critical thinking appraisal: Forms A and B Manual*. U.S.: Harcourt, Brace & Jouvanovich, Inc.

葉玉珠、陳月梅、葉碧玲、謝佳蓁(2001)。測驗年刊, 48(2), 35-50。

The Development of “The Test of Critical-thinking Skills for Adults”

Yu-chu Yeh	Yueh-mei Chen	Chia-chen Hsieh	Pi-ling Yeh
National Sun Yat-Sen University	M. Ed. National Sun Yat-Sen University	M. Ed. National Sun Yat-Sen University	M. Ed. National Sun Yat-Sen University

Abstract

The purposes of this study were (a) to develop a test for measuring adults' critical-thinking ability—"The Test of Critical-thinking Skills for Adults, TCTS-A" and (b) to establish a scoring norm for TCTS-A. The pretest included 187 students (88 undergraduates and 99 graduates). After the pretest and item analyses, TCTS-A included 25 items and comprised five subtests: Recognition of Assumptions, Inferences, Deductions, Interpretations, and Evaluation of Arguments. The time limit for taking TCTS-A was 25 minutes. Test items in TCTS-A had an average difficulty index of .58 (ranged from .24 to .84) and an average discriminate index of .35 (ranged from .21 to .53). The results of reliability analyses and validity analyses revealed that TCTS-A had a median level of difficulty as well as a satisfactory reliability and validity. Moreover, this study used 326 students (191 graduates and 135 undergraduates) to establish a scoring norm for TCTS-A. Gender differences and study level differences on TCTS-A were also analyzed. No gender differences on TCTS-A scores were found among the undergraduates, while two gender differences on TCTS-A subtest scores were found among the graduates. As for study level differences, the graduates did not outperform the undergraduates on TCTS-A.

Keywords: undergraduates, critical-thinking skills, graduates, gender, norm