

教育學誌

(原初等教育學報)

第三十九期

國立臺南大學教育學系 編印

中華民國一〇七年五月

教育學誌第 39 期

(原名初等教育學報，民國 93 年改名為教育學誌)

出版者：國立臺南大學教育學系

地址：臺南市中西區樹林街二段 33 號

網址：<http://www.edu.nutn.edu.tw/>

電話：(06)2133111#610-613

編審：教育學誌編審委員會（任期自 106 年 1 月至 107 年 12 月）

主編：鄭新輝

編審委員：尹玫君 方德隆 吳裕益 李新鄉 林素微 郭丁熒 楊雅婷
詹盛如 劉世雄 鄭中平 羅希哲

執行秘書：林郁馨

封面設計：紀育廷

承印者：泰成印刷廠

出版年月：民國 107 年 5 月

創刊年月：民國 77 年 6 月

刊期頻率：半年刊

本刊同時登載於臺南大學教育學系網站，網址為 <http://www.edu.nutn.edu.tw/>

工本費：新台幣 300 元

GPN：2009304583

ISSN：2071-3126

教育學誌

(原南師初等教育學報)

第三十九期

目次

- 國小高年級學生行動學習準備度量表內涵建構之初探
..... 顏百鴻、歐陽閻.....1
- 學生對於數學教師教學看法與數學焦慮之關聯性
—以 PISA 2012 臺灣學生為例
..... 陳姿螢.....55
- 四個試題反應模式的整體能力與領域能力估計精確性之比較研究
..... 吳俊賢、涂柏原.....87
- 葉朗對《老子》美學詮釋的觀點
..... 葉若潔.....153

Contents

An Exploratory Study on the Construction of Mobile Learning Readiness for Fifth and Sixth Grade Students	Pai-Hung Yen, Yin OuYang.....2
The Relationship between Students' Viewpoints of Math Teachers and Math Anxiety–Take PISA 2012 Taiwanese Students as an Example	Tzu-Ying Chen.....56
Comparing the Accuracy of Examinees' Overall and Domain Ability Estimates of Four IRT Models	Chun-Hsien Wu, Bor-Yaun Twu.....89
Ye Lang's Perspective on The Aesthetic Interpretation of " Lao Zi "	Jo-Chieh Yeh.....154

教育學誌 第三十九期

2018年5月，頁1~53

國小高年級學生行動學習準備度量表內涵建構之初探

顏百鴻

臺南市松林國小總務主任

歐陽閻

國立臺南大學教育學系教授

摘要

本研究旨在發展國小高年級學生行動學習準備度量表，並建構行動學習準備度之內涵，採用德懷術與問卷調查法作為本研究方法，主要研究工具採自編之「國小高年級學生行動學習準備度量表」，並由4位學者、9位具有教學實務經驗專家組成德懷術小組，進行兩回合的德懷術調查研究，以確認國小高年級學童行動學習準備度各向度內涵的適切性與重要性。本研究之重要研究發現有：(1) 國小高年級學童行動學習準備度之內涵可包含使用科技能力、學習態度、溝通能力、學習信念和學習偏好等五大向度。(2) 國小高年級學童行動學習準備度量表中各向度所包含之內涵具有周延性。(3) 國小高年級學童行動學習準備度量表中各向度所包含之內涵具有重要性。

關鍵字：國小高年級學生、行動學習、準備度、德懷術

An Exploratory Study on the Construction of Mobile Learning Readiness for Fifth and Sixth Grade Students

Pai-Hung Yen

Director of General Affairs Division,
Tainan Municipal Sigang District Songlin Elementary School

Yin OuYang

Professor,
Department of Education, National University of Tainan

Abstract

The main purpose of this study was to develop the dimensions and contents of the mobile learning readiness for fifth and sixth grade students. In this study, a comprehensive review of the mobile learning readiness literature was provided, and the Delphi technique was employed to verify the appropriateness of the new contents of the mobile learning readiness for fifth and sixth grade students. This study found that the mobile learning readiness for fifth and sixth grade students should include five dimensions: ability of using technology, learning attitude toward mobile learning, communication ability, learning belief for mobile learning and learning preference for mobile

learning. Furthermore, all items in each dimension are important and comprehensive.

Key words: elementary school students, mobile learning, readiness, Delphi technique

壹、緒論

一、研究背景與動機

數位科技發展已經與各種生活層面產生密不可分的關係，網路科技的運用不斷改變人們傳統在食、衣、住、行、育、樂等各方面的生活習慣與運作模式。根據 Google 在 2013 年 8 月所發佈針對全球智慧型手機行為調查報告指出，亞太地區國家較 2012 年快速成長，其中台灣智慧型手機使用比例，從 2012 年的 32% 提高到 2013 年的 51%；Canalys(2012)的報告也指出，智慧型手機的數量在 2011 年已經超越個人電腦；臺灣自 102-105 年家用電腦持有數連三年下滑，顯示智慧型手機與平板的重要性日增，再者 12 歲以上人口中，擁有行動載具上網的比率也高達 73.1%(國家發展委員會，2016)。由智慧型手持裝置所擁有的入口比例不斷增加的趨勢來看，人們對於新科技產品的接受度逐漸提高。

而教育在此發展過程中也不斷地與數位科技進行各種面向的互動與結合，其目的在於持續增進教師教學與學生學習，讓數位科技成為數位學習輔具；再者也認為科技。陳德懷(2009)指出，智慧型手持裝置廣泛為人們在生活中使用的情況下，未來學習場域中學生每人皆能使用至少一部的智慧型手持裝置作為上課的重要學習輔具；也是有助於教與學過程中的重要裝置(Khaddage, Muller, & Flintoff, 2016)。以現在年輕的學習者為例，早已將科技視為日常生活的一部分，可謂是數位原民(digital natives)(Prensky, 2001)，當智慧型手持裝置數量不斷增加，學生或其父母至少擁有一部智慧型行動裝置，如行動電話、平板電腦等(Nedungadi & Raman, 2012)，這

些科技產品不但是溝通工具，同時也是行動學習載具，學習者能藉由手持行動裝置在具彈性與即時性的情況下，使用豐富的線上學習資源(Cheon, Lee, Crooks, & Song, 2012)，更強調教學和學習的產生過程不受時間與空間的限制(Kukulska-Hulme & Traxler, 2005)，而科技產品的易用性、有用性的提升，也更促進學習者行動學習(Hussin, Manap, Amir & Krish, 2012)。

在這新一波學習變革當中，儘管有一些研究(Corbeil & Corbeil, 2011; Mahat, Ayub, & Wong, 2012; Ng & Nicholas, 2012; Park, Nam, & Cha, 2012; Taleb & Sohrab, 2012)指出，學習者已經準備好面對行動學習；然而，有更多的研究正在探究影響教師和學習者參與行動學習的因素(如 Cheon et al., 2012)，其中一個最重要的變項即是準備度，包含學生的準備度以及學校本身必須更新操作、使用科技有關的各項準備度(Wang, Zhu, Chen, & Yan, 2009)。再者，有學者也指出美國在「學生對行動學習的準備度」這個議題尚未探討完全，值得進一步研究(Cheon et al., 2012)；此外，國內近年來，亦受到這一波行動學習的影響，由教育單位和電子科技業者聯手推動行動學習計畫，結合智慧型手持裝置和數位學習內容，提供學生不同的學習方式與學習經驗。只是，針對國內目前積極推動行動學習的輔導計畫，瞭解學生在面對行動學習過程中所需要的準備與能力，是有其必要性。然而，國內卻缺乏一套可符應現行教育部及各縣市教育局在國民小學積極辦理行動學習的研究工具，以測量行動學習準備度。

有鑑於國內外在行動學習準備度此一議題的相關研究皆有不足的情況下，本研究將從遠距學習、數位學習到行動學習之科技發展與學習的脈絡之中(Georgiev, Georgieva, & Smrikarov, 2004)，先探

討國內外數位學習準備度的相關內涵，進而再參酌行動學習的相關論述後，歸納出行動學習準備度的內涵。同時，也有研究針對學習者本身的數位學習準備度加以探討，如 Saekow 和 Samson(2011)就認為，數位學習成功實施之前，要先評估學習者應用科技的數位學習準備度；再者，目前國內小學推動行動學習的年級以國小高年級為主(劉遠禎，2016)。因此，本研究目的為發展出一套適用於評量國小高年級學生行動學習準備度之量表，以作為教育相關單位未來在發展或推動行動科技導入教學現場之參考依據。

貳、文獻探討

一、行動學習的相關概念

Traxler(2009)認為所謂的行動學習意指在任何教育的環境中，所有單獨或主流的科技皆是可攜式或是掌上型的裝置；或是如 Quinn(2000)、Keegan(2005)強調科技本身的行動性，而將行動學習定義為透過行動運算裝置(如 Palms、智慧型手機)進行數位學習。上述觀點偏重工具性的解釋，而 Geddes(2004)則認為行動學習是使用行動科技獲取任何知識和技能，並且強調在行為上產生改變的過程。

Wagner(2005)進一步指出，行動學習不應該視為由數位學習轉移至行動裝置上，應該強調行動裝置本身所存在的價值與能力，而這些價值與能力可以促使人們在任何時間、地點連結到教育學習情境與人際之中。而 Lan 和 Sie(2010)認為行動學習是一種學習模式，允許學習者在任何地方、任何時間使用行動科技和網路獲得學習內容。因此，可將行動學習視為，使用行動科技所具備的優勢而進行的學習活動，所產生之有意義的學習。

二、行動學習與數位學習準備度之探討

行動學習奠基於數位學習的發展與理論基礎，將先就數位準備度加以說明，再探究行動學習準備度。

數位學習準備度的概念，最早由 Warner·Christie 和 Choy (1998) 所提出，主要分為三方面：(1)學生的學習喜好：有別於傳統面對面的教室教學，學生對於傳遞知識形式的喜好；(2)學生的學習信心：學生對於使用電子媒介溝通、學習、上網的自信心；(3)學生的自我學習：學生個人有能力進行自我學習。有學者(Chapnick, 2000; McVay, 2000; Redmon & Salopek, 2000; Rosenberg, 2001; Smith, 2005)認為數位學習準備度是一種評量工具，可以從中發覺影響一個機構是否成功導入數位學習策略的組織和個人因素，而 Borotis 和 Poulymenakou (2008)也有類似的看法，認為數位學習準備度是指一個組織對於許多數位學習經驗或行動在軟、硬體方面的準備度。

由於數位學習準備度影響的層面十分廣泛，從國與國之間的比較、企業間競爭力的評比和教育層面的現況說明，從國內外相關研究報告(吳美美，2004；趙美聲、陳鏗任、王玉蘭，2007；Economist Intelligence Unit, 2003, 2008, 2010)可以發現，伴隨著科技開發與成長所影響範圍，大至世界國家，小至生活社會，科技的影響無所不在，相關人員紛紛投入研究盼能將各方面與科技結合，進而達到生活、工作、教育數位化的階段，以善用科技特性與功能開啟另一波數位化革命。而在科技變革的初期，在大量科技硬體的規畫與設計之後，使用者接著將面臨這一波不同於以往的操作模式、使用習慣的改變，尤其是行動科技的發展，學習者所面臨更重要的挑戰是，對於使用這些數位化科技工具的準備度是否足夠(Koole, 2009；Ogamba & Peters, 2016)。

為了成功實施數位學習計畫，Saekow 和 Samson(2011)認為首先應該要評估學習者整合科技的數位學習準備度。Guglielmino(2003)提到影響數位學習的因素，主要由科技準備度和自我導向學習準備度所組成，這兩部分皆能用知識、態度、技能和習慣說明，也就是學習者必須要有科技相關的知識(knowledge)、正確使用科技的態度(attitude)、具備所需科技的技能(skills)以及正向積極的習慣(habits)；而 Clarke(2004)提到數位學習所要具備的要素有時間管理、接收與責任、學習規劃、自我評估、問題解決、對抗壓力、動機、反映與研究的能力；Smith(2005)指出數位學習準備度涉及科技技能、如何利用網路學習、學習風格、偏好及策略、學習者的軟硬體能力，以及網路學習溝通的態度。國內的研究也有類似的探討，如徐新逸和黃美蘭(2008)將數位學習準備度歸納為時間管理能力、使用電腦的經驗、電腦操作能力、線上寫作技巧、線上溝通能力、同儕合作能力、學習喜好、學習動機、設定目標與完成時程等九個向度。另外，Watkins(2005)提出數位學習準備度，可以用以下十項重點問題進行評估：(1)知識與技能問題，關於課程中將會使用的特定技能；(2)數位學習先備經驗；(3)相關議題課程內容的先備經驗；(4)對於課程的期望；(5)可參與課程的時間；(6)可同時利用的活動；(7)網路的使用；(8)線上溝通的情況；(9)使用軟體的結構；(10)使用科技支援的情況；而 Tubaishat 和 Lansari (2011)則針對大學生數位學習準備度所進行的研究中，則將數位學習準備度量表分為六項，分別為科技設備、網路使用、學生電腦技能、自信心發展、常用溝通模式和對數位學習的認知等因素。

此外，近年來對於行動學習準備度也有相關研究，如 Cheung、Yuen 和 Tsang (2011)認為行動學習準備度在於技術可行性、學生需

求、教學效益等三向度；而 Christensen 和 Knezek(2017)的將行動學習準備度以可能性、效益性、偏好和外在影響加以建構。Lin、Yeh 和 Wang (2016)在其行動研究準備度發展了三個向度的行動學習準備度量表，分別為三個向度：自我效能、樂觀態度、自我導向學習。

表 1

國內外數位/行動學習準備度相關研究之指標彙整表

作者(年代)	數位學習準備度的指標
Guglielmino & Guglielmino(2003)	科技準備度、自我導向學習準備度
Clarke(2004)	時間管理、接收與責任、學習規劃、自我評估、問題解決、對抗壓力、動機、反映與研究的能力
Watkins(2005)	知識與技能問題、數位學習先備經驗、課程內容的先備經驗、對於課程的期望、可參與課程的時間、可同時利用的活動、網路的使用、線上溝通的情況、使用軟體的結構、使用科技支援的情況
Smith(2005)	科技技能、如何利用網路學習、學習風格、偏好及策略、學習者的軟硬體能力、網路學習溝通的態度
Tubaishat & Lansari (2011)	科技設備、網路使用、學生電腦技能、自信心發展、常用溝通模式和對數位學習的認知
Cheung, Yuen & Tsang (2011)	技術可行性、學生需求、教學效益
Lin, Yeh & Wang (2016)	自我效能、樂觀態度、自我導向學習
Christensen & Knezek(2017)	可能性、效益性、偏好和外在影響
徐新逸、黃美蘭(2008)	時間管理能力、使用電腦的經驗、電腦操作能力、線上寫作技巧、線上溝通能力、同儕合作能力、學習喜好、學習動機、設定目標與完成時程
陳采秀(2012)	學習動機、自我管理、自我調控、自我效能、網路溝通互動、及資訊科技能力

資料來源：研究者自行整理

此外，經由表 1 國內外各研究者針對數位學習準備度所提出的指標，作進一步歸納與整理成表 2 的面向。從表 2 中可以發現，所有準備度的指標都集中於學習者在身心狀態本身所應具備的能力，如自我學習能力、科技能力、學習信念、學習態度、學習目標與目的等指標，也可從中看出在數位學習的需求下，部分能力是具高度需求與所使用之科技有關，據此才能順利完成數位學習活動，從表 2 可得知相關研究內涵累計的狀況。

表 2

國內外數位/行動學習準備度相關研究之內涵彙整表

作者 (年代)	自我 導向學習	人際 溝通能力	使用 科技能力	科技 使用經驗	認知	科技 知識	態度	信念	學習 動機	時間 管理	課程 先備經驗	線上 寫作能力	同儕 合作能力	學習 者喜好	學習 習慣	生活 型態	學習 目標與目的	學習 資源的運用
Mattice & Dixon(1999)	✓	✓	✓															
McVay(2001)	✓	✓	✓															
Schrum (2003)			✓	✓												✓		✓
Guglielmino & Guglielmino(2003)			✓			✓	✓								✓			
Bernard et al.(2004)	✓	✓	✓					✓										
Watkins(2005)		✓	✓							✓	✓							✓
Smith(2005)		✓	✓	✓			✓							✓				
Pillay, Irving, & Tones(2007)			✓	✓			✓							✓				
Valtonen, Kukkonen, Dillon, & Väisänen (2009)			✓					✓		✓								

(續下頁)

(接上頁)

作者 (年代)	自我導向學習	人際溝通能力	使用科技能力	科技使用經驗	認知	科技知識	態度	信念	學習動機	時間管理	課程先備經驗	線上寫作能力	同儕合作能力	學習者喜好	學習習慣	生活型態	學習目標與目的	學習資源的運用
Hung, Chou, Chen, & Own(2010)	✓	✓	✓	✓					✓									
Dray, Lowenthal, Miszkiewicz, Ruiz-Primo, & Marczynski(2011)	✓	✓	✓	✓				✓		✓								
Tubaishat & Lansari (2011)		✓	✓															
Ouma, Awuor, & Kyambo (2013)			✓		✓		✓											
Lin, Yeh, Wang (2016)	✓						✓											
Christensen & Knezek(2017)														✓				✓
徐新逸、黃美蘭(2008)		✓	✓	✓						✓		✓	✓	✓			✓	
洪敏玲(2012)	✓	✓							✓									
陳采秀(2012)		✓	✓						✓									
次數合計	7	11	15	6	1	1	5	3	3	4	1	1	1	4	1	1	2	2

資料來源：研究者自行整理

綜合上述，學者對於數位學習準備度量表的架構可以發現，構成量表的基本要件十分多元，且依不同研究量表的研究對象、需求、內涵有不同的組成要素，大致可以歸納為自我導向學習、人際溝通能力、使用科技能力、認知、態度、時間管理等 18 項指標(如表 2-4)。為篩選以上十八項相關因素，研究者對於單一因素有二個以上的研究所含括者便加以採納，經統計後計有：自我導向學習、人際溝通能力、使用科技能力、科技使用經驗、態度、時間管理、學習者喜好、信念、學習動機與學習目標和目的等十項因素。以下先就各因素的整合加以說明：

以「時間管理」和「學習目標與目的」因素來看，Clarke(2004)提到「時間管理」是學習者必須要負責管理自己的學習時間，並分析自己的學習目標、個人時間規畫與課程執行對於時間管理的能力的養成；而 Martin 和 Osborne(1989)也認為「時間管理」是個人設定目標和優先順序後，考量自身情況妥善規畫所面臨的學習任務，以最短時間內達成目標，這與陳年興、楊錦潭(2006)所提的「學習目標與目的」之內涵皆包含於時間管理範疇之中。而「自我導向學習」因素，Knowles(1975)認為是一種學習者主動參與的歷程，能設定學習目標，且確立相關學習資源；有些學者(Candy, 1991; Donavant, 2009; Eble, 1998; Guglielmino, 1977)也認為自我導向學習具有主動學習、效率學習、熱愛學習、時間管理、設定目標、認為學習是快樂的事與能獨立進行的概念；此因素說明學習者需具備積極的學習動機，因此本研究將「動機」、「學習目標與目的」和「時間管理」合而為一，併入「自我導向學習」因素。

以「使用科技能力」因素而言，Smith 等人(2003)、Smith(2005)、Watkins(2005)、Pillay 等人(2007)皆指出參與線上學習軟體和硬體設

備的使用能力之重要性。而 Schrum 和 Benson(2002)將「科技使用經驗」解釋為使用科技軟體書寫、列印且能定期寄收電子信件、處理簡易電腦軟硬問題的能力，因此兩者因素可視為同一因素「使用科技能力」。此外，Watkins 等人(2004)、Watkins(2005)認為學習者應具備網路溝通的技能，包括利用電子信件、視訊軟體與其他人溝通，此與 Clarke(2004)曾說明之溝通為在課堂上的小組討論相似，因此將「人際溝通能力」更改為「溝通能力」。

綜合上述，依據表 2 所歸納而出的國內外數位學習準備度共有十項，在研究者透過相關研究文獻的探討而將十項因素整併為：自我導向學習、溝通能力、使用科技能力、信念、學習態度、學習偏好等六項。然而，在自我導向學習因素方面，以現有的數位學習與行動學習準備度量表的實測對象大多以大學、高中以上的學生為主的情況下，可以瞭解到高中以上的學生或成人可能具有較佳的自我導向學習能力，亦較能對自我的學習負責任，此因素對於國內國小高年級學童而言，仍需由教師規畫相關課程加以養成和訓練，所以，本研究不將「自我導向學習」納入行動學習準備度的內涵。

三、行動學習準備度之內涵

考量本研究的研究目的、研究對象，針對上述行動學習的定義與行動學習準備度的理論，採用其他相關數位學習準備度的研究取向，歸納出本研究所提之行動學習準備度的內涵，共整理出行動學習準備度的五大內涵，說明如下：

(一) 科技使用能力

Watkins(2005)認為電腦科技的使用是為了針對數位學習的需求，使用自己的電腦軟硬體設備能符合課程所需，

以便能順利進行線上課程；我國教育部所頒訂的九年一貫課程綱要之基本能力中提到，使用科技的能力是指認識資訊科技的範疇與內涵，並能有效地使用電腦進行資訊的搜尋、處理及傳遞作業，進而能應用於生活和學習上(教育部，2008)。因此，本研究定義使用科技能力為，利用平板電腦等行動載具之學習過程中，學習者善用所具備的相關科技能力解決學習的問題，如文書處理能力、學習軟體安裝、網路溝通能力、資訊搜尋能力等。

(二) 學習態度

關於學習態度，張春興(2013)認為是一種透過教與學的歷程，針對特定的方向與目標，進行持久習慣性與一致性的內在心理反應。在 Chen(2013)利用行動裝置作為學習工具所進行的研究指出，學生對於此學習方式和行動裝置的結合，具有良好的學習態度。是以，本研究將學習態度定義為，學習者本身對於利用平板電腦等行動載具進行學習，所持有的行為傾向或其內部反應的準備狀態。

(三) 溝通能力

McVay(2000)曾指出，在網路學習中，創造學習者和其教學者之間互動與溝通的機會是很重要的；而同儕之間或學生與教學者的溝通能力，經常能透過即時的互動促進與激勵學習，因此在數位學習過程中是必備的技能(Watkins, Leigh, & Triner, 2004)。故本研究將溝通能力定義為，學習者在行動學習過程中，能使用電子郵件、社交網站或即時通訊軟體與他人溝通的能力。

(四) 學習信念

Ross(2001)認為信念是一種對事物或命題確信不疑或完全接納的心理傾向或狀態，是對某種事物、對象或命題表示接納、贊成或肯定的態度。在一項行動學習準備度的研究中提到，學生的信念亦會影響他們使用行動學習裝置的傾向(Cheon et al., 2012)。由此可知，學習者的學習信念對於學習具有相當的影響力，故本研究定義學習信念為，學習者本身利用平板電腦等行動載具學習的過程中，對於自我科技使用能力、自我問題解決能力、以及平板電腦等行動載具本身所抱持的態度，進而影響學習行為的看法。

(五) 學習偏好

學習偏好是學習者在跨學科的學習、問題解決過程中，是因人而異的學習方式和經驗(Kruzich, Friesen, & Van Soest, 1986)。Kolb(1985)也認為學習者因個人遺傳、學習經驗與當前環境要求，所培養出選擇學習方式的不同偏好為個人因素。有研究相信學習偏好是個別學習者喜好行為的預測項目(Bostrom, Olfman & Sein, 1993)，同時也能影響學習者的表現和學習結果(Cassidy, 2004)，也能促使成功的行動學習實施(Christensen & Knezek, 2017)的研究因此，本研究將學習偏好定義為，學習者能在利用平板電腦等行動載具學習的過程中，選擇適合自己的學習方式，以完成學習任務。

透過上述的五大內涵說明，作為本研究探討行動學習準備度的重要理論基礎，據以作為量表發展的重要依據。

參、研究方法

本研究旨在建構出能符應國小高年級學生行動學習準備度量表，研究者首先從行動學習文獻和相關研究為基礎，綜整歸納出行動學習準備度之五大內涵，編製成德懷術專家問卷。再經由德懷術(Delphi Technique)邀請專家學者審定修改，建立問卷各向度的效度，以下分別就研究對象、研究工具逐一說明：

一、研究對象

本研究採用德懷術在於行動學習準備度量表發展在國內屬於初步發展階段，為了兼顧理論與實務面的差異，需要從事資訊教育、數位學習或教學科技相關領域的專家學者提供寶貴的意見，強化、建構國小高年級學生行動學習準備度量表的内容效度，最佳化研究工具。

就德懷術專家小組而言，Delbecq、Van de Ven 和 Gustafson(1975)指出，若德懷術專家小組之同質性較高時，小組人數最好是 15 至 30 人。若德懷術專家小組達到 10 人以上時，成員間的群體誤差可降至最低，若專家小組大於 13 人時，其信度將會高於.80 (Dalkey, 1969)，故考量本研究主題與提高信度，選取 4 位國內從事資訊教育、數位學習或教學科技相關領域之專家學者及 9 位教學實際工作者共 13 位組成德懷術專家小組(詳見表 3)。

表 3

本研究德懷術專家委員一覽表

編號	類別	職稱	性別	教學年資	學校規模	職務	專長
01		教授	女				資訊教育、數位學習、 資訊倫理
02	學者專家	副教授	男				行動學習、學習歷程管理、 網路專題學習
03		助理教授	男				電腦輔助教學、程式設計、 網路管理與應用
04		特聘教授	女				數位學習、資訊教育、 遠距教學
05		國小教師	男	15	大型	導師	行動學習、資訊融入教學
06		國小教師	女	20	大型	導師	數位學習、資訊融入教學
07		國小教師	女	8	大型	導師	行動學習、資訊融入教學
08	教學實務專家	國小教師	男	8	大型	科任	數位學習、資訊融入教學
09		國小教師	男	20	大型	科任	行動學習、資訊融入教學
10		國小教師	男	14	小型	科任	行動學習、資訊融入教學
11		國小教師	男	22	中型	科任	行動學習、資訊融入教學
12		國小教師	男	21	中型	導師	行動學習、資訊融入教學
13		國小教師	女	24	大型	導師	數位學習、資訊融入教學

二、研究工具

本研究主要的研究工具為自行發展的「國小高年級學生行動學習準備度量表」，透過文獻探討和分析編製成研究問卷初稿，再經由德懷術專家效度的程序，建構行動學習準備度的基本內涵，茲說明如下：

(一) 研究量表初稿編製

本研究所編製之「國小高年級學生行動學習準備度量表」，包括使用科技能力、學習態度、溝通能力、學習信念、

學習偏好等五個向度，題目內容初稿主要參考 Bernard 等人(2004)、Cheon 等人(2012)、Dray 等人(2011)、Hamat、Embi 和 Hassan (2012)、及 Hung 等人(2010)之問卷編製而成，例如使用科技能力向度方面之題項「我能使用網路搜尋資訊」(Hung et al., 2010)；學習態度向度方面之題項「課堂中使用行動學習是一種愉快的經驗」(Cheon et al., 2012)；溝通能力向度方面之題項「我有信心使用科技工具有效地與他人溝通」(Cheon et al., 2012)；學習信念向度方面之題項「我能享受線上學習課程」(Valtonen, Kukkonen, Dillon & Väisänen, 2009)；學習偏好向度方面之題項「我習慣觀看數位教材的學習方式」(徐新逸、黃美蘭，2008)。本研究量表初稿共有 52 題，包含 12 題「使用科技能力」、13 題「學習態度」、8 題「溝通能力」、10 題「學習信念」和 9 題「學習偏好」。

(二) 德懷術調查問卷

依據「國小高年級學生行動學習準備度量表」初稿，本研究實施二次德懷術問卷調查，第一次德懷術調查問卷主要目的在於徵詢專家委員的意見，依據其專業知識及對數位學習的了解，在調查問卷中以李克特氏(Likert-type)五等量表來判斷其重要性。在重要性評定部分，採用李克特式五等量表來判斷各題項的重要性；而第二次德懷術問卷調查的目的在請專家委員依第一次問卷調查之統計分析結果及專家意見彙整結果，重新評定各題項的重要性，或提供其他意見與觀

點，專家委員可變更或維持原來意見，也請其儘量能說明理由。問卷採用李克特氏五等量表來判斷其重要性。各題項分數從 1-5 分，以「1」分代表「非常不重要」、「2」分代表「不重要」、「3」分代表「還算重要」、「4」分代表「重要」、「5」分代表「非常重要」，題項所得的分數愈高，表示該項目的重要性愈高，反之則重要性愈低。

(三) 德懷術調查問卷收斂標準

在德懷術數回合問卷調查的填答中，專家小組成員在某些題目的評析難免會發生些許的改變，因此有學者提出以意見一致性與穩定度等兩方面為主的統計判斷。首先是一致性，Faherty(1979)、Holden 和 Wedman(1993)認為，當專家群對某項目的意見分佈之四分位差小於或等於.50，即專家群對該題項的意見達到高度的一致性；而當四分位差大於 1.00 時，即專家群對該題項的意見並未達到一致性的共識。而 Shieh(1990)也指出，當該題項後一回合結果之標準差小於或等於前一回合結果的標準差，即可認定此提項之專家小組意見達到共識，當達到共識之累積題項達到全體三分之二以上，則可宣稱此研究之專家意見達到一致性，就可終止研究問卷之進行。

接著是穩定度，Franchack、Desy 和 Norton(1984)認為部分指標未達意見一致標準時，可用問卷的整體穩定程度作為調查是否結束的依據；也就是說，德懷術專家小組成員對於每個題項改變程度小於或等於 15%時，便稱小組成

員對此題項的評估結果達穩定狀態(Linstone & Turoff, 1975)，而問卷中所有題項之穩定度累積大於或等於 70% 時，德懷術的問卷調查即可停止進行(Holden & Wedman, 1993)。

肆、研究結果與討論

一、第一回合德懷術調查問卷結果分析

第一回合德懷術調查問卷，將「國小高年級學生行動學習準備度量表」分為五個向度，共有 52 個題目。填答方式則由德懷術專家小組成員針對調查問卷中各向度的題目，依其重要性勾選及開放性意見陳述。問卷回收後，計分資料為問卷中題目五點量表評定，並求出眾數(Mo)、平均數(M)與標準差(SD)。

專家小組成員提出第一回合德懷術調查問卷指標項目須修正及建議，經內容分析與意見彙整、將「國小高年級學生行動學習準備度量表」各向度之題目數修正為，「使用科技能力」有 12 題、「學習態度」有 13 題、「溝通能力」有 8 題、「學習信念」有 10 題、「學習偏好」有 9 題，共 52 題，以下就各向度之統計分析、委員意見和意見回覆分別說明：

(一)「使用科技能力」向度

1. 「使用科技能力」向度統計分析

第一次德懷術專家意見調查問卷中，「使用科技能力」向度的題目共有 12 題，如表 4 所示之統計分析結果，此向度各題目的眾數，皆為 4(重要)或 5(非常重要)；平均數介

於 4.17 至 5.00 之間，表示委員們對於各題目的重要性為重要至非常重要之間的程度；就重要性離散情形方面，標準差介於 0 至 1.01 間，顯示委員們對於部分題目的意見不太一致。

表 4

第一次德懷術調查問卷「使用科技能力」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差
A1	我會使用平板電腦的筆記功能，記錄課程相關內容。	5	4.50	0.65
A2	我能使用平板電腦的拍照或錄影工具，記錄課程的相關內容。	5	4.42	0.64
A3	我會使用平板電腦，與老師、同學進行學習資料的傳輸。	5	4.67	0.62
A4	我會使用平板電腦自行下載與課程相關的 APP 軟體。	4	4.25	0.72
A5	我具有使用平板電腦，儲存網頁中與課程內容相關的圖片、文字的能力	5	4.75	0.43
A6	我具有使用平板電腦中，利用搜尋引擎，尋找課程內容所需資訊的能力	5	5.00	0.00
A7	我具有使用平板電腦中，文書處理軟體的基本操作能力，例如記事本、備忘錄等軟體。	5	4.42	0.64
A8	我具備使用平板電腦上傳課程學習資料給老師或同學的能力。	5	4.25	1.01
A9	我具備運用平板電腦的 App 軟體，上傳資料或檔案與老師或同學分享。	5	4.00	1.00
A10	我具備運用平板電腦中的即時反饋系統 (IRS)，立即回答學習相關問題的能力。	4	4.17	0.69
A11	我具備操作平板電腦的各項功能，以符合課程需求。	4	4.33	0.62
A12	我認為自己目前所具備的科技能力，能夠應付課堂上使用平板電腦的需求	4	4.50	0.65

2. 「使用科技能力」向度與題目之委員意見彙整

經彙整第一回合德懷術專家調查問卷，將「使用科技能力」向度由原本 12 個題目合併 2 題、修正 10 題，其中 A3 和 A8 題目內容與 A7 雷同，故將合併原題號 A3「我能使用平板電腦，與老師、同學進行學習資料的傳輸」和原題號 A8「我能使用平板電腦上傳課程學習資料給老師或同學的能力」至 A7。

至於修正題目的部分，除了將 A1、A2、A5、A6、A7、A10 等 6 個題目中「我會...」修改為「我能...」的用語修正外，還將 A4「我會使用平板電腦自行下載與課程相關的 APP 軟體。」修正為「我能使用平板電腦下載與老師所教或與課程相關的 App 軟體」、A9「我具備運用平板電腦的 App 軟體，上傳資料或檔案與老師或同學分享。」修正為「我能運用平板電腦的 App 軟體(如：dropbox、google 雲端硬碟)，上傳資料或檔案給老師或同學」、A11「我具備操作平板電腦的各項功能，以符合課程需求」修正為「我能適時操作平板電腦的功能，來符合課程需求」。

(二)「學習態度」向度

1. 「學習態度」向度之統計分析

第一次德懷術專家意見調查問卷中，「學習態度」向度的題目共有 13 題，如表 5 所示之統計分析結果，此向度各題目的眾數皆為 5(非常重要)；平均數介於 4.38 至 4.69 之間，表示委員們對於各題目的重要性為重要至非常重要之

間的程度；就重要性離散情形方面，標準差介於 0.46 至 0.84 間，顯示委員們對於此向度的題目意見頗為一致。

表 5

第一次德懷術調查問卷「學習態度」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差
B1	使用平板電腦上課時，我認為是很愉快的學習經驗。	5	4.69	0.46
B2	使用平板電腦上課時，我會提醒自己要保持良好學習態度。	5	4.38	0.84
B3	使用平板電腦上課時，我會隨時提醒自己要跟老師的進度。	5	4.46	0.63
B4	使用平板電腦上課的過程中，我會主動完成老師指派的作業。	5	4.69	0.46
B5	使用平板電腦上課的過程中，我會很快地利用平板電腦找到學習內容。	5	4.46	0.75
B6	在使用平板電腦上課的過程中，我會專心地閱讀學習內容。	5	4.62	0.74
B7	當使用平板電腦上課遇到問題時，我會儘快尋求老師的協助。	5	4.54	0.63
B8	當使用平板電腦上課遇到問題時，我會儘快尋求同學的協助。	5	4.46	0.75
B9	當使用平板電腦上課遇到問題時，如果要向別人求助，我不會覺得很害羞。	5	4.46	0.75
B10	即使使用平板電腦上課的課程活動較長，我仍會耐心的學習。	5	4.54	0.63
B11	我認為使用平板電腦學習時，自己的學習態度很重要。	5	4.69	0.46
B12	我喜歡使用平板電腦進行學習。	5	4.46	0.75
B13	如果未來還有機會，我仍有興趣使用行平板電腦上課。	5	4.54	0.63

2. 「學習態度」向度與題目之委員意見彙整

經彙整第一回合德懷術專家調查問卷，將「學習態度」向度由原本 13 個題目合併 5 題、修正 7 題，原題號 B11 題不變更。在合併的部分，B2「使用平板電腦上課時，我會提醒自己要保持良好學習態度」，採納委員意見：題意抽象，可由其他題目回答；B5「使用平板電腦上課的過程中，我會很快地利用平板電腦找到學習內容」和 B8「當使用平板電腦上課遇到問題時，我會儘快尋求同學的協助」內容較屬於「A 使用科技能力」向度，故將 B5、B8 與 A6 題合併；也有委員認為 B12、B13 的題意較偏向「E 學習偏好」的題目，故採納委員意見，將 B12、B13 與「E 學習偏好」的題目合併。

至於修正題目的部分，除了將 B1、B4、B6、B7、B9、B10 等 7 個題目，加入「我認為...」於句首的用語修正外，還將原 B3「使用平板電腦上課時，我會隨時提醒自己要跟上老師的進度」修正為「我認為使用平板電腦上課，我可以跟上老師的進度」。

(三)「溝通能力」向度

1. 「溝通能力」向度之統計分析

第一次德懷術專家意見調查問卷中，「溝通能力」向度的題目共有 8 題，如表 6 所示之統計分析結果，此向度各題目的眾數為 4(重要)或 5(非常重要)；平均數介於 4.23 至 4.69 之間，表示委員們對於各題目的重要性為重要至非常

重要之間的程度；就重要性離散情形方面，標準差介於 0.5 至 1.12 間，顯示委員們對於部分題目的意見不太一致。

表 6

第一次德懷術調查問卷「溝通能力」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差
C1	我會使用平板電腦的溝通工具，與老師、同學相互聯絡。	5	4.54	0.50
C2	我會使用平板電腦的溝通工具，主動與同學互相傳遞學習相關的資訊	5	4.62	0.62
C3	我會使用平板電腦的溝通工具，向別人尋求問題的解決。	5	4.38	0.62
C4	使用平板電腦上課時，我會主動參與小組的討論活動。	5	4.69	0.61
C5	我會使用平板電腦的即時通訊工具，協同同學解決問題。	5	4.23	0.97
C6	我會選擇平板電腦中適當的溝通工具，與老師、同學聯絡事情。	4	4.08	1.27
C7	我所具備的平板電腦溝通工具，能幫助我解決學習上的問題。	5	4.23	1.12
C8	我認為自己目前所具備的溝通能力，能夠應付課堂上的需求。	5	4.23	1.12

2. 「溝通能力」向度與題目之委員意見彙整

經彙整第一回合德懷術專家調查問卷，將「溝通能力」向度由原本 8 個題目合併 2 題、修正 6 題。在合併的部分，C2 題意與「A 使用科技能力」向度之 A7 題雷同，故將原題號 C2「我會使用平板電腦的溝通工具，主動與同學互相傳遞學習相關的資訊」與 A7 題合併。而 C6「我會選擇平板電腦中適當的溝通工具，與老師、同學聯絡事情」題意與 C1 雷同，故將原題號 C3 與 C1 題合併。

至於修正題目的部分，除了將 C1、C2、C5 等 3 個題目中將原有的「溝通工具」增加描述為「溝通工具(如：line、google talk 等 App 軟體)」的修正外，採納委員意見將 C4 「使用平板電腦上課時，我會主動參與小組的討論活動」修正語句為「我會使用平板電腦主動參與小組的討論活動」、將 C7 「我所具備的平板電腦溝通工具，能幫助我解決學習上的問題」修正為「我認為具備使用平板電腦的溝通工具能力，能幫助我解決學習上的問題」以符合該向度；而 C8 「我認為自己目前所具備的溝通能力，能夠應付課堂上的需求」的原題意，未能明顯指出與平板電腦有關，故亦採納委員意見，將 C8 修正為「我認為自己目前所具備的應用平板電腦進行溝通的能力，能夠應付課堂上的需求」。

(四)「學習信念」向度

1. 「學習信念」向度之統計分析

第一次德懷術專家意見調查問卷中，「學習信念」向度的題目共有 10 題，如表 7 所示之統計分析結果，此向度各題目的眾數為 5(非常重要)；平均數介於 4.38 至 4.62 之間，表示委員們對於各題目的重要性為重要至非常重要之間的程度；就重要性離散情形方面，標準差介於 0.49 至 0.75 間，顯示委員們對於部分題目的意見頗為一致。

表 7

第一次德懷術調查問卷「學習信念」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差
D1	我相信使用平板電腦上課，能幫助我整合課程內容。	5	4.62	0.49
D2	我相信平板電腦，對我的學習活動很重要。	5	4.62	0.49
D3	我相信使用平板電腦上課，能立即解決課堂中學習的問題	5	4.38	0.62
D4	我相信平板電腦的溝通軟體，是有效的溝通工具。	5	4.46	0.75
D5	我相信平板電腦在學習過程中，扮演重要的角色。	5	4.54	0.75
D6	我相信平板電腦攜帶方便的特性，是個有助於學習的工具	5	4.62	0.62
D7	我相信使用平板電腦學習的過程，也能提升我個人使用軟體的能力。	5	4.54	0.50
D8	我支持使用平板電腦學習。	5	4.54	0.63
D9	平板電腦能幫助我更有效率地完成學習任務。	5	4.62	0.49
D10	使用平板電腦上課，對我而言，不是困難的事。	5	4.46	0.63

2. 「學習信念」向度與題目之委員意見彙整

經彙整第一回合德懷術專家調查問卷，將「學習信念」向度由原本 10 個題目合併 1 題、修正 6 題，另有 3 題(D1、D3、D5)不變更。在合併的部分，委員指出 D10 題意較偏向「A 使用科技能力」向度，故將原題號 D10「使用平板電腦上課，對我而言，不是困難的事」合併至 A10 題。

至於修正題目的部分，委員大都針對題目語句的流暢性和易讀性，提供用語微調的建議，如將 D2「我相信平板電腦，對我的學習活動很重要」修正為「我相信平板電腦，

對我的學習很重要」、將 D4 中的「...溝通軟體」修改並增加舉例說明為「即時溝通工具（如：line、google talk 等 App 軟體）」、將 D6「我相信平板電腦攜帶方便的特性，是個有助於學習的工具」修正為「我相信平板電腦攜帶方便的特性，有助於學習」、D7「我相信使用平板電腦學習的過程，也能提升我個人使用軟體的能力」修正為「...能提升我個人使用軟體的能力」、D8「我支持使用平板電腦學習」修正為「我會支持使用平板電腦來學習」、D9「平板電腦能幫助我更有效率地完成學習任務」修正為「我相信平板電腦能幫助我更有效率地完成學習任務」。

(五)「學習偏好」向度

1.「學習偏好」向度之統計分析

第一次德懷術專家意見調查問卷中，「學習偏好」向度的題目共有 13 題，如表 8 所示之統計分析結果，此向度各題目的眾數為 5(非常重要)；平均數介於 4.54 至 4.69 之間，表示委員們對於各題目的重要性為重要至非常重要之間的程度；就重要性離散情形方面，標準差介於 0.49 至 0.63 間，顯示委員們對於此向度題目之意見頗為一致。

表 8

第一次德懷術調查問卷「學習偏好」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差
E1	我喜歡使用平板電腦上課的方式。	5	4.62	0.49
E2	我喜歡使用平板電腦來解決上課的問題。	5	4.54	0.63
E3	我喜歡使用平板電腦與老師、同學溝通的方式。	5	4.46	0.63
E4	我喜歡使用平板電腦中動畫和影音的學習方式。	5	4.62	0.62
E5	我喜歡利用平板電腦進行小組合作學習。	5	4.69	0.46
E6	我喜歡使用平板電腦進行線上蒐尋資料的方式。	5	4.69	0.46
E7	我喜歡使用平板電腦上課，能滿足我學習的需求。	5	4.54	0.63
E8	我覺得自己能適應將平板電腦作為學習的工具。	5	4.54	0.63
E9	我喜歡將平板電腦的功能應用在學習上。	5	4.62	0.49

2. 「學習偏好」向度與題目之委員意見彙整

經彙整第一回合德懷術專家調查問卷，將「學習偏好」向度原有題目合併 2 題、修正 4 題，另有 3 題(E1、E2、E5)不予變更。在合併部分，第 E8 和第 E9 號題目內容與 E1 題雷同，故將合併原題號 E8「我覺得自己能適應將平板電腦作為學習的工具」和原題號 E9「我喜歡將平板電腦的功能應用在學習上」至 E1 題。

至於修正題目的部分，委員針對向度定義與語句流暢性和易讀性，建議修正和微調 E3、E4、E6、E7 的部分語句，讓題意更為明確。將 E3 題「我喜歡使用平板電腦與

老師、同學溝通的方式」修正為「我喜歡使用平板電腦與老師或同學溝通」、將 E4 題「我喜歡使用平板電腦中動畫和影音的學習方式」修正為「我喜歡使用平板電腦觀看動畫和影音進行學習」、將 E5 「我喜歡使用平板電腦進行線上蒐尋資料的方式」修正為「我喜歡使用平板電腦進行線上蒐尋資料」、將 E7 「我喜歡使用平板電腦上課，能滿足我學習的需求」修正為「使用平板電腦上課，能滿足我學習的需求」。

第一回合德懷術專家問卷分析結果，係研究者依據專家們針對各題項所提出之意見加以修改或合併題項。此回合中，「使用科技能力」向度，修正 10 題、合併 2 題，題數由原本的 12 題減為 10 題；「學習態度」向度，修正 7 題、合併 5 題、1 題不變更，題數由原本的 13 題減為 8 題；「溝通能力」向度，修正 6 題、合併 2 題，題數由原本的 8 題減為 6 題；「學習信念」向度，修正 6 題、合併 1 題、3 題不變更，題數由原本的 10 題減為 9 題；「學習偏好」向度，修正 4 題、合併 2 題、3 題不變更，題數由原本的 9 題減為 7 題。總量表題數由 52 題減少為 40 題，加上合併的題項，作為第二回合德懷術調查的問卷。

二、第二回合德懷術調查問卷結果分析

第二回合德懷術調查問卷，填答方式則由德懷術專家小組成員針對調查問卷中各向度的題目，依其重要性勾選及開放性意見陳述。問卷回收後，計分資料為問卷中題目五點量表評定，並求出眾數(Mo)、平均數(M)與標準差(SD)；意見資料為彙整專家小組成員意見，進行題目內容的修正。

至於刪題的標準，在第二次回合問卷調查中，若題目平均數小於 3.5 且標準差大於 1.00，表示專家認為此題目的重要性不高並且不具高度共識，則該題予於刪除。若經過問卷統計處理，問卷中該題目的四分位差小於或等於.50，即判定專家小組成員對該題目已達高度共識，且專家小組成員對於問卷題目的改變比率低於 15%，且累積 70%的題目均達穩定度狀態時，則本研究德懷術調查工作即停止進行，不再進行第三次德懷術問卷調查；反之，則持續進行。

關於本次第二回合德懷術專家問卷的資料整理與分析，在一致性部分，全部題目的四分位差皆小於或等於.50，表示專家小組對於問卷的題目具有高度的共識；在穩定性的部分，專家小組成員針對題目重要性評比的改變情形低於 15%，且已累積 70%的題目達穩定狀態，如表 4-15 所示。綜合上述結果，本研究之德懷術調查工作便能停止進行，不再進行第三次德懷術問卷調查。

以下依第二回合德懷術問卷各向度之統計分析、以及專家委員針對題目進行意見回覆的部分加以說明，若無意見的向度，僅呈現統計分析結果：

(一)「使用科技能力」向度

第二次德懷術專家意見調查問卷中，「使用科技能力」向度的題目共有 10 題，如表 9 所示之統計分析結果，此向度各題目的眾數，皆為 4(重要)或 5(非常重要)；各題目的四分位差介於 0 至 0.5 之間；平均數介於 3.92 至 5.00 之間，表示委員對於各題目的重要性為有些重要至非常重要之間的程度；就重要性離散情形方面，大部分題目的標準差介於 0.00 至 0.74 間，顯示委員們對於大部分題目的意見頗為一致，僅 A8 題的標準差為 1.07，表示委員們對此題目的意見不太一致。而專家們在此向度並無意見陳述，故此向度所有題目皆維持原狀。

表 9

第二次德懷術調查問卷「使用科技能力」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差	四分位差
A1	我能使用平板電腦的筆記功能，記錄課程相關的內容。	5	4.69	0.46	0.5
A2	我能使用平板電腦的拍照或錄影工具，記錄課程的相關內容。	5	4.62	0.49	0.5
A3	我能使用平板電腦下載與老師所教或與課程相關的 App 軟體。	4	4.38	0.74	0.5
A4	我能使用平板電腦，儲存網頁中與課程內容相關的圖片或文字的能力。	5	4.92	0.27	0
A5	我能使用平板電腦的搜尋引擎，尋找課程內容所需的資訊	5	5.00	0.00	0
A6	我能使用平板電腦的文書處理軟體，例如記事本、備忘錄等	4	4.38	0.49	0.5

(續下頁)

(接上頁)

原題號	題目	眾數	平均數	標準差	四分位差
A7	我能運用平板電腦的 App 軟體 (如：dropbox、google 雲端硬碟)，上傳資料或檔案給老師或同學。	4	4.38	0.49	0.5
A8	我能運用平板電腦中的即時反饋系統(IRS)，立即回答學習的相關問題的能力。	4	3.92	1.07	0.5
A9	我能適時操作平板電腦的功能，來符合課程需求。	5	4.46	0.63	0.5
A10	我認為自己具備課程上使用平板電腦所需的科技能力。	4	4.46	0.50	0.5

(二)「學習態度」向度

第二次德懷術專家意見調查問卷中，「學習態度」向度的題目共有 8 題，如表 10 所示之統計分析結果，此向度各題目的眾數皆為 5(非常重要)；各題目的四分位差介於 0 至 0.5 之間；平均數介於 4.38 至 4.77 之間，表示委員對於各題目的重要性為重要至非常重要間的程度；就重要性離散情形方面，標準差介於 0.42 至 0.62 間，顯示委員們對於此向度的題目意見頗為一致。而專家們在此向度並無意見陳述，故此向度所有題目皆維持原狀。

表 10

第二次德懷術調查問卷「學習態度」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差	四分位差
B1	我認為使用平板電腦上課，會是很愉快的學習經驗。	5	4.62	0.49	0.5
B2	我認為使用平板電腦上課，我可以跟上老師的進度。	5	4.38	0.62	0.5
B3	我認為使用平板電腦上課的過程中，我會主動完成老師指派的作業。	5	4.54	0.50	0.5
B4	我認為使用平板電腦上課的過程中，我會專心的參與學習活動。	5	4.62	0.62	0.5
B5	我認為使用平板電腦上課遇到問題時，我會儘快尋求老師或同學的協助。	5	4.69	0.46	0.5
B6	我認為使用平板電腦上課遇到問題向別人求助時，我不會覺得很害羞。	5	4.62	0.49	0.5
B7	我認為使用平板電腦上課時，我仍會耐心的學習。	5	4.54	0.50	0.5
B8	我認為使用平板電腦學習時，自己的學習態度很重要。	5	4.77	0.42	0

(三)「溝通能力」向度

第二次德懷術專家意見調查問卷中，「溝通能力」向度的題目共有 6 題，如表 11 所示之統計分析結果，此向度各題目的眾數為 4(重要)至 5(非常重要)；各題目的四分位差介於 0 至 0.5 之間；平均數介於 4.15 至 4.85 之間，表示委員對於各題目的重要性為重要至非常重要之間的程度；就重要性離散情形方面，標準差介於 0.36 至 0.70 間，顯示委

員們對於此向度的題目意見頗為一致。而專家們在此向度並無意見陳述，故此向度所有題目皆維持原狀。

表 11

第二次德懷術調查問卷「溝通能力」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差	四分位差
C1	我會使用平板電腦的溝通工具(如: line、google talk 等 App 軟體)，與老師、同學相互聯絡。	5	4.46	0.63	0.5
C2	我會使用平板電腦的溝通工具(如: line、google talk 等 App 軟體)，向別人尋求問題的解決。	4	4.31	0.61	0.5
C3	我會使用平板電腦主動參與小組的討論活動。	5	4.85	0.36	0
C4	我會透過平板電腦的即時通訊工具(如: line、google talk 等 App 軟體)，協助同學解決問題。	4	4.15	0.66	0.5
C5	我認為具備使用平板電腦的溝通工具能力，能幫助我解決學習上的問題。	4	4.23	0.70	0.5
C6	我認為自己目前所具備的應用平板電腦進行溝通的能力，能夠應付課堂上的需求。	4	4.23	0.70	0.5

(四)「學習信念」向度

第二次德懷術專家意見調查問卷中，「學習信念」向度的題目共有 9 題，如表 12 所示之統計分析結果，此向度各題目的眾數為 4(重要)至 5(非常重要)；各題目的四分位差均為 0.5；平均數介於 4.38 至 4.69 之間，表示委員對於各題目的重要性為重要至非常重要間的程度；就重要性離散情形方面，標準差介於 0.46 至 0.63 間，顯示委員們對於部

分題目的意見頗為一致。而專家們在此向度並無意見陳述，故此向度所有題目皆維持原狀。

表 12

第二次德懷術調查問卷「學習信念」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差	四分位差
D1	我相信使用平板電腦上課，能幫助我整合課程內容。	5	4.62	0.62	0.5
D2	我相信平板電腦，對我的學習活動很重要。	5	4.54	0.63	0.5
D3	我相信使用平板電腦上課，能立即解決課堂中學習的問題	4	4.38	0.49	0.5
D4	我相信平板電腦的即時溝通工具(如：line、google talk 等 App 軟體)，是有效的溝通工具。	5	4.46	0.63	0.5
D5	我相信平板電腦在學習過程中，扮演重要的角色。	5	4.46	0.63	0.5
D6	我相信平板電腦攜帶方便的特性，有助於學習。	5	4.62	0.62	0.5
D7	我相信使用平板電腦學習的過程，能提升我個人使用軟體的能力。	5	4.54	0.63	0.5
D8	我會支持使用平板電腦來學習	5	4.54	0.63	0.5
D9	我相信平板電腦能幫助我更有效率地完成學習任務。	5	4.69	0.46	0.5

(五)「學習偏好」向度

第二次德懷術專家意見調查問卷中，「學習偏好」向度的題目共有 8 題，如表 13 所示之統計分析結果，此向度各題目的眾數為 5(非常重要)；各題目的四分位差介於 0 至 0.5 之間；平均數介於 4.54 至 4.77 之間，表示委員對於各題目的重要性為重要至非常重要之間的程度；就重要性離散情

形方面，標準差介於 0.42 至 0.63 間，顯示委員們對於此向度題目之意見頗為一致。

表 13

第二次德懷術調查問卷「學習偏好」之描述性統計分析結果

原題號	題目	眾數	平均數	標準差	四分位差
E1	我喜歡使用平板電腦上課的方式。	5	4.69	0.46	0.5
E2	我喜歡使用平板電腦來解決上課的問題。	5	4.62	0.49	0.5
E3	我喜歡使用平板電腦與老師或同學溝通。	5	4.54	0.50	0.5
E4	我喜歡使用平板電腦觀看動畫和影音進行學習。	5	4.69	0.46	0.5
E5	我喜歡利用平板電腦進行小組合作學習。	5	4.69	0.61	0
E6	我喜歡使用平板電腦進行線上蒐尋資料。	5	4.77	0.42	0
E7	使用平板電腦上課，能滿足我學習的需求。	5	4.54	0.63	0.5
E8	我覺得自己能適應將平板電腦作為學習的工具。	5	4.54	0.63	0.5

經由專家意見的彙整，將原本欲合併至 E1 的題目「我覺得自己能適應將平板電腦作為學習的工具」加以保留，意見如表 14，而專家們對其他題目並無意見陳述，故不予變更或合併。

表 14

第二次德懷術專家問卷「學習偏好」向度之委員意見彙整一覽表

原題號	題目	委員意見
E8	◎我覺得自己能適應將平板電腦作為學習的工具。	這一題是否宜併入第一題，因第一題是針對老師主導的上課方式，這一題重點應該是自學生面向認為自己是否能適應將平板電腦作為學習的工具，關注的焦點應該有不同。

研究者在彙整第二回合德懷術調查問卷之資料，分析結果發現除了 A8 一題外(標準差為 1.07，但平均數為 3.92)，其他所有題目(39 題)的描述性統計平均數皆大於 3.50 且標準差小於 1.00，分析結果符合所設定的標準，故不需刪除題目；經內容分析與意見彙整後，「國小高年級學生行動學習準備度量表」各向度中，除了保留「學習偏好」向度之 E8 題，題數從 7 題增加為 8 題之外，其他向度的題目數維持不變，「使用科技能力」向度有 10 題、「學習態度」向度有 8 題、「溝通能力」向度有 6 題、「學習信念」向度有 9 題，共計五個向度 41 個題項；再者，依據德懷術停止標準，本研究之專家成員具有共識題項已累積達全體三分之二以上，具有一致性，以及專家成員對每個題項改變程度已小於或等於 15%，具有穩定性。因此，研究者整合專家意見後，形成國小高年級學生行動學習準備度量表，如表 15。

表 15

專家意見整合後之國小高年級學生行動學習準備度量表

層面	題目
A 使用 科技 能力	A1 我能使用平板電腦的筆記功能，記錄課程相關的內容。
	A2 我能使用平板電腦的拍照或錄影工具，記錄課程的相關內容。
	A3 我能使用平板電腦下載與老師所教或與課程相關的 App 軟體。
	A4 我能使用平板電腦，儲存網頁中與課程內容相關的圖片或文字的能力。
	A5 我能使用平板電腦的搜尋引擎，尋找課程內容所需的資訊。
	A6 我能使用平板電腦的文書處理軟體，例如記事本、備忘錄等。
	A7 我能運用平板電腦的 App 軟體(如：dropbox、google 雲端硬碟)，上傳資料或檔案給老師或同學。
	A8 我能運用平板電腦中的即時反饋系統(IRS)，立即回答學習的相關問題的能力。
	A9 我能適時操作平板電腦的功能，來符合課程需求。
	A10 我認為自己具備課堂上使用平板電腦所需的科技能力。
B 學習 態度	B1 我認為使用平板電腦上課，會是很愉快的學習經驗。
	B2 我認為使用平板電腦上課，我可以跟上老師的進度。
	B3 我認為使用平板電腦上課的過程中，我會主動完成老師指派的作業。
	B4 我認為使用平板電腦上課的過程中，我會專心的參與學習活動。
	B5 我認為使用平板電腦上課遇到問題時，我會儘快尋求老師或同學的協助。
	B6 我認為使用平板電腦上課遇到問題向別人求助時，我不會覺得很害羞。
	B7 我認為使用平板電腦上課時，我仍會耐心的學習。
	B8 我認為使用平板電腦學習時，自己的學習態度很重要。
C 溝通 能力	C1 我會使用平板電腦的溝通工具(如：line、google talk 等 App 軟體)，與老師、同學相互聯絡
	C2 我會使用平板電腦的溝通工具(如：line、google talk 等 App 軟體)，向別人尋求問題的解決
	C3 我會使用平板電腦主動參與小組的討論活動。
	C4 我會透過平板電腦的即時通訊工具(如：line、google talk 等 App 軟體)，協助同學解決問題
	C5 我認為具備使用平板電腦的溝通工具能力，能幫助我解決學習上的問題。
	C6 我認為自己目前所具備的應用平板電腦進行溝通的能力，能夠應付課堂上的需求。

(續下頁)

(接上頁)

層面	題目
D 學 習 信 念	D1 我相信使用平板電腦上課，能幫助我整合課程內容。
	D2 我相信平板電腦，對我的學習很重要。
	D3 我相信使用平板電腦上課，能立即解決課堂中學習的問題。
	D4 我相信平板電腦的即時溝通工具（如：line、google talk 等 App 軟體），是有效的溝通工具。
	D5 我相信平板電腦在學習過程中，扮演重要的角色。
	D6 我相信平板電腦攜帶方便的特性，有助於學習。
	D7 我相信使用平板電腦學習的過程，能提升我個人使用軟體的能力。
	D8 我會支持使用平板電腦來學習。
	D9 我相信平板電腦能幫助我更有效率地完成學習任務。
E 學 習 偏 好	E1 我喜歡使用平板電腦上課的方式。
	E2 我喜歡使用平板電腦來解決上課的問題。
	E3 我喜歡使用平板電腦與老師或同學溝通。
	E4 我喜歡使用平板電腦觀看動畫和影音進行學習。
	E5 我喜歡利用平板電腦進行小組合作學習。
	E6 我喜歡使用平板電腦進行線上蒐尋資料。
	E7 使用平板電腦上課，能滿足我學習的需求。
	E8 我覺得自己能適應將平板電腦作為學習的工具。

註：題號已重新編碼

伍、結論與建議

一、結論

- (一) 國小高年級學生行動學習準備度之內涵有使用科技能力、學習態度、溝通能力、學習信念、學習偏好等五大向度

本研究透過文獻探討，初步歸納國小高年級學生行動學習準備度之內涵，包括使用科技能力、學習態度、溝通能力、學習信念、學習偏好等五大向度；經過二回合德懷術問卷調查，調整部分向度之題項後，這些內涵和題目皆

獲得德懷術專家小組的共識。因此，本研究所提出的國小高年級行動學習準備度的內涵，不僅具有學理基礎，而且透過德懷術調查法獲得實證支持。

(二) 國小高年級學生行動學習準備度中各向度所包含之內涵具有周延性

本研究透過文獻探討，初步歸納在國小高年級行動學習量表中各向度所包含之主要內涵如下：1.使用科技能力：利用平板電腦等行動載具之學習過程中，學習者善用所具備的相關科技能力解決學習的問題，如文書處理能力、學習軟體安裝、網路溝通能力、資訊搜尋能力等，此研究發也呼應 Hussin 等人(2012)、Iqbal 和 Bhatti(2015)所提，認為學生擁有較高的技能準備度有助於使用行動學習裝置；2.學習態度：學習者本身對於利用平板電腦等行動載具進行學習，所持有的行為傾向或其內部反應的準備狀態，除了 Iqbal 和 Bhatti(2015)認為正向態度準備度對於行動學習助益，Cheon 等人（2012）的研究結果也顯示學習者的學習態度會影響行動學習；3.溝通能力：學習者在行動學習過程中，能使用電子郵件、社交網站或即時通訊軟體與他人溝通的能力。此與 Tubaishat 和 Lansari(2011)的研究結果說明溝通能力的重要性；4.學習信念：學習者本身利用平板電腦等行動載具學習的過程中，對於自我科技使用能力、自我問題解決能力、以及平板電腦等行動載具本身所抱持之態度，進而影響學習行為的看法。此與 Cheon 等人(2012)在行動學習準備度的研究結果相同的是，學習信念對於行動

學習成功的實施具有高度影響；5.學習偏好：學習者能在利用平板電腦等行動載具學習的過程中，選擇適合自己的學習方式，以完成學習任務；在 Christensen 和 Knezek(2017) 的研究也強調學習偏好對於行動學習的推動具有影響性。經過二回合德懷術問卷調查後，這些內涵大多具有適切性，並根據專家學者之意見進行部分內涵小幅度的修正。

(三) 國小高年級學生行動學習準備度中各向度所包含之內容具有相當之重要性

本研究所建構之國小高年級學生行動學習準備度內涵，在第一回合的德懷術中專家認為內涵重要性平均數達 4.49(採五點量表)；在第二回合的德懷術中內涵重要性之平均數達 4.54。表示經過二回合的德懷術問卷調查，專家小組的重要性回覆意見趨於一致性和穩定性，更顯示出此內涵的重要性。

二、建議

(一) 鼓勵推動行動學習之國小教師，先行瞭解與掌握「國小高年級學童行動學習準備度」之內涵

本研究建構國小高年級學童行動學習準備度之各向度和內涵，具有周延性與重要性，由於國內推動行動學習在各級學校的運作仍在鼓勵參與之初期階段，尤其是國民小學階段，若國小教師能先行瞭解高年級學生在行動學習方面的準備度，對其將行動學習融入於課程教學之中，將能適當的指導學生學習，而學習成效將更可期。

(二) 編製國小高年級行動學習準備度量表以評估行動學習準備度之現況

根據結論，本研究建構國小高年級學童行動學習準備度量表具有創新性與重要性，為擴大其實用性，建議未來可參酌本研究所建構之量表，於實施行動學習前可先行評估國民小學針對高年級學生推動行動學習之實際準備程度；同時，其結果亦可做為學校、教育主管單位推動行動學習之參考依據。這份量表編製宜先進行預試，瞭解題目的適切性，並加以修正，同時也要具備良好的信效度，才能真正測出學童行動學習準備度的現況。

(三) 持續研發適合中學學生行動學習準備度之量表以擴大行動學習準備度之現況研究

本研究以建構國小高年級學童行動學習準備度為主，由於中學教育目標、特性、學生基本能力與國民小學不盡相同，所以不能完全適用於中學學生。為了擴大瞭解國內學生行動學習準備度情況，可針對中學教育特性與需求，結合學生實際現況，建構適合中等教育學生之行動學習準備度量表。

(四) 規畫國小高年級行動學習準備度量表於國小場域進行驗證

本研究為加以確認國小高年級行動學習準備度量之信效度、五個向度中題目之適切性、以及教師與學生對於題目之理解程度，將進行量表的後續施測實證研究。

參考文獻

- 吳美美(2004)。數位學習現在與未來發展。《圖書館學與資訊科學》，**30**(2)，92-106
- 徐新逸、黃美蘭(2008)。數位時代數位學習素養之內涵分析。《教育資料與研究》，**80**，147-172。
- 國家發展委員會(2016)。105年個人家戶數位機會調查報告。取自 <https://itaiwan.moe.gov.tw/upfd/newsfiles/NS-170308-001100.pdf>。
- 張春興(2013)。《現代心理學(重修版)》。臺北：東華。
- 教育部(2008)。97年國民中小學九年一貫課程綱要。取自 http://teach.eje.edu.tw/9CC2/9cc_97.php
- 陳德懷(2009)。一對一數位學習的研究。《數位學習科技期刊》，**1**(2)，106-111。
- 趙美聲、陳鏗任、王玉蘭(2007)。數位學習準備度指標之發展。《T&D飛訊》，**64**，1-20。
- 蕭顯勝(2010)。學習用電子書載具導入準備度。取自 http://www.ecotw.com.tw/elq/e_book003.php。
- 蕭顯勝、蔡福興、游光昌(2005)。在行動學習環境中實施科技教育教學活動之初探。《生活科技教育月刊》，**38**(6)，40-57。
- Bernard, R. M., Brauer, A., Abrami, P. C., & Surkes, M. (2004). The development of a questionnaire for predicting online learning achievement. *Distance Education*, *25*(1), 31-47.
- Borotis, S., & Poulymenakou, A. (2008). *Critical success factors for e-learning adoption: Handbook of research on instructional systems and technology*. Greece: IGI Global.

- Bostrom, R. P., Olfman, L., & Sein, M. K. (1993). Learning styles and end-user training: A first step. *MIS Quarterly*, 17(3), 118–120.
- Canalys. (2012). *Smart phones overtake client PCs in 2012*. Retrieved from <https://www.canalys.com/newsroom/smart-phones-overtake-client-pcs-2011>
- Chapnick, S. (2000). *Are you ready for e-learning?* Retrieved from <http://blog.uny.ac.id/nurhadi>
- Chen, X. B. (2013). Tablets for informal language learning: student usage and attitudes. *Language Learning & Technology*, 17(1), 20-36.
- Cheon, J., Lee, S., Crooks, S.M., & Song, J. (2012). An investigation of mobile learning readiness in higher education based on the theory of planned behavior. *Computers & Education*, 59, 1054-1064.
- Cheung, S. K., Yuen, K. S., & Tsang, E. Y. (2011). *A study on the readiness of mobile learning in open education*, Proceedings of the IT in Medicine and Education (ITME), IEEE, 1, 133-136. 10.1109/ITiME.2011.6130801
- Christensen, R., & Knezek, G. (2017). *Readiness for integrating mobile learning in the classroom: Challenges, preferences and possibilities*. *Computers in Human Behavior*, 76, 112-121.
- Christensen, R., & Knezek, G. (2017). *Validating a Mobile Learning Readiness Survey: Assessing Teachers' Dispositions Toward Adoption*. *Journal of Digital Learning in Teacher Education*, 1-12.
- Clarke, A. (2004). *E-learning skills*. New York: Palgrave Macmillan.
- [com/newsroom/smart-phones-overtake-client-pcs-2011](https://www.canalys.com/newsroom/smart-phones-overtake-client-pcs-2011)

- Corbeil, J. E., & Corbeil, M. E. (2011). Are we ready for mobile learning now? 2007mobilelearningpredictions revisited. *Issues in Information Systems, XII*(2),142-152.
- Dalkey, N. (1969). An experimental study of group opinion: the Delphi Method. *Futures, 1*(5), 408-426.
- Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes*. Glenview, IL: Scott, Foresman.
- Dray, B. J., Lowenthal, P. R., Miskiewicz, M. J., Ruiz-Primo, M. A., & Marczyński, K. (2011). Developing an instrument to assess student readiness for online learning: a validation study. *Distance Education, 32*(1), 29-47.
- Faherty, V. (1979). Continuing Social Work Education: Results of a Delphi Survey. *Journal of Education for Social Work, 15*(1), 12-19.
- Franchak, S. J, Desy, J., & Norton, E. L. (1984). Involving business, industry, and labor: Guidelines for planning and evaluation vocational education programs. (Research and Development Series. No. 250). The Ohio State University at Columbus. The National Center for Research in Vocational Education.
- Geddes, S. J. (2004). Mobile learning in the 21st century: Benefit for learners. *The Knowledge Tree e-Journal, 30*(3), 214-228.
- Georgiev, T., Georgieva, E., & Smrikarov, A.(2004). M-Learning - a New Stage of E-Learning. International Conference on Computer Systems and Technologies CompSysTech2004, IV28, 1-5

- Guglielmino, P. J., & Guglielmino, L. M. (2003). "Are you ready for e-Learning?" In Piskurich, G. M., *The AMA handbook of e-learning: Effective design, implementation, and technology solutions*. New York: AMACOM.
- Guglielmino, L. M. (1977). *Development of the Self-Directed Learning Readiness Scale* (Unpublished doctoral dissertation). University of Georgia.
- Hamat, A., Embi, M. A., & Hassan, H. A. (2012). Mobile learning readiness among UKM lecturers. *Procedia-Social and Behavioral Sciences*, 59, 406-410.
- Holden, M. C. & Wedman, J. F. (1993). Future issues of computer-mediated communication: The results of a delphi study. *Educational Technology Research and Development*, 41(4),5-24.
- Huang, Y. M., Lin, Y. T., & Cheng, S. C. (2010). Effectiveness of a mobile plant learning system in a science curriculum in Taiwanese elementary education. *Computers & Education*, 54(1), 47-58.
- Hussin, S., Manap, M. R., Amir, Z., & Krish, P. (2012). *Mobile learning readiness among Malaysian students at higher learning institutes*. *Asian Social Science*, 8(12), 276-283. doi:10.5539/ass.v8n12p276
- Iqbal, S., & Bhatti, Z. A. (2015). *An investigation of university student readiness towards M-learning using technology acceptance model*. *The International Review of Research in Open and Distributed Learning*, 16(4), 83-102.

- Keegan, D. (2004). Mobile learning: the next generation of learning. *In The 18th Asian Association of Open Universities Annual Conference*, Shanghai.
- Keegan, D. (2005). The incorporation of mobile learning into mainstream education and training. *In Proceedings of the 4th World Conference on Mobile Learning*, Cape Town, South Africa, 25-28 October 2005. Retrieved from <http://www.mlearn.org/mlearn2005/>
- Khaddage, F., Muller, W., & Flintoff, K. (2016). Advancing mobile learning in formal and informal settings via mobile app technology: Where to from here, and how? *Educational Technology & Society*, 19(3), 16-26.
- Kolb, D. A. (1985). *Learning style inventory: Self-scoring inventory and interpretation booklet*. Boston: Mcber and Company.
- Koole, M. L. (2009). A model for framing mobile learning. In M. Ally(Ed.), *Mobile learning: Transforming the delivery of education and training*. Athabasca, AB: Athabasca University Press.
- Kruzich, J. M., Friesen, B. J., & Van Soest, D. (1986). Assessment of student and faculty learning styles: Research and application. *Journal of Social Work Education*, 22(3), 22-30.
- Kukulska-Hulme, A. G., & Traxler, J. (2005). *Mobile Learning: A handbook for educators and trainers*. New York: Routledge Falmer.
- Kynaslahti, H. (2003). In search of elements of mobility in the context of education. In H. Kynaslahti & P. Seppala (Eds.), *Mobile learning* (pp. 41-48).

- Lan Y. & Sie Y. (2010). Using RSS to support mobile learning based on media richness theory. *Computers & Education*, 55(2), 723-732.
- Lin, H.-H., Lin, S.-j., Yeh, C.-H., and Wang, Yi-Shun* (2016). Measuring Mobile Learning Readiness: Scale Development and Validation. *Internet Research*, 26(1), 265-287.
- Linstone, H. A., & Turoff, M. (1975). *The Delphi method: Techniques and applications*. Reading, Massachusetts: Addison-Wesley Publishing Company
- Mahat, J., Ayub, A. F. M. & Wong, S. L. (2012). An assessment of students' mobile self-efficacy, readiness and personal innovativeness towards mobile learning in higher education in Malaysia. *Procedia-Social and Behavioral Sciences*, 64(9), 284-290.
- McVay, M. (2000). *Developing a web-based distance student orientation to enhance student success in an online bachelor's degree completion program*. Unpublished practicum report presented to the Ed. D. Program. Florida: Nova Southeastern University.
- Nedungadi, P., & Raman, R. (2012). A new approach to personalization: integrating e-learning and m-learning. *Educational Technology Research and Development*, 60(4), 659-678.
- Ng, W. & Nicholas, H. (2012). A framework for sustainable mobile learning in schools. *British Journal of Educational Technology*, 43(1), 1-21.
- Ogamba, S., & Peters, I. (2016). Factors Affecting mobile Learning Readiness Among Students and Lecturers: A Model for Mobile

- Learning Readiness in Kenyan Universities. *INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY*, 15(1), 6408-6417. Retrieved from <https://cirworld.com/index.php/ijct/article/view/1705>
- Park, S. Y., Nam, M. W., & Cha, S. B. (2012). University students' behavioral intention to use mobile learning: Evaluating the technology acceptance model. *British Journal of Educational Technology*, 43(4), 592-605.
- Premsky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, 9, 1-6.
- Quinn, C. (2000). *mLearning. Mobile, Wireless, In-Your-Pocket Learning*. Retrieved from <http://www.linezine.com/2.1/features/cqmmwiyp.htm>
- Redmon, J., & Salopek, J. J. (2000). A year in the life of an e-learning project. *Training & Development*, 54, 36-41.
- Rosenberg, M. J. (2001). *E-learning: Strategies for delivering knowledge in the digital age*. New York: McGraw-Hill.
- Ross, M. F. (2001). Teacher's attitudes toward change and subsequent improvements in classroom teaching. *Journal of Educational Psychology*, 80(1), 111-117.
- Saekow, A., & Samson, D. (2011). E-learning Readiness of Thailand's Universities Comparing to the USA's Cases. *International Journal of e-Education, e-Business, e-Management and e-Learning*, 1(2), 126-131.

- Shieh, W. L. (1990). *Using the Delphi technique to determine the most important characteristics of effective teaching at junior high school level in Taiwan*. Unpublished doctor dissertation, University of Cincinnati.
- Smith, P. J. (2005). Learning preferences and readiness for online learning. *Educational Psychology*, 25(1), 3-12.
- The Economist Intelligence Unit. (2003). *The 2003 e-learning readiness rankings*. Retrieved from http://www.dokeos.com/doc/thirdparty/The_Economist.pdf
- The Economist Intelligence Unit. (2008). *E-readiness Rankings 2008, Maintaining Momentum: A White Paper from the Economist Intelligence Unit*. London, United Kingdom: Economist Intelligence Unit and The IBM Institute.
- The Economist Intelligence Unit. (2010). *Digital economy rankings 2010 Beyond e-readiness*. A report from the Economist Intelligence Unit. London, United Kingdom: Economist Intelligence Unit and The IBM Institute.
- Traxler, J. (2009). The Evolution of Mobile Learning. In R. Guy (Ed.), *The Evolution of Mobile Teaching and Learning* (pp. 1-14). California: Informing Science.
- Tubaishat, A. & Lansari, A. (2011). Are Students Ready to Adopt E-Learning? A Preliminary E-readiness Study of a University in the Gulf Region. *International Journal of Information*, 1(5), 210-215.

- Valtonen, T., Kukkonen, J., Dillon, P., & Väisänen, P. (2009). Finnish high school students' readiness to adopt online learning: Questioning the assumptions. *Computers & Education, 53*(3), 742-748.
- Wagner, E. D. (2005). Enabling Mobile Learning. *Educause Review, 40*(3), 40-53.
- Wang, Q., Zhu, Z., Chen, L., & Yan, H. (2009). E-Learning in China. *Campus-Wide Information Systems, 26*, 47-61.
- Warner, D., Christie, G., & Choy, S. (1998). *Readiness of VET clients for flexible delivery including on-line learning*. Brisbane: Australian National Training Authority.
- Watkins, R. (2005). Developing interactive-learning activities. *Performance Improvement, 44*
- Watkins, R., Leigh, D., & Triner, D. (2004). Assessing readiness for e-learning. *Performance Improvement Quarterly, 17*(4), 66-79.

教育學誌 第三十九期

2018 年 5 月，頁 55~85

學生對於數學教師教學看法與數學焦慮之 關聯性—以 PISA 2012 臺灣學生為例

陳姿螢

國立成功大學教育研究所碩士生

摘要

本研究資料以 PISA 2012 學生調查問卷為主，研究對象為臺灣 15 歲參加 PISA 2012 之學生，有效人數為 3976 人。以學生對於數學教師教學看法為自變項，數學焦慮為依變項，進行後續研究，目的為探討不同類型看法之學生對於數學焦慮之差異。本研究使用潛在類別分析，將學生對於數學教師教學看法分為三類，分別為：「低幫助低關懷取向」、「高幫助高關懷取向」、「高幫助低關懷取向」，而研究結果顯示「高幫助低關懷取向」看法之學生，其平均數學焦慮程度較高，而「高幫助高關懷取向」觀感之學生，其平均數學焦慮程度較低。希望能夠從此研究了解這兩者之關聯性，並能運用至教學現場中。

關鍵字：PISA、教師風格、數學焦慮、潛在類別分析

The Relationship between Students' Viewpoints of Math Teachers and Math Anxiety –Take PISA 2012 Taiwanese Students as an Example

Tzu-Ying Chen

Master student,

Institute of Education, National Cheng Kung University

Abstract

This research mainly focuses on the PISA 2012 student questionnaire, which targets 15-year-old Taiwanese students who joined the PISA 2012 as the research object, with an effective number of 3976 participants. Taking the students' viewpoints of math teachers as independent variables and mathematical anxiety as the dependent variable to do the following research, the purpose is to explore the differences among students who have different types of viewpoints on mathematical anxiety. This research uses Latent Class Analysis to classify students' viewpoints of math teachers into three categories: "low help low enthusiasm orientation," "high help high enthusiasm orientation" and "high help low enthusiasm orientation".

The results show that students who are treated by "high help low enthusiasm orientation" teacher have higher extent of mathematical anxiety on average, whereas students who are treated by "high help

high enthusiasm orientation" teacher have lower extent of mathematical anxiety on average. We hope to understand the relationship between the mathematical anxieties and the students' viewpoints of math teachers, and thus applied this finding to the teaching scene.

Keyword: PISA, Teaching Style, Math Anxiety, Latent Class Analysis

壹、緒論

一、研究背景與動機

一位教師的職業生涯中會遇到無數多個學生，然而一個學生的生命旅程中，卻可能僅僅遇過那一兩位改變他的教師，教師這個職業與其他行業相比要說偉大可能很多人不會那麼認同，但不可否認的是教師有一定存在的必要性及專業度。

數學是一項工具學科，為科學之母。在現行的十二年國教中，將國小中低年級數學課每周上課節數提升為 4 節；國際公認測驗 PISA 也將數學領域列入評比，由此可見數學的重要性。但在實際的教學場域中，數學卻是許多學生討厭的科目排名前幾名，雖然討厭數學這個科目卻礙於其他種種原因（例如：為了成績、升學、畢業等等因素）不得不努力學習數學，影響數學學習成效的原因有許多不同的層面，其中包括了：學生個人特質、學習態度、家庭背景、課程內容、教師教學風格和師生互動等等因素，然而現今學習數學時會產生不安焦慮的感覺，卻是不少學生都共面臨的問題。

Hadfield 與 Lillibridge (1991) 指出，國小求學階段的學習經驗會導致數學焦慮，這些經驗會降低學生的數學信心，致使進入國中階段後規避數學。另外，Jackson 與 Leffingwell (1999) 調查了學生受教育時期發生焦慮的原因，從中可看出數學焦慮與課堂內學生與教師的互動有直接關聯。然而，2010 年國際教育成就調查委員會（簡稱 IEA）主辦的「數學教師培育跨國研究」調查結果，我國中小學準教師在「數學知識」及「數學教學」的素質指標大幅領先美、德、俄、新加坡等國，但關懷、敬業態度的名次卻落後（陳至

中，2010)。準教師擁有領先他國的素質，卻較不具教學關懷，如何具有教學關懷，使用合宜的教學策略使學生獲得較佳的學習成效，是值得我們好好反思的。

每一位學生對於教師教學的觀感都不盡相同，即使提供一樣的資源，也有可能覺得得到不一樣的幫助，教師上課時所表現出的態度，每位學生所能感受到的也都不一樣。目前國內學生對於教師教學看法及各種對於教師帶給他們的幫助、教學成效等等，大多呈現於大學的教師教學評鑑中，雖然中小學也有推動教師專業評鑑，但卻缺乏學生對於教師教學看法、學生對於教師給予之幫助多寡的想法和學生知覺教師教學關懷此部分。

綜合以上之論述，具有數學焦慮的學生並非少數，而數學焦慮是影響數學學習成效的重要因素之一，師生互動與數學焦慮更是有直接的關聯性，因此研究者希望能藉由 PISA 2012 學生調查問卷中之題項分析不同學生對於數學教師的教學關感與其數學焦慮之關聯性，期望能將研究結果應用於教學現場中，瞭解學生們的想法，從根本原因降低他們對於數學的焦慮減少對於數學學習的排斥，因而本研究將針對學生對於數學教師教學觀感與數學焦慮做相關探討。

二、研究目的

本研究之目的旨在了解學生對於數學教師教學觀感與數學焦慮之現況，以及了解不同類型學生對於數學焦慮之差異性，並進一步探討學生對於數學教師教學觀感與數學焦慮之關聯。本研究研究問題如下：

- (一) 現今學生對於數學焦慮之概況？

- (二) 不同類型觀感之學生對於數學焦慮之差異情況為何？
- (三) 學生對於數學教師教學觀感與數學焦慮之關聯性為何？

貳、文獻探討

本章文獻探討共分兩節。第一節為教師風格及教師關懷；第二節為數學焦慮之定義及成因。

一、教師風格及教師關懷

(一) 教師風格

教學風格源於過去對教學效能之研究，將教師教學時的特徵進行組合，形成所謂的教學風格(silvernail, 1986)。教學風格除了是一種教學行為的偏好外，其本身無法獨立顯現，所以必須藉由教師在教學上表現個人特質來展現教師的教學風格(Heimlich & Norland, 2002)。

然而黃泓霖、劉唯玉(2013)指出對於教學風格學者們提出廣義與狹義兩種定義方式，狹義定義之學者(林生傳, 1988; Felder & Silverman, 1988; Gregorc, 1979)認為，教學風格僅指教師於指導過程中，顯露於外的教學行為或策略；廣義定義之學者(Grasha, 2002; Heimlich & Norland, 1994)則認為，教學風格應包含教師內在之個人特質，如信念、態度、人格特質、教育哲學及師生互動等層面，以及外顯之教學行為，如教學行為與方法、教學策略等，並呈現穩定一致的特性。

陳佩瑩（2009）指出從學生觀點知覺教師教學風格偏好，包含教師主導風格及小組引導風格，教師主導風格型為教師決定上課規範、方式與評量方式，即為教師掌握課堂的主權；而小組引導風格型，則是上課方式以小組方式進行，教師與學生會共同參與討論。

Grasha(2002) 將指導者分類為：專家型、權威型、個人典範型、促進型及授權型五種。

- 1.專家型：擁有豐富之知識與技術，並專注於傳遞資訊，會要求學生提昇能力來跟上自己，並要求其做學習的準備。
- 2.權威型：提供學生正面和否定的回饋，給予明確的指導，確立教學目標、期望及規定，讓學生有架構、方向的學習，強調學生應以正確、標準的方式來學習。
- 3.個人典範型：以個人經歷來指導學生，由自己建立思考與行為的典範，以示範方式指導學生並鼓勵學生去觀察與學習。
- 4.促進型：強調師生互動，透過詢問問題、探討與建議來指導學生，希望學生能獨立、主動且有責任感的去學習，並且在旁給予建議與鼓勵。
- 5.授權型：注重發展學生個人能力，營造自主學習的風氣，會鼓勵學生自行擬定學習計畫，教師角色為資源提供者。

然而研究者認為教學風格有其複雜性及多樣性，教師之教學最直接的影響就是學生，高雄市國小六年級學生，意識教師領導類型以「鼓勵」最常使用，學童知覺教師領導類型和學習態度有高相關，也就是說，國小教師領導類

型愈積極者，學生的學習態度愈佳，而學生良好的學習態度對教師的領導也有正面影響（李清榮，2005）。若教師對學生投入更多的關懷，學生也會更積極投入學習中，負面行為應會較減少，對於學習上的排斥也會有所降低，因此以下將對關懷進行相關探討。

（二）關懷之內涵

李校正（2007）提及關懷的內涵有 4 個：接納、同理、真誠和尊重。

- 1.接納：全然的接受，無論對方說了什麼、做了什麼都不予以批判的接納，並願意關懷、幫助他。
- 2.同理：站在對方的立場，試圖瞭解他人的感受及其內在世界，並試著讓對方之小你對他的感覺、想法和行為有所理解。
- 3.真誠：發自內心的。以真心、愛心、誠心、衷心誠意地關懷對方且不求回報，發揮人性本善的關懷。
- 4.尊重：接納他人，將其視為獨立的個體，每個人都有各自的感受與經驗，不應以自身經歷和觀點去批評，甚至影響對方。

邱士珊（2013）的文章中提及 Mayeroff 於 1971 年對關懷的本質為幫助他人成長，並涉及 8 種元素：瞭解、調整步調、耐心、誠實、信任、謙卑、希望、勇氣。

- 1.瞭解：瞭解可分為三種方式，包括語言、文字與只可意會的瞭解，關懷不光只是出於善意或給予溫暖，更需了解對方的需求並給予適當的回應。
- 2.調整步調：從廣義的角度去調整關懷的方式，付諸行動或不去做，皆是一種關懷的方式。

- 3.耐心:給予對方時間並按其步調成長，包括忍受某種程度的困惑和錯亂。
- 4.誠實:關懷時需誠實地看待一切，即為發自內心去關懷他人時必須說真話，如果只在意自己在對方眼中的形象，是無法完整地回應對方的需求。
- 5.信任:教師將關懷的對象是為獨立的個體，相信對方能以自身方式去成長，且從錯誤中學習。
- 6.謙卑:抱持謙卑的心讓我們隨時準備好且更願意瞭解對方與自己，並互相學習。
- 7.希望:關注本身所擁有的能力、特質可以在未來有何發展，在心中保持希望，相信自己未來發展無窮。
- 8.勇氣:當教師關懷學生，需有勇氣瞭解他的家庭狀況並與家長溝通，採取適當的處理方式。

綜合上述，關懷的內涵包括瞭解、尊重、接納、同理心、調整步驟，為內在的行為表現，關懷的內涵看似簡單，但當自己需發自內心去關懷他人時卻又不是那麼的容易。

(三) 關懷之目的

Mayeroff(1971)明確的指出關懷之目的在「幫助他人成長」，我（關懷者）與他（被關懷者）的關係是:不強迫他接受我的暗示，而是讓對方成長的方向引導我，幫助我決定該如何解決他的需要，同時留意如何展現適當的會應方式。然而 Noddings(1984)則強調關懷的最終目的是透過養育或教育的方式達成對方的自我實現與自我成長。

上述學者之觀點，關懷之目的為經由雙方互動，涉及想法或行為上的改變，促使被關懷者之成長。

二、數學焦慮之定義及成因

(一) 數學焦慮定義

有關數學焦慮的相關概念，最早是由 Dreger 和 Aiken 在 1957 年所提出的數字焦慮，意指學習者對運算算術時一種不安情緒反應，且多數的學習者對其它學科並不會如此，唯獨對數學產生不安與焦慮感（洪秀珍、謝臥龍、駱慧文，2015）。許多學者也提出對於數學焦慮的看法，敘述如下：

甯自強（1983）認為，數學焦慮指的是學生學習數學時，曾遭受挫折和衝突，使得學生自己本身在學習數學、做數學作業，尤其是參加數學科考試時產生的緊張狀態。

魏麗敏（1989）認為數學焦慮是一種狀態焦慮，是個人在處理數字或使用數學概念時，所產生的緊張、不安的情緒狀態。

Slavin(1991)認為數學焦慮乃指個體面對數學問題或接受數學測驗時所引起的情緒緊張或不安狀態。

吳明龍（1997）認為任無數學焦慮是個人在處理數字、使用數學概念、學習數學或參加數學考試時，所產生不安、緊張、畏懼等焦慮狀態。

Gresham(2007)認為數學焦慮是一種非理性恐懼數學的現象，包括無法去思考、學習數學，以及面對數學時有不舒服的感受。

不同學者間之定義略有些微差異，經歸納整理後數學焦慮之意涵為個人在運用數學概念、學習數學和解決數學相關問題時，所產生的不安、心情緊張、畏懼等負面情緒。而本研究將數學焦慮定義為面對數學之相關活動時，會產生焦慮不安等負面之情緒反應，即為數學焦慮。

（二）數學焦慮成因

甯自強（1983）認為數學焦慮的產生，主要是數學學習過程中有挫折出現，或無法融入學習而變成了旁觀者，以致於無法積極投入學習活動，或是自認為學習是沒有價值的。

Hilton（1980）提出數學焦慮和教師的權威式指導教學方法有關聯的想法。Mcleod（1989）認為教師在教學過程中沒有營造一個正向、支持與接納的學習氣氛，將會使學生產生焦慮與不安的情緒。Haladyna、Shaughnessy 與 Shaughnessy（1983）認為與教師的正面互動經驗是降低數學焦慮的重要因素之一，這些正面的經驗包括教師的熱情、對學生的尊重、幫助學生學習的承諾、留意每個人、公平地互動、讚美和加強。

林怡如（2004）認為學生對數學產生焦慮與恐懼的原因亦包括課程與教材內容的安排。此外，許斯琪（2001）指出造成國一學生數學焦慮的原因亦包括數學教材內容太難等。

Norwood (1994) 認為數學焦慮不是僅有單一原因，若學童沒有能力去處理失敗、過度的曠課、缺乏自我概念，且父母與教師對數學的態度是強調學習數學是藉由練習但卻缺少瞭解的觀念，皆有造成學童數學焦慮的可能。

PISA 學生問卷中焦慮部分呈述多為，感覺到焦慮、緊張、無助，或科目太難和怕表現糟等等，透過上述文獻之回顧可了解數學焦慮的原因可歸類為四大類：自我因素、家長因素、教師教學因素和教材內容因素，數學焦慮與數學學習成效有一定的影響性，無法了解造成焦慮的原因，即無法對症下藥解決學生對於數學的害怕，甚至可能因此形成惡性循環，因此本研究將針對教師教學因素與數學焦慮之關聯性做更深入的探討。

三、教師關懷與學生焦慮之關聯

在過往的研究中顯示學生知覺教師關懷、教師肯定與喜悅學業情緒有正相關(Jacob, 1996)。而許多學者也提出學生焦慮預防與解決之方法，其中不少建議提及教師關懷部分。涂金堂(1995)對於數學焦慮的預防之建議為：1.先從具體的教材內容著手；2.給學生持續的正向回饋；3.培養學生良好的數學態度；4.提供學生成功的數學學習經驗；5.避免臨時性的考試；6.加強數學科的補救教學。此外古明鋒(1997)亦指出預防數學焦慮產生的教學應注意：1.教師在教學中要依學生的認知發展水準，給予適當的教學，避免給予不當的過高期望，以避免因挫折而產生的數學無慮。2.在實際的教學中，輔導兒童從日常生活經驗中，獲得有關的數學知識，避免過

度強調算數技巧，或是機械的練習。亦即應避免為計算而計算，只注重答案的獲得，使得數學學習與生活脫節。3.適應個別差異，對不同能力的學生，提供其各自所需的學習時間，讓每一個學生都能達到精熟的學習的地步，亦即讓低成就的學生也能獲得成功的學習機會，減少因一連串的失敗經驗，而產生害怕與焦慮，在面對數學的學習時，能以較為積極和自信態度。4.寓教於樂，教材中應提供遊戲化的數學教學活動，在沒有威脅性學習環境中，藉以引發兒童主動學習數學的意願。

根據以上文獻我們能得知教師關懷與學生學習焦慮有一定程度的關聯性，學生知覺教師關懷其學業情緒較佳，也較學生知覺教師無關懷者，較不易產生焦慮的情況。而本研究採學生自主感受來判斷教師關懷。

參、研究設計與方法

本章主要說明本研究的研究方法與設計，總共分為四節。第一節主要描述研究步驟；第二節將介紹本研究之研究對象；第三節說明研究工具，主要為 PISA 2012 試題；第四節為資料分析，針對本研究所使用的統計分析方法進行說明。

一、研究步驟

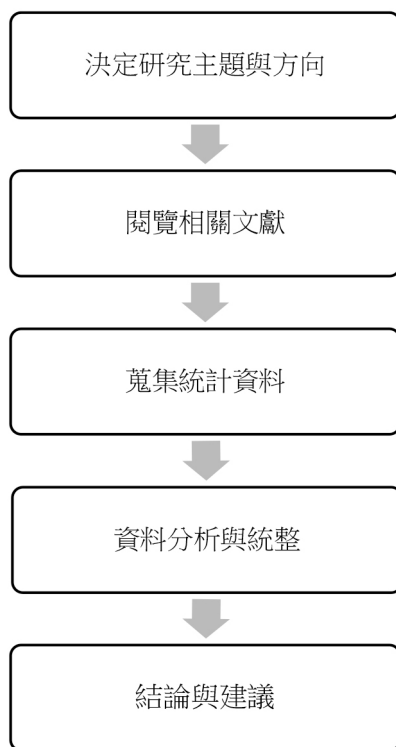


圖 1 研究步驟圖

本研究之研究步驟依序為：決定研究主題與方向、閱覽相關文獻、蒐集統計資料、資料分析與統整、結論與建議。

本研究的統計資料包括學生對於數學教師教學觀感與學生數學焦慮之問卷填答情形，問卷填答結果由 PISA 資料庫獲得。使用 M-plus 軟體進行潛在類別分析，以便適切的將學生進行分類，之後開始進一步使用 SPSS 軟體進行分析，用以檢視不同類別之學生對於數學焦慮之概況。依據分析與統整後的結果，撰寫研究結論與建議。

二、研究對象

每次 PISA 的測驗，各國家約有 4,500-10,000 名學生受測，學生來源係選自各公私立學校介於 15 歲 3 個月至 16 歲 2 個月的學生隨機抽選，並非以年級做選擇。

本研究之資料分析是來自 PISA 2012 臺灣地區資料，施測對象為臺灣九、十年級（15 歲）的學生，即國中三年級和高中（職）一年級，有效樣本人數為 3,976 人。

三、研究工具

本研究主要探討學生對於數學教師教學觀感與數學焦慮的相關研究，以下分別對學生對於數學教師教學觀感和數學焦慮的研究工具進行探討。

（一）學生對於數學教師教學觀感

學生對於數學教師教學觀感以 PISA 2012（試題 ST77）評量結果為主，PISA 量表分成五個題目，一個題目裡包含兩個選項偏向學生知覺數學教師教學較有幫助及較有熱情

的回答編碼為 2，而偏向較無幫助及較無熱情的回答編碼為 1。方便能利用此資料進行潛在類別分析，能獲得學生對於數學教師教學觀感的潛在模式。五個題目分別為：1.老師對每位學生的學習感興趣。2.當學生有需要時，老師會給予額外的幫助。3.老師會幫助學生學習。4.老師不斷教導，直到學生了解為止。5.老師給學生表達的機會。

表 1

PISA 2012 學生對於數學教師教學觀感題目及選項

題目	選項
老師對每位學生的學習感興趣	1.在所有課堂上 2.在大部分的課程上 3.在一些課堂上 4.從來沒有或幾乎不曾
當學生有需要時，老師會給予額外的幫助	1.在所有課堂上 2.在大部分的課程上 3.在一些課堂上 4.從來沒有或幾乎不曾
老師會幫助學生學習	1.在所有課堂上 2.在大部分的課程上 3.在一些課堂上 4.從來沒有或幾乎不曾
老師不斷教導，直到學生了解為止	1.在所有課堂上 2.在大部分的課程上 3.在一些課堂上 4.從來沒有或幾乎不曾
老師給學生表達的機會	1.在所有課堂上 2.在大部分的課程上 3.在一些課堂上 4.從來沒有或幾乎不曾

(二) 數學焦慮

數學焦慮以 PISA 2012 (試題 ST42) 評量結果為主，此題問題題組包含數學自信與數學焦慮兩部分，而本研究僅擷取數學焦慮部分。學生問卷填答的選項採用四點量表，分別為「非常同意」、「同意」、「不同意」、「非常不同意」，非常同意編碼為 4。同意編碼為 3，不同意編碼為 2，非常不同意編碼為 1，藉此了解學生焦慮程度。

表 2

PISA 2012 數學焦慮題目及選項

題目	選項
我常會擔心數學課會很困難	1.非常同意 2.同意 3.不同意 4.非常不同意
當我必須做數學作業時，我會很緊張	1.非常同意 2.同意 3.不同意 4.非常不同意
解數學問題時，我會很緊張	1.非常同意 2.同意 3.不同意 4.非常不同意
解數學問題時，我會很無助	1.非常同意 2.同意 3.不同意 4.非常不同意
我擔心數學拿不到好成績	1.非常同意 2.同意 3.不同意 4.非常不同意

四、資料分析

本研究使用 M-plus 軟體進行潛在類別分析，以便適切的將學生進行分類，獲得合適的學生對於數學教師教學觀感的潛在模式，之後開始進一步使用 SPSS 軟體進行分析，使用一般線性模式單變量分析，以檢視不同類別之學生感受對於數學焦慮之概況。

潛在類別分析是用來探討外顯變項背後的潛在變項，可以了解到樣本的潛在性特質，其目的是利用問卷的試題將學生分類到各個組別中。在探討潛在類別模式的指標有 AIC(Akaike information criterion)、BIC(Bayesian information criterion)、Adjusted BIC、Entropy、Adjusted-LRT(Adjusted Lo-Mendell-Rubin test)這幾個指標。而最佳的類別選擇，要兼顧其解釋力和參數估計值，則是 AIC 和 BIC 值越小越好，Entropy 其值越接近 1 代表正確性越高，簡言之其值則越大越好，在依各個類別的值，尋找最適合的指標來做進一步分析。

本研究使用電腦套裝軟體 SPSS 進行分析，將學生焦慮程度使用一個標準差進行分類，分別為低度焦慮、中度焦慮、高度焦慮，接著將上述的學生對於數學教師教學觀感潛在模式與焦慮程度做比對，以顯示不同類別的潛在模式對於數學焦慮程度的差異為何。

肆、研究結果

一、學生對於數學教師教學觀感潛在類別分析

學生對於數學教師教學觀感潛在類別模型如表 3，在選擇最佳的類別時 AIC、BIC 值要越小為佳，若樣本人數較大時，則 BIC 值比較適合用來選擇最佳的模式，結果顯示，由表 3 可知，2-class 到 3-class 的 BIC 值有明顯的下降，之後趨近平緩甚至有上升的現象，而 3-class 的 BIC 值是最小的，且 ABIC 值又較 4-class 高，ENTROPY 也較 5-class 高，相較之下較趨近於 1，因此，研究者以 3-class 為最佳模式。

表 3

學生對於數學教師教學觀感潛在類別分析模型摘要表

Model	AIC	BIC	ABIC	LRT	p	ENTROPY
2-class	18,020.49	18,089.66	18,054.71	3,480.173	0	0.847
3-class	17,580.14	17,687.04	17,633.02	443.437	0	0.702
4-class	17,546.36	17,690.98	17,617.9	44.883	0	0.745
5-class	17,551.06	17,733.41	17,641.26	7.153	0.3523	0.677

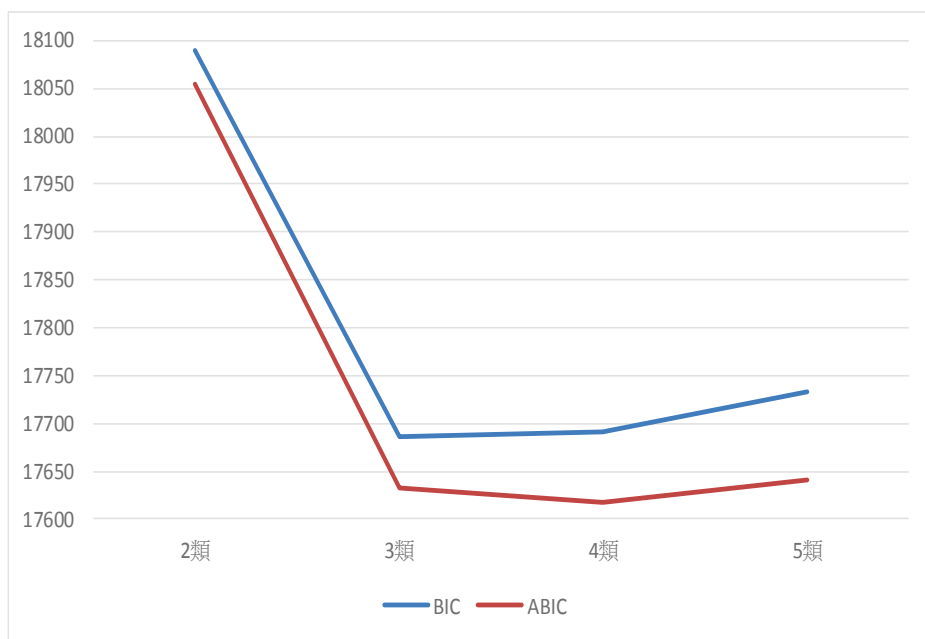


圖 2 學生對於數學教師教學觀感潛在類別 BIC、ABIC 指標線圖

表 4 顯示研究者將學生分為三類，而這三類學生在試題的作答反應機率，這三類學生的反應機率分別為 0.11、0.58、0.31，即為全部的學生約有 11% 被分到 class1，class1 的學生認為數學教師對於他們的學習是較不感興趣的，對於他們的學習幫助也較低，因此將這類學生對於數學教師的觀感命名為「低幫助低關懷取向」；大約有 58% 的學生被歸類到 class2，class2 的學生認為當他們有學習上的需要時，老師是會給予額外的幫助，老師也會很有耐心的教導他們，且給予他們表達的機會，因此將這類學生對於數學教師的觀感命名為「高幫助高關懷取向」；最後的 31% 學生，他們感受到數學教師會幫助他們學習，也會提供額外的幫助，但對他們的學習是較沒有興趣的，因此將這類學生對於數學教師的觀感命名為「高幫助低關懷取向」。

表 4

三類學生試題反應機率

試題	Class1	Class2	Class3
老師對每位學生的學習感興趣	0.05	0.80	0.30
當學生有需要時，老師會給予額外的幫助	0.05	0.99	0.80
老師會幫助學生學習	0.04	1.00	0.91
老師不斷教導，直到學生了解為止	0.06	0.96	0.51
老師給學生表達的機會	0.23	0.97	0.63
人數	445.00	2289.00	1,242.00
類別機率	0.11	0.58	0.31
觀感類別	低幫助低關懷	高幫助高關懷	高幫助低關懷

二、不同類型觀感學生與數學焦慮之關聯

對於不同觀感類別之學生，有效人數為 3,976 人，不同觀感類別學生對於數學焦慮表現趨勢如表 5 和圖 3，學生認為數學教師對於他們而言是「高幫助高關懷」，其平均數學焦慮最低，次之，學生認為數學教師是「低幫助低關懷」，其平均數學焦慮較「高幫助高關懷」高一些，最後，若學生認為數學教師為「高幫助低關懷」，其平均數學焦慮程度高於前述兩者。然而學生認為數學教師是「低幫助低關懷」，其標準差為 1.07，相較另外兩種觀感之學生，其焦慮程度差異較大。

表 5

不同觀感之學生數學焦慮平均數與標準差

學生對數學教師的觀感	平均值	標準差	N
低幫助低關懷	.31	1.07	445.00
高幫助高關懷	.30	.95	2,289.00
高幫助低關懷	.36	.89	1,242.00
總計	.32	.95	3,976.00

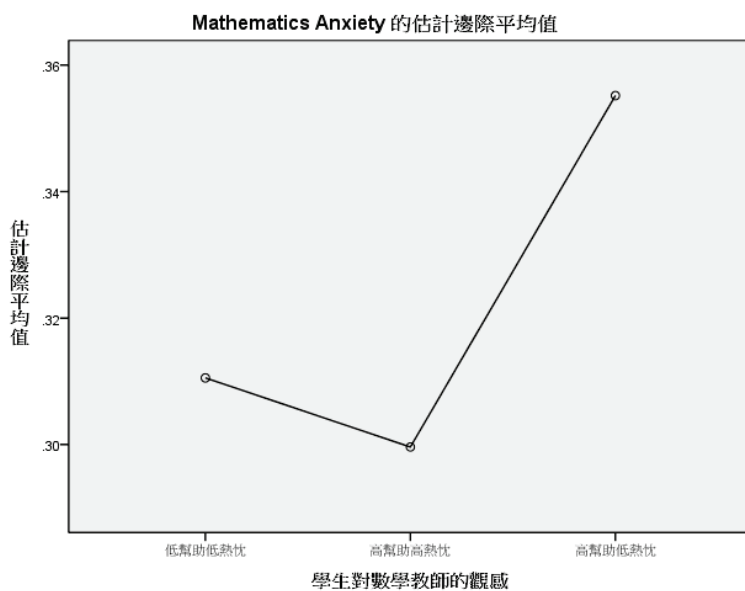


圖 3 不同觀感之學生數學焦慮折線圖

三、焦慮程度之差異

由於前述三項觀感分類雖有焦慮較高或較低之趨勢，但焦慮程度差異不大，因此研究者再進一步將數學焦慮做分類，分為：低度焦慮、中度焦慮、高度焦慮。由表 6 發現「低幫助低關懷」、「高幫助高關懷」、「高幫助低關懷」此三類觀感之學生大多屬於中度焦慮。圖 4 可更明顯觀察到「低幫助低關懷」感受之學生，偏向於低度焦慮或高度焦慮，「高幫助低關懷」與「高幫助高關懷」感受之學生偏向於中度焦慮。

表 6

學生對數學教師的觀感焦慮分組交叉列表

		焦慮分組			總計
		低度焦慮	中度焦慮	高度焦慮	
學生對數學教師的觀感	低幫助 計數	68	321	56	445
	低關懷 學生對數學教師的觀感內的 %	15.3%	72.1%	12.6%	100.0%
	焦慮分組內的 %	14.0%	10.5%	13.4%	11.2%
	佔總計的百分比	1.7%	8.1%	1.4%	11.2%
教師	高幫助 計數	288	1759	242	2,289
	高關懷 學生對數學教師的觀感內的 %	12.6%	76.8%	10.6%	100.0%
	焦慮分組內的 %	59.1%	57.3%	57.9%	57.6%
	佔總計的百分比	7.2%	44.2%	6.1%	57.6%
觀感	高幫助 計數	131	991	120	1,242
	低關懷 學生對數學教師的觀感內的 %	10.5%	79.8%	9.7%	100.0%
	焦慮分組內的 %	26.9%	32.3%	28.7%	31.2%
	佔總計的百分比	3.3%	24.9%	3.0%	31.2%
總計	計數	487	3,071	418	3,976
	學生對數學教師的觀感內的 %	12.2%	77.2%	10.5%	100.0%
	焦慮分組內的 %	100.0%	100.0%	100.0%	100.0%
	佔總計的百分比	12.2%	77.2%	10.5%	100.0%

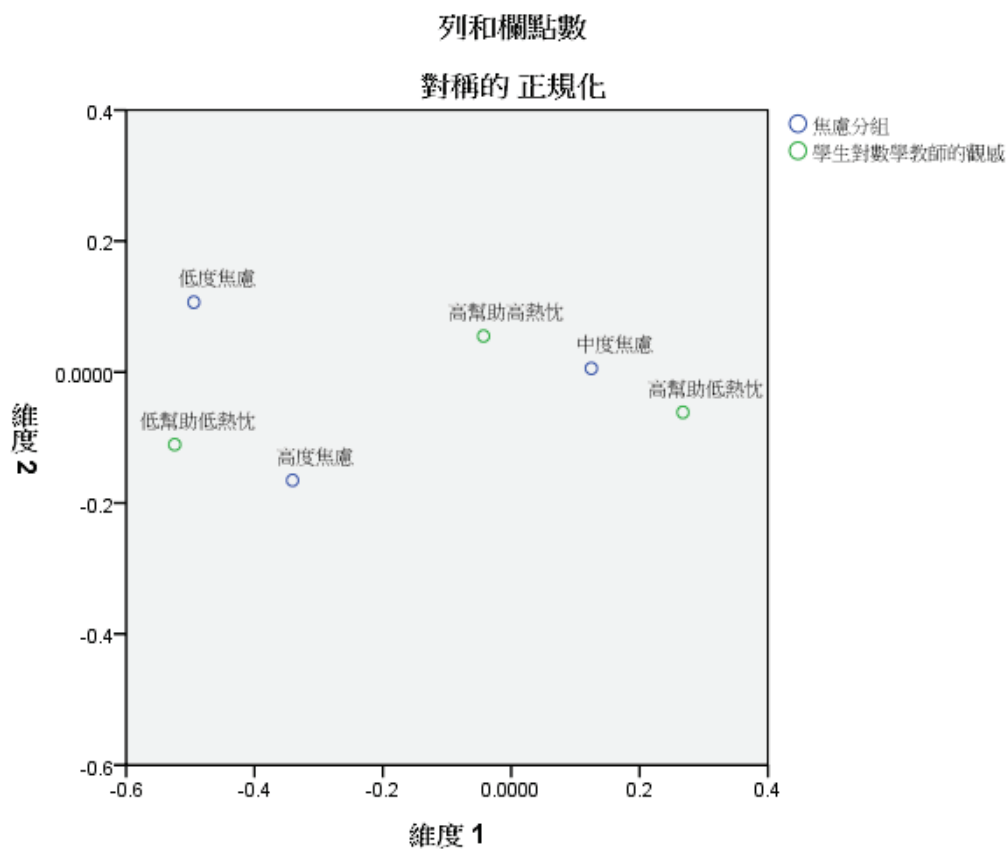


圖 4 三類觀感與焦慮分組圖

伍、結論與建議

本章共分為兩節，第一節為本研究之結論；第二節為本研究之建議，以利後續研究之參考與使用。

一、結論

本研究利用潛在類別分析將學生對於數學教師教學觀感進行分類，可以從中觀察到學生對於數學教師教學的想法，因此，利用潛在類別分析找出合宜的模式是有其意義的。

綜合以上研究的結果，學生對於數學教師教學觀感可以分成三個類別，分別為：「低幫助低關懷」、「高幫助高關懷」、「高幫助低關懷」。然而在這三種類別當中，「高幫助高關懷」的平均數學焦慮程度是較低的，「低幫助低關懷」平均數學焦慮程度有高有低，而「高幫助低關懷」的平均數學焦慮程度為較高。

由此研究可以發現，學生對於數學教師教學觀感類型不同，是會影響到其數學焦慮，所以學生對於數學教師教學觀感與數學焦慮是有關連性的。

二、建議

本研究發現，學生知覺到數學教師的教學態度有差異，對於學生本身的數學焦慮也會有所影響，若學生認為教師對於他們而言是有幫助且有關懷的，數學焦慮就會較低，但如果僅僅只是有幫助卻較無關懷，反而數學焦慮就會提升不少。

(一) 認為數學教師為低幫助低關懷

此類學生數學焦慮程度有高有低，推測其原因可能為，在一般課堂中他們並未投入到學習中，教師上課時可能因為跟不上進度，就開始沉浸在自己的世界，對於學習也較無興趣，認為自己就是學不好的學生，就算認真的上課還是聽不懂，因此認為教師對於他們是低幫助的，自己無心在於學習，進而投射到覺得教師也無教學熱情，數學焦慮程度不高可能為源自於自己可能已經放棄自己，認為學習對於自己而言不是重要的事，數學焦慮程度也就不會高，又或者是因為自己一直學不會，因此而造成高焦慮。

這類型學生給予教師的建議是，他們是很需要被關懷的孩子，不要心急按部就班的來，首先先找出他們的興趣所在，那怕是很簡單的小問題，也要先讓他們體會到成功的愉快，才會有後續繼續學習的動機，才不會一而再再而三地從學習中逃走。

(二) 認為數學教師為高幫助高關懷

此類學生可能為教室之中程度中上的孩子，由於數學上的表現相較之下較優異，成就感來自於教師上課時的鼓勵或是考試成績上的回報等等，亦或者是跟教師無太大過節，個性所趨，在作答時就會偏向認為教師對於他們是高幫助高關懷的，因為有對於教師的信心，相信在遇到困難時，教師都會給予他們幫助，且因為教師具有教學關懷，也不會因此就對他們感到不耐煩，數學焦慮程度也就最低。

這類型學生給予教師的建議是，持續保持他們對於教師教學的觀感，不可以因為他們焦慮程度最低，在教學方面就有所偏頗，教師用心的付出學生也真的有感受到，師生關係才不會越來越疏遠。

(三) 認為數學教師為高幫助低關懷

此類學生數學焦慮程度是最高的，或許是因為之前較不佳的學習經驗導致而成，學生可能認為教師上課時該講述的重點都有提到，但考試成績往往不如預想，而有問題卻又不敢提問，怕教師會為此對自己留下不好的印象，此時教師可能因為要趕進度或種種原因，而沒有發現此現象，就會造成惡性循環，學生不會但又不敢問，那不會的題目或觀念永遠都不會。由於學生認為教師真的有認真在授課，但似乎又沒有因此使自己成績有所提升，就會認為教師是高幫助低關懷的，認真的學習卻又拿不到自己理想的分數，也就是焦慮形成的原因所在。

這類型學生給予教師的建議是，了解學生真正的問題所在，拉近與學生之間的距離，讓學生覺得其實教師也是可以成為朋友的，也就是所謂的亦師亦友，有問題時教師都會教到學生懂，也不會感到不耐煩，使學生可以接近教師時可以不用這麼的害怕。

參考文獻

- 古明峰（民 86）。**數理教育：孩子為什麼害怕數學-談數學焦慮**。國教世紀，第 175 期，29-33。
- 吳明隆（1997）。**國小學生數學學習行為與其電腦焦慮、電腦態度關係之研究**。（為出版博士論文）。國立嘉義大學，嘉義縣。
- 李清榮（2005）。**高雄市國小教師領導類型、班級氣氛與學習態度之研究**。（未出版之碩士論文）。國立高雄師範大學，高雄市。
- 林生傳（1988）。**新教學理論與策略**。臺北市：五南。
- 林怡如（2004）。**綜合高中學生數學焦慮、數學自我效能與數學學業成就之關係研究**。國立雲林科技大學碩士論文，雲林縣。
- 邱士珊（2013）。**新店區國小教師對新移民子女關懷行動之個案研究—從學生學習面向來看**。國立臺北教育大學課程與教學研究所學位論文，1-167。
- 洪秀珍、謝臥龍、駱慧文（2015）。科技大學女學生「數學領域認同」、「數學性別刻板」、「性別角色刻板」、「情境訊息」與「數學焦慮」之研究。**科學與人文研究**，3(3)，30-54。
- 涂金堂（民 85）。「數學焦慮」學生的輔導。**諮商與輔導**，第 128 期，21-23。
- 許斯琪（2001）。**數學教師降低國中學生數學焦慮之行動研究**。國立彰化師範大學碩士論文，彰化縣。
- 郭倩琳、李校正、王碧玲（2007）。護理關懷課程之發展歷程。**護理雜誌**，54（4），53-60。
- 陳至中（2010）。我數學師資領先國際 關懷不足，**中國時報**。取自 http://tedsm.math.ntnu.edu.tw/news/20100417_3.pdf

- 陳佩瑩 (2009)。國小學童對自然科教師教學風格偏好模式之研究。未出版碩士論文，國立嘉義大學，嘉義縣。
- 陳冠芸 (2016)。初探學生評鑑教師教學之適當性。臺灣教育評論月刊，5(1)，201-203。
- 甯自強 (1983)。數學恐懼症的治療與預防。教育資料文摘，11(6)，119-127。
- 黃泓霖、劉唯玉 (2013)。我的麻吉實習輔導老師：實習輔導教師教學風格與實習教師學習風格關係之研究。師資培育與教師專業發展期刊，6(1)，1-24。
- 黃建智 (2014)。國民小學教師思考風格類型、情緒智力對正向管教作為之關係探討：以新北市五股地區為例。國立臺北教育大學教育經營與管理學系學位論，1-95。
- 魏麗敏 (1989)。國民中小學一般焦慮、數學焦慮、數學態度之比較研究。台中師院學報，5(1)，129-153。
- Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engr. Education*, 78(7), 674-681.
- Grasha, A. F. (2002). *Teaching with styles: A practical guide to enhancing learning by understanding teaching and learning styles*. Pittsburgh, PA: Alliance.
- Gregorc, A. F. (1979). Learning/teaching styles: Their nature and effects. In J. Keefe (Ed.), *Student learning styles: Diagnosing and prescribing programs* (pp. 19-26). Reston, VA: National.
- Gresham, G. (2007). *An Invitation into the Investigation of the Relationship between Mathematics Anxiety and Learning Styles in*

- Elementary Preservice Teachers. *Journal of Invitational Theory & Practice*, 13, 24-33.
- Hadfield, O. D., & Lillibridge, F. (1991). *A hands-on approach to the improvement of rural elementary teacher confidence in science and mathematics*. Paper presented at the Annual National Rural Small Schools Conference, Nashville, TN.
- Haladyna, T., Shaughnessy, J., Shaughnessy, M. (1983). A causal analysis of attitude toward Mathematics. *JRME*, 14 (1), 19-29.
- Heimlich, J. E., & Norland, E. (1994). *Developing teaching style in adult education. The Jossey-Bass higher and adult education*.
- Heimlich, J. E., & Norland, E. (2002). Teaching Style: Where Are We Now ? *New Directions for Adult and Continuing Education*, 93, 17-25.
- Hilton, P. (1980). Math anxiety: Some suggested causes and cures. *Two-Year College Mathematics Journal*, 11 (3), 174-188.
- Jackson, C. D. & Leffingwell, R. J. (1999). *The role of instructors in creating math anxiety in students from kindergarten through college*. *The Mathematics Teacher*, 92 (7), 583-586.
- Mayeroff, M. (1971). *On Caring*. New York: Harper Perennial.
- Mcleod, D. B., & Adam, V.M. (1989). *Affect and mathematical problem solving*. New York: Berlin Heidelberg.
- Noddings, N. (1984). *Caring: a feminine approach to ethics and moral education*. Berkeley: University of California Press.

- Norwood, K. S. (1994). The effect of instructional approach on mathematics anxiety and achievement. *School Science and Mathematics, 94*(5), 248-254.
- Peterson, K. D. (1995). *Teacher evaluation: A comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin Press, Inc. series. Retrieved from ERIC database. (ED375322)
- Silvernail, D. L. (1986). *Teaching styles as related to student achievement*. Washington, DC: National Education Association.
- Slavin, R. (1991). Achievement effects of ability grouping in secondary schools: A best evidence synthesis. *Review of Educational Research, 60*, 471-499.

教育學誌 第三十九期

2018 年 5 月，頁 87~151

四個試題反應模式的整體能力與領域能力估計精確性之比較研究

吳俊賢

桃園市桃園區中山國民小學教師

涂柏原*

國立臺南大學教育學系副教授

摘要

本研究旨在比較廣義次向度模式 (Brandt, 2012)、多向度隨機係數多項 logit 模式 (MRCMLM; Adams, Wilson, & Wang, 1997) 及高階試題反應模式 (HO-IRT; de la Torre & Song, 2009) 在各種模擬條件下對考生在測驗上的表現所估計得到的整體能力值及領域能力值估計的精確性。模擬資料時，領域向度數 (3、5 個)、各領域的題數 (20、40 題) 以及各領域能力之間的相關 (0.2、0.5 及 0.8) 等三個變項被操弄。在 12 個條件的組合之下，每一個皆重複模擬 30 次，每次模擬 1000 個考生的作答反應。這些資料以 MCMC 法分別以上述三個模式來估計考生的整體能力以及各領域的能力。以 $r_{\zeta\zeta}^2$ 、RMSE 及 BIAS 來評估各試題反應模式的整體能力及領

*通訊作者：涂柏原，聯絡方式：e-mail：bortwu@gmail.com

域能力估計的精確性。同時也以 98 年第一次國中基本學力測驗五個科目 1000 名考生的作答反應資料進行整體能力及領域能力的估計，並比較估計能力參數差異情形。

研究發現如下：

- (1) 模擬研究結果顯示，當領域向度數增加、各領域能力之間的相關係數變大，各領域內題數增加時，三個試題反應模式的整體能力和領域能力估計精確性均相當不錯。同時，領域能力估計值之間的相關亦均高於模擬時所設定的真值。
- (2) 以貝氏估計的 DIC 指標作為評估三種試題反應模式對 98 年第一次國中基本學力測驗資料的適配情形，結果發現廣義次向度模式的適配情形最好，其次為 MRCMLM 模式和 HO-IRT 模式。
- (3) 利用廣義次向度、MRCMLM 及 HO-IRT 等三種模式來估計 98 年第一次國中基本學力測驗的資料，發現三種模式所估計得到的領域能力之間的關介於.86 與.92 之間，屬於高度相關；三個模式的相關平均數分別為 0.88、0.90、0.90，而根據 DETECT 分析的結果，推測 98 年第一次國中基本測驗的資料應接近本質上的單一向度。

關鍵字：多向度隨機係數多項 logit 模式、高階試題反應模式、領域能力、廣義次向度模式、整體能力

Comparing the Accuracy of Examinees' Overall and Domain Ability Estimates of Four IRT Models

Chun-Hsien Wu

Teacher,

Taoyuan Municipal Chung Shan Elementary School

Bor-Yaun Twu*

Associate Professor,

Department of Education, National University of Tainan

Abstract

The purpose of this study is to compare the accuracy of overall and domain ability estimates given by four multidimensional item response models, including Generalized Sub-dimensional Model (GSM; Brandt, 2012), MRCMLM (Adams, Wilson, & Wang, 1997), and Higher-Order IRT Model (de la Torre, & Song, 2009). The numbers of domain abilities, and numbers of items in each domain as well as the size of correlation coefficients were manipulated, and under each combination of conditions, 1,000 examinees' response vector were generated to be a multidimensional data set with simple structure. Ability and item parameters were calibrated by using the MCMC algorithm. Additionally, the empirical data given by the 1000

*Corresponding Author: Bor-Yaun Twu ;E-mail: bortwu@gmail.com

examinees' responses on the 2009 Basic Competence Test (BCTEST) were also analyzed. The results are as the following.

1. In the simulation study, it was found that the accuracy of estimation of overall ability will be better when the numbers of domain abilities are larger, the correlation among the domain traits are larger, or there are more items in each domain. In addition, the correlation coefficients between estimated domain abilities are higher than those between the true abilities.
2. The values of the DIC index indicated that the GSM model fit the BCTEST data best.
3. The average correlation coefficients between the domain abilities were 0.88, 0.90 and 0.90 for the GSM, MRCMLM and HO-IRT, respectively. This result suggested that the BCTEST data is nearly essentially unidimensional.

Keywords: domain abilities, generalized subdimensional model, HO-IRT, MRCMLM, overall abilities

壹、緒論

心理與教育測驗的範疇中，常見一份測驗是由多個內容領域的分測驗所組成，例如全民英檢就包括了聽力、口說、閱讀理解以及寫作等四個分測驗，美國 ACT 的測驗科目英文、數學、科學和閱讀等四個科目。考生在該測驗上的表現，通常是以整份測驗中答對題數的總分或數個分測驗所得量尺分數加以平均來作為考生的加權總分。然而，如果將考生在各領域內容所得到的原始分數加以平均作為考生的整體能力，會忽略考生在不同內容領域上有不同原始分數的事實，況且各領域能力值間可能有相關存在。因此，本研究將利用模擬的測驗資料以及實徵資料，來探討不同的試題反應理論模式所得到的整體能力值及領域能力值之間的差異，比較各模式間能力估計的精確性以及模式資料之適配性，了解各試題反應模式對測驗總分計分的優缺點。

一、整體能力與領域能力參數估計之議題

測驗分析的首要目標在於透過有限的考生作答資料，能夠有效且可靠地估計考生的能力（包括描述考生整體能力及領域能力），以提供考生回饋。前項對考生在測驗試題表現的整體成就或精熟度提供完整的報告（Yao, 2010），可以作為教育決策的指標或入學門檻的標準；後者可代表考生在各內容領域上的表現強弱，作為補救教學或學習診斷的依據（Sinharay, Puhon, & Haberman, 2011）。

就實務上來說，國內外的大型測驗計畫都會將考生在各領域測驗(分測驗)的得分予以加權，來得到一個加權分數(weighted score)以作為考生在該份測驗上的整體表現（Peterson, Kolen & Hoover,

1989)。例如，美國法學院入學許可測驗（Law School Admission Test, LSAT）的總分是由閱讀理解、分析推理、與邏輯推理等三個分測驗分數的總和；CBT/McGraw-Hill 的 TerraNova 測驗之內容計有數學、閱讀、語言、科學以及社會學科等五大內容領域；施測對象為 1 至 12 年級學生，施測的結果報告在五個內容領域測驗上皆有領域量尺分數，考生在五個內容領域測驗上的得分之平均數就是整體量尺分數。ACT 測驗的總分亦為英文、數學、科學和閱讀等四個分測驗的領域量尺分數之平均數。此種對內容領域量尺分數進行加權計算所得到的測驗總分，稱為組合分數（composite score）。

關於組合分數議題的研究，在實務上最常被拿來計算組合分數的方法有信度加權法、效度加權法以及迴歸分析法等。而近年來以試題反應理論（item response theory, IRT）作為試題參數與考生能力參數分析在心理及教育測驗的應用上已是十分普遍。Rudner（2001）在討論組合分數的成份加權議題時，將計算組合分數的方法分為兩大取向：內隱取向（implicit approach）與外顯取向（explicit approach），內隱取向包括原始分數加總（adding raw scores）及 IRT 兩種；而外顯取向包括難度加權（weighting by difficulty）、信度加權（reliability weighting）和效度加權（validity weighting）等。

上述之組合分數估算的方法除了 IRT 理論中能力參數的估計是由非線性方式得到的，其他方法都是採用線性的方法。單向度 IRT 模式估計參數時，基本上已經同時解決「加權」和「量尺化」的問題，所估計得到的考生能力（ability）值參數 θ 已是個最佳解（涂柏原、盧思丞，2012; Childs, Elgie, Gadalla, Traub, & Jaciw, 2004; Wainer & Thissen, 1993）。比如採單向度二參數 IRT 模式（2PL）進行能力參數估計時，無需再考慮權重的問題，因為 2PL 模式中的鑑別度（ a 參數）是計算加權組合分數時的最佳權重。

縱使單向度 IRT 模式有這個優勢，但包含不同內容領域的測驗在內容領域專家和學者的見解中，本質上是多向度的 (multidimensional)，然而實務上在估計考生能力參數時，整份測驗卻常被視為是單一向度的 (unidimensional) 或是多重單一向度的 (multi-unidimensional)。僅用單向度 IRT 模式來估計考生的能力，會忽略考生領域能力值之間的相關，而對能力參數估計的精確性產生影響；而多向度試題反應模式 (multidimensional item response theory, MIRT) 可將領域向度間的相關性納入估計程序中，提昇各領域向度能力估計的精確性 (de la Torre & Song, 2009)。

過去幾年來可以同時計算考生整體能力值和領域能力值的多向度試題反應模式陸續被發展出來，Sheng 與 Wickle (2008) 基於整體能力與領域能力不同線性組合觀點提出的兩個二階常態肩型多向度試題反應模式，可以估計試題鑑別度參數、難度參數，並可同時估計考生整體能力與領域能力。Brandt (2008, 2012) 針對測驗理論中單向度與多向度的議題以為及降低測驗試題間產生局部試題依賴 (local item dependence, LID) 對考生能力參數估計準確性造成影響，提出次向度模式 (Rasch subdimensions model) 以及更為一般化的廣義次向度模式 (generalized subdimension model)，除了可以估計考生在整份測驗上的整體能力值外並可將考生整體能力轉換至各個領域向度上計算出領域能力值。de la Torre 及 Song (2009) 認為考生的潛在能力間會有階層關係存在；個人在測驗中的特定領域表現受到領域能力影響，多個領域能力則是被單一的高階能力所解釋，此高階能力就是考生整體能力，因此高階試題反應模式也同時可以估計考生整體能力值及領域能力值。郭伯臣、謝典佑、吳慧珉、林佳樺 (2012) 曾探討當資料具備一因子高層次 IRT

模式結構時，誤用單一向度 IRT 模式估計整體分數或用 MIRT 模式估計領域分數之影響。但針對這些 MIRT 模式估計考生整體能力、領域能力精確度進行比較的研究仍屬少數。

二、試題反應理論相關模式

試題反應理論（IRT）主要概念為考生在一個測驗試題上的表現可以用一個連續且單調遞增的數學函數來描述考生的表現，此函數稱為試題特徵曲線（item characteristic curve, ICC）。這個函數通常包含一個或多個參數來描述考生的潛在特質，以及一個或多個參數來描述試題的特性，且考生的潛在特質與試題的難度參數是在相同的連續量尺上。如果考生的潛在特質只有一個，依試題參數的個數可分為單參數模式或 Rasch 模式（Rasch, 1960）、二參數模式與三參數模式（Birnbaum, 1968; Lord, 1952）。

（一）多向度試題反應模式

當測驗的主要目的是獲得考生單一能力時，使用單向度 IRT 模式進行能力估計是很自然的。然而人類心理運作過程卻是十分複雜，而且許多試題所需的知識或技能，非考生單一向度的能力所能滿足，需整合多個向度的能力才可順利解題；在這種情形下，採用單向度 IRT 模式進行能力參數估計就顯得不實際。

多年來 MIRT 模式的發展已有許多成果（Adams, Wilson, & Wang, 1997; Christoffersson, 1975; McDonald, 1997; Reckase & Mckinley, 1991; Rijmen & Briggs, 2004; Sympson, 1978; Whitely, 1980）。Reckase（1985）提出多

向度二參數模式 (multidimensional two-parameter logistic model, M2PL), 其公式如下：

$$P(Y_{ni} = 1) = \frac{\exp(\sum_{k=1}^m a_{ik}\theta_{nk} + d_i)}{1 + \exp(\sum_{k=1}^m a_{ik}\theta_{nk} + d_i)} \quad (1)$$

其中, $Y_{ni}=0$ 或 1 代表考生 n 在試題 i 的得分, a_{ik} 為試題 i 在第 k 個向度上的鑑別度參數, d_i 為一純量, 是與試題 i 的難度有關的參數, θ_{nk} 為考生 n 在第 k 個向度上的領域能力。這個模式一般稱為補償性模式, 原因是指數部分為考生在各向度的能力值與鑑別度相乘加總後, 再與難度參數作加法運算, 因此若考生的某一領域能力較低時, 可透過其他領域之能力值給予補償來提昇答對該題的機率。

相對於補償模式, 另一類的模式為非補償性模式 (或稱為部分補償模式), 是 Sympson (1978) 所提出的, 為單向度三參數試題反應模式的多向度化, 其公式可表示為：

$$P(Y_{ni} = 1) = c_i + (1 - c_i) \prod_{k=1}^m \frac{\exp[a_{ik}(\theta_{nk} - b_{ik})]}{1 + \exp[a_{ik}(\theta_{nk} - b_{ik})]} \quad (2)$$

由公式 (2) 可看出非補償性多向度試題反應模式可視為多個單向度試題反應模式機率相乘積, 使得各向度能力間幾乎沒有補償效果存在。

Adams、Wilson 與 Wang (1997) 以 Rasch 模式為基礎，提出一個更廣義的多向度隨機係數多項羅吉模式 (multidimensional random coefficients multinomial logit model, MRCMLM)。如果在考生的反應底下需要一組 D 個特質始能答對試題，則其模式公式如下：

$$P(Y_{ik} = 1) = \frac{\exp(\mathbf{b}_{ik}\boldsymbol{\theta} + \mathbf{a}'_{ik}\boldsymbol{\xi})}{\sum_{k=1}^{K_i} \exp(\mathbf{b}_{ik}\boldsymbol{\theta} + \mathbf{a}'_{ik}\boldsymbol{\xi})} \quad (3)$$

其中， $Y_{ik} = 1$ 表示考生在第 i 題選擇第 k 個選項， K_i 為第 i 題的計分類別數； \mathbf{b}_{ik} 為第 i 題在第 k 個反應類別上的計分向量 (scoring vector)； $\boldsymbol{\theta}$ 代表考生能力參數向量； \mathbf{a}_{ik} 為第 i 題在第 k 個反應類別上的設計向量 (design vector)； $\boldsymbol{\xi}$ 為試題參數向量。因為 MRCMLM 是一個廣義的模式，多數 Rasch 族系的模式如 Fischer (1973) 的羅吉斯潛在特質模式 (LLTM)、Andrich (1978) 評定量表模式 (RSM)、Masters (1982) 部分給分模式及 Rasch 題組反應模式皆屬於 MRCMLM 的特例。

(二) 高階試題反應模式

de la Torre 及 Song (2009) 提出高階試題反應模式 (one-factor higher-order item response theory, HO-IRT)，該模式的主要概念認為考生的潛在能力間有階層的關係存在。個人在測驗中的特定領域表現受到領域能力 (domain ability) 所解釋，而領域能力間的關係則為一個整體能力

(overall ability) 所解釋，這個高階能力就是考生的整體能力，因此高階試題反應模式包含整體能力以及領域能力。

一份測驗若包含 k 個領域， $k = 1, 2, 3, \dots, K$ ， $\theta_n^{(k)}$ 表示考生 n 在領域 k 上的能力值；若各個內容領域測量相同的能力，則整份測驗就是單向度測驗；若不同領域的能力間有相關存在，他們之間的相關就由高一階的整體能力所解釋，。因此可將 HO-IRT 模式視為多重單一向度模式，其整體能力 $\theta_n \sim N(0,1)$ 。

HO-IRT 模式中第一層為第 n 個考生在領域內容 k 第 i 題的作答反應，以 $X_{ni}^{(k)}$ 來表示；第二層為考生的作答反應透過 IRT 模式連結到領域量力 $\theta_n^{(k)}$ 及試題參數 β 上；第三層為考生的領域能力透過迴歸參數 $\lambda^{(k)}$ 連結到整體能力上。也就是

$$\theta_n^{(k)} = \lambda^{(k)} \theta_n + \varepsilon_{nk} \quad (4)$$

就數學上看來，公式 (4) 中的 $\lambda^{(k)}$ 的值可以為負數，但在測驗應用上，領域能力與整體能力之間的相關是正的，因此 $0 < \lambda^{(k)} \leq 1$ 。 ε_{nk} 為誤差項，與其他誤差項間為彼

此獨立，假設具有平均數為 0，變異數為 $1 - \lambda^{(k)2}$ 之常態分配，亦即，

$$\varepsilon_{id} \sim N(0, 1 - \lambda^{(k)2}) \quad (5)$$

在整體能力已知的情況下，領域能力彼此相互獨立，領域能力 k 與 k' 之間的相關為 $\lambda^{(k)}\lambda^{(k')}$ ，而迴歸參數的限制也使得整體能力與領域能力具有相同的能力量尺。關於參數估計部份，de la Torre 與 Song (2009) 描述了試題參數已知時，同時估計整體能力、領域能力和迴歸參數 $\lambda^{(k)}$ 的結果，若迴歸係數 $\lambda^{(k)}$ 已知，可以採取最大概似估計法來估計考生整體能力及領域能力；然而實際情況中 $\lambda^{(k)}$ 也需要估計，模式也更加複雜，因此他們採用 Markov Chain Monte Carlo (MCMC) 方法來進行參數估計。在階層貝氏架構下，HO-IRT 模式可以表示為：

$$\theta_n \sim N(0, 1) \quad (6)$$

$$\lambda^{(k)} \sim Uniform(-1.0, 1.0) \quad (7)$$

$$\theta_n^{(k)} | \theta_n, \lambda^{(k)} \sim N(\lambda^{(k)}\theta_n, 1 - (\lambda^{(k)})^2) \quad (8)$$

(三) 廣義次向度模式

Brandt (2008) 針對測驗理論中單向度與多向度的議題以及為要降低試題間產生局部試題依賴 (LID) 對測驗準確性造成的影響，提出了 Rasch 次向度模式 (Brandt, 2008)，該模式基本假定為每個考生在測量的向度上有一個

主要的能力 (general ability) (以下稱為整體能力), 以及在領域向度 (或稱次向度) 上有著或強或弱的領域能力 (specific abilities)。這個模式的公式如下：

$$P(Y_{ni} = 1) = \frac{\exp(\theta_n + \gamma_{nk(i)} - b_i)}{1 + \exp(\theta_n + \gamma_{nk(i)} - b_i)} \quad (9)$$

其中 $P(Y_{ni} = 1)$ 表示考生 n 正確回答試題 i 的機率, b_i 為試題難度, θ_n 為考生整體能力, $\gamma_{nk(i)}$ 為考生 n 的能力在特定領域 k 上之相對強弱, $k(i)$ 表示試題 i 屬於次向度 k , 而 $\gamma_{nk(i)} = d(i)\gamma_{nk}$, 當試題 i 所測量的內容是屬於領域向度 k 時, 則 $d(i) = 1$, 否則 $d(i) = 0$ 。

實際上測驗內容經常是多向度的, 但在能力估計及成績報告時, 經常假定測驗是單向度的情況, 為解決此問題以及因測驗內容而產生的局部試題依賴, Brandt (2012) 將 Rasch 次向度模式延伸為廣義次向度模式, 其模式為：

$$P(y_{ni} = 1) = \frac{\exp[d_{k(i)}(\theta_n + \gamma_{nk(i)}) - b_i]}{1 + \exp[d_{k(i)}(\theta_n + \gamma_{nk(i)}) - b_i]} \quad (10)$$

廣義次向度模式與 Rasch 次向度模式的差異在於多了 $d_{k(i)}$ 這個轉換參數 (translation parameter), 其作用為將分測

驗量尺能力轉換至共同量尺上， $d_{k(i)}$ 表示試題 i 測量第 k 個領域。為了模式辨認的問題，需要加上以下四項限制：

$$\sum_{k=1}^K \gamma_{nk} = 0, \text{ 對所有的 } n = 1, 2, \dots, N \quad (11)$$

$$\text{Cov}(\theta_n, \gamma_{nk}) = 0, \text{ 對所有的 } k = 1, 2, \dots, K \quad (12)$$

$$\sum_k d_{k(i)} = K \quad (13)$$

每一個次向度的平均試題難度為 0，或是整體能力和每一個次向度能力之平均數為 0。 (14)

廣義次向度模式中各個領域能力定義為 $\theta_{nk} = d_k(\theta_n + \gamma_{nk})$ ，由公式 (13) 所推導出的整體能力等同於各個領域能力轉換至共同量尺上（除以 d_k ）的平均，如公式 (15) 所示。

$$\begin{aligned}
\theta_n &= \theta_n + \frac{\sum_k \gamma_{nk}}{K} && \text{(根據公式 11, } \sum_k \gamma_{nk} = 0 \text{)} \\
&= \frac{K\theta_n}{K} + \frac{\sum_k \gamma_{nk}}{K} \\
&= \frac{\sum_k (\theta_n + \gamma_{nk})}{K} \\
&= M((\theta_{n1} + \gamma_{n1}), (\theta_{n2} + \gamma_{n2}), \dots, (\theta_{nk} + \gamma_{nk})) \\
&= M\left(\frac{\theta_{n1}}{d_1}, \frac{\theta_{n2}}{d_2}, \dots, \frac{\theta_{nk}}{d_k}\right) \tag{15}
\end{aligned}$$

其中 $M(\cdot)$ 表示括號內各引數之平均數， θ_{nk}/d_k 為相對應的能力估計轉換至單向度變異數成份，此時各個領域能力的變異數就都會相等。

廣義次向度模式中次向度間的共變數如公式(16)所示，

$$\begin{aligned}
Cov(\theta_{n1}, \theta_{n2}) &= Cov[d_1(\theta_n + \gamma_{n1}), d_2(\theta_n + \gamma_{n2})] \\
&= d_1 d_2 Cov(\theta_n + \gamma_{n1}, \theta_n + \gamma_{n2}) \\
&= d_1 d_2 [Cov(\theta_n, \theta_n) + Cov(\theta_n, \gamma_{n2}) + Cov(\theta_n, \gamma_{n1}) + Cov(\gamma_{n1}, \gamma_{n2})] \\
&= d_1 d_2 [Var(\theta_n) + Cov(\gamma_{n1}, \gamma_{n2})] \tag{16}
\end{aligned}$$

公式(16)倒數第二行中， $Cov(\gamma_{n1}, \theta_n) = Cov(\gamma_{n2}, \theta_n) = 0$ ，由此可見 Rasch 多向度試題反應模式各個向度間的共變數與廣義次向度模式的一樣。

Brandt 與 Duckor (2013) 將廣義次向度模式延伸為「廣義次向度多元計分模式」，使得模式可以處理多元計分的資料。他們以「教室評量素養量表」(classroom assessment literacy, CAL; Duckor, 2006; Duckor, Draney, & Wilson, 2009) 的資料利用廣義次向度多元計分模式進行分析，發現廣義次向度模式對資料的適配優於 Rasch 單向度試題反應模式，同時廣義次向度模式所估計考生整體能力之標準誤低於 Rasch 單向度模式，降低的幅度約為 1.8%。

三、整體能力與領域能力研究的相關議題

Gulliksen (1950)、Wang 與 Stanely (1970)、Lord 與 Novick (1968)、Feldt 和 Brennan (1989) 以及 Wainer 與 Thissen (1993) 等學者曾詳細地探討組合分數計算的方法與相關的議題，整體能力的估計其實也是組合分數的計算，然而在本文中，將僅聚焦在與 IRT 有關聯的部分。

雖然可以同時計算整體能力與領域能力的 IRT 模式陸續被發展出來，然而這些模式整體能力與領域能力估計的精確性比較之研究仍屬少數。Yao (2010) 以模擬產生的資料，探討以單向度三參數試題反應模式 (3PL)、高階試題反應模式 (HO-IRT)、題內多向度試題模式 (MIRT) 及雙因素模式 (bi-factor model) 等四個模式在整體能力及領域能力上表現情況；並以均方根誤差 (root mean

squared error, RMSE)、絕對偏誤(absolute bias, ABS)及偏誤(BIAS)來比較估計結果。Yao 所操弄的模擬條件包含樣本數(500、1000、2000 人)、領域能力間的相關(0、0.3、0.5、0.7、0.9)，試題參數是由某一個大型的八年級數學成就測驗所取得，該測驗包含四個內容領域，各內容領域的題數分別為 15 題、15 題、12 題與 18 題；Yao 分別以高階試題反應模式和題內多向度試題反應模式來模擬作答反應，各種條件組合下均模擬 20 次。

Yao 發現在試題參數及領域能力參數的回覆性上，MIRT 模式在所有模擬情境上表現均略優於 HO-IRT 模式。當考生領域能力之間為低度相關時，整體能力在參數回覆性上，利用最大訊息法(maximum information method)估計考生整體能力的 MIRT 模式和 HO-IRT 模式表現優於 3PL 模式及 bi-factor 模式；而領域能力間相關大於 0.8 時，四種試題反應模式的表現皆相當接近。根據其研究的結果，Yao (2010) 指出如要採用 MIRT 模式或 HO-IRT 模式來報告整體能力及領域能力且要求測驗信度大於 0.8 時，領域能力間的相關必須大於 0.7 或是領域能力與整體能力間的相關要大於 0.9。

de la Torre 及 Song (2009) 以 HO-IRT 模式為「真實模式」，探討不同領域向度數、領域向度的題數、領域向度之間相關程度等三個條件對試題反應模式估計值的影響。參數估計採是 MCMC 演算法，而用來模擬作答反應資料的試題參數是由一個國際標準數學題庫所選出的。他們所操弄的條件包括：能力向度(2、5 個)、題數(10、20 及 50 題)、領域能力之間的相關(0、0.4、0.7 及 0.9)，評估的標準為 Pearson 相關及均方誤差(mean square error, MSE)。他們發現在整體能力的部分，整體能力的真值與估計值之間的相關會隨著向度間相關的增加、題數增加而變大，而整體能力後驗變異

數的估計，3PL 試題反應模式所得到的結果較 HO-IRT 模式為佳。

在領域能力的估計方面，當領域向度之間的相關變大時，領域能力的真值與估計值之間的相關也隨之變大(由 0.82 至 0.96)。de la Torre 和 Song (2009) 指出相較於 3PL 模式，HO-IRT 模式能提供較佳的領域能力估計，但若領域能力間相關過低，則參數估計改善的程度並不大。在領域能力後驗變異數方面，當固定其他變項而增加測驗長度時，估計的精準度會越高，而只有在分測驗的相關為 0.7 以上時，領域向度數的增加才會使得後驗變異數的精準度提升。

四、研究目的與問題

(一) 研究目的

由前面的敘述可知，本研究分為模擬資料與實徵資料兩部分，主要的研究目的如下：

- 1.以多向度測驗結構產生模擬資料，探討廣義次向度模式、MRCMLM 模式與 HO-IRT 模式等三種模式的能力估計的精確性。
- 2.在貝氏估計的架構下，探討廣義次向度模式、MRCMLM 模式與 HO-IRT 模式等三個 IRT 模式分析 98 年第一次國中生基本學力測驗之適配情形。
- 3.以 98 年第一次國中基本學力測驗實徵作答資料的分析結果，探討廣義次向度模式、MRCMLM 模式、HO-IRT 模式等三種模式所估計得到的考生領域能力值之間的相關情形，同時檢測國中生基本學力測驗的多向度程度。

(二) 研究問題

根據上述研究目的，本研究將探討以下的研究問題：

- 1.不同的領域向度數、各領域向度題數、及領域能力之間的相關等條件組合之下，所產生具有簡單結構的多向度模擬資料，利用廣義次向度模式、MRCMLM 模式與高階試題反應模式等三種模式來進行分析，所估計得到的考生整體能力及領域能力的精確性為何？
- 2.在貝氏統計的架構下，三個試題反應模式對實徵資料的適配程度為何？
- 3.以廣義次向度模式、MRCMLM 模式、及高階試題反應模式分析 98 年第一次國中基本學力測驗所估計得到的考生之整體能力及各領域能力為何？以及由三個模式所得到的各個領域能力之間的相關為何？

貳、研究設計與方法

本研究主要探討廣義次向度模式 (Brandt, 2012)、MRCMLM 模式 (Adams, Wilson, & Wang, 1997) 以及 HO-IRT 模式 (de la Torre & Song, 2009) 等三個試題反應模式對考生整體能力及領域能力估計的精確性。研究資料包含模擬產生以及實徵的作答反應，期望透過由模擬資料所得到的結果，作為實徵資料分析之基準。

模擬資料時所操弄的條件，包含領域向度數 (3 和 5 個)、各內容領域之題數 (20 及 40 題) 以及向度能力值之間的相關 (0.2、0.5 和 0.8) 等三個條件，受試人數為 1000 人。以簡單結構或題間多向度的形式來產生模擬資料，亦即，在每一個內容領域中的各題之作答反應資料是以單向度 Rasch 模式產生的，而實徵資料則是採用實際參加 98 年第一次國中基本學力測驗的 1000 人樣本在五個測驗科目 (國語、英文、數學、自然及社會) 之實際作答資料來進行估計。研究架構如圖 1 所示，以下將詳細的說明。

一、研究設計

(一) 模擬資料設計

1. 模擬資料之條件設定

本研究在產生模擬資料時，操弄領域向度數、各領域向度的題數以及領域向度間考生能力相關的程度等三個變項。模擬人數設定為 1000 人，重覆模擬受試作答反應 30 次，以探討在領域向度個數 (D)、領域向度內題數 (n) 及各領域能力值相關程度 (r) 等條件下，各 IRT 模式對整

體能力及領域能力估計精確度的差異情形。之所以僅重複模擬作答反應資料 30 次，主要是因為採用 WinBUGS 軟體進行參數估計時，MCMC 演算法會耗去相當多的時間，以 1000 名考生、五個內容領域、200 個試題為例，以廣義次向度模式估計考生整體能力及領域能力，電腦硬體配備採用 AMD Phenom 9600B Quad-Core Processor 2.3GHz 處理器、8GB 記憶體需花 10 個小時的時間。文獻上使用貝氏估計法來估計較複雜的試題反應模式時，重複模擬作答反應資料的次數常只為 10 次或 10 次以下 (Bolt & Lall, 2003; Cao & Stokes, 2008; Klein Entink, Fox, & van der Linden, 2009; Li, Bolt, & Fu, 2006; van der Linden, Klein Entink, & Fox, 2010)。底下針對所操縱變項進行說明。

- (1) 考生整體能力及領域能力：每個模擬情境的受試人數為 1000 人，每位考生在領域向度上皆有一單向度潛在能力，考生之各個領域能力集合起來具有多變量標準常態分佈，即 $\theta \sim MVN(0,1)$ 。領域能力個數為 3 個或 5 個，各領域能力值之間的相關分別設定為 0.2、0.5 與 0.8。實際產生考生在各個領域上的能力時，是分別由 $N(0,1)$ 隨機抽樣得到 3 個或 5 個值，並利用 Cholesky decomposition 對於領域向度能力間之相關矩陣進行分解，例如，若領域個數為 3，且各個領域能力之間的相

關為 .2 時，領域能力間的相關矩陣 $\begin{bmatrix} 1 & .2 & .2 \\ .2 & 1 & .2 \\ .2 & .2 & 1 \end{bmatrix}$ 利用

Cholesky decomposition 進行分解，得到一個上三角形或下三角形的矩陣，這個矩陣乘上前面所提到由 $N(0,1)$ 隨機抽樣得到 3 個值所組成的矩陣，然後得到用來產生模擬資料時考生的真正的能力值。利用這些能力真值，以及模擬產生的試題參數，在各個領域向度內以單向度 Rasch 模式來產生作答反應資料，而跨各個內容領域，整體就成為具有簡單結構的多向度的資料。而以三個領域能力真值的平均數，作為考生整體能力的真值，以供之後研究比較及模擬考生作答反應之用；同樣地模擬五個領域能力值 $\theta \sim N(0,1)$ ，亦是取五個領域能力值的平均值作為整體能力的真值，用以與其他試題反應模式估計之整體能力做為比較之用。

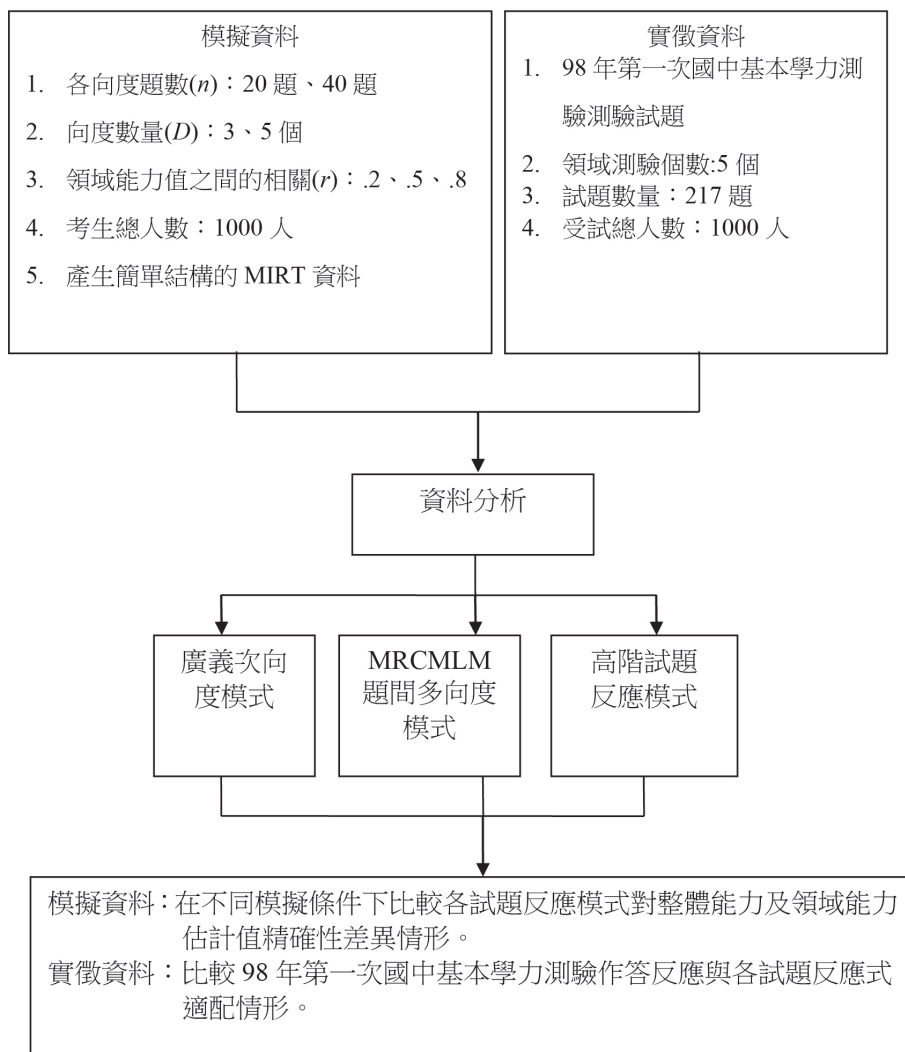


圖 1 研究架構圖

(2) 領域能力題數及試題難度參數設定：研究中各領域的試題之難度參數由均勻分布 $U(-2, 2)$ 產生；內容領域能力為 3 個和 5 個，各內容領域的題數設定為 20 題及 40 題，

因此所產生的測驗總題數組合有 60、100、120 及 200 題四種。

2. 模擬步驟

本研究中以 R 軟體來產生考生反應資料，以下為模擬資料產生之步驟。

步驟 1：界定領域向度數、領域向度能力值相關係數及題數。

步驟 2：試題難度參數為 $-2\sim 2$ 的均勻分布。

步驟 3：設定受試人數及考生各領域能力分佈 $\theta \sim N(0,1)$ 。

步驟 4：根據步驟 2 之試題參數和步驟 3 所產生的考生能力值，產生具有簡單結構的多向度資料，計算考生答對該試題的機率值。

步驟 5：每個試題的答對機率值 $P(\theta)$ 與由均勻分佈 $U(0,1)$ 產生的隨機值作比較，若 $P(\theta)$ 大於產生的隨機值，則視為答對，計作 1；反之，則視為答錯，計作 0，以此產生考生作答反應。

步驟 6：根據步驟 1~5 程序，各種模擬情境均重複 30 次，總共產生 12 組模擬情境。

(二) 實徵資料

實徵資料為實際參加九十八年第一次國中生基本學力測驗的考生中所抽取的 1000 人隨機樣本資料，是由「國中基本學力測驗委員會」所提供的，男女生各 500 人。國中生基本學力測驗是台灣教育體系中，學生由國中升學至高中的升學考試，是一種學科成就取向的測驗型式，測驗的

內容包括國文科、英文科、數學科、社會科和自然科等五個科目。各科目試題數量如下：國文科 48 題、英語科 45 題、數學科 34 題、自然科 58 題及社會科 63 題，合計試題數量為 248 題。實際分析時，將這些 248 題視為是同屬一個測驗的題目，而五個科目則視為是五個內容領域。

二、資料分析方法

(一) 試題反應模式參數估方法

本研究旨在比較三種 IRT 模式的整體能力和領域能力估計的精確性，因此模式的試題參數及考生能力參數的估計，最好都採用相同的方法，以免其間的差異是由不同方法之差異所造成。廣義次向度模式及 HO-IRT 模式要估計的參數較多，不易採用邊際最大似估計法 (marginal maximum likelihood estimation, MMLE)，故研究中三種試題反應模式皆採用 MCMC，估計軟體採用 WinBUGS (Lunn, Thomas, Best, & Spiegelhalter, 2000)。

貝氏估計中所有要估計的參數皆須設定先驗分佈 (prior distribution)，因此，考生整體能力均設定為 $\theta_n \sim N(0,1)$ ，單參數 Rasch 試題反應模式的難度參數 $b_i \sim N(0,1)$ ，Rasch 題組反應模式的題組隨機效果 $\sigma_{d(j)}^2 \sim InverseGamma(1,1)$ 。HO-IRT 中迴歸係數 $\lambda^{(d)} \sim U(-2.0, 2.0)$ ，領域能力 $\theta_n^{(d)} | \theta_n, \lambda^{(d)} \sim N(\theta_n \lambda^{(d)}, 1 - \lambda^{(d)2})$ 。

廣義次向度模式中領域能力
 $\gamma_{nk} \sim \text{inverse-Wishart}(R[1;5,1;5], 5)$ ，轉換參數
 $d_k \sim N(0,1)I(0,)$ 。

(二) 比較考生整體能力值與領域能力估計值之差異情形

1. 模擬資料部分

本研究採用均方根誤差 (root mean square error, RMSE)、偏誤 (bias) 以及相關係數 $r_{\zeta\hat{\zeta}}$ 來評估整體能力及領域能力參數回覆的正確，公式如下：

$$r_{\zeta\hat{\zeta}} = \frac{1}{R} \sum_{r=1}^R \text{Cor}(\zeta, \hat{\zeta}) \quad (21)$$

$$\text{RMSE}(\hat{\zeta}) = \sqrt{\sum_{r=1}^R (\hat{\zeta} - \zeta)^2 / R} \quad (22)$$

$$\text{Bias}(\hat{\zeta}) = \sum_{r=1}^R (\hat{\zeta} - \zeta) / R \quad (23)$$

其中 R 為重複的次數， ζ 、 $\hat{\zeta}$ 分別是能力值之真值與估計值。

2. 實徵資料部分

因為實徵資料的參數真值不可知，各個模式估計精確性的比較以下列兩種方式進行，第一使用貝氏估計中模式比較常用的 DIC 準則，作為模式比較標準，DIC 越小代表模式與資料的適配度越佳，以此指標做為判斷來選取 98 年第一次國中生基本學力測驗最適配的試題反應模式；其二，將由 DIC 指標所選出最適配的試題反應模式所估計得

到的整體能力值視為基礎線，與其他試題反應模式估計所之整體能力值作精確性比較。所用的指標包括（1）以基礎線模式所估計得到的整體能力值與其他模式所估計得到的整體能力值之間的相關；（2）上述的 BIAS 以及 ABS 表示整體能力估計的精確性。因為實徵資料僅有一筆，前述 BIAS 公式中的 R 改代表人，而不是隨機樣本重複的次數；而且

$$ABS = \sum_{r=1}^R |\hat{\zeta} - \zeta| / R \quad (24)$$

其中 R 亦代表受試樣本的人數。

（三）比較不同試題反應模式所估計之領域能力值間相關

1. 模擬資料部分

計算研究所採用的各個 IRT 模式所估計得到的考生各個領域能力彼此間的相關，以及與這些領域能力與模擬時所用的考生真正能力值之間的相關。

2. 實徵資料部分

計算各個 IRT 模式在國中生基本學力測驗 1000 人反應資料上所估計得到的領域能力彼此之間的相關，為要對此項分析的效度有更清楚的瞭解，筆者進一步以 R 軟體 sirt PACKAGE 中 DETECT 程序檢定 98 年第一次國中基本學力測驗是否為多向度的測驗或是本質單一向度的測驗。

參、結果與討論

本研究主要目的為比較三種 IRT 模式對考生整體能力與領域能力估計的差異情形。底下分成兩個小節來呈現分析的結果，第一小節是模擬資料分析的部分，而第二小節呈現實徵資料的結果。

一、模擬資料分析結果

關於模擬資料的結果，以下分兩個部分說明：首先為比較各多向度 IRT 模式的整體能力與領域能力估計值的相關係數 $r_{\zeta\zeta}$ 、RMSE 及 BIAS，其次為比較不同模擬情境下，各個試題反應模式（以下簡稱為模式）所估計的領域能力值之間的相關情形。

（一）各試題反應模式資料適配性

從表 1 可發現當領域向度數為 3 時，MRCMLM 模式的 DIC 值最小，HO-IRT 模式最大，表示 MRCMLM 模式的適配度最佳，HO-IRT 模式最差，然而差距不會太大。當領域向度數為 5 時，廣義次向度模式的 DIC 值最小，MRCMLM 與 HO-IRT 模式的結果相近，皆稍微大一點點，差距也不會太大，表示廣義次向度模式的模式資料適配度最佳。由此可見，在多向度測驗結構中，當向度數較多時，廣義次向度模式對於模擬資料的適配情形最佳，當向度數較少時，MRCMLM 模式對於模擬資料有最好的適配度。無論向度數的多寡，HO-IRT 模式的適配接比最佳的模式稍遜色一點點。

表 1

各試題反應模式資料適配 DIC 值

向度數	領域能力 相關係數	領域內 題數	廣義次向度 模式	MRCMLM 模式	HO-IRT 模式
3	.3	20	59000	58472	59543
		40	117444	106835	117919
	.5	20	59188	58648	59709
		40	117821	106843	118243
	.8	20	58913	58900	59314
		40	117600	117616	117980
5	.3	20	97257	98141	98320
		40	190418	190982	191367
	.5	20	97302	99097	98180
		40	190394	191301	191207
	.8	20	96366	96968	96914
		40	190743	190787	191003

(二) 整體能力估計精確性之比較

利用同一組模擬條件產生 30 筆測驗作答資料 (每筆資料 1000 名受試作答反應)，進行各模式的能力參數估計之分析，並分別計算各模式所估計考生整體能力與領域能力估計值與真值間的 BIAS 值、RMSE 值及相關係數 $r_{\zeta\zeta}$ ，作為精確度的指標，分析結果如表 2 至表 4。

1. 整體能力估計值與真值之相關 $r_{\zeta\zeta}$

領域向度數為 3 個時，無論領域能力之間的相關係數為 0.2、0.5 或 0.8，領域內題數為 20 或 40 題，MRCMLM 模式所得到的整體能力估計值與真值之相關係數 $r_{\zeta\zeta}$ 皆最大。

當領域向度數為 5 個，整體能力估計相關係數 $r_{\zeta\zeta}$ 值以廣義次向度模式最佳；在領域向度相關為 0.8 時，無論領域題數為 20 題或 40 題時，基本上三個模式的表現相近，而廣義次向度模式與 HO-IRT 模式所得到的相關係數 $r_{\zeta\zeta}$ 非常接近，也稍稍優於其他二個模式。而領域能力相關係數為 0.2 時且領域內題數為 20 題時，HO-IRT 模式的表現較明顯不如其他二者。

由表 2 綜合各模式在兩種向度資料上的表現來看，當各領域向度的題數增加時， $r_{\zeta\zeta}$ 的值也隨之變大，可以推知當題數變多時，各個試題反應模式整體能力的估計表現會變得更好，而當各領域能力之間的相關係數變大時， $r_{\zeta\zeta}$ 的值也隨之提高。這個傾向是合理的，因為當領域能力之間若是獨立沒有相關存在的話，整體能力就沒有心理計量上多大的意義。

若以個別的試題反應模式的表現來看，由表 2 的數字可以發現當向度數為 3 時，MRCMLM 模式的表現最佳，廣義次向度次之，但與 MRCMLM 相差不大；而當向度數為 5 時，以廣義次向度模式的表現最佳，其次是 MRCMLM 和 HO-IRT 模式。因此，就整體能力估計的目的來說，廣義次向度模式和 MRCMLM 模式二者皆是可以選擇的。當領域能力之間的相關大於 0.5 時，HO-IRT 模式亦是個選項。

表 2

試題反應模式估計的整體能力值與真值之相關 ($r_{\hat{\theta}}$)

向度數	領域能力 相關係數	領域內題數	廣義次向度 模式	MRCMLM 模式	高階試題反 應模式
3	.2	20	.877	.894	.810
		40	.907	.918	.829
	.5	20	.919	.924	.913
		40	.946	.951	.943
		20	.941	.942	.938
		40	.963	.964	.959
5	.2	20	.916	.913	.871
		40	.953	.946	.928
	.5	20	.948	.941	.947
		40	.973	.961	.971
		20	.964	.960	.964
		40	.980	.977	.980

表 3

各試題反應模式整體能力值估計的 RMSE

向度數	領域能力 相關係數	領域內題數	廣義次向度 模式	MRCMLM 模式	高階試題反 應模式
3	.2	20	.353	.329	.394
		40	.287	.281	.287
	.5	20	.339	.329	.374
		40	.277	.265	.295
		20	.329	.328	.347
		40	.260	.255	.271
5	.2	20	.274	.291	.298
		40	.205	.218	.267
	.5	20	.278	.313	.284
		40	.207	.243	.241
		20	.275	.288	.281
		40	.203	.223	.214

表 4

各試題反應模式整體能力值估計的 BIAS

向度數	領域能力 相關係數	領域內題數	廣義次向度 模式	MRCMLM 模式	高階試題反 應模式
3	.2	20	-.046	.019	-.102
		40	-.041	.030	.070
	.5	20	-.003	-.003	-.031
		40	-.006	-.002	-.047
	.8	20	-.004	-.003	-.041
		40	-.005	-.004	-.012
5	.2	20	-.003	-.004	-.004
		40	-.002	-.000	-.002
	.5	20	-.002	-.006	.002
		40	-.004	.003	-.001
	.8	20	-.004	.006	-.001
		40	-.001	-.009	.002

2. 整體能力值估計的 BIAS 及 RMSE

只要各個考生估計得到的整體能力值與真值之排序情形是一致的，二者之間的相關就會很高，因此整體能力值與真值之間的相關 $r_{\zeta\zeta}$ 並無法反應估計精確性的全貌，還需要 RMSE 和 BIAS 來補足。

由表 3 的結果來看，就 RMSE 來說，當向度數為 3 時，以 MRCMLM 所得到的值最小，是表現最佳的模式，而廣義次向度模式次之，與 MRCMLM 的結果很接近；而就 5 個向度的資料來說，廣義次向度模式所得到的結果為最佳，MRCMLM 次之，這個部分的組型與前面的 $r_{\zeta\zeta}$ 結果一樣。說明了在這三個模式在整體能力估計的表現，以

MRCMLM 和廣義次向度模式較佳。另外，當各領域的題數增加時，RMSE 值變小，各領域能力之間的相關變大時，RMSE 的值也隨之變小。

三個模式所得到的整體能力估計之 BIAS 部分，由表 4 的結果可以發現，當向度數為 3 時，以 MRCMLM 所得到的結果最佳（即 BIAS 之絕對值最小），廣義次向度模式次之，但基本上與 MRCMLM 的結果接近，當向度數等於 5 時，三個模式的 BIAS 值相近，大致上以廣義次向度的模式最優。而領域能力之間的相關係數以及各領域之題數對於 BIAS 的影響比較不明確，只有在 3 個向度的資料中，當領域能力之間的相關變大時，可以看到 BIAS 有隨之變小的情形。由表 4 可以看到 BIAS 的值小於 0 的較多，可知這三個模式大致上皆低估了真實的整體能力。

然而，BIAS 是個整體性的指標，若要更清楚瞭解各個模式對於整體能力的高低估情形，必須察看估計的整體能力值與真值之散佈圖。在此，筆者以五個向度的資料所估計得到的整體能力及其相對應真值繪製散佈圖為例，以節省篇幅。當各個領域能力之間的相關為 0.8 且各領域題數為 40 時，由圖 2 可以看到廣義次向度模式（左上）與 HO-IRT 模式（右上）的結果相近，他們與 MRCMLM 的結果差異不大，而 MRCMLM 可看到低能力被高估而高能力被低估的現象。而當領域間的相關為 0.2 且各領域題數為 20 時，廣義次向度模式和 HO-IRT 模式二者的組型相似（見圖 3），

雖然真值與估計值之間的相關比前一項低，高低估的情形尚不是十分明顯，但是 MRCMLM 模式的高低估現象較圖 2 的明顯。

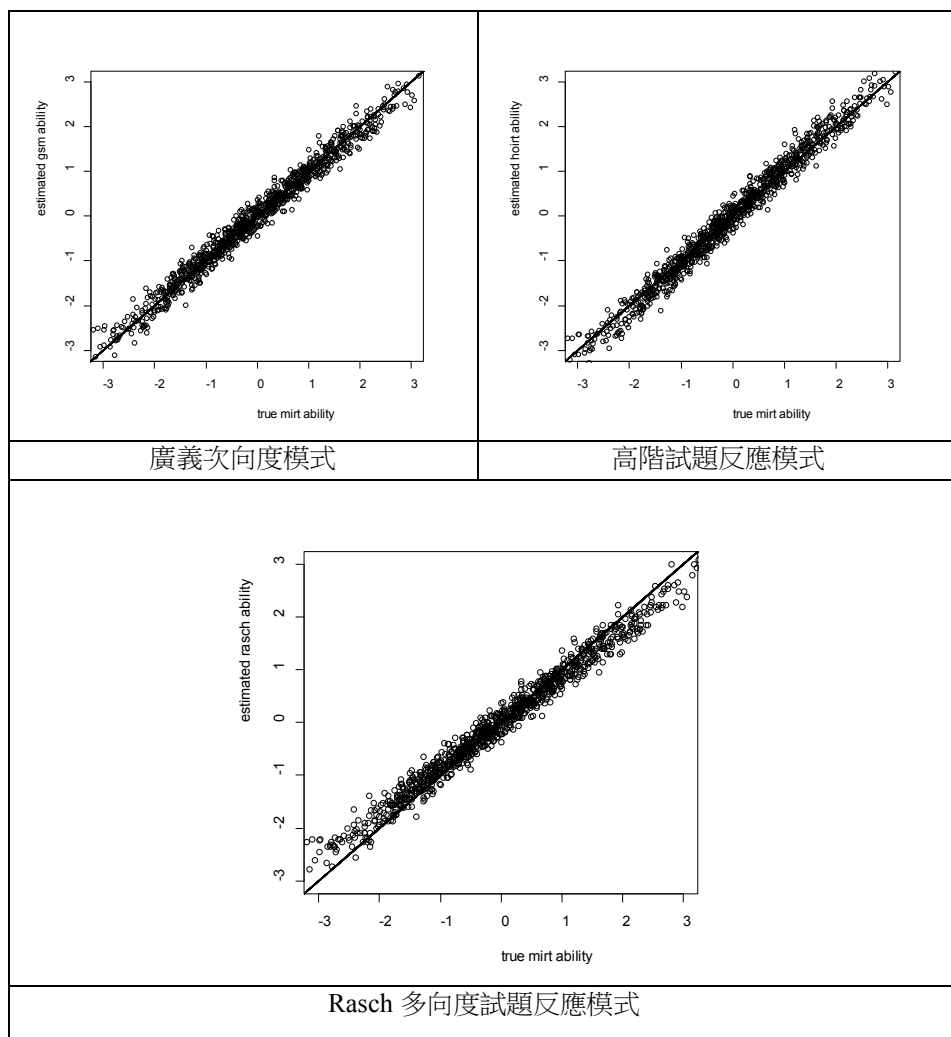


圖 2 三個模式在五個領域向度資料之整體能力估計值與真值之散佈圖

(領域能力之間的相關 0.8、各領域題數 40)

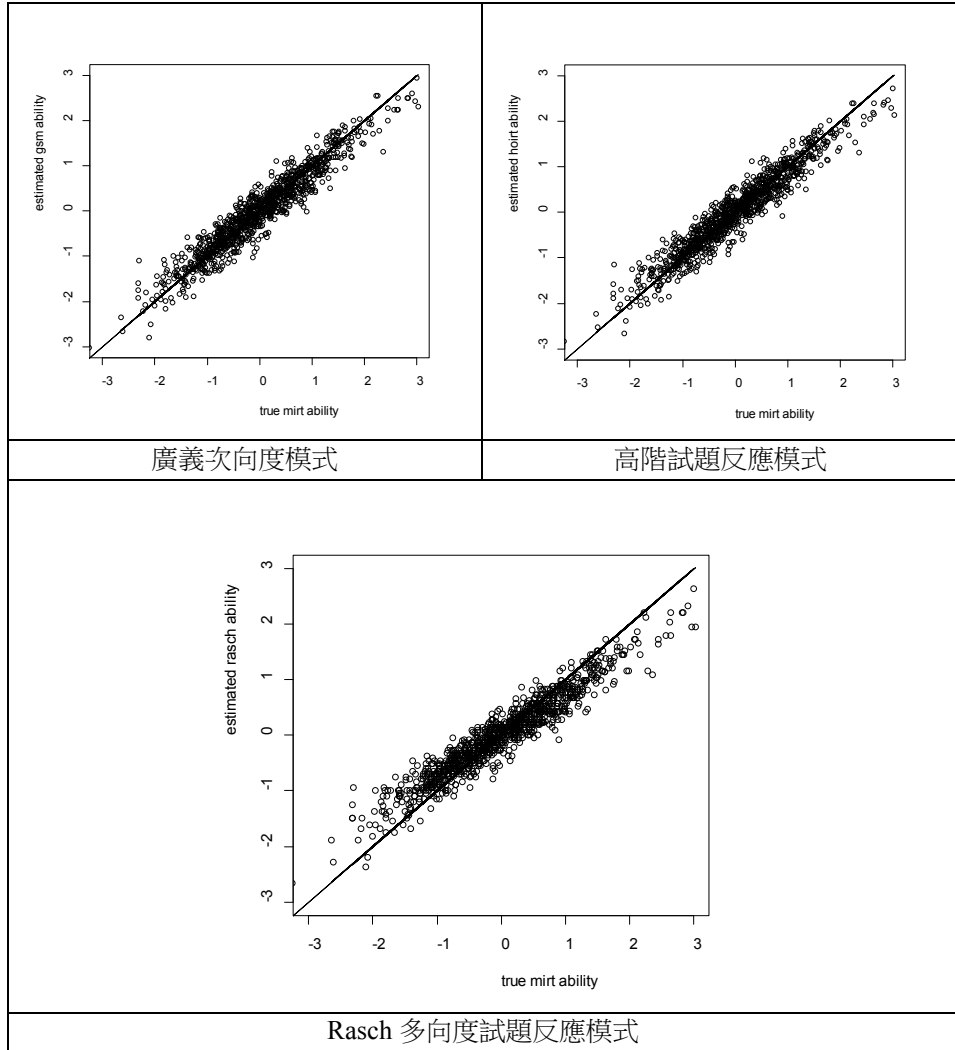


圖 3 三個模式在五個領域向度資料之整體能力估計值與真值之散佈圖
(領域能力之間的相關 0.2、各領域題數 20)

(三) 領域能力估計精確性比較

當領域向度數為 3 時，由表 5 中的數據可以看到，當題數增加時，各個模式所估計得到各領域的能力 (θ_1 、 θ_2 及 θ_3) 與真值之間的相關 ($r_{\zeta\zeta}$) 有隨之變大的趨勢；而當各領域能力之間的相關變大時，三個模式之 $r_{\zeta\zeta}$ 的值也有隨之變大的趨勢。就各個模式來看，整體而言，以 MRCMLM 所得到的 $r_{\zeta\zeta}$ 值最大，廣義次向度模式次之。

就 RMSE 部分而言，當題數變多時，各個模式所得到的值都有變小的趨勢，當領域能力之相關變大時，各個模式所估計的各個領域能力之 RMSE 值皆有變小的趨勢，因此題數變多時，或是領域間的相關變大時，各領域能力的估計精確度都變得較佳。就各個模式來看，MRCMLM 各個領域能力估計的表現最佳，廣義次向度模式次之。

以 BIAS 來看，大致上隨著題數的增加，BIAS 有變小的趨勢，而當領域能力之間的相關變大時，大致上 BIAS 的值也隨之變小。就所有情境來看，MRCMLM 模式的表現最好，廣義次向度模式次之。綜合 $r_{\zeta\zeta}$ 、RMSE 和 BIAS 的結果來看，可知對於領域向度能力的估計，在三個向度的情形下，MRCMLM 模式是表現最好的模式，其次為廣義次向度模式。

當領域向度數為 5 時，由表 6 中的數據可以到，當題數增加時，無論各領域之間的相關為何，各個模式所估計得到各領域的能力（ θ_1 、 θ_2 、 θ_3 、 θ_4 及 θ_5 ）與真值之間的相關（ $r_{\zeta\zeta}$ ）有隨之變大的趨勢；而當各領域能力之間的相關變大時，三個模式之 $r_{\zeta\zeta}$ 的值也有隨之變大的趨勢。而就各個模式來看，整體而言，以廣義次向度模式得到的 $r_{\zeta\zeta}$ 值最大，而 HO-IRT 模式次之，然而二者的結果非常接近。

RMSE 的部分，當題數增加時，各個模式所得到的值都有變小的趨勢，當各領域能力之間的相關變大時，各個模式領域能力估計的 RMSE 值大致上有變小的趨勢。就各個模式來看，廣義次向度模式各個領域能力估計的 RMSE 最小 HO-IRT 模式次之。

由表 6 最後的部分來看，題數的增加對於領域能力估計的 BIAS 似乎沒有明顯的影響力，而各領域能力之間的相關變大時，大致上 BIAS 有隨之降低的現象，雖然不是那麼明顯。就所有情境來看，廣義次向度模式、MRCMLM 與 HO-IRT 模式的表現十分接近。

表 5

各試題反應模式估計領域能力值與真值之 $r_{\zeta\zeta}$ 值、RMSE 及 BIAS (D=3)

領域能力 相關	領域題數	廣義次向度模式			MRCMLM 模式			高階試題反應模式		
		θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
$r_{\zeta\zeta}$										
.2	20	.812	.890	.861	.813	.894	.866	.804	.810	.857
	40	.886	.940	.917	.886	.940	.921	.804	.841	.856
.5	20	.818	.893	.876	.821	.897	.878	.806	.840	.860
	40	.890	.938	.916	.890	.940	.921	.890	.939	.920
.8	20	.857	.909	.892	.857	.911	.892	.856	.905	.893
	40	.907	.946	.924	.908	.947	.925	.909	.943	.924
RMSE										
.2	20	.486	.611	.529	.486	.608	.526	.488	.791	.573
	40	.379	.458	.412	.368	.458	.409	.367	.452	.404
.5	20	.477	.598	.512	.473	.589	.508	.474	.768	.547
	40	.375	.459	.402	.370	.458	.402	.370	.562	.431
.8	20	.424	.556	.471	.423	.548	.465	.426	.723	.492
	40	.339	.428	.394	.339	.420	.394	.341	.545	.404

(續下頁)

(接上頁)

領域能力 相關	領域題數	廣義次向度模式			MRCMLM 模式			高階試題反應模式		
		θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
		BIAS								
.2	20	.025	.029	-.030	-.003	.026	-.018	-.047	.057	-.035
	40	.008	.031	.038	-.001	.028	.034	-.045	.057	-.044
.5	20	.003	.019	-.027	-.001	.012	-.015	-.034	.039	-.035
	40	-.008	.020	-.022	-.002	.016	-.012	-.004	.019	-.020
.8	20	-.001	.012	-.019	.002	.009	-.007	-.002	.007	-.019
	40	-.002	.010	-.016	.001	.007	-.005	-.001	.009	-.010

表 6

各試題反應模式估計領域能力值與真值之 $r_{\zeta\zeta}$ 值、RMSE 及 BIAS (D=5)

領域 能力 相關	領域 題數	廣義次向度模式					MRCMLM 模式					高階試題反應模式				
		θ_1	θ_2	θ_3	θ_4	θ_5	θ_1	θ_2	θ_3	θ_4	θ_5	θ_1	θ_2	θ_3	θ_4	θ_5
		$r_{\zeta\zeta}$														
.2	20	.819	.900	.846	.875	.893	.789	.873	.807	.845	.872	.817	.892	.844	.873	.886
	40	.896	.942	.918	.940	.935	.827	.900	.859	.886	.887	.896	.937	.917	.940	.930
.5	20	.835	.901	.860	.885	.899	.812	.884	.838	.868	.883	.834	.895	.858	.899	.893
	40	.904	.944	.922	.941	.936	.848	.913	.884	.909	.910	.904	.939	.922	.939	.933
.8	20	.870	.918	.891	.908	.915	.860	.912	.883	.900	.911	.868	.910	.888	.905	.909
	40	.921	.952	.930	.947	.947	.910	.941	.920	.938	.940	.920	.949	.931	.947	.946

(續下頁)

(接上頁)

領域 能力 相關	領域 題數	廣義次向度模式					MRCMLM 模式					高階試題反應模式				
		θ_1	θ_2	θ_3	θ_4	θ_5	θ_1	θ_2	θ_3	θ_4	θ_5	θ_1	θ_2	θ_3	θ_4	θ_5
RMSE																
.2	20	.486	.594	.566	.569	.606	.529	.682	.644	.350	.667	.492	.777	.610	.647	.783
	40	.368	.453	.412	.394	.467	.482	.605	.548	.564	.639	.370	.587	.444	.453	.590
.5	20	.464	.586	.540	.543	.582	.502	.663	.602	.580	.632	.465	.743	.570	.607	.731
	40	.355	.440	.401	.390	.461	.457	.547	.502	.480	.552	.354	.555	.420	.435	.555
.8	20	.414	.529	.479	.484	.531	.429	.555	.500	.502	.545	.428	.660	.492	.526	.648
	40	.323	.408	.383	.366	.478	.351	.455	.417	.397	.453	.328	.409	.387	.371	.496

(續下頁)

(接上頁)

領域 能力 相關	領域 題數	廣義次向度模式					MRCMLM 模式					高階試題反應模式				
		θ_1	θ_2	θ_3	θ_4	θ_5	θ_1	θ_2	θ_3	θ_4	θ_5	θ_1	θ_2	θ_3	θ_4	θ_5
BIAS																
.2	20	-0.007	-0.069	-0.028	.077	.017	-0.043	-.146	-.104	.168	.072	-0.003	-.072	-.019	.077	.008
	40	-0.005	-0.070	-0.020	.081	.011	.015	.099	-0.004	.083	.050	-0.007	-.067	-.020	.077	.007
.5	20	-0.005	-0.060	-0.022	.058	-.003	-.080	-.186	-.152	-.034	-.060	-0.004	-.057	-.015	.061	.006
	40	-0.004	-0.044	-0.012	.066	.007	-.037	-.047	-.098	-.009	-.037	-0.007	-.059	-.019	.060	.004
.8	20	-0.001	-.034	-.009	.042	.004	-.028	-.086	-.050	-.031	-.041	-0.004	-.037	.037	.002	.002
	40	-0.001	-.045	-.010	.029	-.003	-.045	-.074	-.048	-.009	-.033	-0.005	-.038	-.013	.036	.003

(四) 不同試題反應模式所估計的領域能力值之間的相關

在全部的 12 種模擬情境下，採用三種試題反應模式得到考生領域能力估計值，進行相同模式下領域能力估計值之相關分析，因每種模擬情境均重複 30 次，因此以 30 次相關的平均數來進行比較，結果呈現在表 7 至表 12 之中。

表 7 為考生各領域能力值之間的相關皆設為 0.2，領域能力向度數為 3，各領域題數為 20 及 40 題時，以各模式估計得到的領域能力值之相關。由表中的結果可以看到，廣義次向度模式、MRCMLM 模式及 HO-IRT 模式所得到的領域能力彼此間的相關組型皆非常近似，無論各領域的題數是 20 題或 40 題。各領域能力之間的相關在 0.14~0.24 之間，大致上滿足原先所設定的 0.2。

表 8 為考生各領域能力值之間的相關皆設為 0.2，領域能力向度數為 5，各領域題數為 20 及 40 題時，以各模式估計得到的領域能力值之相關。由表中的結果可以看到，廣義次向度、MRCMLM 及 HO-IRT 等三種模式所得到的領域能力彼此間的相關組型皆非常近似，各領域能力之間的相關大致上也滿足原先所設定的 0.2，而且當題數增加時，各領域能力之間的相關係數有較接近 0.2 的傾向。

表 7

領域能力估計值之間的相關 ($r = .2$, $D=3$)

模 式	廣義次向度模式			MRCMLM 模式			高階試題反應模式		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
20 題									
D1	1			1			1		
D2	.15**	1		.14**	1		.15**	1	
D3	.24**	.20**	1	.23**	.21**	1	.24**	.22**	1
40 題									
D1	1			1			1		
D2	.15**	1		.15**	1		.15**	1	
D3	.22**	.23**	1	.22**	.23**	1	.22**	.23**	1

註：D1 為向度一，D2 為向度二，D3 為向度三。

** $p < .01$

表 8

領域能力估計值之間的相關 ($r = .2$, $D=5$)

模式	廣義次向度模式					MRCMLM 模式					高階試題反應模式				
	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
20 題															
D1	1					1					1				
D2	.24**	1				.26**	1				.25**	1			
D3	.26**	.20**	1			.28**	.20**	1			.25**	.23**	1		
D4	.24**	.26**	.18**	1		.27**	.29**	.21**	1		.25**	.28**	.21**	1	
D5	.20**	.20**	.22**	.23**	1	.26**	.20**	.22**	.26**	1	.22**	.23**	.23**	.25**	1
40 題															
D1	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
D1	1					1					1				
D2	.21*	1				.21*	1				.21*	1			
D3	.22*	.19*	1			.27*	.23*	1			.23*	.21*	1		
D4	.20*	.24*	.19*	1		.18*	.28*	.18*	1		.20*	.25*	.20*	1	
D5	.16*	.17*	.22*	.22*	1	.19*	.20*	.24*	.21*	1	.17*	.19*	.22*	.23*	1

註：D1 為向度一，D2 為向度二，D3 為向度三，D4 為向度四；D5 為向度五。

* $p < .01$

表 9 為考生各領域能力值之間的相關皆設為 0.5，領域能力向度數為 3，各領域題數為 20 及 40 題時，以各模式估計得到的領域能力值之相關。由表中的結果可以看到，廣義次向度模式、MRCMLM 模式及 HO-IRT 模式所得到的領域能力彼此間的相關組型皆非常近似，無論各領域的題數是 20 題或 40 題。各領域能力之間的相關在 0.50~0.58 之間，大致上滿足原先所設定的 0.5。

表 10 為考生各領域能力值之間的相關皆設為 0.5，領域能力向度數為 5，各領域題數為 20 及 40 題時，以各模式估計得到的領域能力值之相關。與 3 個向度資料稍有不同的是，MRCMLM 模式所估計得到的各領域能力之間的相關係數偏離原先所設定的 0.5 的程度比廣義次向度模式大。

表 9

領域能力估計值之間的相關 ($r = .5$, $D=3$)

模式	廣義次向度模式			MRCMLM 模式			高階試題反應模式 D2		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
20 題									
D1	1			1			1		
D2	.53**	1		.53**	1		.51**	1	
D3	.59**	.54**	1	.58**	.54**	1	.56**	.56**	
40 題									
D1	1			1			1		
D2	.50**	1		.50**	1		.51**	1	
D3	.54**	.54**	1	.55**	.54**	1	.56**	.56**	1

註：D1 為向度一，D2 為向度二，D3 為向度三，D4 為向度四，D5 為向度五。

** $p < .01$

表 10

領域能力估計值之間的相關 ($r=.5$, $D=5$)

模式	廣義次向度模式					MRCMLM 模式					高階試題反應模式				
	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
20 題															
D1	1					1					1				
D2	.56**	1				.57**	1				.58**	1			
D3	.59**	.55**	1			.63**	.57**	1			.59**	.59**	1		
D4	.56**	.56**	.56**	1		.60**	.58**	.58**	1		.58**	.61**	.59**	1	
D5	.54**	.54**	.56**	.57**	1	.58**	.56**	.58**	.59**	1	.57**	.60**	.60**	.61**	1
40 題															
D1	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
D1	1										1				
D2	.54**	1				.59**	1				.55**	1			
D3	.55**	.53**	1			.62**	.55**	1			.55**	.55**	1		
D4	.54**	.55**	.52**	1		.58**	.57**	.53**	1		.54**	.58**	.55**	1	
D5	.50**	.51**	.53**	.53**	1	.55**	.54**	.56**	.55**	1	.52**	.54**	.55**	.55**	1

註：D1 為向度一，D2 為向度二，D3 為向度三，D4 為向度四，D5 為向度五。

** $p < .01$

表 11 為考生各領域能力值之間的相關皆設為 0.8，領域能力向度數為 3，各領域題數為 20 及 40 題時，以各模式估計得到的領域能力值之相關。由表中的結果可以看到，廣義次向度模式、MRCMLM 模式及 HO-IRT 模式所得到的領域能力彼此間的相關組型皆非常近似，無論各領域的題數是 20 題或 40 題。各領域能力之間的相關在 0.83~0.89 之間，大致上滿足原先所設定的 0.8。

表 12 為考生各領域能力值之間的相關皆設為 0.8，領域能力向度數為 5，各領域題數為 20 及 40 題時，以各模式估計得到的領域能力值之相關。由表中的結果可以看到，廣義次向度及 HO-IRT 等二種模式所得到的領域能力彼此間的相關組型皆非常近似，MRCMLM 模式稍微各領域能力之間的相關大致上也滿足原先所設定的 0.8，而且當題數增加時，各領域能力之間的相關係數有較接近 0.8 的傾向。整體而言，廣義次向度模式與 HO-IRT 的相近，比 MRCMLM 稍微好一些。

表 11

領域能力估計值之間的相關 ($r = .8$, $D=3$)

模式	廣義次向度模式			MRCMLM 模式 D3			高階試題反應模式 D2		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
20 題									
D1	1			1			1		
D2	.86**	1		.86**	1		.86**	1	
D3	.87**	.85*	1	.88**	.86**	1	.88**	.89**	1
40 題									
D1	1			1			1		
D2	.84**	1		.84**	1		.83**	1	
D3	.86**	.86**	1	.87**	.87**	1	.85**	.86**	1

註：D1 為向度一，D2 為向度二，D3 為向度三，D4 為向度四，D5 為向度五。

** $p < .01$

表 12

領域能力估計值之間的相關 ($r = .8$, $D=5$)

模式	廣義次向度模式					MRCMLM 模式					高階試題反應模式				
	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
20 題															
D1	1					1					1				
D2	.88**	1				.89**	1				.88**	1			
D3	.89**	.88**	1			.91**	.89**	1			.87**	.90**	1		
D4	.88**	.87**	.87**	1		.90**	.88**	.88**	1		.87**	.90**	.89**	1	
D5	.88**	.87**	.87**	.87**	1	.88**	.87**	.88**	.88**	1	.87**	.91**	.90**	.90**	1
40 題															
D1	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
D1	1					1					1				
D2	.86**	1				.87**	1				.84**	1			
D3	.86**	.84**	1			.90**	.85**	1			.84**	.85**	1		
D4	.84**	.85**	.84**	1		.88**	.86**	.84**	1		.83**	.86**	.85**	1	
D5	.83**	.83**	.83**	.83**	1	.85**	.84**	.85**	.83**	1	.82**	.85**	.85**	.85**	1

註：D1 為向度一，D2 為向度二，D3 為向度三，D4 為向度四，D5 為向度五。

** $p < .01$

由表 7 至表 12 中可得知在所有模擬情境下，無論領域向度數、領域能力之間的相關大小及各領域內題數為何，廣義次向度向模式、MRCMLM 模式、以及 HO-IRT 模式所估計得到的考生領域能力之間的相關係數均較模擬條件中所設定之相關係數的值高。研究中考生能力參數估計採用 MCMC 演算法，因貝氏估計中須設定所估計參數之先驗分佈，導致領域能力估計值之間的相關係數高於真值間的相關，相關係數增加的幅度為 0.0 至 0.09，此情況與 Yao(2012) 的研究結果類似。

二、實徵資料分析結果

如前所述，研究中分析的實徵資料是取自 98 年第一次國中基本學力測驗的受試之作答反應，包含國語、英語、數學、自然及社會等五科，其中國語科試題數量為 48 題，英文科 45 題，數學科 34 題，社會科 63 題及自然科 58 題，合計共 248 題。因此若以多領域內容的觀點來看，國中基本學力測驗可視為具有 5 個領域向度的能力測驗。這個小節分兩個部分敘述：其一在貝氏估計架構下，比較各試題反應模式的 DIC 值，作為模式選擇的依據。第二為瞭解各試題反應模式估計考生整體能力估計的精確性以及各試題反應模式估計之領域能力值間的相關情形。

(一) 貝氏估計架構下各多試題反應模式 DIC 值比較

各種模式的適配結果如下：廣義次向度模式的 DIC 值為 216387，MRCMLM 模式的 DIC 值為 216827，HO-IRT 模式的 DIC 值為 216464，這些數據呈現在表 13 之中。由

DIC 的值可看出適配度最佳的是廣義次向度模式，其次為 HO-IRT 模式，再者為 MRCMLM 模式。

表 13

實徵資料各試題反應模式之適配結果

	Dbar	Dhat	pD	DIC
廣義次向度模式	212881	209376	3505.41	216387
MRCMLM 模式	213044	209260	3783.40	216827
高階試題反應模式	212918	209372	3546.36	216464

(二) 各試題反應模式估計 98 年第一次國中基本學力測驗考生整體能力值精確性比較

將 1000 名參與 98 年第一次國中基本學力測驗的作答反應資料以廣義次向度模式、MRCMLM 模式及 HO-IRT 模式估計考生整體能力值，並以廣義次向度模式估計之整體能力作為「真值」，計算其他試題反應模式所估計整體能力值與真值間相關 ($r_{\hat{\theta}\theta}$)，結果呈現在表 14 中。由表中可知各試題反應模式所估計之整體能力值與廣義次向度模式估計整體能力值間相關都非常地高，皆達到 0.99 以上。由表 14 可知以廣義次向度模式作為真值情況下，各試題反應模式的 ABS 及 BIAS 最小的是 HO-IRT 模式，其次為 MRCMLM 模式。

表 14

實徵資料各試題反應模式整體能力估計值精確度比較

試題反應模式	$r_{\xi\xi}$	ABS	BIAS
高階試題反應模式	.998	.124	-.037
MRCMLM 模式	.994	.132	-.127

註：以廣義次向度模式估計得到的整體能力值為真值。

(三) 不同試題反應模式下考生在領域能力值的相關

表 15 為各模式分析 98 年第一次國中基本學力測驗 1000 筆作答反應資料，所估計得到的考生領域能力值之間的相關係數，由表 15 的結果可知廣義次向度模式、MRCMLM 模式及 HO-IRT 模式所估計領域能力值之間的相關係數在 0.86 ~ 0.92 間。

由表 12 模擬資料的結果中，可以看到當領域能力值之間的相關為 0.8，無論各領域內的題數是 20 或 40 題時，廣義次向度模式、MRCMLM 模式及 HO-IRT 模式皆會高估考生領域能力值間的相關。以此推論，呈現在表 15 國中生基本學力測驗中考生各領域能力之間的相關，實際上應在 0.8 附近。

為瞭解九十八年第一次國中基本學力測驗中所包含的五個學科的測驗其為單一向度或本質上單一向度 (essentially dimensional) 的強度如何？筆者再以 DETECT 程序 (Kim, 1994; Zhang & Stout, 1999) 檢定資料的向度性。DETECT 為探索性無母數統計法，目的為評估領域向度的數量以及偏離單向度的程度。其原理為利用條件共變理論

(conditional covariances)，將測驗中試題分為幾個群集 (clusters)，因此同屬一個群集的試題會產生正的條件共變數；而分屬不同群集的試題則會產生負的條件共變數 (Nandakumar & Ackerman, 2004)。其多向度指標 DETECT 的値之意義如下：當 DETECT 的値大於 1.00 時，代表強烈程度的多向度存在；當 DETECT 的値在 0.4~1.00 之間，代表中等程度多向度存在；當 DETECT 的値在 0.2~0.4 之間，代表些微程度的多向度存在；如果 DETECT 的値小於 0.2 時，代表測驗在本質上是單一向度。本研究以 R 軟體 sirt Package 中的 conf.detect 函數進行向度性檢測，得到國中基本學力測驗資料的 DETECT 値為 0.22，很接近本質上為單一向度的臨界值 0.2，因大於 0.2，故可推論由五個科目的國中基本學力測驗的確有些微程度的多向度存在，但接近本質上為單一向度之測驗，這或許也是各個領域能力之間的相關那麼高的原因。

表 15

實徵資料各試題反應模式估計的領域能力值之間的相關係數

模式	廣義次向度模式					MRCMLM 模式					高階試題反應模式				
	國文	英文	數學	自然	社會	國文	英文	數學	自然	社會	國文	英文	數學	自然	社會
國文	1					1					1				
英文	.88**	1				.90**	1				.86**	1			
數學	.86**	.87**	1			.89**	.89**	1			.87**	.87**	1		
自然	.91**	.88**	.88**	1		.92**	.89**	.88**	1		.91**	.89**	.95**	1	
社會	.90**	.87**	.89**	.90**	1	.91**	.89**	.89**	.90**	1	.92**	.88**	.92**	.91**	1

註：Rasch 題組反應模式估計得到的領域能力與本研究領域能力值之定義不是很相符，故其結果未在表中呈現。

肆、結論與建議

本節根據研究目的與研究問題，歸納出主要的研究結論，並對各試題反應模式的採用與未來研究方向提出相關的建議。

一、研究結論

- (一) 無論資料是以何種試題反應模式來進行分析的，隨著領域能力向度數的增加，各領域能力值之間的相關變大，以及各領域題數的增加，整體能力值估計的精確性會提升， $r_{\zeta\zeta}$ 值會逐漸變大，RMSE 及 BIAS 會漸次下降。當領域向度數為 3 個，由整體能力估計之 $r_{\zeta\zeta}$ 、RMSE 及 BIAS 值來看，大體上以 MRCMLM 模式表現為最佳，其次為廣義次向度模式。當領域向度數為 5 時，由整體能力估計值之 $r_{\zeta\zeta}$ 、RMSE 及 BIAS 值來評估，以廣義次向度模式的表現最佳。當領域能力間的相關係數為 .8 時，廣義次向度模式與 HO-IRT 模式在 $r_{\zeta\zeta}$ 、RMSE 及 BIAS 值的表現都相當地接近。因此在無法確定測驗資料多向度的強度之情況下，選擇廣義次向模式來估計考生整體能力與領域能力似乎是可接受的選擇。
- (二) 模擬研究中領域能力間的相關係數分別為 0.2、0.5 及 0.8，無論領域的向度數、各領域的題數為何，因為是利用 WinBUGS (Lunn, Thomas, Best, & Spiegelhalter, 2000) 來完成各種模式的估計，可能受到貝氏估計中 MCMC 演算法需

要設定各參數的先驗分布，以致於四種模式所估計的領域能力值之間的相關皆高於模擬資料時所設定的相關真值。

- (三) 根據貝氏估計中 DIC 指標的值，可以發現對於 98 年第一次國中生基本學力測驗的資料之分析，廣義次向度模式是適合度最高的模式，其次為 MRCMLM 模式和 HO-IRT 模式。
- (四) 由實徵資料分析的結果，發現廣義次向度模式、MRCMLM 模式與 HO-IRT 模式等三種者所估計得到的考生領域能力值之間的相關介於 0.86 與 0.92 之間，屬於高度相關，三個模式的相關平均數分別為 0.88、0.9、0.9。此外，以 DETECT 法檢測結果為 0.22，已接近單一向度的臨界值 0.2，呈現出微弱多向度程度之測驗，因此推測由五個不同領域向度科目所組成的國中基本學力測驗應接近本質上單一向度之測驗。

二、研究建議

- (一) 由模擬資料進行試題反應模式能力參數估計精確性比較的結果得知，廣義次向度模式在所有的模擬條件中，其能力參數估計精確性整體來說都是最高或次高，因此以廣義次向度模式分析具有領域能力的測驗試題是一個不錯的選擇。此外由實徵資料的分析結果也看到，廣義次向度模式對資料的適配情形優於 MRCMLM 模式及 HO-IRT 模式。然而因廣義次向度模式的分析目前是使用 WinBUGS 來完成的，它採用 MCMC 法來進行參數估計，估計能力參數時花費的時間由 3 個小時至 10 個小時不等，頗為耗時，如可利用邊際最大概似估計法 (marginal maximum likelihood, MML) 或期望後驗估計法 (expected a posteriori, EAP) 等演算法發展新的軟體，將可大大縮短參數估計的時間，而提升此模式是實務上的可用性。

- (二) 本研究中採用的試題反應模式均未考量鑑別度參數以及猜測度參數，現階段測驗的情境多採用選擇題型式，為更符合實際測驗情境的要求，對於未來研究方向建議可採用二參數試題反應模式、三參數試題反應模式，並將鑑別度、猜測度參數納入高階試題反應模式及多向度試題反應模式，進行研究分析與比較。
- (三) 以模擬資料時所操弄的變項來看，模擬條件仍有擴展的空間。本研究只操弄領域能力向度（3 個、5 個）、領域能力值之間的相關程度（0.2、0.5、0.8），各個領域向度內題數（20 題、40 題）等三個條件，產生十二種模擬資料情境，來對各試題反應模式進行整體能力及領域能力估計精確性的比較，但模擬條件無法窮盡所有可能的情境。因之，未來進行研究時，可以考慮不同的樣本人數（不是固定為 1000 人），領域向度數也可以增加（兩個或六個以上），其他的領域向度內題數亦可考慮（10 題、30 題或 50 題以上）。進一步比較各試題反應模式對能力參數估計的精確性，並釐清廣義次向度模式在這些模擬條件下是否仍具有優勢。
- (四) 研究中僅採用九十八年第一次國中基本學力測驗作答反應做為實徵資料，建議日後研究可以納入其他的測驗資料，比如大學指定入學考試或全民英檢（GEPT）資料等，檢驗在不同的測驗資料中，各個試題反應模式的表現，是否有所不同。

參考文獻

- 涂柏原、盧思丞 (2012)。成就測驗組合分數議題探討。 *教育研究學報*，**46** (1)，119-137。
- 郭伯臣、謝典佑、吳慧珉、林佳樺 (2012)。一因子高層次試題反應理論模式之評估。 *測驗學刊*，**59**(3)，329-348。
- Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied psychological measurement*, *21*(1), 1-23.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*(4), 561-573.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (pp. 395-479). Reading, MA: Addison-Wesley Publishing Company.
- Bolt, D. M., & Lall, V. F. (2003). Estimation of compensatory and noncompensatory multidimensional item response models using Markov chain Monte Carlo. *Applied Psychological Measurement*, *27*(6), 395-414.
- Brandt, S. (2008). Estimation of a Rasch model including subdimensions. *IERI monograph series: Issues and methodologies in large-scale assessments*, *1*, 51-69.
- Brandt, S. (2012). *Definition and classification of a generalized subdimension model*. Paper presented at the 2012 annual

- conference of the National Council on Measurement in Education (NCME), Vancouver, BC.
- Brandt, S., & Duckor, B. (2013). Increasing unidimensional measurement precision using a multidimensional item response model approach. *Psychological Test and Assessment Modeling*, 55(2), 148.
- Cao, J., & Stokes, S. L. (2008). Bayesian IRT guessing models for partial guessing behaviors. *Psychometrika*, 73(2), 209-230.
- Childs, R. A., Elgie, S., Gadalla, T., Traub, R., & Jaciw, A. P. (2004). IRT-linked standard errors of weighted composites. *Practical Assessment, Research & Evaluation*, 9(13), Retrieved from <http://PAREonline.net/getvn.asp?v=9&n=13>.
- Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika*, 40(1), 5-32.
- de Ayala, R. J. (2013). *Theory and practice of item response theory*. New York, NY: Guilford Publications.
- de la Torre, J., & Hong, Y. (2010). Parameter estimation with small sample size a higher-order IRT model approach. *Applied Psychological Measurement*, 34(4), 267-285.
- de la Torre, J., & Song, H. (2009). Simultaneous estimation of overall and domain abilities: A higher-order IRT model approach. *Applied Psychological Measurement*, 33(8), 620-639.
- de la Torre, J., Song, H., & Hong, Y. (2011). A comparison of four methods of IRT subscore. *Applied Psychological Measurement*, 35(4), 296-316.

- Duckor, B. M. (2006). *Measuring measuring: An item response theory approach*. University of California, Berkeley, CA.
- Duckor, B., Draney, K., & Wilson, M. (2009). Measuring measuring: toward a theory of proficiency with the constructing measures framework. *Journal of applied measurement, 10*(3), 296–319.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105-146). New York, NY: Macmillan.
- Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta psychologica, 37*(6), 359-374.
- Gulliksen, H., & Wilks, S. S. (1950). Regression tests for several samples. *Psychometrika, 15*, 91-114.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory principles and applications*. Boston, MA: Kluwer-Nijhoff Publishing.
- Kim, H. (1994). *New techniques for the dimensionality assessment of standard test data*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign, IL.
- Klein Entink, R. H., Fox, J. P., & van der Linden, W. J. (2009). A multivariate multilevel approach to the modeling of accuracy and speed of test takers. *Psychometrika, 74*(1), 21-48.
- Li, Y., Bolt, D. M, & Fu, J. (2006). A comparison of alternative models for testlets. *Applied Psychological Measurement, 30*(1), 3-21.

- Lord, F. M. (1952). A theory of test scores. *Psychometric Monograph*, No. 7.
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley Publishing Company.
- Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and computing*, 10(4), 325-337.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- McDonald, R. P. (1997). Normal-ogive multidimensional model. In *Handbook of modern item response theory* (pp. 257-269). New York, NY: Springer.
- Nandakumar, R., & Ackerman, T. (2004). Test modeling. In A. Kaplan (Ed.), *The sage handbook of quantitative methodology for the social sciences* (pp. 93-105). Thousand Oaks, CA: Sage Publications.
- Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 221-262). Phoenix, AZ: The Oryx Press.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danmarks Paedagogiske Institut, Copenhagen.
- Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement*, 9(4), 401-412.

- Reckase, M. D., & McKinley, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement, 15*(4), 361-373.
- Rijmen, F., & Briggs, D. C. (2004). Multiple person dimensions and latent item predictors. In P. De Boeck, & M. Wilson (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 247-265). New York, NY: Springer-Verlag.
- Rudner, L. M. (2001). Informed test component weighting. *Educational Measurement: Issues and Practice, 20*(1), 16-19.
- Sheng, Y., & Wikle, C. K. (2008). Bayesian multidimensional IRT models with a hierarchical structure. *Educational and psychological measurement, 68*, 413-430.
- Sinharay, S., Puhan, G. and Haberman, S. J. (2011), An NCME instructional module on subscores. *Educational Measurement: Issues and Practice, 30*(3), 29-40.
doi:10.1111/j.1745-3992.2011.00208.x
- Swygert, K. A, McLeod, L. D, and Thissen, D (2001). Factor analysis for items or testlets scored in more than two categories. In D. Thissen & H. Wainer (Eds.), *Test scoring* (pp. 217-250). Mahwah, NJ: Erlbaum.
- Sympson, J. B. (1978). A model for testing with multidimensional items. In *Proceedings of the 1977 computerized adaptive testing conference* (pp. 82-98). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program.

- van der Linden, W. J., Klein Entink, R. H., & Fox, J. P. (2010). IRT parameter estimation with response times as collateral information. *Applied Psychological Measurement, 34*(5), 327-347.
- Wainer, H., & Thissen, D. (1993). Combining multiple-choice and constructed-response test scores: Toward a Marxist theory of test construction. *Applied Measurement in Education, 6*(2), 103-118.
- Wang, M. W., & Stanley, J. C. (1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research, 40*(5), 663-705.
- Whitely, S. E. (1980). Multicomponent latent trait models for ability tests. *Psychometrika, 45*(4), 479-494.
- Yao, L. (2010). Reporting valid and reliable overall scores and domain scores. *Journal of Educational Measurement, 47*(3), 339-360.
- Yao, L. (2012). Multidimensional CAT item selection methods for domain scores and composite scores: Theory and applications. *Psychometrika, 77*(3), 495-523. doi: 10.1007/s11336-012-9265-5
- Zhang, J. M. & Stout, W. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika, 64*(2), 213-249.

教育學誌 第三十九期

2018年5月，頁153~182

葉朗對《老子》美學詮釋的觀點

葉若潔

臺北市立大學教育學系碩士班學生

摘要

本研究探討葉朗對《老子》美學的詮釋觀點，葉朗視《老子》為中國美學的起點，將「先秦、兩漢」視為中國美學發展發端期。並將中國美學的範疇界定在「審美意象」，而非「美的理論」或「審美意識」。至於《老子》的美學，葉朗以「道、氣、象」視為中心，並將「道」處理出「原始混沌」、「產生萬物」、「沒有意志」、「自己運動」、「無與有的統一」五種特性。再由「道、氣、象」產生「有、無、虛、實」等變化，還有「美、妙、味」等發展，最後以「滌除玄鑒」作為人們觀照事物的最高境界。

關鍵字：葉朗、老子、道德經、中國美學

Ye Lang's Perspective on The Aesthetic Interpretation of " Lao Zi "

Jo-Chieh Yeh

Master student,

Program in Education, University of Taipei

Abstract

This research deliberates about the interpretations of Ye Lang on Lao Zi. Ye regards Lao Zi as the birth of Chinese aesthetics. In his opinion, the pre-Qin period, Western Han and Eastern Han dynasties are the starting point of the development of Chinese aesthetics. He defines the category of Chinese aesthetics as "aesthetic intention" rather than "aesthetic theory" or "aesthetic concept".As regards Lao Zi, Ye Lang focuses on the concepts of "Tao, Qi, Xiang", defining the five properties of "Tao", which are "primeval chaos", "creation of all", "non-existence of will", " self-motion", and "unification of nothing and ens". He also explains the fact that " Tao", "Qi", and "Xiang" generate changes such as " with", "without ", "false" and "true", as well as the development in "beauty", "delicacy" and "flavor". In the end, one reaches the highest level of observation is to get rid of one's ignorance, bias and superstition beliefs to clear the mind and allow it be tranquil and serene, just the clean surface of a mirror without any impurities or pollution.

Keywords: Yelang, Lao Zi, Tao De Ching, Chinese Aesthetics

壹、緒論

一、研究動機

葉朗是中國當代的美學權威，著有《中國美學大綱》、《中國小說美學》、《現代美學體系》等書，《中國美學史》以時代的整體進展為軌範，講述中國傳統美學發展一個大致的輪廓，內容豐富、前後串連。而葉朗所撰《中國美學史》的主編李澤奉、劉仲如在總序講述：

中國的文化源遠流長，光輝燦爛，它是在自己的民族生活土壤中生長壯大的，因此它具有強烈的民族性。數千年來，中國文化也受到過外來文化的衝擊，如漢唐時期的佛教文化、明清時期的西方文化等，但它並未被其征服，始終保持著自己的民族特點，從價值觀到穿衣吃飯，都顯示出了自己的民族特性。但是，中國的文化又不是一成不變，它仍在其歷史的進程中，不斷吸入新的成份，使其繁榮興盛。（葉朗，1996：總序2）

葉朗本身也在該書自序中說明：

現代美學應當是站在二十世紀九十年代的高度¹，吞吐東西方文化的全部精華建設起來的、具有多種文化視野的真正國際性的學科。如果輕視、忽略、拋棄中國美學（以及整個東方美學），就不能實現東西方美學的融合，從而也就不能真正建設現代美學。（葉朗，1996：自序1）

¹ 葉朗執筆當時為西元1996年，二十世紀。

李澤厚與劉綱紀在《中國美學史》提到：「西方常常是由道德而走向審美，以審美境界為人生最高境界。因為中國哲學所追求的天人合一，絲毫不捨棄自然，不否定人的自然生命，不是到超自然的天堂中去尋找不朽。對中國哲學來說，生死是自然的不可避免的變化，否定現世的生存而追求彼岸世界中的永恆不朽，被看作是不現實的幻想。所謂不朽不在彼岸世界中，不在人的靈魂不滅，而就在現世的人間關係中，就在人的積極有為的社會活動中，在個體對後世所作的不可磨滅的貢獻中。」（李澤厚與劉綱紀，2004：38-39）以此分析中得知，中國哲學上的永恆乃是於世的價值程度，老子說：「不失其所者久，死而不亡者壽」²人若想要長長久久的生存，就不能失去所依附之「所」，對於魚，其「所」是水，對於人的「所」，老子認為是「道」。還有一種超越生死的長壽，就是在死後，其精神足以為天下式，其典範亦能為後人法。孟子說：「夭壽不貳，修身以俟之，所以立命也。」³孟子認為在世的時候，不論性命或長或短，仁人君子當居其所、盡其本分，身心便能牢牢安住。老子與孟子雖分屬道家與儒家二個思想系統，在永恆的觀念上，卻有異曲同工之妙。

西方知識界對於中國美學確實十分隔膜和十分無知，這是美學這門學科始終未能突破西方文化的局限的一個重要原因。（葉朗，1996：2）

從哲學上看，中國古代哲學思想把人擺在極高的位置，認為人是天地萬物中最寶貴的東西，如何求得人與自然的統一，個體

² 《老子》第三十三章

³ 《孟子》盡心上篇

與社會的和諧發展，始終是中國哲學最為重視的問題，也是中國美學最重視的問題。（李澤厚與劉綱紀，2004：37）雖然葉朗的一家之言，不能代表整個西方美學界，但葉朗認為現代美學應該涵蓋東方、西方，我們若偏執一方，就難能看到美學全貌，甚至產生對立，失去美學的初衷一使人愉悅。這都致使美學停留在學術界，無法走進生活，成為真正的學問。對於中國美學的核心理論，葉朗再提他的見解：

對於中國傳統美學的基本精神和理論內核，我們還缺乏認識，至少還沒有準確地把握。在這方面，我們應該有一種緊迫感和危機感。當我們建設現代美學的時候，我們不要忘記：中國傳統美學的本來面目，中國傳統美學的豐富性和深刻性，中國傳統美學在當代的價值，等等，這一切都還有待於我們用現代眼光重新發現和重新研究。（葉朗，1996：2）

西方、東方自有始以來，有著極大的文化差異，然而對於美的追求與渴望卻是相同的。所以當我們有心研究中國美學時，不宜把拿西方的美學標準直接套用，這可能會造成規格不符的狀況。基於天人合一的觀念，中國美學總是從人與自然的統一中去找美，不僅不同於西方中世紀的神學的美學，而且在很大程度上避免了西方唯心主義美學和機械唯物主義的缺陷。（李澤厚與劉綱紀，2004：34）所以中國美學應該有一套自己的脈絡系統，有別於其他學者在中國美學架構上各有所見，範疇的界定上有所不同，葉朗在這方面有著非常深入的著墨，值得加以探究，此即為研究者的主要研究動機。

二、研究方法及目的

本研究採理論思辨研究法，關於研究目的，有以下二點：

- (一) 界定《老子》的中國美學地位
- (二) 了解葉朗對《老子》美學思想詮釋的觀念

老子是葉朗列為首位的中國美學家，葉朗認為《老子》的美學思想對後代的影響，淵遠流長。葉朗以何種觀念將《老子》的美學思想進行綱目式的分類、詮釋？是本研究的主要探討的重點。

貳、中國美學的重要發展

一、先秦美學具發端地位

葉朗將近代以前的中國美學史分成三個時期，分別是「發端」的先秦、兩漢，「展開」的魏晉南北朝至明代，還有「總結」的清代前期

先秦是我國由奴隸制向封建制過渡的時代，是一個社會變動的時代，是一個產生新的經濟、新的人物、新的思想的時代。殷周的宗教神學統治發生動搖，舊的觀念紛紛瓦解，禮壞樂崩，出現一個思想解放、百家爭鳴的局面，理論思維十分活躍。這樣一種社會條件，使得先秦成為中國美學史上第一個黃金時代。（葉朗，1996：6）

就思想的形成，不論是東方、西方，在變動的社會中，較易激起人們求新、求變、求好的意念。所以在東周的戰國時期，狼煙四起，戰火連天，整個社會儼然形成一個大講堂，各家學說、流派不

斷的湧出，人們在其中謀名位、謀富貴，也謀安生。就如葉朗所說，先秦時期是中國美學史的第一個黃金時代。

先秦理性精神，從美學上看，有多方面的關係：1.決定了朝廷美學體系由三代（夏商周）類型轉為秦漢類型。2.為士人美學奠定了文化、哲學、思想基礎。3.為中國美學以後的體系性展開作了結構性的構築。（張法，2000：50）

這個時期的美學家如老子、孔子、易傳葉朗、莊子、荀子等都是大哲學家。他們的美學是他們哲學體的一部分。他們提出了「道、氣、象、妙、味、美、大、興、觀、群、怨、滌除玄鑒（覽）、心齋、坐忘、厲與西施，道通為一、象罔可以得道、觀物取象、立象以盡意、化性起為而成美…」等範疇和命題，為整個中國古典美學的發展奠定了哲學基礎。後代美學家所探討的理論問題差不多都以萌芽的形式包含在先秦美學之中。（葉朗，1996：6）

張法（2006）認為中國社會歷史包含著兩個相輔相成的方面，一是共時性上的人道與天道，二是歷時性上的盛世與衰世，孔子從人道與盛世立論，老子從天道與衰世立論。葉朗說「這是中國古典美學的發端時期。這個發端是哲學的發端。」尤其是《老子》的歷史影響深遠，後代的審美客體、審美觀照、藝術創造和藝術生命的一系列特殊看法，《老子》、《易傳》和《莊子》，是這些學說的思想發源地。

二、老子為美學思想起點

對於中國思想的起點，源自於老子或孔子，各個學者持有不同的見解，例如張法說：「從人物的年代講，老子早於孔子，但從二子的思想凝結成成熟的著作體系看，《論語》先於《老子》。」（張法，2000：53）高輝陽（1983）以三點跡象證明「老子」成書年代晚於孔子的戰國時代。首先，他以「大道廢，有仁義；智慧出，有大偽」⁴、「法令滋彰，盜賊多有」⁵認為《老子》書中具有強烈的懷疑和批判精神，因此不論老子其人是誰，但《老子》一書可確定是在儒、墨等「顯學」之後出現的。第二點，再引司馬遷在《史記》說到「道家使人精神專一，動合無形，瞻足萬物。其為術也，因陰陽之大順，采儒墨之善，撮名法之要。與時推移，應物變化」⁶認定道家採納了當時各家的主張，因此是後起的。

此說法確實有其立基點，然而就時代背景來說，在孔子集儒家大成之前，老子是專對儒家思想進行辯駁，或是單就社會上崇尚虛名、忽略實質的現象進行反證？更何況在百家爭鳴的先秦時期，以孔子為首的儒家思想並未受到整體社會的看重，一直得到漢武帝獨崇儒術後，儒家思想才對整個中國社會發生鋪天蓋地的影響。因此將老子「絕聖棄智、絕仁棄義」的主張，視為儒家思想的對立方，或是補儒墨之不足，研究者對這個說法持保留的態度。而且司馬談口中的「道家」是以後世人的角度縱觀戰國時期，其涵蓋「老子、莊子」及其後人，不應以此斷定老子成書在儒、墨之後。而且《史

⁴ 《老子》第十八章

⁵ 《老子》第五十七章

⁶ 此文為司馬談〈論六家要旨〉，收錄在《史記》的〈太史公自序〉與《漢書》的〈司馬遷傳〉當中。

記》中亦有「孔子適周，將問禮於老子」⁷之說，高輝陽並未解釋，使其觀點前後不一貫。

第三點，高輝陽認為孔墨講求「名」、「實」的統一，「正名」是為了「正實」，一直到辯者出現，才離開「實」來說「名」。而《老子》提出「先天地生」的「道」來做為宇宙萬物的根本起源和原理，因此《老子》的出現是在儒墨等「顯學」之後。研究者亦不支持這個分析，何以因為孔子沒直接談到「道」，就斷定「道」的觀念，必須在孔子之後，世人才能開始談呢？儒家與道家本就不同思想系統，不應視其為因果承襲關係。

而葉朗（1996）在《中國美學史》直接以一句話表明立場「中國美學史的起點是老子的哲學和美學」，查其附錄說明「老子的生卒年已不可詳考。一種說法是老子約生於周簡王六年（前五八〇年），約死於周敬王二十年（前五〇〇年）。」其他再無佐證來支持老子先於孔子。再探葉朗（1986）在《中國美學史大綱》所述，才有完整的說明，研究者整理重點如下：

- （一）引詹劍峰《老子其人其書及其道論》肯定《老子》書是出自老子本人的手筆。
- （二）1973年發掘馬王堆漢墓出土的兩種帛書《老子》，甲種本不避劉邦名諱，乙種本僅避劉邦名諱，不避文帝名諱，可以說明《老子》一書在秦之前已流傳。
- （三）《韓非子》、《莊子》、《墨子》皆援引老子的話，推翻梁啟超、楊榮國認為《老子》書成於孟、莊、墨之後。
- （四）墨子說的「尚賢」與老子的「不尚賢」意思不同，不可以此來斷先後關係。

⁷《史記》老子韓非列傳。

- (五) 周公早在公元前十一世紀就制禮作樂，非待孔子才提出禮樂之說，故《老子》的「夫禮者，忠信之薄而亂之首」為何要等孔子之後才能提出呢？
- (六) 《老子》雜有戰國時期的用語，可能在成書後受到竄改、加工，此現象在先秦書籍皆出現，並非《老子》一書才有的。
- (七) 《老子》著重宇宙觀和認識論，有其自己的體系及特點，哲學味要比《論語》濃，而且中國古典美學一系列獨特的理論，都發源於老子哲學及美學。

因此我們有充份的理由，證明老子思想先於孔子，也可以說老子的思想對中國美學發展造成極大的影響。

三、審美意象為中國美學主要核心

學術界對於中國美學史的對象和範圍有以下三種見解：

表一

葉朗對中國美學史的界定

對象和範圍	關於美的理論	審美意象	審美意識
舉例	孔子論美、王充論美…	道、氣、象、意等概念	美學理論著作、各朝代的文學、藝術作品
葉朗評斷	過於狹窄	以此為中心	過於寬廣

葉朗認為僅把中國美學史侷限在古代思想家談到美的段落：孔子、孟子、墨子等人怎麼論美，太狹窄了。因為：

這涉及對美學對象的看法。美學不限於研究「美」。美學研究的對象是人類審美活動的本質、特點和規律。隨著社會的發

展，美學的範圍日益擴大，美學的分支學科越來越多。美學的這些分支學科，有的在古代就已得到相當的發展，有的在古代僅僅有某種萌芽。我們研究美學史，應該站在現代的高度，把自己的視野放寬一些。（葉朗，1996：1-2）

基於以上理由，葉朗認為光就美學理論來研究，無法進入中國美學的核心。其實不光是美學，就連哲學、心理學、社會學等其他學科，都是後世逐漸整理歸納的，中國哲人鮮少直接提出一套完整的理論學說，多是在實行中見真章。儒家思想是「坐而言，不如起而行。」「先行其言，而後從之。」⁸任何的學說思想，必須要當事人先貫徹執行了，說明是真的可行的範疇，而非天馬行空的論調，其學說才會被世人接受、流傳。這與西方思想，將哲人的學說與品格區分為二的概念很不相同。

有別於狹義的美學史，以哲學家、文藝家、文學理論批評家著作中已形成系統的美學理論或觀點，作為主要研究對象，廣義的研究對中國美學史是更為適合重要的，因為它將各個歷史時代的文學、藝術以至社會風尚中的審美意識進行全面的考察。（李澤厚與劉綱紀，1986：4-5）

中國古典美學體系是以審美意象為中心的。在中國古典美學體系中，「美」並不是中心範疇，也不是最高層次的範疇。「美」這個範疇在中國古典美學中的地位遠不如在西方美學中那樣重要。如果僅僅抓住「美」字來研究中國美學史，或者以「美」這個範疇為中心來研究中國美學史，那麼一部中國美學史就將變得十分單調、貧乏，索然無味。（葉朗，1996：2）

⁸ 《論語》為政篇

葉朗認為既然研究美學，就要研究每個時代的「審美意識」，而這些審美意識多集中在一些大思想家的美學思想中。美學範疇和美學命題是他們的理論結晶，所以「一部美學史，主要就是美學範疇、美學命題的產生、發展、轉化的歷史。」如果脫離了這些，就很難將美學史符合邏輯並做到統一。審美意識史與藝術史、美學史是有關聯的，加以研究能夠協助我們更理解、分析各每時代的美學理論。然而葉朗一再提醒我們，不宜將藝術作品，包括它們表現出的審美意識作為研究的中心對象，將其視為背景材料來處理，更為合適。

以上就是葉朗界定的美學主要範疇，以審美意識為核心，並要研究每個時代的相關歷史，佐以藝術品為背景材料，這樣就能掌握大部分的中國美學。

參、評《老子》美學之詮釋觀點

一、葉朗評《老子》美學

《老子》又稱《道德經》，共有八十一篇，上半段以「道」為本體，不斷探討「道」的本質、「道」的意義，還有「道」生成演化出的種種現象。下半段闡述「德」的作用，有修身之方、養性之則，也有治國之策、天下安平之象。各篇文義雖然獨立，實則由幾個重要觀念在反覆延伸。

葉朗不同於其他學者，他認定《老子》是中國思想源頭，並以審美意識為主軸，將《老子》分成四個類別，分別是：「道、氣、

象」、「有、無、虛、實」、「美、妙、味」、「滌除玄鑒」，屬於層次上的漸次分析。以下以葉朗區分的這四個類別，進行《老子》美學詮釋之觀念探討。

（一）道、氣、象

《老子》的哲學和美學，以葉朗來看，是中國古典美學史的起點。他說：「中國美學體系的中心範疇並不是『美』。同樣，《老子》美學中最重要的範疇也不是『美』，而是『道、氣、象』這三個互相聯結的範疇。」所以我們應該先從「道、氣、象」來研究《老子》的美學。以《老子》來看，葉朗整理出老子心中的道，有以下五種特質：

第一，「道」是原始混沌。「有物混成，先天地生。寂兮寥兮，獨立而不改，周行而不殆，可以為天下母。吾不知其名，字之曰道。強為之名曰大。」⁹（葉朗，1996：10-11）

葉朗以第二十五章的內容來說明，老子思想裡的「道」，在還沒有天地之前就有了，所以不是靠外力所生，是一切的源頭，包含形成萬物的可能性。

第二，「道」產生萬物。「道沖而用之，或不盈。淵兮，似萬物之宗。」¹⁰「道生一，一生二，二生三，三生萬物。萬物負陰而抱陽，沖氣以為和。」¹¹（葉朗，1996：11）

⁹ 《老子》第二十五章

¹⁰ 《老子》第四章

¹¹ 《老子》第四十二章

這二段經文都有提到「道」是萬物的宗主，是一切的源頭，道由無極推向太極，太極能生陰陽，陰陽之中開始調和後，出現了萬事萬物。此處亦可與《易經》八卦的演化相呼應。道由無極推向太極後，有了陰、陽二儀，各是一個爻，即是「道生一」；二儀生四象，有了少陰、少陽、太陰、太陽，各是二個爻，即是「一生二」；四象生八卦，有了乾、坤、震、巽、坎、離、艮、兌，各是三個爻，即是「二生三」，接著八卦能推演出六十四卦，包羅宇宙天地之各種變化，即是「三生萬物」。

第三，「道」沒有意志。「道法自然」¹²「道常無為而無不為」¹³
「道生之，德畜之，物形之，勢成之。是以萬物莫不尊道而貴德。道之尊，德之貴，夫莫之命而常自然。故道生之，德畜之。長之，育之，亭之，毒之，養之，復之。生而不有，為而不恃，長而不宰，是謂玄德。」¹⁴（葉朗，1996：11）

葉朗舉出上列三段經文，認為「道」是自然的，沒有什麼特意的作為，雖然「道」使四時寒暑運行、陰陽日月交替，但是對於萬物是自然的使之生長、養育，沒有目的性，沒有情感、偏好，也不會去佔有或主宰萬物。

不同於儒家將自然當作文明世界的借鏡，道家完全將自然視為美的典範，亦即作為學習的榜樣，從直觀自然到整個生命映照自然的方式生活，這即是一種迴向作用。（彭美鑑，2011：524）老子著重道的本體應任自然，對於心的善惡看法源於對自然的看法是無善

¹² 《老子》第二十五章

¹³ 《老子》第三十七章

¹⁴ 《老子》第五十一章，「亭之、毒之」亦作「成之、熟之」

無惡。(朱梅詔，1994：148) 老子的美學思想深深地植根於他的哲學思想之中，他繼承了《易經》的生命意識傳統，從宇宙論的角度看待人生和社會諸問題，要求順任自然，在美學上體現為崇尚自然，以自然為美，追求樸拙的審美境界。(朱志榮，2007：84)

第四，「道」自己運動。「有物混成，先天地生。寂兮寥兮，獨立而不改，周行而不殆，可以為天下母。吾不知其名，字之曰道。」¹⁵ (葉朗，1996：11)

「道」非靜止不動，而是獨立不改、周行不殆，沒有間斷的在流轉。葉朗認為，因為「道」有這樣的運動，才能構成宇宙萬物的生命。

第五，「道」是「無」與「有」的統一。「道可道，非常道。名可名，非常名。無名天地之始，有名萬物之母。故常無欲，以觀其妙；常有欲，以觀其徼。」¹⁶「天下萬物生於有，有生於無。」¹⁷「迎之不見其首，隨之不見其後。」¹⁸ (葉朗，1996：11-12)

整部《老子》充滿著這樣的字句，將「有」與「無」相提，而且沒有規定性、限制性，也沒有具體形象，不能單憑感覺能把握。葉朗以此評判，老子認為千差萬別的事物都是由「道」產生，然而「有」與「無」並不是兩個相對的東西，「有」與「無」都是道，「道」具備「有」與「無」的雙重性。

¹⁵ 《老子》第二十五章

¹⁶ 《老子》第一章

¹⁷ 《老子》第四十章

¹⁸ 《老子》第十四章

葉朗否定「道」是精神性的東西，若是將之視為柏拉圖的「理念」是不充分的。根據老子說過：「吾不知誰之子，象帝之先。」¹⁹用這個沒有意志、沒有目的的「道」，否定了有意志的「上帝」、「天命」。葉朗認為老子的哲學體系，在總體上帶有唯物論的傾向。

葉朗接著探討老子思想中，「氣」和「象」的範疇。他以《管子·內業》：「精也者，氣之精者也。」，而說：「所謂『精』，就是『氣』。」研究者在此要提出修正，以管子所言，『精』是氣之精髓、精華，例如陽陰之氣的流動，而凝結成生命，所以『精』是氣的總精華，而非『精』等同『氣』。

「道之為物，惟恍惟惚。惚兮恍兮，其中有象。恍兮惚兮，其中有物。窈兮冥兮，其中有精。其精甚真，其中有信。」²⁰葉朗以此段經文說明「道」看似恍惚窈冥，但不是絕對的虛無。「道」包含了「象」、「物」、「精」，所以「道」是真實存在的。研究者再修正葉朗的說法，研究者認為萬物的本體是「道」，「氣」和「象」是「道」的產物，而非「道」與「氣」同時是「象」的根源。「氣」和「象」不能脫離「道」，否則成為毫無意義的東西。研究者補充說明，以「人」來說，人身上也有「道」、「氣」、「象」，「道」是人的靈性，「氣」是人的氣息，「象」是人的身軀，人一但失去了「道」，就不會有「氣息」，只剩下一個空的形軀，已經沒有意義了。

葉朗說：「老子關於『道、氣、象』的論述，對中國古典美學的發展產生了極為深遠的影響。」

從魏晉南北朝開始，中國古典美學關於審美客體、審美觀照以及藝術生命形成了一系列特殊的看法。美學家們認為，審美客

¹⁹ 《老子》第四章

²⁰ 《老子》第二十一章

體並不是孤立的、有限的「象」。「象」必須體現「道」，體現「氣」，才成為審美對象。審美觀照也不是對於孤立的、有限的「象」的觀照。審美觀照必須從對於「象」的觀照進到對於「道」的觀照。(葉朗，1996：14)

中國審美方式來源於文化的觀照方式，這就是在先秦文獻中普遍使用的「觀」。觀，是一種游目，相對於西方文化的焦點，顯出了中國特色。(張法，2000：42)在時空動態體系的文化中，仰觀俯察、遠近游目是為了獲得宇宙的真諦。這就是《老子》的觀天道：「萬物並作，吾以觀其復」。同樣，理性化以後的詩文的仰觀俯察、遠近游目也是要在有限的時空裡體會出宇宙意識。(張法，2000：43)

因此，葉朗在此做一結論：「在中國古典美學中，藝術家在自己的作品中必須表現宇宙的本體和生命（道、氣），只有這樣，藝術品本身才有生命力。」

(二) 有、無、虛、實

宗白華(2006)對於虛、實之間的結構概念，提出與葉朗相契的見解：虛和實的問題，這是一個哲學宇宙觀的問題。這可以分成兩派來講。一派是孔孟，一派是老莊。老莊認為虛比真實更真實，是一切真實的原因，沒有虛空存在，萬物就不能生長，就沒有生命的活躍。儒家思想則從實出發，如孔子講「文質彬彬」²¹，一方面內部結構好，一方面外部表現好。孟子也說「充實之謂美」²²。但是孔孟也並不停留於實，而是要從實到虛，發展到神妙的意境：「充

²¹ 《論語》雍有篇

²² 《孟子》盡心下篇

實而有光輝之謂大，大而化之之謂聖，聖而不可知之謂神。」聖而不可知之，就是虛：只能體會，不只能欣賞，不能解說，不能摹仿，謂之神。所以孟子與老莊並不矛盾。他們都認為宇宙是虛和實的結合，也就是《易經》上的陰陽結合。（宗白華，2006：19-20）

葉朗再提：「天地萬物都是『無』和『有』的統一，『虛』和『實』的統一，有了這種統一，天地萬物才能流動、運化，才能生生不息。」

「天地之間，其猶橐籥乎？虛而不屈，動而愈出。」²³「三十輻，共一轂，當其無，有車之用。埴埴以爲器，當其無，有器之用。鑿戶牖以爲室，當其無，有室之用。故有之以爲利，無之以爲用。」²⁴

葉朗講明：「一切因為有了『虛空』，才能有萬物的流動、運化，才有不竭的生命。」人們常追求滿，而不知虛的可貴及必須。常常「以實為得，以虛為失」，其實如同經文的舉例，若是有沒「虛」，那個「實」的本體是無法運用的。

老子的這種思想，對中國古典美學的發展影響也很大。「虛實結合」成為中國古典美學一條重要的原則，概括了中國古典藝術的重要的美學特點。這條原則認為，藝術形象必須虛實結合，才能真實地反應有生命的世界。（葉朗，1996：16）

因此，葉朗說：「中國古代詩、畫的意象結構中，虛空、空白有很重要的地位。沒有虛空，中國詩歌、繪畫的意境就不能產生。」而漢寶德也曾提及老子虛實相應的思想，在現代建築藝術的

²³ 《老子》第五章

²⁴ 《老子》第十一章

影響力。²⁵葉朗在此舉了宗白華所言，中國藝術因為有虛實之的概念，與西方藝術有極大的差別。

宗白華：「埃及、希臘的建築、雕刻是一種團塊的造型。米開朗基羅說過：『一個好的雕刻作品，就是從山上滾下來也滾不壞的，因為他們的雕刻是團塊。』中國就很不同。中國古代藝術家要打破這團塊，使它有虛有實，使它疏通。」（引自葉朗，1996：17）

有別於西方藝術，中國國畫、書法、文學、詩詞歌賦之種種藝術作品，無處不見留白之美、含蓄之意，這也是《老子》美學潛移默化的效果。

（三）美、妙、味

「天下皆知美之為美，斯惡已；皆知善之為善，斯不善已。故有無相生，難易相成，長短相較，高下相傾，音聲相和，前後相隨。」²⁶

葉朗以此章為指標，說明老子將「美」與「善」區分開來。之後又提到論語：「子謂韶，盡美矣，又盡善也。謂武，盡美矣，未盡善也。」²⁷孔子亦將「美」與「善」有明顯的劃分。葉朗認為老

²⁵ 漢寶德：「建築原是一種實用的藝術，而且建築的精神在空間。空間是看不到、摸不到的東西，如果談空間的美，就是抽象中之抽象了。二次大戰之後，有位張一調先生，自普林斯頓大學建築系念了博士學位，他的論文就是討論空間不可見的美學，並引用了老子的哲學。論文不長，我做學生的時候為了精讀，就翻譯了一遍。一時之間道家的『無』成為成大建築系熱門的討論話題。所以建築界不會不接受抽象的幾何美感，甚至可以接受無形的美感，在思想上是極前衛的。」（漢寶德，2004）

²⁶ 《老子》第二章

²⁷ 《論語》八佾篇

子在此給了「美」二個規定，一是與「善」有所區別，另一則是相對於「惡」存在的。

對於老子而言，個體不同，時過境遷，美的可以變成惡的；善的可以變成不善的，它們的價值是相對的，美與善沒有什麼絕對的意義，但是這樣，並不表示老子否定美，或者老子沒有美感精神。事實上老子對美感的要求是超於世俗眼光的。（朱梅詔，1994：144）問題的實質在於，美和醜一樣都是現世的，是有限的，是變化的。問題的關鍵在於，不是追求現世的美，避免現世的醜，而是超越現世的美醜，而追求真正的美。（張法，2000，62）老子深明由於每個人的認知不同，相對的也會產生差異性而無法統一，所以老子認為唯有脫離為物所控制的現實形下之審美機制，進而追求不受物限之形上精神美，無為無欲、悶悶昏昏，不與俗眾群舞，成為其特殊之處。（郭益悅，2011：89）

「五色令人目盲，五音令人耳聾，五味令人口爽，馳騁畋獵令人心發狂，難得之貨令人行妨。是以聖人為腹不為目，故去彼取此。」²⁸

葉朗將此段經文解釋為老子「對美和藝術採取一種簡單的否定的態度。」既然「聖人為腹不為目」，所以應該在社會生活中排除美和藝術。「為腹」是實用的東西，能產生快感；「為目」就是美的東西，它產生美感。研究者亦提出不同的見解，老子在此否定的五色、五音、五味皆是外在的、表層的，屬於感官上的享受，對於生命不具有實質幫助，反正會激起人的欲求。所以聖人並不追求這些表層的快感，所以「不為目」，追求的是內在的、精神的，

²⁸ 《老子》第十二章

所以「為目」。此處不可將「快感」與「美感」強加解釋為一件東西。而漢寶德的著作²⁹，亦以此章句，判定老子否定美感，這是誤解老子之意。

這裡的「五色令人目盲，五音令人耳聾」，指的是對於聲色的感官愉快的放肆的無節制的追求，其結果會使人失去正常的理智感覺，以致眼睛看不清事物，耳朵聽不清聲音，也就是魏源注所謂「視久則眩，聽繁則惑」（《老子本義》），這是老子對無止境地沉溺於聲色感官享樂所引起的感官麻木狀態的一種說明，絕不是說看五色就會使人目盲，聽五音就會使人耳聾。（李澤厚與劉綱紀，2004：241）更何況適當的田獵運動使人身心舒暢，會令人心發狂的應該是過度的奔馳在田獵之中。而生活環境是不可能脫離五色、五音、五味，若因眼能見五色、耳能聽五音、口能嚐五味就導致人們目盲、耳聾、口爽，這樣究竟要人們如何措其耳目手足？因此老子此處絕非否定五色之美及五味之感，而是講求凡事不可過度。

老子對審美和藝術的活動絕沒有從根本上完全否定，他只是要求審美和藝術不應產生「令人目盲」、「令人耳聾」這樣有害的結果，而應該同人的生命的自由相一致。人類雖然需要美，但卻不應當因為美的追求而犧牲自己的生命的自由，不應把美看做是人的外在目的，使美成為反過來支配人的東西。（李澤厚與劉綱紀，2004：244）

「古之善為士者，微妙玄通，深不可識。夫唯不可識，故強為之容：豫兮，若冬涉川；猶兮，若畏四鄰；儼兮，其若客；渙兮，若冰之將釋；敦兮，其若樸；曠兮，其若谷；渾兮，其若

²⁹ 漢寶德：「古人有『目迷五色』這句話，是把美視為善行的障礙。」（漢寶德，2010）

濁。孰能濁以澄靜之，徐清？孰能安以久動之，徐生？保此道者不欲盈，故能敝不新成。」³⁰

古之善為士者所以「妙」，就在於他取法於體現了道的無的一面，所以他深不可識，只能勉強用慎重、警覺、端莊、融和、淳厚等形象加以形容。葉朗分析這種對於體道之士的心態和客觀的描述，可視為老子對於人格美的理想呈現。

（四）滌除玄鑒

「滌除玄鑒，能無疵乎？」³¹葉朗解釋為「洗去人們的各種主觀慾念、成見和迷信，使頭腦變得像鏡子一樣純淨清朗。『玄鑒』就是對道的觀照。」

台灣「道德經」通行版本，亦作「滌除玄覽」。葉朗在註釋說明：「玄鑒」，《老子》通行本作「玄鑒」，帛書乙種本作「玄監」。高亨、池曦朝說：「『監』字即古『鑒』字。」「『鑒』與『鑑』同，即鏡子。」「後人不懂『監』字本義，改作『覽』字。」³²本研究採用葉朗用字。

老子用「滌除玄鑒」來說明，主體只有在內心極度寧靜、純粹時才可體悟道的玄機。這種觀道的過程是一種超功利的、超越感覺器官束縛的審美體悟。人們之所以不能隨時直觀道的奧妙，是因為他們蔽於自身各種主觀慾念、偏見，只有「棄聖絕知」，掙脫這一切束縛，才能於去物忘我的雙重否定中獲取審美觀照之永恆，實現與自然萬物的真正溝通和對話，進入理想的審美境界。（朱志榮，2007：87）

³⁰ 《老子》第十五章

³¹ 《老子》第十章

³² 《試論馬王堆漢墓中的帛書老子》，《文物》一九七四年十一期

葉朗在此給了「滌除玄鑒」二個含義。一個是把觀照「道」作為認識的最高目的。另一個是要求人們排除主觀欲念和主觀成見，保持內心的虛靜。並舉出「致虛極，守靜篤。萬物並作，吾以觀復。」³³作為這層含義的發揮。他說：「人心只有保持虛靜的狀態，才能觀照宇宙萬物的變化及其本原。」在老子之後的歷朝歷代的思想家，開始對老子「虛靜」的認識論有所繼承和發揮，例如「虛壹而靜」³⁴及「心齋坐忘」³⁵，還有魏南北朝的「澄懷味象」³⁶、「澄懷道觀」³⁷。甚至在文學創作的領域，出現陸機³⁸、劉勰³⁹、劉禹錫⁴⁰、蘇軾⁴¹等大量文人的引用。葉朗最後總結：「在中國美學史上，審美心胸的理論是一個影響很大的理論。這種審美心胸的理論雖然是由莊子建立的，但它的最早的源頭，卻是老子『滌除玄鑒』的命題。」

二、其他學者評《老子》美學

朱志榮（2007）在《中國美學簡史》中，認為《老子》的美學思想主要體現在以下幾個方面：「自然觀、大音希聲說、虛靜說、

³³ 《老子》第十六章

³⁴ 《荀子》解蔽篇

³⁵ 《莊子》人間世

³⁶ 南朝劉宋畫家宗炳的《畫山水序》：「聖人含道應物，賢者澄懷味象。」

³⁷ 《南史·隱逸傳》：宗炳「好山水，愛遠遊，西涉荆、巫，南登衡岳，因結宇衡山，欲懷尚平之志，有疾還江陵，嘆曰：『老疾將至，名山恐難遍睹，唯澄懷觀道，卧以游之。』凡所游履，皆圖之於室，謂之「撫琴動操，令山水皆響」。

³⁸ 陸機《文賦》：「佇中區以玄覽，頤情志於典墳。遵四時以嘆逝，瞻萬物而思紛。」

³⁹ 劉勰《文心雕龍·神思篇》：「是以陶鈞文思，貴在虛靜，疏淪五藏，澡雪精神。」

⁴⁰ 劉禹錫《秋日過鴻舉法師寺便送歸江陵引》：「梵言沙門，猶華言去欲也。能離欲，則方寸地虛。虛而萬景入，入必有所洩，乃形乎詞。」

⁴¹ 蘇軾《送參寥師》：「欲令詩語妙，無厭空且靜。靜故了群動，空故納萬境。」

有無相生。」「自然觀」講的是「道」本源自然；「大音希聲」是指「道」在真實世界的最高極至的展現是瀰漫於天地之間；「虛靜說」說明存養「道」的法則，不是填滿一切，而是放空一切；「有無相生」是指世間一切，乃由相互對立的現象構成。「具體藝術」對老子而言是世俗的。（朱梅詔，1994：147）老子的美的理想與道的理想相一致，表現為絕現世的聖與智，得天道的平、淡、素、樸。老子從宇宙之道的高度，告誡人們不要被具體的美迷惑，而失去更根本的欣賞美的本性。（張法，2000：63）

此外，賴錫三在其論文中說明：陳榮灼找出道家思家三個關鍵字：道、生、自然，並運用海德格的存有學，一一重新詮釋它們的原初意義外，進一步分析彼此間的優位性和分際關係，更將「道—生—自然」的三位一體性，和海德格之“logos—poiesis—Ereignis”三者對應比較。（賴錫三，2008：57）賴錫三認為陳氏明顯用海德格的存有論來詮解老子的形上觀念。（賴錫三，2008：58）這是學術界的另一種面向的分類法則。

若非專注於中國美學的研究，一般學者對於《老子》美學的探究較少，例如周芬伶的《美學課》一書，僅在美學的四階段、五里程碑裡提到「儒道思想，儒家提出以善為美，美善統一，道家則提出自由為美。」此一小段話。美學的詮釋進路將老子之道，帶入眼前這個具體的生活世界，它將不斷提醒人們要用「藝術」而非「技術」的角度來感受這個世界，顯然，從美學的角度來切入，將帶領我們轉向從「物」的角度來論「道」，結果將形成一種「即物而道」、「即身體而道」、「即空間而道」、「即世界而道」的景觀。（賴錫三，2008：82）道家從自然的永恆性和無限性，以及它的合目的性和合規律性的天然合理的統一中去追求人類生活的理想，主張「法自

然」，過一種純任自然，不為功利得失而苦心勞神的生活，擺脫外物對人的役使，達到精神上的絕對自由。而這種生活的實現，在道家看來就是最高的美，絕對的美。（李澤厚與劉綱紀，2004：249）在古代，人們多半對自然表現畏懼崇敬之感，老子能夠跳脫觀念的窠臼，透過自然本性發現相應於人生的真理，這種直觀迴向的方式可說是中國審美意識的發端。（彭美鑑，2011：524）

李澤厚與劉綱紀理解「無為而無不為」是以「無為」的原則代替「有為」的原則，事事純任自然，不要人為地去干預它。這在李澤厚與劉綱紀看來是荒唐的，因為老子根本取消了人的自覺的努力和能動的作用。另一方面又解讀老子看到人只有在順應自然規律的情況下才能達到自己的目的，如果企圖人為地改變和破自然規律，那就只能到失敗。（李澤厚與劉綱紀，1986：236）研究者認為此說不能充份解釋老子「無為」之意。探其原文「為無為，則無不治」⁴²、「道常無為而無不為」⁴³道是順應萬事萬物的本性去運行，沒有任何執著，有如「上善若水」⁴⁴中，水有七個德性「居善地，心善淵，與善仁，言善信，政善治，事善能，動善時」能利萬物。若老子摒棄人的積極作為，又何來讚揚水的德性呢？又何以接著說「功成身退，天之道」？若老子不要人積極的作為，又何來功成之說？老子只是人們在功成的時候，如水不爭般的欠身而退，方是天地平衡之道。

⁴² 《老子》第三章

⁴³ 《老子》第三十七章

⁴⁴ 《老子》第八章

再看，李澤厚與劉綱紀引「故失道而後德，失德而後仁，失仁而後義，失義而後禮。夫禮者，忠信之薄而亂之首。」⁴⁵理解為老子認為禮教是禍亂的根源，欲除禍亂，必廢禮教。老子對國家、法律、禮教等一切典章制度以及物質文明的成就統統給以猛烈的批判。（李澤厚與劉綱紀，1986：235）實為對經文本體不夠了解。為何「禮」的出現代表著人民的忠信之心很薄弱，意味著禍亂的源頭呢？因為人們已經不明白生天生地生萬物的道是什麼，玄牝之門又是什麼，失去源頭的人慢慢不重視德性，如此再講仁義，已是言不由衷的話了。所以人們急需要找回的自身上的道，而非廢除禮教。由此來理解「大道廢，有仁義；智慧出，有大偽」⁴⁶自然一氣呵成，而「法令滋彰，盜賊多有」⁴⁷是因為「民多利器，人多伎巧」，人們應該尋回本性良心，而非取消文明。

肆、結論與建議

一、結論

《老子》一書直接論述有關審美和藝術的立論不多，但有關哲學和社會人生問題的某些議題，對後世的美學發展有獨特的貢獻，並且被後人申引至文藝上，形成有關美學之特殊命題。（郭益悅，2011：88）老子從道法自然出發，認為只有在遵循自然這一普遍法則的前提下，萬物才能夠和諧共存。這對後世的藝術觀和審美觀產生了重要而深遠的影響。（朱志榮，2007：85）老子主張自然、無

⁴⁵ 《老子》第三十八章

⁴⁶ 《老子》第十八章

⁴⁷ 《老子》第五十七章

為，反對人為、有為，主施妄為，由順其自然達到純真素樸、淡然超越的境界，就是要把自然真樸作為藝術的最高生命、審美的最高標準。這對後世藝術追求自然風格有深遠影響，同時，也成為評價藝術作品高下優劣的一個重要標準。（朱志榮，2007：86）

葉朗推崇《老子》為中國美學的起點，再將「先秦、兩漢」視為中國美學發端期，觀看中國歷代美學作品，《老子》美學的影響範圍的確是最廣也最深，因為無論在時間點或影響度，《老子》美學堪稱中國美學之源頭。而他認為中國美學的範疇界定在「審美意象」，而非「美的理論」或「審美意識」，因為只探討「美的理論」，會將中國美學限縮在很窄狹的領域，若將「審美意識」也納入中國美學範疇，也顯得過於廣泛，不易凸顯中國美學之特質，應當聚焦於「審美意象」，方能探討出核心價值。

葉朗以「道、氣、象」視為《老子》思想的中心，並以「道」為起源，整理歸類出「道」有「原始混沌」、「產生萬物」、「沒有意志」、「自己運動」、「無與有的統一」五種特性。再由「道、氣、象」產生「有、無、虛、實」等變化，還有「美、妙、味」等發展，最後以「滌除玄鑒」作為人們觀照事物的最高境界。「道、氣、象」乃組成自然界的三元素，「道」是萬事萬物的根源，「氣」是道在運行後產生的無形力量，「象」是氣的流通後，生發出有形的物質，這是以宇宙為核心，做為第一層分類。「有、無、虛、實」即「道」在世間的現象，一但有一個「有」存在，必有一個「無」與之相呼應；有一個「實」存在，必也存在「虛」的情況，這是葉朗的第二層分類。「美、妙、味」是「道」與人結合後，在人心中產生的感受。人在追求「美」的愉悅，另一個角度是否就有「惡」的作為？當人們能夠做到「味無味」時，就趨近於老子最高的美感。「妙」是人們內心貼近於「道」後，方得的奧妙、玄妙之境，這是第三層

分類。找到自然界的生成元素「道、氣、象」了，又知道「有、無、虛、實」是世間確實存在的現象後，人與道結合產生「美、妙、味」的愉悅感受，還得將內心維持在「滌除玄鑒」之最佳觀道狀態，方能摒除知見、偏失，沒有瑕疵的觀道，進而體現出物我的本質。以上是葉朗對《老子》美學思想詮釋的觀念。

欲詮釋《老子》美學，必須先了解《老子》的思想哲學。比較其他學者對《老子》美學的看法，由於李澤厚與劉綱紀對於「無為」的理解已存有一家之見，接著再談「道」的意涵，也無法貫穿老子的思想。他們認為老子拋棄所得的成就而向後倒退的想法，是違背歷史進程的，因此，由他們探究的《老子》美學，雖然在局部分析上的能提出不錯的闡述，但是若要進行全面又深入的詮釋，則不及葉朗能掌握住《老子》的整體性及完成度。至於大部份的學者僅停留在「審美意識」來看《老子》美學，導致論述較為分散，未如葉朗專就「審美意象」的探討來得具體，這個地方亦能體現葉朗探討「審美意象」之高明。由以上各點可知葉朗詮釋的《老子》，在近代中國美學研究上，立有重要的地位。

二、建議

本研究發現，葉朗的對中國美學的研究極為深入，不但旁徵博引，也能就單一特性進行反覆的探討。然而在探討「道、氣、象」的時候，葉朗雖佐以管子言，將「氣」與「精」視為一體，但未去思索「氣」、「精」二種在本質上的差異。另外葉朗認為「道」與「氣」同時是「象」的根源，缺少了在層次上的遞嬗。尚有以「為腹不為目」去解釋老子在「五色令人目盲」的篇章是對美和藝術採取一種簡單的否定態度，其在分析上產生疏鬆之處。

《老子》一書闡明許多人與道相契合的境界、人與萬物互平衡的狀態，例如：「和光同塵」⁴⁸、「守中」⁴⁹、「見素抱樸」、「少私寡欲」⁵⁰、「抱一」⁵¹、「塞兌閉門」⁵²、「以身觀身」⁵³、「治人事天莫若嗇」⁵⁴、「持保慈、儉、不敢為天下先」⁵⁵等等觀念，就何以「滌除玄鑒」為觀道的最佳解釋，其他觀念並不足以超越「滌除玄鑒」呢？還是這些觀念與「滌除玄鑒」是一樣的意思？葉朗並未在其諸多著作中加以陳述、比較，此為可惜之處，但也提供一個未來在研究中國美學的研究方向。綜言之，中國美學需要不斷的深究，以及更多學者的投入，以提供不同的見解、分析、思辨，才能築起堅實的架構及思想脈絡，此亦是研究中國美學學者需要重視的領域。

48 《老子》第四章、第五十六章

49 《老子》第五章

50 《老子》第十九章

51 《老子》第二十二章

52 《老子》第五十二章、第五十六章

53 《老子》第五十四章

54 《老子》第五十九章

55 《老子》第六十七章

參考文獻

- 朱志榮 (2007)。中國美學簡史。北京：北京大學出版社。
- 朱梅詔 (1994)。哲學美學命題--談老子「大音希聲」的音樂美學。
問學集，4，141-149。
- 李澤厚與劉綱紀 (2004)。中國美學史。台北：漢京文化事業有限公司。
- 周芬伶 (2016)。美學課。臺北市：九歌。
- 宗白華 (2006)。中國美學史論集。安徽：安徽教育出版社。
- 高輝陽 (1983)。老子美學思想述評。**藝術學報**，33，215-231。
- 張法 (2000)。中國美學史。上海：上海人民出版社。
- 郭益悅 (2011)。從道家思想論魏晉王羲之書法美學。**明新學報**，37，87-100。
- 彭美鑑 (2011)。先秦美學思想。**有鳳初鳴年刊**，7，513-535。
- 葉朗 (1986)。中國美學史大綱。台北市：滄浪出版社。
- 葉朗 (1996)。中國美學史。台北市：文津出版社。
- 漢寶德 (2004)。漢寶德談美。台北：聯經。
- 漢寶德 (2010)。如何培養美感：落實台灣生活美學第一本。台北：聯經。
- 賴錫三 (2008)。當代學者對《老子》形上學詮釋的評論與重塑——朝向存有論、美學、神話學、冥契主義的四重道路。**清華學報**，38(1)，35-84。

國立南大學教育學系「教育學誌」徵稿辦法

2004.03.17 系法規小組訂定
2004.05.21 系務會議通過
2004.06.29 系務會議修訂通過
2007.09.14 系務會議修訂通過
2008.06.26 系務會議修訂通過
2008.09.22 系務會議修訂通過
2008.11.03 系務會議修訂通過
2011.10.14 系務會議修訂通過
2015.12.28 系務會議修訂通過
2017.06.07 系務會議修訂通過

- 一、徵稿內容：本學誌以倡導學術研究風氣為目的，徵稿範圍涵蓋以基礎理論或運用學科撰寫的教育議題，研究內容包括理論性論述分析，以及實徵性研究。非學術性稿件、報導性文章、教學講義、進修研習活動報告、翻譯稿件恕不接受。
- 二、截稿日期：本學誌採隨到隨審方式，每年五月、十一月出版。
- 三、撰寫原則：其他相關事宜請參考「教育學誌撰稿格式」
 - 1.文稿字數：中文以 10,000 字至 25,000 字為限（含題目、中英文摘要、中英文關鍵字、註釋、參考書目、附錄、圖表等）。英文以 5,000 字至 10,000 字為限（含題目、中英文摘要、中英文關鍵字、註釋、參考書目、附錄、圖表等）。中文摘要、英文摘要篇幅以不超過 300 字為原則、中英文關鍵字各為二至五個。
 - 2.本學誌之規格為 A4 大小，稿件之版面請以 A4 紙張電腦打字。
 - 3.來稿之編排順序為中文摘要、英文摘要、正文（註解請採當頁註方式）、參考書目、附錄。
 - 4.英文稿件請參照美國心理學會(APA)之寫作格式(第六版)。
 - 5.有關本學誌之「教育學誌投稿者基本資料表」、「著作權授權

同意書」、「教育學誌撰稿格式」，請逕至國立臺南大學教育學系網站查詢，網址為 <http://www.edu.nutn.edu.tw/>。

- 四、審查方式：本學誌採匿名審查制度，先進行形式審查，再由「教育學誌編審委員會」聘請有關專家至少二人擔任審查工作。
- 五、文責版權：稿件以「未曾出版」之學術性期刊論文為限，並不得一稿多投。來稿如有一稿多投，違反學術倫理、無故撤稿、或侵犯他人著作權者，除由作者自負相關的法律責任外，二年內本學誌不再接受該位作者投稿。
- 六、稿件交寄：投稿者請詳填並備妥「教育學誌投稿者基本資料表」、「書面稿件」和「電子全文檔案」(word 格式，採傳送電子檔方式)各 1 份，以掛號交寄。
- 七、投稿地址：投稿地址為「臺南市（郵遞區號 700）樹林街二段三十三號，收件者：國立臺南大學教育學系教育學誌編審委員會」。稿件請自備副本，本學誌一概不退還稿件。
- 八、聯絡人：有任何事項逕行與本學誌執行秘書聯絡。
連絡電話：(06) 2133111#613；e-mail：
sinja8336@mail.nutn.edu.tw。
- 九、通知錄用與否：稿件寄出後，請以 e-mail 方式知會本學誌執行秘書，本學誌將在收稿後兩週內通知投稿者收到稿件訊息。本學誌將在收稿後，儘快回覆審查結果。來稿若經採用，請填具「著作權授權同意書」，將發給「接受刊登證明」。
- 十、校正與抽印本
 - (一) 來稿若經採用，本期刊因編輯需要，保有文字刪修權。
 - (二) 作者應負論文排版完成後的校對之責。
 - (三) 本學誌出版後將致贈當期學誌三本。如需抽印本者，請自行至本系網頁下載全文。
- 十一、本辦法經系務會議通過後實施，修正時亦同。

教育學誌撰稿格式

壹、來稿請用 A4 格式電腦打字，四界邊界為 2.5 公分，並以 word 文字檔存檔。

貳、不論中、英文稿均需具備中、英文題目與作者中、英文姓名、職稱及服務單位。

參、請附中、英文摘要，以不超過 300 字為原則。請在摘要下增加中英文關鍵字，以二至五個為原則。

肆、內容層次

壹、

一、

□(一)

□□1.

□□□(1)

□□□□①

伍、字型及格式

一、題目：標楷粗體 18 號字，置中。

二、作者姓名職稱及服務單位：標楷體 14 號字，置中。

三、摘要(標題)：標楷粗體 16 號字，置中。摘要內容：標楷體 12 號字。(文的左緣和右緣需調整切齊)關鍵字：標楷體 14 號字。英文請用 Times New Roman 字體。

四、第一層標題：標楷體 16 號字。

第二層標題：標楷體 14 號字。

第三層標題：標楷體 12 號字。

五、內文、內文接續：新細明體 12 號字，分段落，左右對齊。

六、參考文獻：「標題」標楷體 16 號字，「內容」新細明體 12 號字。

七、圖表：置中，內容新細明體 10 號字。編號以阿拉伯數字撰寫。表之編號與標題在表「上方」，圖之編號與標題在圖「下方」。

八、行距：以「單行間距」為原則。

陸、參考文獻標註格式

依 APA 手冊(第六版)(American Psychological Association, 2009)所訂格式。

一、文中引註格式

本節「引用」一詞係指參考(reference)，作者、年代之後「不必」加註頁碼，(參見下文說明)。倘係直接引用(quotation)，則直接引用部分需加引號(40字以內時)，或全段縮入兩格(40字以上時)，並在作者、年代之後加註頁碼，如：(艾偉，1955，頁3)，或(Watson, 1918, p.44)。

(一) 中文資料引用方法

1. 引用論文時：

(1) 根據艾偉(2007)的研究……

(2) 根據以往中國學者(艾偉，2007)的研究……

2. 引用專書時：

(1) 艾偉(2007)曾指出……

(2) 有的學者(艾偉，2007)認為……

3. 如同一作者在同年度有兩本書或兩篇文章出版時，請在年代後用 a、b、c 等符號標明，例如：(艾偉，2007a)，或(Watson, 2007a)。文末參考文獻寫法亦同。

(二) 英文資料引用方式

1. 引用論文時：

(1) 根據 Johnson(2007)的研究……

(2) 根據以往學者(Johnson, 1990; Lin, 1999)的研究……

2. 引用專書時：

(1) Johnson (1990)曾指出……

(2) 有的學者(Lin, 1995)認為……

二、文末參考文獻列註格式

(一) 如中英文資料都有，中文在前，英文(或其他外文)在後。

(二) 中文資料之排列以著者姓氏筆劃為序，英文則按姓氏之字母先後為準。

(三) 無論中外文資料，必須包括下列各項：

1. 著者。

2. 專題全名(或書名)。

3. 期刊名稱及卷、期數。

4. 出版年度。

5. 頁碼。

(四) 請在中文書名、中文期刊論文名稱與卷數下面畫一橫實底線或採用黑體。請參閱(九)實例 1.(1)，2.(2)和 3.(1)。

(五) 外文書名與論文名稱，其全名之第一字母須大寫外，其餘皆小寫。請參閱(九)實例 1.(2)，和 2.(2)。

(六) 請在外文書名下畫一橫實底線，或排印成斜體字。請參閱(九)實例 1.(2)，和 2.(2)。

(七)外文期刊須寫全名，重要字母均須大寫，並請在期刊名稱及卷數下，畫一橫實底線或排印成斜體字。請參閱(九)實例 3.(2)和(3)。

(八)關於編輯及翻譯的書籍之列註體例，請參考(九)實例 4、5、6。

(九)實例

1. 書籍的作者僅一人時

(1)蘇薌雨 (1960)。心理學新論。台北：大中國。

(2)Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

2. 書籍的作者為二人或二人以上時

(1)楊國樞等 (1978)。社會及行為科學研究法。台北：東華。

(2)Mussen, P. H., Conger, J. J., & Kagan, J. (1974). *Child development and personality*. New York: Harper & Row.

3. 期刊論文

(1)蘇建文 (1978)。親子間態度一致性與青少年生活適應。師大教育心理學報，11，25-35。

(2)Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 5 (20), 158-177.

(3)Lehman, I. J., & Phillips, S. E. (1987). A survey of state teacher – competency examination programs. *Educational Measurement: Issues and Practice*, 7(1), 14-18.

4. 編輯的書籍

(1)林清江主編 (1981)。比較教育。台北：五南。

(2) Letheridge, S., & Cannon, C. R. (Eds.). (1980). *Bilingual education: Teaching English as a second language*. New York: Praeger.

5. 編輯書籍中之一章/篇

(1)黃光雄等 (1992)。英國國定課程評析。載於國立台灣師範大學教育研究所主編，教育研究所集刊 (34 期，頁 181-201)。台北：編者。

(2) Kahn, J. V. (1984). Cognitive training and its relationship to the language of profoundly retarded children. In J. M. Berg (Ed.), *Perspectives and progress in mental retardation*

(pp.211-219). Baltimore, MD: University Park.

6.翻譯的書籍

(1)黃光雄編譯 (1989)。教育評鑑的模式 (D. L. Stufflebeam 和 A. J. Shinkfield 原著，1985 年出版)。台北：師大書苑。

(2) Habermas, J. (1984). *The theory of communicative action* (T. McCarthy, Trans.). Boston: Beacon Press. (Original work published 1981).

7.網路資料：當不知出版年代時，中文以（無日期）英文以(n.d.)標示：

(1)林清江（無日期）。國民教育九年一貫課程規劃專案報告。取自

<http://www.mihjh.cyc.edu.tw/wwwsearch/%E4%B9%9D%E5%B9%B4%E4%B8%80%E8%B2%AB/9class.htm>

(2)Newman, K. (n.d.). *A pilot systematic review and meta-analysis of the effectiveness of problem based learning, learning, teachingsupport network-01 special report 2*. Retrieved from http://www.ltsn-01.ac.uk/docs/pbl_report.pdf

教育學誌投稿者基本資料表

投稿日期	年 月 日	投稿序號	
			(免填)
字數	(字數請用電腦字數統計)	語文類別	<input type="checkbox"/> 中文 <input type="checkbox"/> 英文 <input type="checkbox"/> 其他(_____)
論 文 名 稱	中文：		
	英文：		
作者資料	姓	名	服務單位及職稱(全銜)
第一作者	中文：		中文：
	英文：		英文：
共同作者 A	中文：		中文：
	英文：		英文：
共同作者 B	中文：		中文：
	英文：		英文：
通訊作者	中文：		中文：
	英文：		英文：
聯絡電話及聯絡地址	(O) (H) e-mail：(請務必填寫) 通訊處： (含郵遞區號)		FAX： 行動電話：
論文遞送方式	郵寄論文一份紙本(含教育學誌投稿者基本資料表)、電子檔務必 e-mail 傳送至 sinja8336@mail.nutn.edu.tw		

- 投稿地址：70005 台南市樹林街二段 33 號 電話：06-33111#613 執行秘書
國立臺南大學教育學系「教育學誌編審委員會」

