

教育學誌

(原初等教育學報)

第四十期

國立臺南大學教育學系 編印

中華民國一〇七年十一月

教育學誌第 40 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 11 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)

第四十期

目次

- 高雄市教保服務人員教保專業成長現況與需求之調查研究
..... 董怡君、歐陽閻.....1
- 家庭社經地位與學業成就關係之研究—
以學業發展自我效能為中介變數
..... 邱仕凱.....59
- 測驗向度數評估方法的比較
..... 楊彥文、涂柏原.....121
- 學校組成特色和閱讀優勢對學生閱讀成績的跨層次影響：
以台灣 2011 年 PIRLS 資料進行多層次研究
..... 林政達、莊雙喜.....181

Contents

The Current Status and Needs of Educare Professional
Development for Preschool Educators in Kaohsiung City
.....I-Chun Tung, Yin OuYang.....3

A Study of the Relationship between
Family Socio-Economic Status(SES)and Academic Achievement :
Using Academic Development Self-Efficacy as Mediator Variable
..... Shih-Kai Chiu.....60

A Performance Comparison of Test Dimensionality
Assessment Methods
..... Yen-Wen Yang, Bor-Yaun Twu.....123

The Cross-Level Effects of School Characteristics and
Perceived Reading Advantage on Student Reading Achievement:
Multilevel Study of Taiwan PIRLS 2011
..... Cheng-Ta Lin, Shuang-Shii Chuang.....183

教育學誌 第四十期

2018年11月，頁1~58

高雄市教保服務人員教保專業成長現況與需求之調查研究

董怡君

國立臺南大學教育學系課程與教學教學碩士班碩士生

歐陽閻*

國立臺南大學教育學系教授

摘 要

本研究旨在探討高雄市公立幼兒園教保服務人員教保專業成長現況與需求情形，並比較不同背景之教保服務人員在教保專業成長需求上之差異。本研究採問卷調查之方式，針對高雄市公立幼兒園教保服務人員進行問卷普查，共發出950份問卷，回收有效問卷681份，有效回收率為72%。本研究的主要發現如下：(1) 高雄市公立幼兒園教保服務人員教保專業成長需求屬中上程度，對於「兒童健康與照護」相關課程需求最高。(2) 不同年齡、擔任職務及幼兒園規模之教保服務人員，在教保專業成長各層面之需求有部分差異。而不同婚姻狀況與地理位置之教保服務人員，在教保專業成長各層面之需求並無差異。

*通訊作者:歐陽閻，聯絡方式:e-mail:ouyang@mail.nutn.edu.tw

關鍵字：教保專業、專業成長現況、專業成長業需求、教
保服務人員

The Current Status and Needs of Educare Professional Development for Preschool Educators in Kaohsiung City

I-Chun Tung

Master student,

The Master Program of Curriculum and Instruction,
Department of Education, National University of Tainan

Yin OuYang*

Professor,

Department of Education, National University of Tainan

Abstract

The purposes of this study were to investigate the current status and needs for educare professional development among preschool educators in Kaohsiung City and to compare the differences among the preschool educators across different background variables. This Study conducted the survey method and the questionnaires were administered to preschool educators in Kaohsiung City. 950 copies were distributed, and 681 valid responses were collected, resulting in a valid response rate of 72%. The main findings are as follows: (1) Preschool educators in Kaohsiung City show an intermediate-to-high level of need for educare professional development, with the need for courses on

“children’s health and care” being the strongest. (2) Preschool educators’ needs in some aspects of educare professional development vary by age, administrative position and school size, but their needs in any aspect of educare professional development do not vary by marital status and geographical location.

Keywords: educare profession, current status of professional development, needs for professional development, preschool educators

壹、前言

2011年6月29日，總統頒布了幼兒教育及照顧法（簡稱幼照法），自此幼稚園及托兒所統稱為幼兒園，指對幼兒提供教育及照顧服務之機構。幼照法之實施目的為保障幼兒接受適當教育及照顧之權利，確立幼兒教育及照顧方針，健全幼兒教育及照顧體系，以促進其身心健全發展。課程大綱揭示幼兒教育是各教育階段的基礎，幼兒園教保服務之實施，須與家庭及社區密切配合，以達成維護幼兒身心健康、養成幼兒良好習慣、豐富幼兒生活經驗、增進幼兒倫理觀念、培養幼兒合群習性、拓展幼兒美感經驗、發展幼兒創意思維、建構幼兒文化認同、啟發幼兒關懷環境等目標（教育部，2015）。為落實幼照法之精神及達到課程大綱之目標，教保服務人員是重要關鍵，而專業成長課程能維持教保服務人員專業能力之持續發展，藉由教保專業成長活動補充職前教育之不足，拓展視野、進行教師增能，增進研究能力，將有效提升教保品質。因此本研究欲了解幼兒園教保服務人員教保專業成長之現況，此為本研究主要動機之一。

再者，由於幼照法與課程大綱之實施，幼兒園隨之改變的層面包括法令、師資、教材、課程、教學、財務等，教保服務人員勢必得要熟悉新的教保服務法規；需適應與教保員一師一保的合作模式；學習從幼兒園、家庭及其社區取材，提供幼兒多樣化的經驗；要理解課程大綱內涵劃分為身體動作與健康、認知、語文、社會、情緒和美感六領域；進行教保活動時須試著以幼兒為主體，時時省思自己，觀察幼兒的需求；同時得配合扶助弱勢政策協助辦理補助。當面對此一改革，教保服務人員為了適應新的變革將衍生出何

種教保專業成長課程需求，此為本研究主要動機之二。

公立幼兒園教保服務人員之師資與私立幼兒園比較是相對穩定，根據林佩蓉（2010）與教保服務行動聯盟（2010）之研究發現，由於低薪、低福利、超時工作等因素，造成私立幼兒園教保人員的流動率高。其中有心致力於幼兒教育且具教保資格者，則努力考取公幼教師或公立幼兒園契約進用人員（教保員），多數無法在身心俱疲的工作環境中繼續奉獻者，則轉業另謀高就，甚是可惜。林惠娟（2005）的研究即發現，任職於公立機構者有較充裕的時間進行專業成長，研習次數和時數都較多，並對進修和研習機會感到充足，於是對進修和研習經驗也就有較高的滿意度。鑒於公立幼兒園教保服務人員有較充裕的時間進行專業成長且服務年限相當長，能將所學應用於實務現場，發揮長遠的影響力，因此有必要針對公立幼兒園教保服務人員進行教保專業成長現況與需求之調查，此為主要動機之三。

最後，位處南臺灣的高雄縣市各有其特色，高雄市為發展成熟之都會型城市，高雄縣幅員遼闊，涵蓋原住民鄉及沿海鄉鎮，兩者於2010年12月25日合併成高雄市，是具有38個行政區之大都市，幅員遼闊，轄區內有一般地區、偏遠地區、特偏地區，部份學校位在偏遠及特偏地區，交通不便、路途遙遠、時間無法配合，有礙教保服務人員參與專業成長。因此擬以合併後之大高雄地區公立幼兒園教保服務人員為主要研究對象，以進一步了解其教保專業成長之現況與需求，此為主要研究動機四。

基於上述，本研究採問卷調查之方式，希冀深入了解高雄市公立幼兒園教保服務人員之教保專業成長現況與需求，同時進一步分析不同個人背景因素的教保服務人員在教保專業成長之需求是否

有差異存在，希冀能提出具體建議，供師資培育機構、教育主管當局、學校行政人員作為規畫幼兒園教保服務人員教保專業成長活動之參考。為達上述之研究目的，本研究之主要研究問題有以下三項：

- (一) 高雄市公立幼兒園教保服務人員教保專業成長之現況為何？
- (二) 高雄市公立幼兒園教保服務人員教保專業成長之需求為何？
- (三) 不同背景變項的高雄市公立幼兒園教保服務人員，在教保專業成長需求之差異情形為何？

貳、文獻探討

一、幼兒園教保服務人員教保專業成長需求之內涵

教師專業成長已成為全世界學前教育改革的重點（OECD, 2005）。而所謂教師專業成長是一個逐步形成的概念，是有關教師的學習—學習如何應用他們的知識來幫助學生學習（He & Ho, 2013）。Jenlink 與 Kinnucan-Welsch（1999）提出專業成長涵蓋了個人之看法、理論及實務之改變，因此專業成長是伴隨著專業訓練活動之參與實施而讓個人之專業看法、理論、與實務變得更成熟，專業技術發展得更完整，專業知識量增多，專業自主權增強，提高服務利他的專業精神，專業倫理規範加深，顯現出後一個階段比前一階段的個體更高的層次發展。而 Clement 與 Vandenberghe（2000）亦認為專業成長是連續不間斷的過程，包括職前的準備到在職的導入、更新及精緻化，唯有連續的進修才能達成向上提升。為符應終身學習之教育趨勢，所謂專業成長是朝向專業知能的終身成長，特別強調其發展性，是終身不斷學習成長的歷程（黃淑苓，2005）。由此可知專業成長是一個不停止的動態過程，不間斷地進行專業發

展讓本身之專業知識、技能、倫理、精神、自主性趨向一種茁壯成熟的狀態。幼兒園教保服務人員如同每種專業人員，在工作生涯的不同階段存在各種專業知能上的需求，有賴不中斷的專業成長方能滿足現實工作的需要及內心驅動的需求。

張淑雲（2002）認為教師專業成長需求是指教師在教學生涯中，有感在教學知識、技能、態度上有所不足，想要進一步充實成長，所以透過各種活動以彌補現有的不足，漸達成理想狀態之過程。而林桂垣（2010）指出教師專業成長需求是教師在教學生涯過程中，覺知到個人在某些部分有所不足，產生學習動機，參與所需的成長活動，達成自我實現的目標。廖金貴（2014）則將幼兒園教師專業成長需求定義為在教師之專業工作生涯期間，自覺其所需的專業知識、技能、態度有所缺乏，藉由積極參與有助專業成長的活動來充實自我，以彌補現有的不足及改善教學品質，追求專業理想的過程。

綜上所述，幼兒園教保服務人員教保專業成長需求係為幼兒園教保服務人員在教學生涯歷程中，希望由感受到教保專業知能之不足的狀態逐漸轉變至理想狀態而產生的驅動力，或是追求人生存在的價值之動力。

而有關幼兒園教保服務人員教保專業成長需求之內涵，過去均會著眼於一位稱職的幼兒園教保服務人員所需具備的專業知能來加以界定，例如：教育部（1994）訂有《幼稚園教師手冊》，文中規範幼教老師應具備三種專業內涵：（一）發揮專業知能：包含教學原理和方法、了解幼兒發展情形、能設計和實施課程、懂得班級經營、做好幼兒保育和輔導；（二）堅守專業人員的規範：包含服務的理念和客觀、公正的態度、專業道德；（三）秉持優良的專業

精神。

葉郁菁在 1999 年訪談了 18 位任職於幼教機構的現職教師，以了解他們認為幼教師資應具備的專業能力，研究結果除教學、管理、溝通、專業及照護五大知能外，另提出專業教育信念及運用科技的能力。陳淑琴（2000）整理各學者的看法，則提出幼兒園教師須具有健全的人格、適性課程發展之理論與實務、幼兒發展與學習理論與實務、教室團體經營知能、幼兒觀察與評量知能、規劃與維持健康安全與支持性學習環境之能力、建構良好親師關係能力、特殊幼兒教育的理論與實務、正向溝通與人共處合作的能力、態度開放樂於接收新知勇於嘗試改變等十項專業知能。

美國幼兒教育學會（National Association for the Education of Young Children, NAEYC）發展的《2001 年學士或初級幼教專業人員標準》可作為我國幼兒園教保服務人員關鍵能力之參考，五項關鍵能力包含：（一）促進幼兒發展與學習；（二）與家長和社區建立良好關係；（三）藉由檔案文件、觀察和評量以支持幼兒和家長；（四）實踐有效教學；（五）成為幼教專業教師（引自賴春金，2003）。之後，2002 年美國幼兒教育學會又公佈《NAEYC 幼教專業準備的標準—進階證照綱要》強調進階幼教人員在程度上應比初階幼教人員能有較高層次且專門化的表現，更精準掌握初階幼教人員所應具備的能力。更進一步地，進階的幼兒教師應具備下列幾項專業技能，包括了解不同文化之能力、熟悉相關理論與研究、辨認及使用專業資料的能力、具有運用倫理原則的知識、優良的溝通技巧、了解研究方法並發展研究能力、與他人合作共同教學的能力、領導技巧及為政策辯駁的能力等（引自林春妙、楊淑朱，2005）。

林春妙和楊淑朱（2005）透過研究調查發現，幼教師應具備

幼教專業知識、教學能力、溝通能力、保育能力、經營班級能力、專業成長能力、園務行政能力等七個向度之專業知能。而范熾文與張瑜（2008）統整學者之觀點，歸納出幼兒園教師專業成長的內涵應有五個層面：通識素養、幼教知能、班級經營、研究發展、行政管理。

依據 2013 年教育部公布實施之《幼兒園教保專業知能研習實施辦法》規定教保專業知能課程範圍如下：（一）學前教保政策及法令；（二）幼兒園主管人員領導及行政管理；（三）幼兒園課程及教學；（四）幼兒園空間規劃及環境設計；（五）幼兒學習評量及輔導；（六）幼兒觀察解析及應用；（七）幼兒健康及安全；（八）學前融合教育；（九）教保專業倫理；（十）幼兒園園家互動及親師關係；（十一）其他有助於教保專業知能發展之課程。此法規明確定義了幼兒園教保服務人員教保專業之關鍵能力，成為幼兒園教保服務人員專業成長課程規劃之重要指標。

陳彥奇（2015）探討幼兒園教保服務人員發展與專業能力的關係時，綜合歸納幼兒園教保服務人員專業能力為：（一）熟知幼兒發展與學習；（二）整合學校、社區與家庭的關係；（三）實施合宜的學習評量；（四）照顧幼兒安全與健康；（五）進行專業進修與發展；（六）了解個別差異及提供需求；（七）設計多元適性的課程活動；（八）執行與實施學習活動。

檢視各方學者及機構對幼兒園教保服務人員教保專業知能之內涵雖有不同的分類方法，但在內容方面有極大的共識，重點在於教師必須要遵守專業倫理精神，了解幼兒之身心發展，作為設計適性、適齡、多元學習活動之基礎，並將教育專業知識落實在班級經營，實施有效教學，與同事、家長、行政之間維持良好的溝通與合作，且不忘進行專業成長，學習資訊科技、行政管理、研究施作等，

以因應時代所需。

本研究針對研究目的，參考 2014 年教育部國民及學前教育署臺教國署國字第 1030119578 號函，將教保研習主題規劃為教保課程與幼兒學習、政策與法令、園長（含負責人、校長及園主任）領導、兒童健康與照護、教師的情緒管理及人際關係與問題本位導向等六大主題課程，以六大主題課程為架構，融合上述各學者及機構對教保專業知能之研究成果，以作為本研究探討之幼兒園教保服務人員教保專業知能之內涵，以進一步作為規畫教師專業成長需求之依據，茲說明如下：

- (一) 教保課程與幼兒學習：此層面涵蓋甚廣，包含幼兒之學習活動及教師實施有效教學的所有環節，如具備廣度知識、熟悉教育理論、了解特殊幼兒及弱勢家庭幼兒教育、佈置學習環境、設計課程活動、運用班級經營技巧、進行行為輔導、觀察評量施作、運用科技的能力、了解不同文化之能力、觀摩學習之能力、融入重要議題之能力、掌握教育趨勢、注重反思與創新之能力等。
- (二) 政策與法令：此為重要的通識素養，幼兒園教保服務人員應了解學前教保政策及法令，並有為政策辯駁的能力，相關教保政策及法令包括幼照法與相關子法、性別相關法令（含性別平等、性侵害防治、性騷擾防治、性侵害或性騷擾事件處置及責任通報等）、法治教育、兒童及少年福利與權益保障法及家庭暴力防治法、網路安全、智慧財產權、個人資料保護法、學前弱勢學生就學及福利權益保障等、教保服務人員工作倫理、家長參與知能、勞動基準法概論及勞動權益保障、家庭教育、主管人員教保專業知能等。
- (三) 園長（含負責人、校長及園主任）領導：是指處理園所行政事務、推動與領導教保活動課程、了解幼兒園評鑑、良好的

領導與溝通、行政管理能力，注重專業發展並能妥善整合社區、家長與教師，營造美好和諧關係，經營高品質之園所。

- (四) 兒童健康與照護：須具備幼兒發展與保育知識、關注兒童保護、了解簡易醫療常識和基本救命術（心肺復甦術、基本創傷救命術、哈姆立克急救法）、落實生活教育、照顧幼兒營養與健康，而照護幼兒安全包括了用藥安全、遊戲安全、環境衛生安全、飲食安全、交通安全、防災教育等。
- (五) 教師的情緒管理及人際關係：主要是指認識自我與他人的感覺與情緒、說話的技巧及表達情緒的方法、時間管理及情緒管理。幼兒園教保服務人員能夠照顧好自己的情緒，以合宜的方式抒發心情，同理他人的情緒，運用良好的說話技巧來與行政人員、幼兒、家長、同事進行有效的溝通。
- (六) 問題本位學習：主要是指教保人員研究發展之能力，能檢核自己的實際問題與需求，察覺工作的現場問題，再以問題本位學習思考模式尋找出問題的因果脈絡，建構解決問題本位導向的研習主題與課程，實踐有效教學，目前規劃有幼兒園的策略領導及經營行銷系列、親師溝通系列、協同教保活動課程系列等相關內涵。

最後，本研究以上述幼兒園教保服務人員教保專業知能之內涵為基礎架構，同時參酌高雄市、臺南市、臺中市、臺北市、新北市、花蓮縣、澎湖縣等縣市所辦理之 104 年度教保專業研習課程彙整表，整理出有助幼兒園教保服務人員專業成長之課程項目與內容如表 1，以作為本研究設計幼兒園教保服務人員教保專業成長需求之主要內涵，總計六個面向共 37 個細項，特此說明。

表 1

教保服務人員教保專業成長面向與相關研習課程表

教保專業成長面向	相關研習課程
一、教保課程與幼兒學習	<ol style="list-style-type: none"> 1. 幼兒園課程模式介紹（華德福、瑞吉歐取向、主題與學習區、統整課程、方案教學法等） 2. 幼兒園優質經驗分享與傳承（優質幼兒園課程分享、教學卓越幼兒園參訪） 3. 幼兒閱讀（繪本深度閱讀、創作與演出故事） 4. 特殊需求幼兒的課程規劃與輔導（特殊需求幼兒概論、評量、教學和輔導策略、家庭溝通與協助） 5. 弱勢家庭幼兒的課程規劃與輔導（弱勢家庭及幼兒概論、教學和輔導策略、家庭溝通與協助） 6. 重要議題融入課程（如性別平等、品德、法治、生命、安全、防災教育等） 7. 幼兒評量與輔導及實務案例（學習評量概念、學習評量指標說明、指標分析與實例探討） 8. 幼兒觀察解析與應用及實務案例（幼兒發展之概念、幼兒學習與觀察紀錄） 9. 幼兒學習環境設計、規劃與實行及實務案例：學習區的運用、布置與經營、如何與角落活動融合
二、政策與法令	<ol style="list-style-type: none"> 1. 幼兒教育及照顧法與相關子法 2. 性別相關法令（性別平等、性侵害防治及性騷擾防治相關法令規章、性侵害或性騷擾事件處置及責任通報、性別意識建構） 3. 法治教育（兒童及少年福利與權益保障法、家庭暴力防治法、責任通報及輔導、網路安全、智慧財產權、個人資料保護法、學前弱勢學生就學及福利權益保障等） 4. 教保服務人員工作倫理（工作倫理之意涵、守則、實例分析、專業價值與實踐） 5. 家長參與知能：幼兒園如何推動家長參與 6. 勞動基準法概論及勞動權益保障實務案例研討 7. 學校附設幼兒園主管人員教保專業知能（課程與教學發展、教師專業發展、教學實務分享） 8. 家庭教育（親子衝突與溝通、家長如何瞭解幼兒情緒狀態、情緒發展並關懷陪伴）

(續下頁)

(接上頁)

教保專業成長面向	相關研習課程
三、園長（含負責人、校長及園主任） 領導	1. 領導與溝通（園主任和組長之領導與溝通、女性的溝通） 2. 教保活動課程領導（課程監督、帶領省思） 3. 園長（含負責人及校長）專業發展（學前教保政令、幼兒學習與課程、學習環境規劃等）
四、兒童健康與照護	1. 幼兒的飲食行為與發展 2. 幼兒期的營養需求及幼兒餐點的設計與烹煮 3. 幼兒健康飲食行為之教學策略（含不良飲食行為之導正） 4. 地球永續發展認識及其對幼兒教育的啟示 5. 二到六歲幼兒常見疾病與意外的預防與處理 6. 二到六歲幼兒的運動和休閒 7. 二到六歲幼兒發展與輔導 （幼兒語言與身體動作發展及其輔導策略） 8. 學步兒自理能力的培養 9. 二到六歲幼兒的遊戲、活動與規劃 10. 基本救命術訓練（心肺復甦術、基本創傷救命術、哈姆立克急救法等） 11. 安全教育（交通安全、飲食安全、遊戲安全、環境衛生安全、照護安全、防災教育及緊急救護情境演習 1 次以上）
五、教師的情緒 管理及人際關係	1. 認識自我與他人的感覺與情緒 2. 說話的技巧及表達情緒的方法 3. 時間管理及情緒管理
六、問題本位導向	1. 幼兒園的策略領導及經營行銷系列（幼兒園特色的建立、社區網絡的營造、團隊的營造系列） 2. 親師溝通系列（幼兒園策略行銷與親師溝通、幼兒園的內部與外部行銷系列） 3. 協同教保活動課程系列（協同教保活動課程的督導系列、教保服務人員的協同教保活動課程系列）

資料來源：1. 花蓮縣政府教育處（2015）。花蓮縣 104 年度教保研習彙整表。取自：
http://210.240.39.100/repair_log.asp?rid=A201503181532514958
2. 高雄市政府教育局（2015a）。高雄市 104 年度教保研習彙整表。取自：
<http://www.kh.edu.tw/forms/getDirectory/395>
3. 新北市幼教研習暨資訊中心（2015）。新北市 104 年度教保研習彙整表。取自：
<http://tpcpsis.kidedu.ntpc.edu.tw/files/11-1003-166.php>
4. 臺中市政府教育局（2015）。臺中市 104 年度教保研習彙整表。取自：
<http://www.tc.edu.tw/m/402>
5. 臺北市教師在職研習網（2015）。臺北市 104 年度教保研習彙整表。取自：
<http://insc.tp.edu.tw/index/ClsShowLstB.aspx>
6. 臺南市教保輔導團（2015）。臺南市 104 年度教保研習彙整表。取自：
<https://sites.google.com/site/tncetainan/you-ding-chang-ci>
7. 澎湖縣教保資源中心（2015）。澎湖縣 104 年度教保研習彙整表。取自：
http://kid.magong.com.tw/news_data.php?id=261

二、幼兒園教保服務人員教保專業成長之相關研究

綜合相關研究之結論及統整個人背景變項後，發現幼兒園教保服務人員的年齡、婚姻狀況、教學年資、擔任職務、幼兒園規模、學校地理位置等背景變項因素，對其教保專業成長產生某些程度的影響，因此本研究之自變項擬就此六個變項加以研究，以下針對相關文獻分析進一步探討此六個變項對教保服務人員教保專業成長的影響。

(一) 年齡

在年齡與教保專業成長的相關研究上，張媛甯與岳美秀（2012）研究發現不同年齡的教師在教師專業成長部分構面，年齡較高的教師得分較高，達顯著差異，與林亞萱（2011）的研究結果一致，年齡越長在教師專業發展的整體表現越好。吳靜惠（2005）透過問卷調查及訪談調查，發現不同年齡之教師對專業成長相關課程有不同之需求，其中在 25 歲以下、36-40 歲與 41 歲以上的教師對幼教專業知能相關課程需求為最低，年齡在 26-30 歲的教師對其他相關課程需求最低，而年齡在 31-35 歲的教師則對行政管理能力相關課程需求最低。陳彥奇（2011）的研究發現年齡會影響專業發展的類型，且幼教師年齡介於 40-44 歲者其差異比其他年齡層更為顯著，而廖金貴（2014）在澎湖縣幼兒園教師專業成長現況與需求之調查研究亦發現年齡不同的幼兒園教師在教保專業成長之班級經營層面上有顯著差異。由上可知，年齡越大者在教保專業成長上之得分有越高之趨勢，然就需求面向，則因年齡層不同而有異。

但在另一方面，陳惠珍（2003）的研究則發現不同年齡之幼兒園教師透過網路在職進修的態度並無顯著差異；曾曉苓（2005）研究結果為不同年齡之苗栗縣教保人員對在職進修方式或類型之需求未達到顯著差異；楊淑娟（2006）亦發現不同年齡的幼教師在整體的專業成長動機及其各層面未達到顯著差異；林亞萱（2011）研究發現幼兒教師的年齡和參與學士後在職進修的意願是沒有關聯性的。

（二）婚姻狀況

在婚姻狀況與教保專業成長的相關研究上，家庭因素是教保專業成長會面臨到的困難，有相當多老師在接受訪談時提及考量家庭因素，要照顧小孩及家中長輩，所以會選擇放棄專業成長活動，或是必須早早把成長活動的日期納入生活規畫中，提早安頓好家中老小，也因為家庭因素所以一學期參與專業成長活動的次數也不能太多（林惠娟，2005；范熾文、張瑜，2008）。

在婚姻狀況與教保專業成長的相關研究上，陳美齡（2004）研究發現已婚有小孩之幼兒教師較未婚教師在專業精神方面達顯著水準；林亞萱（2011）指出已婚之幼兒教師在整體的教師專業發展表現上較佳，但未婚教師在整體的進修意願高於已婚教師；吳靜惠（2005）的研究發現不同婚姻狀況的教師對班級經營相關課程都有極高的需求，但未婚教師對行政管理能力相關課程需求為最低，及

已婚教師則對幼教專業知能相關課程需求最低。另一方面，楊淑娟（2006）研究發現不同婚姻狀況其專業成長動機在他人影響層面上呈現不同的差異，但在整體的進修動機及其他層面未達到顯著差異；陳彥奇（2011）、廖金貴（2014）的研究亦發現不同婚姻狀況在專業發展、專業成長需求各層面無顯著差異。

（三）教學年資

林惠娟（2005）研究結果發現資深者傾向滿意進修和研習經驗；陳美齡（2004）的研究發現不同服務年資之教師最常參與的研習活動在實作分享及教學參觀達到顯著。張媛甯與岳美秀（2012）之研究結果為不同年資的教師對教師專業成長，在園務管理知能和專業信念態度上皆有顯著差異，且在教育專業知能層面，年資 8 到 14 年得分均高於 7 年以下和 15 年以上教師之得分；此研究發現與吳靜惠（2005）、陳彥奇（2011）研究結果相似，不同教學年資之幼教師在選擇專業發展類型及需求時會有所差異。而林亞萱（2011）經由問卷調查法研究發現資深之幼教師在教師專業發展的能力與表現、幼教專業態度、幼教教學技能上為佳，此發現與陳雪芳（2008）之研究結果相似。此外，部分學者研究發現，不同年資的幼教師對教師在職進修態度、專業成長動機、參與學士後進修之意願是未達顯著的（陳美齡，2004；陳惠珍，2003；曾曉苓，2005；楊淑娟，2006；廖金貴，2014）。

(四) 擔任職務

在擔任職務與教師教保專業成長的相關研究上，陳彥奇（2011）研究結果發現幼教師擔任不同職務，將影響其專業發展類型的選擇；林亞萱（2011）研究發現園所長在專業態度的表現上較佳。廖金貴（2014）之研究亦發現不同職務之幼兒園教師在教保專業成長需求各層面有顯著差異。然而，曾曉苓（2005）的研究則發現不同職稱之教保人員對在職進修方式或類型之需求未達到顯著；張媛甯與岳美秀（2012）研究結果為不同職務的教師在所有層面均無顯著差異；楊淑娟（2006）研究發現幼教師因職務不同，其在整體的進修動機及其他層面未達到顯著差異。綜整歸納上述研究有關教保服務人員擔任職務之分類，可分為：班級教師、班級教師兼行政（組長、主任、園長）及專任行政（主任、園長），本研究將參考此一分類方式，進行相關問卷之設計與分析。

(五) 幼兒園規模

部分學者研究發現幼兒園規模與教師教保專業成長需求與表現上有顯著差異。曾曉苓（2005）研究發現園所規模為 101~150 人的教保人員對行為觀察與評量、當前幼教相關政策、言談與肢體禮儀、各式活動之舉辦四個項目上之需求顯著高於園所規模為 50 人以下的教保人員。林亞萱（2011）研究發現園所幼兒數 31~60 人之幼兒教師在整體向度、幼教專業知能、幼教教學技能表現上，較 121 人以上為佳；園所幼兒人數在 30 人以下、31~60 人、91~120 人

的幼兒教師在幼教知能的表現上優於 121 人以上。張媛甯與岳美秀（2012）研究發現不同幼兒園規模的教師在多數構面有顯著差異存在，在普通素養知能、教育專業知能構面，3~4 班得分高於 1~2 班；在專業信念態度構面，3~4 班得分高於 1~2 班、5~6 班、7 班以上得分，1~2 班得分也高於 7 班以上得分。

另一方面，陳雪芳（2008）探討幼稚園教師專業成長與教學效能，研究發現不同規模之幼兒園在教師專業成長的三個層面上都未達顯著差異，表示不同規模幼兒園之教師其各層面的專業成長大致相同，與楊淑娟（2006）探討幼稚園教師在職進修與教學效能關係之研究結果相似，不同規模之幼兒園的幼教師在整體的專業成長動機及各層面未達到顯著差異。

（六）學校地理位置

在學校位置與教師教保專業成長之相關研究上，陳雪芳（2008）研究發現，彰化縣教師在通用知能與幼教專業領域及教育專業知能兩個層面上的得分高於臺中縣與南投縣；翁素雅（2004）研究發現不同任教地區的幼兒園實習輔導老師，在專業成長的途徑上達顯著差異；林亞萱（2011）研究發現服務於臺中市地區之幼兒教師在整體的教師專業發展表現上較彰化、南投為佳；陳彥奇（2011）研究發現屏東和高雄縣市幼兒園學校地理位置，會影響其專業發展類型需求。另一方面，陳美齡（2004）、楊淑娟（2006）分

別探討屏東、高雄縣市之幼兒園教師專業成長，研究結果皆為不同學校地理位置在整體專業成長及動機上無顯著差異。

綜整上述相關研究結果可知，不同年齡、婚姻狀況、教學年資、擔任職務、幼兒園規模及學校地理位置之幼兒園教保服務人員，在整體教保專業成長及各個層面上，是否存在差異性，尚無一致性的看法與結果，且以高雄地區為研究範圍的研究各有不同的研究結果，故本研究擬針對這些背景變項作進一步的探討。

參、研究設計與實施

一、研究架構

本研究之目的是以幼兒園教保服務人員專業成長為研究主題，主要在探討背景變項與教師專業成長需求之差異情形。其研究架構圖如圖 1 所示：

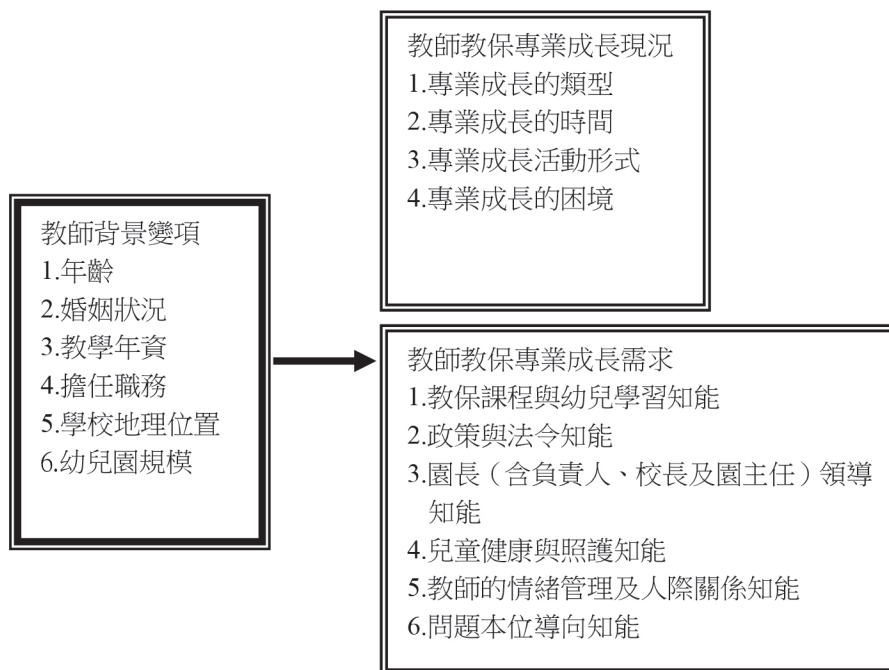


圖 1 研究架構圖

二、研究對象

本研究以高雄市公立幼兒園所有現職教師為正式對象，依據高雄市教育局（2015b）所公告之高雄市 104 學年度第一學期公立幼兒園名錄之資料，公立幼兒園共有 214 所（一般地區 128 所、偏遠地區 15 所及特篇地區 19 所），教師總數為 950 人。本研究針對任教於高雄市公立幼兒園之教保服務人員（含代理、代課教師）、班級教師兼任行政、教保員、專任行政（主任、園長）共 950 人，採用普查之方式，進行調查問卷之施測。根據回收問卷，扣除填答不全之問卷，總計得到 681 份有效問卷，有效回收率為 72%。

三、研究工具

本研究以問卷調查法進行研究，以「高雄市公立幼兒園教保服務人員教保專業成長現況與需求調查問卷」為研究工具，係依據文獻分析及研究架構所得，並參考國內學者（陳雪芳，2008；曾曉苓，2005；楊淑娟，2006；廖金貴，2014）所做的教師專業成長之調查問卷加以設計，編製而成。問卷內容分為三個部份，分別為教保服務人員背景資料、教保專業成長現況量表、教保專業成長需求量表，茲說明如下：

- （一）教保服務人員背景資料：年齡、婚姻狀況、教學年資、擔任職務、幼兒園規模、學校地理位置等六項。
- （二）教保專業成長現況量表：目的在了解幼兒園教保服務人員參加教保專業成長活動的類型、時間、活動形式及困境層面。其中除「參加教保專業成長活動理想時間」為單選題外，其餘均為複選題。
- （三）教保專業成長需求量表：主要在了解幼兒園教保服務人員對教保專業成長相關課程的需求，包括教保課程與幼兒學習知能、政策與法令知能、園長（含負責人、校長及園主任）領導知能、兒童健康與照護知能、教師的情緒管理及人際關係知能、問題本位導向知能六大層面。其填答計分方式採李克特式（*Likert-type Scale*）四點量表，由受試者依其個人知覺，針對問卷之各題項勾選適合選項填答選項分別為「非常需要」、「有些需要」、「不太需要」、「極不需要」，依序給予 4 分、3 分、2 分、1 分。量表之得分愈高者表示受試者對教保專業成長需求層面之知覺度越高。

為增加內容效度，首先函請 8 位幼教領域之學者專家針對本問卷初稿內容之適切性詳加審查，並且提供修正意見。修正後之問卷進行預試後，首先經過項目分析，採用決斷值與相關分析兩種方

法，所有題項均全數保留；接下來將保留的題目則依各層面進行主成分分析，以考驗其建構效度。為了解量表題目的「因素分析適合性」，於進行因素分析前，先進行 KMO(取樣適切性量數)與 Bartlett's 球形檢定，經 KMO 與 Bartlett's 球形檢定分析各向度之因素分析性極佳(顯著性均達.01 顯著水準)後，本量表之因素分析採主成分分析法，以分析各題目在該因素之因素負荷量，以特徵值大於 1 及因素負荷量大於.30 者為選入因素參考標準，統計分析結果所有題目皆符合此一標準，故全數保留，而六個因素的解釋變異量分別是「教保課程與幼兒學習」42.95%、「政策與法令」57.48%、「園長(含負責人、校長及園主任)領導」88.24%、「兒童健康與照護」62.03%、「教師的情緒管理及人際關係」87.11%、「問題本位導向」77.53%。最後，依各分量表之保留題目進行 Cronbach α 信度係數考驗，以確定其內部一致性。各向度的 α 係數分別為「教保課程與幼兒學習」.83 (9 題)、「政策與法令」.89 (8 題)、「園長(含負責人、校長及園主任)領導」.93 (3 題)、「兒童健康與照護」.93 (10 題)、「教師的情緒管理及人際關係」.93 (3 題)、「問題本位導向」.86 (3 題)，均大於 .70 以上，顯示本量表之內部一致性佳。

四、資料處理與分析

本研究以次數分配、百分比、平均數、標準差等描述性統計方式，分析與呈現高雄市公立幼兒園教保服務人員之教保專業成長現況及專業成長需求之情形。另以 t 考驗及單因子變異數分析以呈現高雄市公立幼兒園教保服務人員不同背景變項在教保專業成長需求上的差異情形。

肆、研究結果與討論

一、高雄市公立幼兒園教保服務人員教保專業成長之現況分析

(一) 教保專業成長的類型

高雄市公立幼兒園教保服務人員主要參加的教保專業成長的類型，以從事校內外研習(636人，佔93.4%)居多，其次是線上研習(163人，佔23.9%)及學分或學位進修者(93人，佔13.7%)，而填其他的有6人(佔0.9%)。另外，沒有參加任何專業成長活動的有10人(佔1.5%)，可看出仍以一般專業成長類型為主、其次是正式專業成長類型，而自我專業成長類型則有待加強。

(二) 教保專業成長議題

高雄市公立幼兒園教保服務人員參與教保專業成長內容屬於教保課程與幼兒學習議題的最多，有648位(95.2%)；其次依序為兒童健康與照護，有448位(65.8%)；政策與法令，有301位(44.2%)；情緒管理與人際關係，有162位(23.8%)；園長領導，有76位(11.2%)；問題本位導向，有27位(4.0%)；填其他的有26位(3.8%)。

(三) 園所安排參與教保專業成長的方式

高雄市公立幼兒園所對於教保服務人員參加教保專業成長的安排方式，主要採自願者參加為最多，共650位

(95.4%)；其次為教師自行輪派，有 13 位 (1.9%)；學校指定參加，有 11 位 (1.6%)；其他，有 5 位 (0.7%)。

(四) 參加教保專業成長之實際與理想時間

教保服務人員目前參加教保專業成長活動的主要時間以學期中週末為最多，有 575 位 (84.4%)；暑假次之，有 422 位 (62.0%)；寒假第三順位，有 312 位 (45.8%)；第四順位是學期中週一至週五白天，有 39 位 (5.7%)；第五順位是學期中週一至週五晚上，有 31 位 (4.6%)；最少的是其他，有 21 位 (3.1%)，包含有空時或不定時。

而教保服務人員認為較理想之時間以暑假為最多，有 447 位 (65.6%)；學期中週末次之，有 295 位 (43.3%)；寒假第三順位，有 208 位 (30.5%)；第四順位是學期中週一至週五白天，有 52 位 (7.6%)；第五順位是學期中週一至週五晚上，有 32 位 (4.7%)；最少的是其他，有 19 位 (2.8%)，其中有多位認為理想時間為週三下午固定進修或為開學準備週。

(五) 教保專業成長主要活動形式

教保服務人員目前進行教保專業成長主要的活動形式，依序為專題演講居多，有 508 位 (74.6%)；學術研討會/座談會，有 255 位 (37.4%)；專業討論/交換工作經驗，有 201 位 (29.5%)；個人閱讀教育書刊，有 136 位 (20.0%)；校外教學觀摩，有 129 位 (18.9%)；運用教育網路平臺進行線上學習，有 126 位 (18.5%)；工作坊，有 125 位 (18.4%)；實作演練，有 105 位 (15.4%)；參與輔

導計畫，有 86 位 (12.6%)；撰寫省思札記，有 81 位 (11.9%)；校內教學觀摩，有 75 位 (11.0%)；參加成長團體，有 74 位 (10.9%)；讀書會共讀，有 32 位 (4.7%)；從事行動研究，有 18 位 (2.6%)；其他，有 9 位 (1.3%)。

(六) 一學期參加教保專業成長活動之時數

教保服務人員一學期參加教保專業成長活動時數以 18~36 小時居多，有 483 位 (70.9%)；其次為 37 小時以上，有 106 位 (15.6%)；17 小時 (含) 以下，有 92 位 (13.5%) 為最少。

(七) 教保專業成長的困境

教保服務人員參加教保專業成長活動的困境，分別是參加名額有限為最多，有 388 位 (57.0%)；其次是進修地點距離太遠，有 372 位 (54.6%)；研習排在上班時間，有 260 位 (15.3%)；家庭因素無法配合，有 209 位 (12.3%)；教保工作繁重，有 200 位 (11.8%)；課程內容不符個人需求，有 175 位 (10.3%)；成長制度未與教師績效考評相結合，有 31 位 (1.8%)；園所不支持，有 14 位 (0.8%)。值得重視的是選擇其他項的有 31 位 (1.8%)，多位提到的困境為週末研習無法休息，其次是研習內容太少未多元化、研習日期重疊、代課老師之身分較為受限、無法補休、幼兒園無固定進修日、交通不方便、不能跨縣市參加、名額場次太少、校內研習不開放給校外教師、研習形式以講座為主，少實務指導等。

二、高雄市公立幼兒園教保服務人員教保專業成長之需求分析

(一) 教保專業成長之相關課程整體需求程度

教保專業成長課程包含「教保課程與幼兒學習」、「政策與法令」、「園長（含負責人、校長及園主任）領導」、「兒童健康與照護」、「情緒管理與人際關係」、「問題本位導向」等六個層面，本研究工具為四點量表，中間值為 2.5 分，若得分超過 2.5 分越多，表示對該層面的需求越高。

高雄市公立幼兒園教保服務人員對於教保專業成長六個層面的平均數分布在 3.20~3.4 之間，整體來說對於教保專業成長課程的需求屬於中上程度（詳見表 2）。就平均數高低排序可知，需求最高的教保專業成長課程，依序為兒童健康與照護、教保課程與幼兒學習、情緒管理與人際關係、問題本位導向、政策與法令，而園長（含負責人、校長及園主任）領導需求屬最低。

表 2

教保專業成長需求之各分量表分析

類別	題數	各量表總分	各量表平均數	各量表標準差
教保課程與幼兒學習	9	30.68	3.41	0.38
政策與法令	8	26.04	3.26	0.47
園長（含負責人、校長及園主任）領導	3	9.60	3.20	0.66
兒童健康與照護	10	34.47	3.45	0.46
情緒管理與人際關係	3	9.94	3.31	0.61
問題本位導向	3	9.85	3.28	0.55

(二) 教保專業成長之相關課程各類別需求

1. 教保課程與幼兒學習相關課程之需求

高雄市公立幼兒園教保服務人員對教保課程與幼兒學習相關課程之需求，各題項平均數介於 3.29~3.53，其中以幼兒閱讀（繪本深度閱讀、創作與演出故事）的平均數為最高（ $M=3.53$ ），其次是幼兒園優質經驗分享與傳承（優質幼兒園課程分享、教學卓越幼兒園參訪）與幼兒學習環境設計、規劃與實行及實務案例：學習區的運用、布置與經營、如何與角落活動融合（ $M=3.47$ ）；而重要議題融入課程（如性別平等、品德、法治、生命、安全、防災教育等）的平均數為最低（ $M=3.29$ ）。

2. 政策與法令相關課程之需求

高雄市公立幼兒園教保服務人員對政策與法令相關課程之需求，各題項平均數介於 3.09~3.36，其中以家庭教育（親子溝通、家庭對幼兒發展之影響）的平均數為最高（ $M=3.36$ ），其次是幼兒教育及照顧法與相關子法（ $M=3.33$ ），第三順位是幼兒園主管人員教保專業知能（課程與教學發展、教師專業發展、教學實務分享）（ $M=3.32$ ），而以勞動基準法概論及勞動權益保障實務案例研討的平均數最低（ $M=3.09$ ）。

3. 園長（含負責人、校長及園主任）領導相關課程之需求

高雄市公立幼兒園教保服務人員對園長（含負責人、校長及園主任）領導相關課程之需求，各題項平均數介於 3.16~3.25，其中以教保活動課程領導（課程領導、帶領省思）

的平均數為最高 ($M=3.25$)，其次是園長 (含負責人及校長) 之專業發展 (學前教保政令、幼兒學習與教保課程、學習環境規劃等) ($M=3.19$)，而以領導與溝通 (園主任和組長之領導與溝通技巧) 的平均數最低 ($M=3.16$)。

4. 兒童健康與照護相關課程之需求

高雄市公立幼兒園教保服務人員對兒童健康與照護相關課程之需求，各題項平均數介於 3.27~3.57，其中以基本救命術訓練 (心肺復甦術、基本創傷救命術、哈姆立克急救法等) 的平均數為最高 ($M=3.57$)，其次是安全教育 (交通安全、飲食安全、遊戲安全、環境衛生安全、照護安全、防災教育及緊急救護情境演習) ($M=3.53$)，而以幼兒期的營養需求及幼兒餐點的設計與烹煮的平均數最低 ($M=3.27$)。

5. 情緒管理與人際關係相關課程之需求

高雄市公立幼兒園教保服務人員對情緒管理與人際關係相關課程之需求，各題項平均數介於 3.28~3.38，其中以說話的技巧及表達情緒的方法的平均數為最高 ($M=3.38$)，其次是認識自我與他人的感覺與情緒、及時間管理與人際互動 ($M=3.28$)。

6. 問題本位導向相關課程之需求

高雄市公立幼兒園教保服務人員對問題本位導向相關課程之需求，各題項平均數介於 3.22~3.32，其中以協同教保活動課程系列 (協同教保活動課程的領導系列、教保服務人員的協同教保活動課程系列) 的平均數為最高 ($M=3.32$)，其次是親師溝通系列 (幼兒園策略行銷與親師溝通、幼兒園

的內部與外部行銷系列) ($M=3.30$)，而以幼兒園的策略領導及經營行銷系列(幼兒園特色的建立、社區網絡的營造、社區網絡的營造、團隊的營造系列)的平均數最低($M=3.22$)。

三、不同背景變項之幼兒園教保服務人員在教保專業成長需求之差異分析

(一) 年齡

由表 3 中可知，不同年齡的幼兒園教保服務人員在專業成長需求之教保課程與幼兒學習、政策與法令、園長(含負責人、校長及園主任)領導、兒童健康與照護、問題本位導向等層面未達顯著差異，僅在情緒管理與人際關係層面($F=3.79, p<.05$)達顯著差異，經以 Scheffé 法進行事後比較，結果發現年齡在 51 歲以上的教師，其得分顯著高於 30 歲以下、31-40 歲及 41-50 歲之教師。

表 3

不同年齡教師在教保專業成長需求之單因子變異數分析摘要表

層面	年齡	人數	平均數	標準差	F 值	事後比較
教保課程與 幼兒學習	1. 30 歲以下	116	3.43	0.34	2.04	
	2. 31-40 歲	211	3.44	0.39		
	3. 41-50 歲	277	3.37	0.38		
	4. 51 歲以上	77	3.44	0.38		
政策與法令	1. 30 歲以下	116	3.22	0.41	1.43	
	2. 31-40 歲	211	3.27	0.48		
	3. 41-50 歲	277	3.23	0.49		
	4. 51 歲以上	77	3.34	0.51		

(續下頁)

(接上頁)

層面	年齡	人數	平均數	標準差	F 值	事後比較
園長(含負責人、校長及園主任)領導	1. 30 歲以下	116	3.18	0.59	2.34	
	2. 31-40 歲	211	3.20	0.59		
	3. 41-50 歲	277	3.16	0.70		
	4. 51 歲以上	77	3.38	0.74		
兒童健康與照護	1. 30 歲以下	116	3.44	0.47	2.51	
	2. 31-40 歲	211	3.43	0.46		
	3. 41-50 歲	277	3.40	0.48		
	4. 51 歲以上	77	3.57	0.42		
情緒管理與人際關係	1. 30 歲以下	116	3.27	0.55	3.79*	4>1
	2. 31-40 歲	211	3.27	0.63		4>2
	3. 41-50 歲	277	3.30	0.62		4>3
	4. 51 歲以上	77	3.53	0.56		
問題本位導向	1. 30 歲以下	116	3.26	0.51	2.05	
	2. 31-40 歲	211	3.29	0.52		
	3. 41-50 歲	277	3.25	0.60		
	4. 51 歲以上	77	3.42	0.53		

* $p < .05$

(二) 婚姻狀況

由表 4 中可知不同婚姻狀況的幼兒園教保服務人員在教保專業成長需求各層面上之差異性皆大於 .05 顯著水準，顯示不同婚姻狀況的幼兒園教保服務人員在教保專業成長需求各層面上並無顯著差異。

表 4

不同婚姻狀況教師在教保專業成長需求之 t 考驗摘要表

層面	婚姻	人數	平均數	標準差	t 值
教保課程與幼兒學習	1. 未婚	236	3.44	0.35	1.59
	2. 已婚	445	3.39	0.39	
政策與法令	1. 未婚	236	3.27	0.46	0.41
	2. 已婚	445	3.25	0.48	
園長（含負責人、校長及園主任）領導	1. 未婚	236	3.21	0.62	0.23
	2. 已婚	445	3.20	0.68	
兒童健康與照護	1. 未婚	236	3.41	0.48	-1.13
	2. 已婚	445	3.45	0.46	
情緒管理與人際關係	1. 未婚	236	3.29	0.62	-0.61
	2. 已婚	445	3.32	0.60	
問題本位導向	1. 未婚	236	3.28	0.52	0.02
	2. 已婚	445	3.28	0.57	

（三）教學年資

由表 5 中得知，不同教學年資的幼兒園教保服務人員在專業成長需求之政策與法令、園長（含負責人、校長及園主任）領導、兒童健康與照護、情緒管理與人際關係、問題本位導向等層面均未達顯著差異。不同教學年資的幼兒園教保服務人員雖在教保課程與幼兒學習層面（ $F=3.28, p<.05$ ）達顯著差異，但經以 Scheffé 法進行事後比較發現，兩兩組別之間並未達顯著差異。

表 5

不同教學年資教師在教保專業成長需求之單因子變異數分析摘要表

層面		年齡	人數	平均數	標準差	F 值	事後比較
教保課程與 幼兒學習	1.	5 年以下	157	3.45	0.35	3.28*	n.s.
	2.	6-10 年	105	3.46	0.37		
	3.	11-15 年	118	3.44	0.37		
	4.	16-20 年	134	3.32	0.39		
	5.	21 年以上	167	3.38	0.39		
政策與法令	1.	5 年以下	157	3.25	0.47	0.22	
	2.	6-10 年	105	3.27	0.44		
	3.	11-15 年	118	3.27	0.46		
	4.	16-20 年	134	3.22	0.49		
	5.	21 年以上	167	3.27	0.50		
園長(含負責 人、校長及園 主任)領導	1.	5 年以下	157	3.14	0.63	1.10	
	2.	6-10 年	105	3.29	0.60		
	3.	11-15 年	118	3.15	0.63		
	4.	16-20 年	134	3.24	0.67		
	5.	21 年以上	167	3.21	0.73		
兒童健康與 照護	1.	5 年以下	157	3.45	0.47	0.46	
	2.	6-10 年	105	3.48	0.45		
	3.	11-15 年	118	3.42	0.46		
	4.	16-20 年	134	3.44	0.46		
	5.	21 年以上	167	3.41	0.48		
情緒管理與 人際關係	1.	5 年以下	157	3.28	0.59	0.60	
	2.	6-10 年	105	3.35	0.59		
	3.	11-15 年	118	3.26	0.60		
	4.	16-20 年	134	3.34	0.62		
	5.	21 年以上	167	3.34	0.62		
問題本位 導向	1.	5 年以下	157	3.24	0.55	0.47	
	2.	6-10 年	105	3.34	0.53		
	3.	11-15 年	118	3.28	0.48		
	4.	16-20 年	134	3.29	0.52		
	5.	21 年以上	167	3.28	0.64		

* $p < .05$

註：n.s.表示事後比較結果各組間並無顯著差異

(四) 擔任職務

由表 6 中得知，不同職務的幼兒園教保服務人員，在教保課程與幼兒學習、政策與法令、兒童健康與照護、情緒管理與人際關係等層面均未達顯著差異，但在專業成長需求之園長（含負責人、校長及園主任）領導（ $F=4.80$ ， $p<.01$ ）和問題本位導向（ $F=4.80$ ， $p<.01$ ）等層面達顯著差異，經以 Scheffé 法進行事後比較，其結果分析如下：

1. 就園長（含負責人、校長及園主任）領導層面而言，擔任職務為專任行政（主任、園長）的教師其得分顯著高於擔任班級教師（含代理、代課教師）及教保員之教師。
2. 就問題本位導向層面而言，擔任班級教師兼任行政（組長、主任、園長）的教師其得分顯著高於擔任教保員之教師。

表 6

不同職務教師在教保專業成長需求之單因子變異數分析摘要表

層面	職務	人數	平均數	標準差	F 值	事後比較
教保課程與 幼兒學習	1. 班級教師（含代理、 代課教師）	438	3.41	0.38	1.80	
	2. 班級教師兼任行政 （組長、主任、園長）	64	3.48	0.36		
	3. 教保員	167	3.38	0.37		
	4. 專任行政 （主任、園長）	12	3.56	0.38		
政策與法令	1. 班級教師 （含代理、代課教師）	438	3.23	0.47	2.05	
	2. 班級教師兼任行政 （組長、主任、園長）	64	3.36	0.52		
	3. 教保員	167	3.28	0.47		
	4. 專任行政 （主任、園長）	12	3.39	0.30		

（續下頁）

(接上頁)

層面	職務	人數	平均數	標準差	F 值	事後比較
園長(含負責人、校長及園主任)領導	1. 班級教師 (含代理、代課教師)	438	3.17	0.65	4.80**	4>1 4>3
	2. 班級教師兼任行政 (組長、主任、園長)	64	3.34	0.64		
	3. 教保員	167	3.19	0.69		
	4. 專任行政 (主任、園長)	12	3.81	0.33		
兒童健康與照護	1. 班級教師 (含代理、代課教師)	438	3.42	0.46	0.62	
	2. 班級教師兼任行政 (組長、主任、園長)	64	3.48	0.47		
	3. 教保員	167	3.46	0.49		
	4. 專任行政 (主任、園長)	12	3.55	0.36		
情緒管理與人際關係	1. 班級教師(含代理、 代課教師)	438	3.30	0.60	1.20	
	2. 班級教師兼任行政 (組長、主任、園長)	64	3.39	0.58		
	3. 教保員	167	3.31	0.64		
	4. 專任行政 (主任、園長)	12	3.58	0.47		
問題本位導向	1. 班級教師 (含代理、代課教師)	438	3.28	0.54	4.80**	2>3
	2. 班級教師兼任行政 (組長、主任、園長)	64	3.48	0.50		
	3. 教保員	167	3.21	0.58		
	4. 專任行政 (主任、園長)	12	3.56	0.46		

** $p < .01$

(五) 學校地理位置

由表 7 中可知不同地理位置的幼兒園教保服務人員，在教保專業成長需求各層面上之差異性皆大於 .05 顯著水

準，顯示不同地理位置的幼兒園教保服務人員在教保專業成長需求各層面上之差異情形均無顯著差異。

表 7

不同地理位置教師在教保專業成長需求之單因子變異數分析摘要表

層面	地理位置	人數	平均數	標準差	F 值
教保課程與幼兒學習	1. 一般地區	612	3.41	0.37	0.89
	2. 偏遠地區	43	3.47	0.39	
	3. 特偏地區	26	3.35	0.38	
政策與法令	1. 一般地區	612	3.26	0.47	1.25
	2. 偏遠地區	43	3.33	0.58	
	3. 特偏地區	26	3.14	0.46	
園長（含負責人、校長及園主任）領導	1. 一般地區	612	3.19	0.66	1.47
	2. 偏遠地區	43	3.36	0.66	
	3. 特偏地區	26	3.12	0.62	
兒童健康與照護	1. 一般地區	612	3.44	0.47	2.26
	2. 偏遠地區	43	3.56	0.43	
	3. 特偏地區	26	3.32	0.42	
情緒管理與人際關係	1. 一般地區	612	3.31	0.60	0.50
	2. 偏遠地區	43	3.38	0.68	
	3. 特偏地區	26	3.23	0.54	
問題本位導向	1. 一般地區	612	3.28	0.55	0.21
	2. 偏遠地區	43	3.33	0.62	
	3. 特偏地區	26	3.26	0.51	

（六）幼兒園規模

由表 8 中可知，不同幼兒園規模的教師，在教保課程與幼兒學習、園長（含負責人、校長及園主任）領導、兒童健康與照護、問題本位導向等層面均未達顯著差異，但在教保專業成長需求之政策與法令層面（ $F=6.21, p<.01$ ）

及情緒管理與人際關係層面 ($F=8.60, p<.001$) 均達顯著差異，經以 Scheffé 法進行事後比較，其結果分析如下：

- 1.就政策與法令層面而言，幼兒園規模 5 班以上的教師，其得分顯著高於 1-2 班及 3-4 班之教師。
- 2.就情緒管理與人際關係層面而言，幼兒園規模 5 班以上的教師，其得分顯著高於 1-2 班及 3-4 班之教師。

表 8

不同幼兒園規模教師在教保專業成長需求之單因子變異數分析摘要表

層面	幼兒園規模	人數	平均數	標準差	F 值	事後比較
教保課程與幼兒學習	1. 1-2 班	293	3.40	0.37	1.11	
	2. 3-4 班	195	3.39	0.37		
	3. 5 班以上	193	3.44	0.39		
政策與法令	1. 1-2 班	293	3.22	0.48	6.21**	3>1 3>2
	2. 3-4 班	195	3.21	0.46		
	3. 5 班以上	193	3.36	0.47		
園長(含負責人、校長及園主任)領導	1. 1-2 班	293	3.16	0.66	2.55	
	2. 3-4 班	195	3.17	0.60		
	3. 5 班以上	193	3.29	0.70		
兒童健康與照護	1. 1-2 班	293	3.40	0.46	2.57	
	2. 3-4 班	195	3.44	0.47		
	3. 5 班以上	193	3.50	0.48		
情緒管理與人際關係	1. 1-2 班	293	3.25	0.63	8.60***	3>1 3>2
	2. 3-4 班	195	3.26	0.57		
	3. 5 班以上	193	3.46	0.58		
問題本位導向	1. 1-2 班	293	3.25	0.57	1.90	
	2. 3-4 班	195	3.26	0.51		
	3. 5 班以上	193	3.35	0.57		

** $p < .01$, *** $p < .001$

四、綜合討論

(一) 教保專業成長的現況

1. 教保專業成長的類型

由上述統計結果發現，高雄市公立幼兒園教保服務人員參加教保專業成長的類型以一般專業成長中的「校內外研習」和「線上研習」最多，主要原因是校內外研習活動大多為教育局主辦，輔以公私立幼兒園承辦，一學年規劃近百場次，所有研習資訊皆公開於全國教師進修網，教師可依自己的興趣選擇合適的時間、地點前往參加，且幾乎都是免費，在研習會場可與其他園所教保服務人員，進行人際互動與教學經驗交流，為教保服務人員最主要參與專業進修的類型。此研究結果與陳惠珍（2003）、陳美齡（2004）、曾曉苓（2005）、廖金貴（2014）相同，即幼兒園教保服務人員專業成長多以參加教育局辦理的研習居多。

排序第二的「線上研習」，是教師們在實體研習外的最佳選擇，一來是沒有人數名額的限制，二來是時間自由又彈性，第三是網際網路發達，提高線上研習的便捷性，同時教育當局亦投注心力架設線上研習網站，如教育部數位學習服務平臺、地方行政研習 e 學中心、港都 e 學苑、公務人員終身學習網、臺北 e 大、文官 e 學苑...等，網站將課程分門別類，教師可依需求點選閱覽，相當便利，特別是長期推動的主題，如環境教育、性平教育，局端都會公文來函要求教師完成規定之線上研習時數。事實上，線上研習被視為是能提

供幼兒園教保服務人員接觸專業發展、進而改善其教學品質的機會，尤其是當面對面實體訓練或研習機會受限於經費時，線上及科技為中介的學習將可開創一永續發展的教育機會（Stone-MacDonald & Douglass, 2015）。根據 Lavonen、Jauhainen、Koponen 與 Kurki-Suonio（2004）的看法，傳統的教師專業成長課程常由專家學者進行數小時的演講，傳授知識、策略或技術，較難推動教師進行教學變革與成長；Bransford、Brown 與 Cocking（2000）在探討人類如何學習時，提出老師如同學生，老師的學習同樣需要考慮到學習者的需求與興趣。為了符應學習者多元的需求並提高學習的興趣，教師專業成長活動之形式宜多樣化，其中線上研習是一可考量的方式。

2.教保專業成長議題

本研究的調查結果發現，高雄市公立幼兒園教保服務人員主要參加的教保專業成長的議題，其排序前三名依次為：兒童健康與照護、教保課程與幼兒學習議題、政策與法令。由此顯示，目前教師最關注的是攸關兒童健康與照護的專業知能與技術，例如基本救命術訓練（心肺復甦術、基本創傷救命術、哈姆立克急救法等）、安全教育（交通安全、飲食安全、遊戲安全、環境衛生安全、照護安全、防災教育及緊急救護情境演習），亦是目前各縣市教育局規定每位教師都要參與的課程。其次關注的是教保專業中之教與學的歷程，思考新課程大綱之六大領域教保課程之內容、實施方

式、學習指標，致力於創新每日的課程設計與教學活動，提升自我的專業教學知能。本研究與曾曉苓在 2005 年的研究發現有所不同，她的研究指出，教師在親師合作、幼兒發展與學習、特殊教育等面向有較高的需求，究其原因可能與調查的時間點有關，幼照法 2012 年 1 月 1 日開始實施，幼托進行整合，於是教師所關注的面向轉為新課程大綱、兒童健康照護內涵及相關法令措施所致。而本研究與近期之廖金貴（2014）的研究結果相似，她的研究發現教師專業成長內容以教學知能居多，顯示幼教師最專注投入的面向仍為課程與教學知能。

3. 園所安排參與教保專業成長的方式

由問卷調查結果發現，高雄市公立幼兒園所安排參與教保專業成長的方式依序為自願者參加、教師自行輪派、學校指定參加、其他、抽籤決定。可見幼兒園對於教師參與教保專業研習多採取依教師意願，自由彈性的做法，甚少介入干預，至於少數研習規定要派員參加的研習，通常也是鼓勵教師參加，若無人自願，則採用輪派的方式。本研究結果與楊淑娟（2006）之研究發現相同，園所多採自願參加之方式。

4. 參加教保專業成長的實際與理想時間

（1）參加教保專業成長之實際時間

根據問卷調查結果，高雄市公立幼兒園教保服務人員參加教保專業成長的主要時間，其排序前三名依次為學期中週末、暑假、寒假為最多。究其原因，是因教師認為於週一至週五上班日參與研習，需事先尋找合適的代課老

師、自付代課費及使用個人事假離校，較為不便，因此大部分教師選擇週末之研習，較不會影響班務及幼兒。此研究結果與吳靜惠(2005)、曾曉苓(2005)、廖金貴(2014)相符，即周末是幼兒園教保服務人員最主要參加教保專業成長的時間。

(2) 參加教保專業成長之理想時間

由統計高雄市公立幼兒園教保服務人員理想之參加教保專業成長主要時間結果發現，教師認為最理想的時間為依序為暑假、學期中週末、寒假。究其原因，是因為幼兒園教保服務人員利用以上三項非上班時間參加專業成長，一來可避免影響到正常教學工作，二來若在上班日的夜晚進修，又覺得疲累無法休息，同時在較長的假日中，有兩個整天以上的時間亦可參與時程較長，且深入主題的研習。

教保服務人員參加教保專業成長的實際與理想時間方面略有出入，雖然教師認為最理想的專業成長時間是暑假，然則實際參與時間卻是以學期中周末為主。究其原因，可能是教保專業研習時數計算的週期為1月到12月，到了暑假已是年度的中後期，有所準備的教保服務人員會在前半年就開始利用周末時段進修，累積研習時數，以免心裡牽掛專業成長時數事宜。

5. 教保專業成長之活動形式

由問卷調查結果發現，高雄市公立幼兒園教保服務人員參與教保專業成長活動之形式，其排序前三名依次為專題演講居多、學術研討會/座談會、專業討論/交換工作經驗。究其原因，大多數的研習，都是以專題演講的方式呈現，邀請在教保領域學有專精的專家學者或是在幼教職場有傑出表現的實務工作者，針對教保重要議題進行講演，與會者除了聆聽亦可提問互動；學術研討會/座談會多辦是由大專院校主辦，幼兒園教保服務人員若有至大專院校在職進修者，都會參與教育議題之研討會，聆聽論文發表；專業討論/交換工作經驗之形式常見於同事間的分享、工作社群的對話，甚至是網路專業社群的相互請益，更不受時空的限制，是新興的成長方式，排名第三；第四名是個人閱讀教育書刊，目前有多種幼兒教育月刊、雜誌、教育書籍發行，探討教育走向與趨勢，是老師們成長的好資源；第五名則是相當受老師們喜愛的校外教學觀摩，前往參觀優質園所的軟硬體設施，觀察教學實況與成果，從中反思與學習，讓老師們受益良多，反觀僅占 3.9% 的校內教學觀摩，排名第十一名，同樣是教學觀摩，由於主客易位的緣故，成為老師們較為抗拒的成長形式。另外值得關注的是教育當局力推的「幼兒園參與輔導計畫」僅佔 4.3%，排名第九，各幼兒園參與並不踴躍，因為教保服務人員擔心教授入園指導可能要求園方改變舊有教學模式，或是增加會議紀錄、教學省思與學生觀察

紀錄等負擔，甚為可惜。此研究結果與吳靜惠（2005）、廖金貴（2014）相同，即專題演講為高雄市公立幼兒園教保服務人員參與教保專業成長活動之最常出現之形式。

6.一學期參加教保專業成長活動之時數

根據問卷調查結果，高雄市公立幼兒園教保服務人員一學期參加教保專業成長活動時數累計以 18~36 小時居多，佔 70%；累計為 37 小時以上者，佔 15%；累計為 17 小時（含）以下者，佔 13.5%。由此顯示有八成五的教師利用一學期的時間就能達成幼照法第 15 條第 5 項規定，教保服務人員每年至少參加教保專業知能研習 18 小時以上，顯示公立幼兒園教保服務人員積極參與教保專業成長，促進教保專業能力的保持與提升，但多以滿足法令規定的最低研習時數為主。

7.教保專業成長的困境

本研究的調查結果發現，高雄市公立幼兒園教保服務人員參與教保專業成長的困境，其排序前三名依次為參加名額有限、進修地點距離太遠、研習排在上班時間，其他開放選項則有多位教師反應教師研習辦在周末，假日無法好好休息，感到疲憊。究其原因，大多數的研習皆因場地大小、印製講義份數、統計學員便當、參訪車輛座位等因素，而限定錄取人數，因此越是熱門的研習，越容易因名額有限而失之交臂；而高雄縣市合併之後，教育局在各區皆會舉辦研習，有的進修地點離家太遠，則降低老師參與的意

願；另有研習為工作坊型態須連續三日，會佔用到上班日，或是研習時間排在周末，假日生活去參與成長課程感覺太疲累，成為老師參與教保專業成長的困境。多項研究結果顯示幼兒園教保服務人員參與專業成長的困境為時間無法配合（例如范熾文、張瑜，2008；陳美齡，2004；陳惠珍，2003；廖金貴，2014），此研究發現與本研究結果不同。究其原因，教育部整合教師專業成長資源訊息之全國教師在職進修網於 2005-2006 年開始推廣，提早將整學期的研習資訊提早公佈於該網站，以利教師規劃個人研習時間，再者幼照法自 2012 年開始施行，明確定義教保服務人員應在一年參與 18 小時之專業研習，並在近年執行檢視幼兒園教保服務人員是否達成規定之研習時數，且幼托整合之後新興議題如雨後春筍般冒出，教保服務人員確實感受到專業成長的急迫性，但苦於參加名額有限、進修地點距離太遠、研習排在上班時間等因素，影響其專業成長。

（二）在幼兒園教保服務人員專業成長相關課程之需求

本研究進一步將教保專業成長課程各層面之題項依平均數高低加以排序，歸納幼兒園教保服務人員教保專業需求排名前 10 項，總計有 7 項落在兒童健康與照護層面，3 項落在教保課程與幼兒學習層面，顯示教師最重視與最需要的專業成長課程，是關於兒童安全與健康照護，包含基本救命術訓練、交通安全、飲食安全、遊戲安全、環境衛生安全、照護安全、防災教育及緊急救護情境演習、

遊戲活動設計與規劃、疾病與意外的預防與處理，幼兒發展與輔導策略，幼兒運動和健康、自理能力的培養等；第二需求度高的是教保課程與幼兒學習，包括教保服務人員如何帶領幼兒深入閱讀、學習區的規劃設計、觀摩優質幼兒園有著較高的需求。

幼兒園教保活動課程暫行大綱之總目標揭示，幼兒教育是所有教育階段的基礎，教保服務的實施須達成維護幼兒身心健康、養成幼兒良好習慣、豐富幼兒生活經驗...等目標，可見首要目標是幼兒的身心健康，如此一來，基本救命術訓練（心肺復甦術、基本創傷救命術、哈姆立克急救法等）就顯得格外重要，依據幼照法第 32 條規定，幼兒園新進用之教保服務人員，應於任職前最近一年內接受基本救命術訓練八小時以上；任職後每二年應接受基本救命術訓練八小時以上、安全教育相關課程三小時以上及緊急救護情境演習一次以上。加上幼兒園基礎評鑑項目包含環境設備維護、衛生保健、緊急事件處理、交通安全，皆與安全教育有密切相關，故幼兒園教保服務人員對於攸關幼兒身心健康之兒童健康與照護層面相關課程極為重視，其中以基本救命術訓練和安全教育之需求最高。

另一方面，教保課程與幼兒學習層面也是教保服務人員極為關注的，主要是幼托整合之後，新課綱的實施，對於教保服務人員的課程與教學是一個大的變革，在教學與學習方法上均需重新調整與學習，尤其是如何帶領幼兒

深入閱讀、學習區的規劃設計、觀摩優質幼兒園等教保服務人員更有迫切需求。

本研究結果與曾曉苓(2005)、范熾文和張瑜(2008)、廖金貴(2014)之研究結果不同，其研究結果分別為親師合作、班級經營、幼兒藝術教育為教師專業成長需求最高之課程。究其原因，主要與本研究對象之專業背景與訓練有關，在本研究中教保員有 167 人，幼教師有 438 人，主任、園長有 76 人，由於公立幼兒園的主任與園長大多具有幼教師身分，因此就本研究分析的對象而言，其中教保員約佔 20%，幼教師約佔 80%，而傳統幼教學程中並沒有健康照護類的課程，故在幼托整合制度實施之後，公立幼兒園幼教師會特別覺得需要強化健康照護類的研習，故成為高雄市公立幼兒園教保服務人員在教保專業成長課程之需求最高的項目。

(三) 不同背景變項之幼兒園教保服務人員在專業成長需求上部分有顯著差異

1. 年齡

本研究結果顯示，在情緒管理與人際關係層面，年齡在 51 歲以上的教師，其得分顯著高於 30 歲以下、31-40 歲及 41-50 歲之教師。本研究與曾曉苓(2005)、林亞萱(2011)發現不同年齡的教保人員對在職進修內容之需求沒有差異有所不同。究其原因，可能 51 歲以上之教師雖為全園最資深教師群，累積了長期的工作實務經驗，但因社會發展與變遷快速、少子化等因素，其所面臨的社會工作壓力亦日漸增大，因此對

於「認識自我與他人的感覺與情緒」、「說話的技巧及表達情緒的方法」、及「時間管理及情緒管理」等有關之情緒管理與人際關係的專業成長有高度需求。

2.婚姻

本研究發現不同婚姻狀況的幼兒園教師，在教保專業成長需求各層面上均無顯著差異。此與陳彥奇（2011）、林亞萱（2011）、廖金貴（2014）的研究結果相同。但與陳美齡（2004）、吳靜惠（2005）不同，究其原因，可能是2012年起幼托進行整合，新課程大綱正值推展之際，不同婚姻狀況之教師皆對各項教保專業，有一致的成長需求。

3.教學年資

本研究結果顯示不同教學年資的幼兒園教師僅在教保課程與幼兒學習層面的需求是有差異的，但兩兩組別之間並未達顯著差異。此一結果和曾曉苓（2005）、廖金貴（2014）的研究結果不同，她們的研究指出不同年資之教保人員對在職進修內容之需求沒有差異。究其原因，不同教學年資之教師，雖然每個人經歷過不同年代的幼兒教育，累積了相異的教學經驗，在教保課程與幼兒學習層面之專業成長需求，基礎上是有差異的，但當大家同時面對新的政策時，不同年資的教師都產生了程度不一的學習需求與動力，此雖有差異存在，卻不明顯。

4. 擔任職務

本研究結果發現擔任不同職務的幼兒園教師在園長（含負責人、校長及園主任）領導層面和問題本位導向層面達到顯著差異。擔任職務為專任行政（主任、園長）的教師在園長（含負責人、校長及園主任）領導層面其需求顯著高於班級教師（含代理、代課教師）及教保員；而在問題本位導向層面，擔任班級教師兼任行政（組長、主任、園長）者其需求顯著高於擔任教保員之教師。此一結果與陳彥奇（2011）、廖金貴（2014）相同；而與曾曉苓（2005）相異，其研究結果為不同職務之教保人員在專業成長內容之需求沒有顯著差異。探究其可能原因為近年國教署規劃之教保專業知能研習主題，特將園長（含負責人、校長及園主任）領導獨立為一層面，正符應擔任學校領導者之需求，而教保員為以契約進用之人員，有較高的流動率，故對學校之問題本位之需求較低。除此之外，主任與園長對於最新教育政策與改革居於教學現場的第一線，為能發揮領頭羊的功能，必須先行了解教保專業成長的各層面，故其需求會最高。

5. 學校地理位置

本研究結果顯示任職於不同地理位置的幼兒園教師在教保專業成長需求各層面上均無顯著差異。此與陳美齡（2004）研究屏東縣、楊淑娟（2006）研究高雄市和高雄縣之結果相同，表示教師無論身處一般地區、偏遠地區或特偏地區，在教保專業成長之各層面的需求並無不同。探其原因，由於教

師法已明文規定進修為教師之權利與義務，教師皆具備參與進修，促進自我專業成長之工作態度，因此，服務於一般地區、偏遠地區、特偏地區之幼兒園教師，在教保專業成長需求各層面上均無顯著差異。

6. 幼兒園規模

本研究結果發現，任職於不同幼兒園，規模 5 班以上的教師在政策與法令和、情緒管理與人際關係層面之專業成長需求高 1-2 班及 3-4 班。本研究與曾曉苓在 2005 年的研究結果相似，幼兒園規模較大之幼兒園教師在當前幼教相關政策上之需求高於規模較小之幼兒園教師。但與陳雪芳（2008）、楊淑娟（2006）之研究結果不同，她們的研究結果為不同幼兒園規模之教師在專業成長需求上均無顯著差異。究其原因，2012 年幼照法實施，規定公立學校附設幼兒園置主任，其達一定規模者，應為專任；幼兒園達一定規模，得分組辦事，5 班以上大型幼兒園，班級數較多，園務行政亦更為複雜，於是對政策與法令有較高之需求；另外，大型幼兒園教師人數達 10 人以上，加上學生人數達 150 人以上，再加入全園家長人數，彷彿一個大家庭，人際相處之間更需彼此尊重與協調，以利校務推動，故在情緒管理與人際關係層面上有較高的需求。

伍、結論與建議

一、結論

有鑑於國內教師專業成長之相關研究，文獻資料雖然豐富，但研究對象多以國民中小學為主，鮮少涉及幼兒園（許屏霞、翁銘鴻，2009；引自劉乙儀、張瑞村、范雨潔，2015），故本研究透過問卷調查法針對高雄市公立幼兒園教保服務人員之教保專業成長現況與需求，同時進一步分析不同個人背景因素的教保服務人員在教保專業成長之需求之差異情形，相信所提供之研究結果不管是在學術面或實務面應能對幼兒園教保專業成長此一領域有所貢獻。

本節依據問卷調查結果、綜合討論及主要研究發現，歸納出關於高雄市公立幼兒園教保服務人員教保專業成長之現況包含：教保專業成長的類型以校內外研習為最多、且以「專題演講」之形式居多，而線上研習次之，仍以一般專業成長類型為大宗。

就目前高雄市公立幼兒園教保服務人員所參與的專業成長內容多偏重在「教保課程與幼兒學習」，然而受訪人員需求最殷的是兒童健康與照護層面之課程，其次依需求高低排序為教保課程與幼兒學習、情緒管理與人際關係、問題本位導向、政策與法令及園長（含負責人、校長及園主任）領導之需求，各層面之需求均在中上程度，顯見未來針對高雄市公立幼兒園教保服務人員所提供之專業成長內容宜更多元化以滿足其實際之需求。

高雄市公立幼兒園所安排參與教保專業成長以自願參加者居多、教保專業成長之實際時間以學期中週末為最多；而暑假是最理想之專業成長時間。教保專業成長一學期累計時數達 18~36 小時

者佔多數、最大困境為參加名額有限，是未來企待解決的重要問題之一。

而就不同個人背景變項之高雄市公立幼兒園教保服務人員對教保專業成長的需求而言，年齡越長者、教學年資在 6-10 年者、擔任職務為專任行政（主任、園長）者、學校位於偏遠地區者、以及 5 班以上大型幼兒園者，其在教保課程與幼兒學習、政策與法令、園長（含負責人、校長及園主任）領導、兒童健康與照護、情緒管理與人際關係、問題本位導向等六大層面之需求最高。

二、建議

（一）增加熱門教保專業成長研習課程之場次或名額，亦可思考採用線上研習之方式

本研究發現高雄市公立幼兒園教保服務人員教保專業成長之最大困境為參加名額有限，教師在本身的時間與交通各項因素都允許的條件之下，竟受限於名額有限無法參與，甚是可惜。越受歡迎的研習課程，一方面顯示出教師教保專業成長之真實需求，一方面也令人感受到老師們強烈的學習動機，候補者眾多，人人都期待能錄取參與研習，若能增加場次或是增額錄取，更能促成教師樂在專業成長。因此建議高雄市教育局及幼兒園針對熱門研習課程可增加研習場次或提高錄取人數，如此一來，既可以讓眾多有心學習的教師參與，同時也讓不遠千里而來的講師在同樣的時間可以產生更大的影響力，是教育當局可以參考的

方法。此外，熱門研習課程，亦可採用線上研習之方式加以進行，以突破任數限制，亦可樽節研習之經費。

(二) 規劃階段性質之研習課程

本研究發現不同背景變項的高雄市公立幼兒園教保服務人員在教保專業成長需求有所不同，縱然教保專業之重點項目是不變的，但學科總有深淺之別，同一門課程中，除了要有廣度，期望發展出深度、連續性與差異化，更因考量不同需求之幼兒園教保服務人員依個人學習興趣及學進度來選擇或參與專業成長之研習活動，如同現在進行的課綱研習，有分為 1-2、1-2、1-3 等。因此建議高雄市教育局及幼兒園宜編排階段性質之研習課程，教師可依個人學習的速度與興趣，循序漸進，厚實各項專業能力。

(三) 增加「兒童健康與照護」與「教保課程與幼兒學習」兩層面研習並多採實務形式進行

本研究發現高雄市公立幼兒園教保服務人員反應對「兒童健康與照護」與「教保課程與幼兒學習」此兩方面的教保專業成長課程需求度最高，同時研習形式多以專題演講及研討會形式呈現，擔任講師者大多為學者，且在教保專業成長困境之其他項目中有多位教保服務人員提及課程大多為理論性質，希望多邀請現場實務工作者擔任講師，進行教學演示或技巧的示範，最能引起教保服務人員的共鳴及馬上運用到教學現場，因此建議高雄市教育局及幼兒園宜增加「兒童健康與照護」與「教保課程與幼兒學習」兩層面之相關研習並以實務形式呈現。

(四) 善用教育資源，主動參加課綱輔導

由調查結果發現，教育部推廣之課綱輔導在本調查結果僅佔 4.3%，在 15 項專業成長形式中排行第 9 名。對於教育部推廣之課綱輔導，園所可能擔心輔導教授會要求改變既有的教學模式，或是會有衍生而來的觀察、省思紀錄、經費及成果報告等負擔，再三考量之下仍然不參加。其實園所教保服務人員若想要精進成長，但又覺得周末研習相當疲累的話，建議高雄市公立幼兒園教保服務人員可以參加課綱輔導，這是一個貼近園所實務需求的成長方式，較能針對各園所的特色發展實施差異化的協助，由學有專精的教授入園輔導，以入班觀察、小組或團體分享、教學示範演練、實施讀書會等多元方式進行專業成長，可以省去教師們交通往返奔波之苦，亦可在上班日完成專業精進，並由教育部負責相關經費，是難得的教保專業成長資源，建議幼兒園可以多加利用。

(五) 對未來研究的建議

針對研究對象而言，本研究的對象主要以高雄市公立幼兒園教保服務人員，建議未來的研究對象可增加教育行政人員及課程講師，以充分掌握彼此對教保專業成長課程的期許目標與實際執行之後的成效比對，了解理想與實際的差距，降低差距便是提高效果。此外，亦可進一步針對不同縣市間的實施現況與差異進行跨縣市或全國性的研究，由於不同縣市的教育主管不同，在規劃教保研習時也有不同的考量，未來可以此為方向，進行不同縣市幼兒園

教保服務人員教保專業成長之需求分析或做差異性比較，將有助於更全面性的了解幼兒園教保服務人員在教保專業成長的需求與實際推動情形。

就研究工具而言，本研究工具針對高雄市公立幼兒園教保服務人員的教保專業成長現況與需求兩部分係蒐集相關之學術期刊、論文及縣市政府教保研習課程表彙整其內涵而得，有其相當之理論基礎。然而，教保專業成長的內涵可能會因政策修正、社會需求或最新教育發展趨勢而有所異動，因此建議未來相關之研究工具可因應需求作修正，以符應時代所需。

參考文獻

- 吳靜惠（2004）。臺中市私立幼稚園教師自我專業成長需求之研究（未出版之碩士論文）。靜宜大學，臺中市。
- 林亞萱（2011）。幼兒教師參與學士後在職進修與教師專業發展之調查研究（未出版之碩士論文）。國立臺中教育大學，臺中市。
- 林春妙、楊淑朱（2005）。幼兒教師專業知能之研究。《兒童與教育研究》，1，55-84。
- 林桂垣（2010）。臺北市國中小學教師專業成長內涵與需求之研究（未出版的碩士論文）。國立臺北教育大學，臺北市。
- 林惠娟（2005）。南投縣幼兒教保人員在職進修現況之研究。《朝陽學報》，10，229-247。
- 花蓮縣政府教育處（2015）。花蓮縣 104 年度教保研習彙整表。取自：

- http://210.240.39.100/repair_log.asp?rid=A201503181532514958
- 范熾文、張瑜（2008）。幼稚園教師專業成長需求與相關問題之研究。《學校行政雙月刊》，58，94-118。
- 翁素雅（2004）。《幼稚園實習輔導教師專業成長之研究》（未出版之碩士論文）。國立屏東師範學院，屏東縣。
- 高雄市政府教育局（2015）。《高雄市 104 年度教保研習彙整表》。取自：<http://www.kh.edu.tw/forms/getDirectory/395>
- 高雄市政府教育局（2015）。《高雄市立案幼兒園名冊》。取自：<http://www.kh.edu.tw/forms/getDirectory/101>
- 教育部（1994）。《幼稚園教師手冊》。臺北市：作者。
- 教育部（2013）。《幼兒園教保專業知能研習實施辦法》。取自：<http://edu.law.moe.gov.tw/SearchAllResultList.aspx?Keyword=%E5%B9%BC%E5%85%92%E5%9C%92%E6%95%99%E4%BF%9D%E5%B0%88%E6%A5%AD%E7%9F%A5%E8%83%BD%E7%A0%94%E7%BF%92%E5%AF%A6%E6%96%BD%E8%BE%A6%E6%B3%95&Cur=L>
- 教育部（2015）。《幼兒教育及照顧法》。取自：<http://edu.law.moe.gov.tw/LawContentDetails.aspx?id=GL000542&KeywordHL=>
- 教育部國民及學前教育署（2014年10月27日）。《臺教國署國字第 1030119578 號函》。臺北市：作者。
- 張淑雲（2002）。《國小教師面對九年一貫課程工作壓力、因應策略與專業成長需求之研究》（未出版的碩士論文）。國立臺南師範學院，臺南市。

- 張媛甯、岳美秀（2012）。臺南市公立幼兒園教師覺知專業成長與教學效能之研究。《學校行政雙月刊》，82，70-89。
- 陳彥奇（2015）。《幼稚園教師專業發展與專業能力之調查研究》（未出版的碩士論文）。國立屏東教育大學，屏東縣。
- 陳美齡（2004）。《屏東縣公私立幼兒教師在職進修現況與其專業成長之研究》（未出版之碩士論文）。國立屏東師範學院，屏東縣。
- 陳淑琴（2000）。《幼托師資合流教育與我國主要幼教師資培育課程知識基礎分析研究》。載於國立臺北師範學院（主編），八十八學年度師範學院教育學術論文發表會論文集（頁371-401）。臺北市：國立臺北師範學院。
- 陳雪芳（2008）。《幼稚園教師專業成長與教學效能之研究》（未出版之碩士論文）。朝陽科技大學，臺中市。
- 陳惠珍（2003）。《嘉義市幼稚園教師透過網路在職進修之需求與態度之研究》（未出版之碩士論文）。國立嘉義大學，嘉義市。
- 曾曉苓（2005）。《教保人員在職進修現況之研究——以苗栗地區為例》（未出版之碩士論文）。朝陽科技大學，臺中市。
- 黃淑苓（2005）。《教師專業發展的應然與實然》。《教育科學期刊》，5(2)，155-168。
- 新北市幼教研習暨資訊中心（2015）。《**新北市 104 年度教保研習彙整表**》。取自
<http://tpcpsis.kidedu.ntpc.edu.tw/files/11-1003-166.php>
- 楊淑娟（2006）。《幼稚園教師在職進修與教學效能關係之研究——以高雄地區公立幼稚園為例》（未出版的碩士論文）。樹德科技大學，高雄市。

- 葉郁菁 (1999)。從幼教師資的專業自省檢討我國幼稚園教師所需之教育專業課程。載於中華民國師範教育學會 (主編), **教育改革、師資培育與教學科技：各國經驗國際學術研討會論文集** (頁 271-283)。臺北市：臺灣書店。
- 廖金貴 (2014)。澎湖縣幼兒園教師專業成長現況與需求之調查研究 (未出版的碩士論文)。國立臺南大學，澎湖縣。
- 劉乙儀、張瑞村、范雨潔 (2015)。臺灣幼兒園教師專業發展之困境與展望。**學校行政雙月刊**，96，196-215。
- 臺中市政府教育局 (2015)。臺中市 104 年度教保研習彙整表。取自 <http://www.tc.edu.tw/m/402>
- 臺北市教師在職研習網 (2015)。臺北市 104 年度教保研習彙整表。取自：<http://insec.tp.edu.tw/index/ClsShowLstB.aspx>
- 臺南市教保輔導團 (2015)。臺南市 104 年度教保研習彙整表。取自：<https://sites.google.com/site/tnecetainan/you-ding-chang-ci>
- 澎湖縣教保資源中心 (2015)。澎湖縣 104 年度教保研習彙整表。取自：http://kid.magong.com.tw/news_data.php?id=261
- 賴春金 (2003)。幼稚園教師的關鍵能力。**教育研究**，52，75-83。
- Bransford, J. D., Brown, A. L., & Cocking, R. (2000). *How people learn: Brain, mind experience and school*. Washington, DC: National Academy Press.
- Clement, M., & Vandenberghe, R. (2000). Teachers' professional development: A solitary or collegial adventure? *Teaching and Teacher Education*, 16, 81-101.

- He, P., & Ho, D. (2013). Rethinking the professional development policy for early childhood education in Hong Kong. *Australasian Journal of Early Childhood*, 38(2), 95-102.
- Jenlink, P. M., & Kinnucan-Welsch, K. (1999). Learning ways of caring, learning ways of knowing through communities of professional development. *Journal for a Just and Caring Education*, 5(4), 367-385.
- Lavonen, J., Jauhiainen, J. Koponen, T. I., & Kurki-Suonio, K. (2004). Effect of a long-term in-service training program on teachers' beliefs about the role of experiments in physics education. *International Journal of Science Education*, 26(3), 309-328.
- OECD. (2005). *Teacher matter. Attracting, developing and retaining effective teachers*. Paris: OECD.
- Stone-MacDonald, A., & Douglass, A. (2015). Introducing online training in an early childhood professional development system: Lessons learned in one state. *Early Childhood Education Journal*, 43(3), 241-248.

教育學誌 第四十期

2018年11月，頁59~120

家庭社經地位與學業成就關係之研究 －以學業發展自我效能為中介變數

邱仕凱

國立臺灣師範大學教育學系博士候選人

摘要

本研究探討不同家庭社經地位的國中生，是否透過學業自我效能對學業成就產生影響。採取調查研究法並以立意取樣方式選擇臺北市某完全中學國中部九年級男生53人，女生51人，共104人進行施測。統計方法以SPSS 20.0版進行描述性統計、皮爾森積差相關與點二系列相關、獨立樣本T檢定、多元迴歸與徑路分析。

研究結果發現：(1)家庭社經地位與學業發展自我效能，和學業成就達顯著正相關；(2)不同家庭社經地位對學業成就存在差異，其中第四等級對第二等級達顯著差異；(3)家庭社經地位會直接影響學業成就，也會透過學業發展自我效能間接影響學業成就，但直接效果遠大於間接效果。最後針對研究結果，提出相關的建議。

關鍵字：家庭社經地位、學業發展自我效能、學業成就、
中介效果

A Study of the Relationship between Family Socio-Economic Status (SES) and Academic Achievement : Using Academic Development Self-Efficacy as Mediator Variable

Shih-Kai Chiu

Doctoral Candidate,

National Taiwan Normal University, Education Department

Abstract

This study examined whether academic self-efficacy has influence over the academic achievements of middle school students of different family social statuses. The research methodology adopted in this study was the survey research method, while the research participants were chosen based on purposive sampling. 53 boys and 51 girls (a total of 104 students) were drawn from Grade 9 of the junior high school section of a six-year high school in Taipei City for the study. The statistical methods used in this study were descriptive statistics using SPSS 20.0, Pearson correlation and Point-Biserial correlation coefficient, Independent Sample T-Test, multiple regression, and path analysis.

The research results revealed that: 1. there is a significant positive correlation between family socioeconomic status and academic development self-efficacy and educational achievement; 2. there is a

difference between the different family socioeconomic status and academic achievement, more specifically there is a significant difference in the fourth level and the second level; 3. family socioeconomic status directly influences the academic achievement, as well as indirectly influence their academic achievement through academic development self-efficacy, but the direct effect much greater than the indirect effect. Finally, the study will provide relevant recommendations based on its research findings.

Keywords: family socioeconomic status, academic development self-efficacy, academic achievement, mediator effect

壹、研究動機與研究目的

一、研究動機

過去關於影響學業成就的研究，以 Coleman (1966) 的「教育機會均等報告書」中提到家庭社經背景對學業成就的影響高於學校因素，最為人所熟知，因此家庭社經地位成為研究學業成就的重要變項。眾多研究亦證實家庭社經地位對學業成就的重要性（吳慧瑛，2007；林俊瑩、黃毅志，2008；張芳全，2006；張芳全，2017；陳順利、黃毅志，2015；曾明基，2016；詹秀雯、張芳全，2014；蔡政中，2015；謝孟穎，2003；龔心怡、林素卿、張馨文，2009；Coleman, 1966; Howie & Tjeerd, 2003; Kiamanesh, 2005; Sewell & Hauser, 1976），但有部分研究呈現出無顯著的影響（吳素媛，2004；李文益，2004；李文益、黃毅志，2004；郭丁熒、許竣維，2004；黃瓊儀、吳怡慧、游錦雲，2015；劉劍華、謝百亮，2015；謝亞恆，2004；簡美玲，2006）。後續以後設分析進行的研究則顯示社經地位與學業成就，其實只有低至中度關係（孫旻儀、蔡明學，2007；許崇憲，2002；鄒浮安，1994；Sirin, 2005; White, 1982）。顯示家庭社經地位對學業成就之影響，仍未有一致的共識。因此家庭社經背景與學業成就之間尚存有未明的中介因素與作用機制，有必要持續探索，為本研究的第一個研究動機。

古典社會學家馬克思（Marx）從經濟學的觀點出發，對於階級（class）的認同採取客觀認定，以個人在社會經濟結構中的位置，及其展開的社會關係，把社會分成三種階級，分別是「資本家」、「勞工」與「地主」。將擁有生產工具的人視為是「資本家」，接受資本

家雇用的人為「勞工」。馬克思的階級概念僅考慮經濟因素，被視為是經濟決定論，以階級鬥爭為最終結果，忽略社會變遷與社會流動的可能性。而 *Boureidu* 則延續馬克思的階級概念，提出不同階級會形塑不同的習性、品味與生活形態，稱為「慣習 (*habitus*)」。

不同階級家庭在休閒方式上存在差異，富人的休閒以炫耀性消費為主，達到凸顯自身財富與地位的功能（李華夏，2001），勞工階級則以上酒館及玩賓果遊戲等低消費的休閒活動為主。顯示家庭財富的多寡，不但可用來界定社會階層，也能預測子女的社會流動（陸洛，1997）。*Bowles* 和 *Gintis*（1976）則指出學校體系的結構與資本主義社會的結構符應，中上階層學生多學習抽象知識，勞工階層則以實用技能為主，使得學校教育成為資本主義進行社會再製的重要工具（李錦旭，1989）。故不同階層家庭因為經濟資本的落差，使子女在學校課業表現上不盡相同（周新富，2008）；*Bourdieu* 與 *Passeron*（1977）則認為階級再製主要透過文化層面進行，提出「文化資本 (*cultural capital*)」的概念，以形體化、客觀化與制度化的形式，進行文化的專斷，認為統治階級將對自己有利的意識形態納入學校教育中，使學校成為優勢階級文化的主要場域，勞工階級的學生因為文化差異導致教育失敗，成為文化再製的受害者（鄭英傑，2017）。而文化資本高的學生在入學甄試時因為其特質與教師相近，更容易吸引教授的目光，其入學機會也較大（張訓譯，2017）。少部分文獻未能證實文化資本與教育成就的關係（鄭耀男、陳怡靖，2000；*Katsillis & Rubinson*, 1990）。因此不同階層家庭，在個人秉性、文化財及相關證照與學歷的多寡，存在文化資本上的差異。

另外有研究發現不同階層中，存在相異的友誼來源（陸洛，1997），而教育程度高者則擁有更廣泛而深入的社會網絡（李宗義、許雅淑，2016），因此不同階層家庭在社會連結與社會關係的產生上，出現資源取得的不均（林祐聖、葉欣怡，2005），導致社會資本的階層落差，使子女在學校學習中居於劣勢。

另一方面，Willis（1977）在研究勞工階級子女為何繼續選擇當勞工的問題時，發現小伙子所處的家庭及社區文化存在陽剛與父權思想，與學校中呈現的中產階級文化不一致（黃庭康，2017），因此在結構限制下，勞工階級的學生發展出反學校文化，試圖抗拒中產階級形塑的社會結構，產生對抗權威、學校體制、教室管控、鄙視乖學生等自主性行為，即個體的「能動性（agency）」。

但是過去在研究社會階層時，往往忽略了心理因素的重要性（陸洛，1997）。從心理學的層面來看，個體能動性在心理上的展現，與 Bandura 所謂的「自我效能感」，是個體、行為與環境三者交互作用下的產物，其實有異曲同工之處。而不同階層的家庭，在經濟、文化與社會資本的差異下，導致子女在學校中面對結構上的限制，展演出的個體能動性，轉化成對學業成就表現的「自我效能感」。

「自我效能感」是由 Bandura 提出，後續更發展出群體自我效能、學業自我效能、各學科自我效能等概念，並建構出相關的測量表。而過去關於自我效能的研究也顯示會對學業成就產生影響（王振宏，1999；王凱榮、辛濤、李瓊，1999；李蒼、辛濤、谷生華、申繼亮，1998；徐靜，2013；Bandura, 1977, 1986; Catalina, Stanescu, & Mohorea, 2012; Galyon, Blondin, Yaw, Balls, & Williams, 2012; Weiser & Riggio, 2010; Young, Hyuck, Sunyoung, & You, 2012）。因此

若不同階層子女其學業自我效能不同，將可能對其學業成就產生差異，所以自我效能可視為是家庭社經地位與學業成就間的中介因素，值得深入探究其功能性與影響機制，為本研究的第二個研究動機。

二、研究目的

本研究欲探討家庭社經地位影響學業成就的中介機制，瞭解家庭社經地位是否會透過學業發展自我效能，間接影響學校的學業成就。

三、研究問題

依據上述的研究動機與目的，研究的問題如下：

- 問題一：不同家庭社經地位的學生，學業成就是否具有差異？
- 問題二：不同家庭社經地位的學生，學業發展自我效能是否具有差異？
- 問題三：不同學業發展自我效能的學生，學業成就是否具有差異？
- 問題四：家庭社經地位是否透過學業發展自我效能，對學業成就產生影響？

四、研究假設

依據上述的研究問題，研究的假設如下：

- 假設一：家庭社經地位對學業成就會產生影響
- 假設二：家庭社經地位對學業發展自我效能會產生影響
- 假設三：學業發展自我效能對學業成就會產生影響
- 假設四：家庭社經地位會透過學業發展自我效能影響學業成就

貳、文獻探討

一、社會經濟地位和自我效能的關連性

不同階層的家庭，因為所擁有資本在質量上的不同，因此對子女產生不同程度的影響。其中 Lareau (2003) 的研究發現中產階級父母對子女的學習採取「精心栽培 (concerted cultivation)」的策略，會在特定的時間點參與子女在校的學習，勞工階級父母多是以「自然成長 (natural growth)」方式對待子女的成長 (李怡慧，2015；張旭，2010)。國內學者駱明慶 (2001、2002、2004) 的研究則發現父母親的教育程度為大專畢業、職業為軍公教 (特別是老師)、居住地為臺北市 (特別是大安區)，其子女升學及就讀台大的機會皆顯著大於其他各類別與地區，即便駱明慶 (2016) 的後續研究雖發現多元入學政策能降低家庭背景的影響力，但過去十五年來進入台大就讀的子女仍舊以臺北市大安區的機率最高，且比例最大。由此可知，處在不同社會階層的家庭子女，其成長的過程中因為資源與資本的階層性差異，在心理與生涯歷程上便已產生自信心與效能感上的相對剝奪感，被迫成為機會不平等的階級世代 (李宗義、許雅淑，2016)。

其中關於家庭社會經濟地位與自我效能的相關研究，主要從不同家庭社經地位對子女教育期望或教育抱負為主。所謂的「教育抱負」指個人在追求成就時，主觀評估自己能達到目標的水準，進而影響未來教育年數的取得 (黃毅志、楊雅鈞，2007)。研究發現不同家庭社經地位的子女，其教育期望存在差異，家庭社經地位愈高的子女，教育期望也愈高 (何秋蓮，2007；許哲耀、鄧家駒、陳郁

婷，2014；羅淑苑、黃毅志，2016）。其他的研究亦顯示家庭社經地位愈高，子女的教育抱負也愈高（李文益、黃毅志，2004；張芳全，2011；張芳華，2015）。也有文獻提到家庭社經地位對自我效能具有正向影響（黃建皓，2011；Coleman & Karraker, 1997; Gecas & Schwalbe, 1983; Lin, 2003）。

特別是 Willis（1977）針對勞工階級子女的研究，發現學校裡中產階級的文化與小伙子平時所處的家庭與社區文化不一致時，會形成一種反學校文化，反抗權威、鄙視乖學生，呈現出與勞動階級及工廠文化相似的陽剛特質與父權思想，從中洞察出學校與資本主義的對應，並發展出特定的生存方式，卻也因此受到限制，最終再製其勞工階級的結果（黃庭康，2017）。這種低社經地位家庭子女在社會結構限制下所展現出的個體能動性，在反抗學校教學典範時所具備的自我效能感，反而成為另一種面向的發展。

二、社會經濟地位和學業成就的關連性

過去關於社會經濟地位（Socio-Economic Status，簡稱「社經地位（SES）」）的定義，牽涉到概念型與操作型定義兩種，在研究實務上多採用操作型定義進行，將社經地位予以數量化，以「職業」、「教育程度」與「收入」三個項目做為指標。在 1947 年由國外學者 North 與 Hatt 在國家民意研究中心（National Opinion Research Center，簡稱 NORC）的協助下，針對全國樣本進行職業聲望的調查（李明，1990），後來 Hollingshed（1958）以職業與教育加權得出的社會地位指數，是過去最常用的社經地位測量方式，Duncan（1961）則設計社會經濟指標（Socio-Economic Index，簡稱 SEI），以教育和收入計算社會地位指數的因子（任春容，2010）；

Treiman (1977) 則根據 1968 年的國際標準職業分類，建立跨國比較研究可使用的「國際標準職業聲望量表」；Ganzeboom 與 Treiman (1996) 則根據 1988 年的國際新標準職業分類，建構一套國際「新職業聲望與社經地位量表」，此量表受到國內大型調查研究的使用（黃毅志，2009）。

但上述的量表皆由國外學者所研發，其分類方式不能完全適用臺灣社會，因此國內學者林生傳（2005）曾參照 Hollingshed 的職業與教育指數，分別將職業與教育分成五個等級，將父母的教育程度較高者乘以 4，加上職業指數較高者乘以 7，加總成為二因子社會地位指數。黃毅志（1998）則參酌 1992 年行政院主計處與 1988 年的國際新標準職業分類為基礎，設計「社會變遷調查新職業分類」量表，2003 年更進一步建構「臺灣地區新職業聲望與社經地位量表」，發展出更適合臺灣本土職業且實證研究證實具有良好建構效度的測量工具。

而過去的研究也顯示，社經地位與學業成就密切相關（吳慧瑛，2007；林俊瑩、黃毅志，2008；張芳全，2006；張芳全，2017；陳順利、黃毅志，2015；曾明基，2016；黃芳玫、吳齊殷，2010；詹秀雯、張芳全，2014；蔡政中，2015；謝孟穎，2003；龔心怡、林素卿、張馨文，2009；Coleman, 1966; Howie & Tjeerd, 2003; Kiamanesh, 2005; Sewell & Hauser, 1976)。但部分研究則是呈現無顯著的影響（吳素媛，2004；李文益，2004；李文益、黃毅志，2004；郭丁熒、許竣維，2004；黃瓊儀、吳怡慧、游錦雲，2015；劉劍華，謝百亮，2015；謝亞恆，2004；簡美玲，2006），因此影響的效果眾說紛紜。若以後設分析方式進行相關研究的統合，卻發現社經地位與學業成就，只有低度到中度關係（孫旻儀、蔡明學，2007；許

崇憲，2002；鄒浮安，1994；Sirin, 2005; White, 1982)。因此家庭社經地位雖然是影響學業成就的重要變項，但其中仍存在複雜而未明的機制，家庭社經地位透過哪些中介變數間接影響學業成就，而用來解釋影響學業成就的相關模型也持續在建構與發展中（周新富、王財印，2006；陳瑋婷，2011；張芳全，2006；黃瓊儀、吳怡慧、游錦雲，2015；詹秀雯、張芳全，2014；趙珮晴、余民寧，2012）。

三、自我效能的定義與類型和學業成就的關連性

「自我效能」一詞起源於 Bandura 的自我效能理論，該理論以社會認知理論做為架構，Bandura 在 1986 年時將「自我效能感（self-efficacy）」定義為「一個人判斷對完成目標的行為過程之組織與執行能力」，後來於 1989 年又重新描述為「人們對自己能掌握和控制可能影響生活的事件能力之信念」（蔡順良，2008），強調自我效能是「三元交互作用論」的產物，即個體、行為和環境的交互作用（李恒，2010）。Bandura（1977）更定義了「學業自我效能感」為「學生對自己形成能達到既定學習目標的行動過程之能力判斷（李波，2009）」。Bandura（1977）指出四種形成自我效能的訊息，包括成就經驗（enactive attainment）、替代性經驗（vicarious experience）、口頭說服（verbal persuasion）與生理狀態（physiological state）（周學霽、黃英哲，2007）。

至於自我效能與學業成就的關連性，Bandura（1977,1986）認為自我效能感會影響學習策略的使用，且在學習中表現出更努力、持續性夠強，更少的負面情緒，進而影響學生學業的表現（王振宏，1999；王凱榮、辛濤、李瓊，1999；李薈、辛濤、谷生華、申繼亮，1998；徐靜，2013；Bandura, 1977,1986; Catalina, Stanescu, & Mohorea,

2012; Galyon, Blondin, Yaw, Balls, & Williams, 2012; Weiser & Riggio, 2010; Young, Hyuck, Sunyoung, & You, 2012)。

其中最接近自我效能的心理結構是「自我概念」，即「個體將自己視為一種客觀物體所產生的知覺與評價」。Shavelson (1976) 首先把自我概念區分成「學業」和「非學業」自我概念，後續的研究則將自我概念的結構模型大致分成以下三種：第一種「多維層次性自我概念結構模型」是把自我概念分成「學業」與「非學業」自我概念兩種；第二種「Song-Hattie 模型」則是修正第一種模型，再細分出「班級」、「能力」與「成就」自我概念三部分；最後一種「Marsh/Shavelson 模型」是以第一種模型為基礎，將學業自我概念直接以「語言」與「數學」自我概念取代（沈靜、姚本先，2006）。

而學業自我概念的形成過程，目前的研究以（一）內外參照模型（I/E 模型）與（二）大魚小池效應（BFLPE）兩種為主。所謂的 I/E 模型，即學生會進行內部（與自己比）與外部（與同儕比）的兩種比較，確定自己的能力，以形成對應的自我概念；後者則是指出就讀高成就學校的學生其學業自我概念會下降（Marsh & Parker, 1984）。Marsh 將大魚小池效應納入內外參照模型進行檢驗，發現具有跨文化的適應性，也發現學業自我概念與學業成就存在顯著正相關，相關研究亦證實這項結果（宋曜廷、黃瓏瑩、郭念平、曾芬蘭，2012；侯雅齡，2010；簡晉龍、任宗浩、張淑婷，2008；Möller, Pohlmann, Köller, & Marsh, 2009）。

由此可知，自我概念與自我效能具有高度的相關性，皆與學業成就有關，不管是自我概念還是自我效能都是影響學業成就的重要因素，但兩者之間仍存在不少的差異（Bong & Skaalvik, 2003）。本研究採取 Bandura 的自我效能理論，以「學業發展自我效能」做為

研究變數，主要原因在於該理論內容的完整性與應用性較高，且研究顯示自我效能對學業成就的預測力高於自我概念（Breanne & Negel, 2013）。

四、中介效果的定義與檢測步驟

Baron 與 Kenny（1986）將中介變數定義為「解釋預測與效標變數間關係程度」的變數，需符合三種條件：1.自變數能顯著預測依變數；2.自變數能顯著預測中介變數；3.當自變數被控制時，中介變數能顯著預測依變數（張再名、曾迎新，2014）。

本研究的中介變數為「學業發展自我效能」。依照 Baron 與 Kenny（1986）對於中介變項的因果步驟分析，以多元迴歸及徑路分析的方式驗證中介效果（方杰、張敏強、邱皓政，2012；李德治，2016），分析程序分成三大步驟：

第一步驟：檢驗自變數與依變數的徑路是否存在？

以迴歸分析的方式檢驗徑路係數是否顯著。若顯著則表示該徑路存在，又稱為自變數對依變數的「總效果（total effect）」。

第二步驟：檢驗自變項與中介變數的徑路是否存在？

以迴歸分析檢驗徑路係數是否顯著。若顯著則表示該徑路存在，可以進入第三步驟；若未達顯著，則進行 Sobel 檢定，但樣本數需達兩百個以上，方能使用（溫福星，2012）。若 Sobel 檢定達顯著，表示存在間接效果（Sobel, 1982）。

第三步驟：檢驗自變數是否透過中介變數影響依變數？

以多元迴歸的方式檢驗徑路係數是否顯著。若自變數與依變數的徑路係數不顯著且接近零，則稱為「完全中介」或「直接效果 (direct effect)」，其中自變數對中介變數的效果，稱為「一階效果」，中介變數對依變數的效果，又稱為「二階效果」；若自變數與依變數的徑路係數顯著，且係數小於第一步驟的徑路係數，則稱為「部分中介」或「間接效果 (indirect effect)」。

參、研究方法

一、研究架構

本研究在探討家庭社經地位對學業成就的影響，並以學業發展自我效能做為中介變數。研究架構關係如下：

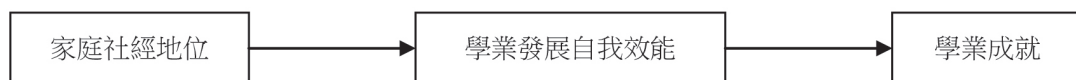


圖 1 研究架構關係圖

二、研究工具

本研究的測量工具是由香港學者袁文德、艾思柏、許錦屏、梁健文、劉兆瑛、陳茂釗、余文基與柯新豔（2004）出版的「學業發展自我效能量表 (Academic Development Self-Efficacy Inventory，簡稱 ADSEI)」，共有 20 個題目。曾以 6776 名香港高中生為施測對

象，利用驗證性因素分析（Confirmatory Factor Analyses，簡稱 CFA）顯示具有五個共同因子，並產生一個適配模型（NNFI = .91；GFI = .918；SRMR = .49；RMSEA = .70），總量表的內部一致性係數（Cronbach's α ）為.93，五個因子的內部一致性係數（Cronbach's α ）分別為「時間管理」因子的.80；「學習和應試技巧」的.80；「向朋友學習」的.84；「升學計畫」的.81 與「做負責任的學習者」的.77。

施測後總量表的建構效度利用因素分析中的主成分分析進行維度縮減，確定所產生的因子數量是否與原始量表相同。若固定為五個因子數，則五個因子可以解釋總變異量的 77.91%，若未限制因子個數，總共萃取出三個共同因子，解釋變異量為 68.64%。顯示若依照原始量表所產生的五個因子可以解釋的變異量較大，因此本研究將此量表用於國中生仍具有一定程度的效度。

至於施測後總量表的內部一致性係數（Cronbach's α ）為.95，各分量表的內部一致性係數（Cronbach's α ）分別為「時間管理」因子的.87；「學習和應試技巧」的.89；「向朋友學習」的.86；「升學計畫」的.83 與「做負責任的學習者」的.85，不論總量表或各分量表皆具有良好的信度。

本研究因為在同一時間點進行量表的施測，且由單一來源的受測者進行填答，在面對李克特式的題項時，易產生「共同方法變異（Common Method Variance，簡稱 CMV）」的問題，因此採取 Harman（1967）的「單因子測試法（one-factor test）」，所得資料採取未轉軸的因素分析，以主成分法進行萃取，發現並未產生一般性因子（general factor），因此未出現共同方法變異的現象（彭台光、高月慈、林鈺琴，2006）。

三、研究對象

以立意取樣的方式選擇臺北市某完全中學國中部九年級學生，施測時間為 105 學年度國中教育會考結束後至畢業離校之前（106 年 5 月 22 日至 6 月 12 日），扣除拒絕施測的班級，與施測當日請假、參與畢聯會及抽離至學習資源中心上課的學生，並刪除其中的無效問卷（填答不全、規律作答等），有效問卷數為 104 份，其中男生 53 人，女生 51 人。

四、研究變項

（一）家庭背景變項

包括「性別」、「家庭結構」、「雙親教育程度」、「雙親職業」與「每月收入」等五項（如附錄二表一）。性別部分男生編碼 1，女生 0；家庭結構部分分成單親與雙親兩種，分別編碼 1 與 2；教育程度方面則依接受教育的年數進行轉換，不識字 0、國小畢業為 6、國中畢業 9、高中或高職畢業 12、二專或五專畢業 14、大學、科技大學或二技畢業 16、碩士畢業 18、博士畢業編碼 22。

職業地位指數的計算，是將職業的分類以「臺灣高等教育資料庫之建置及相關議題之探討（簡稱「高教調查」）」大一新生問卷的分類方式，將職業分成 16 種。選擇「高教調查」的職業分類方式，原因在於「高教調查」與「教育部數位學習國家型計畫-中小學數位學習現況調查（簡稱「數位落差調查」）」的分類方式十分雷同，但「高教調查」問卷特別將專業人員區分為高層與一般專業人員、職業軍

人細分成軍官與士官兵兩種，相較於「數位落差調查」對於職業的分類更為詳細與清晰。依黃毅志（2005）的研究顯示兩種問卷皆由「臺灣地區新社經地位量表」簡化而來，且所轉換成的五等第社經地位測量，皆有良好的再測信度與建構效度。因此本研究使用「高教調查」問卷的職業分類方式，並採取黃毅志（2009）的建議，以「臺灣新職業社經地位量表（New Taiwan Occupational Socio-Economic Scale, 簡稱 NTSEI）」的分類，將職業分成五個等級，等級愈高職業地位愈高，進行職業地位指數的計算。

由蒐集到的資料發現，有半數以上的學生並不清楚父母親的每月收入，若將這些遺漏值逕自插補，將對內在效度造成嚴重的威脅，因此將父母親每月收入這個變項排除在家庭社經地位測量的指標之中。另外分別將父親與母親的教育年數加上各自職業地位指數，與模擬會考總積分進行相關分析，發現母親的教育年數與職業地位指數的加總與模擬會考總積分的相關係數達顯著（ $p < .01$ ），與黃毅志（2009）的研究結果相同。因此本研究所稱「家庭社經地位」是將母親的教育年數與職業地位指數的數值加總而來。

（二）學業發展自我效能

本研究的學業發展自我效能是採用香港學者袁文德等人（2004）所發展的「學業發展自我效能量表（Academic Development Self-Efficacy Inventory, 簡稱 ADSEI）」，採用李克特式六點量表，將「非常沒有信心」編碼 1，到回答「非

常有信心」的編碼 6 (如附錄二表二)，總共 20 個題項，經由固定因子數的探索性因素分析 (EFA)，五個因子可以解釋的變異量達到 77.91%，而且總量表與各分量表皆有良好的信效度。因此本研究所稱「學業發展自我效能」，是將 20 個題項的得分加總而來。

(三) 學業成就

受限於個資法的關係，無法取得學生國中教育會考的成績，因此採用距離會考最近的一次模擬測驗成績做為本研究的學業成就。該模擬測驗由坊間出版社於 105 學年度國中教育會考前夕進行模擬測驗，選擇該次模擬測驗的成績主要原因是範圍與國中教育會考相同外，參與該次模擬測驗的學生總數高達 197092 人，佔該年度國中教育會考報名總人數 243564 人的 80.92%，具有足夠的代表性，且計分方式與國中教育會考相同，採用標準參照，並將五科成績事先制定出各科能力表現標準，分成「精熟 (A)」、「基礎 (B)」與「待加強 (C)」三等級，其中數學的非選擇題，加權比重佔該科的 15%，英文聽力測驗的加權比重佔該科的 20%，作文則是獨立於國文科，並分成最高六等級依序遞減至一等級。

編碼方式採用基北區所制訂的換算方式，分別將精熟 (A) 等級再細分出 A++、A+ 與 A 三種標示，基礎 (B) 等級再細分出 B++、B+ 與 B，待加強 (C) 等級則不進行細分，因此會有三個等級共七種標示，再分別給予 7 至 1

分，作文則從一級分的 0.1 分、兩級分的 0.2 分、三級分的 0.4 分、依序增加至六級分的 1 分。滿級分為五科共 35 分，再加上作文六級分的 1 分，總分為 36 分（如附錄 2 表 3）。

肆、研究結果

研究樣本的描述統計如表 1，男女生所佔比例相當（51% vs.49%），家庭結構部分幾乎皆為雙親家庭，父親教育程度的平均年數為 13.42 年，標準差為 3.115 年，若換算成教育階段，約為二專或五專畢業；母親教育程度的平均年數為 13 年，標準差為 3.196 年，若換算成教育階段，同樣約為二專或五專畢業，顯示父親接受教育的年數略高於母親，但大致上仍處於相同的教育階段之內。

父親的職業地位指數平均為 3.23，標準差為 1.443，母親的職業地位指數平均為 2.77，標準差為 1.476，依照黃毅志（2009、2015）的分類，職業指數最高為等級 5，依序遞減，發現父親的職業地位明顯高於母親，至於若將母親接受教育年數與職業地位指數相加後，與學業成就進行相關性分析，達到顯著正相關，因此做為家庭社經地位的主要測量指標，其平均值為 15.77，標準差為 4.137，若將家庭社經地位由高至低分成四個等級（如附錄 2 表 4），本研究樣本的平均家庭社經地位約處於第二至三等級，約屬於社會學分類中的上層與下層中產階級家庭（吳羽茜、郭盛哲、呂明哲，2014）。

表 1

研究樣本的描述統計

	分類	個數 (百分比)	平均數	標準差
性別	男生	53 (51)		
	女生	51 (49)		
家庭結構	雙親	103 (99)		
	單親	1 (1)		
教育程度	父親		13.42	3.115
	母親		13.00	3.196
職業地位	父親		3.23	1.443
	母親		2.77	1.476
家庭 社經地位	母親教育年數 + 母親職業地位		15.77	4.137

一、不同性別學業成就與學業發展自我效能的描述統計及平均數差異考驗

若依不同的性別分別針對學業成就與學業發展自我效能進行描述性統計，表 2 的結果發現，男生的平均學業成就高於女生，但未達統計上的顯著；若依學科進行分析，亦發現男生在數學、社會與自然三科的平均得分高於女生，接著進行獨立樣本 T 檢定，發現僅有數學科的差異達到統計上的顯著 ($t=2.958, p < .01$)，效果量 Cohen's $d = .58$ ，平均數的差異達到中度效果量；但是在學業發展自我效能的平均得分上，則是女生高於男生，但未達統計上的顯著差異。

顯示過去認為的男生理科強，女生文科強的刻板印象，本研究的結果只能證實男生的數學平均得分顯著高於女生，其餘科目皆未達顯著；在學業成就與學業發展自我效能的平均得分也未因性別的

不同而有統計上的差異。若從教育機會均等的角度來說，本研究的結果顯示男女生在教育結果(產出)的層面上，已接近實質的均等，與過去關於學業成就存在性別差異的研究結果不同(Fischer, Schult, & Hell, 2013)。

表 2

不同性別學業成就與學業發展自我效能的描述統計

	學業成就						學業發展自	
	總積分	國文	英文	數學**	社會	自然	我效能	
	平均數 (標準差)	平均數 (標準差)	平均數 (標準差)	平均數 (標準差)	平均數 (標準差)	平均數 (標準差)	平均數 (標準差)	
男生	17.00 (7.37)	3.23 (1.63)	3.13 (1.68)	3.19 (1.59)	3.77 (1.56)	3.23 (1.63)	0.46 (0.19)	89.68 (17.47)
女生	16.00 (5.81)	3.39 (1.47)	3.31 (1.60)	2.39 (1.59)	3.31 (1.27)	2.98 (1.30)	0.60 (0.17)	90.69 (13.85)

* $p < .05$ 、** $p < .01$ 、*** $p < .001$

二、家庭社經地位、學業發展自我效能與學業成就的相關分析

表 3 呈現的是家庭社經地位、學業發展自我效能、學業成就，及各學科的相關性分析。可以發現家庭社經地位與學業成就的相關性達顯著水準($r = .313$, $p < .01$)，屬於低度正相關；而學業發展自我效能與學業成就(總積分)的相關性亦達顯著水準($r = .322$, $p < .01$)，亦屬於低度正相關，但略高於社經地位與學業成就的相關性。

若考慮各科目的相關程度，發現家庭社經地位與作文平均得分未達顯著相關，其餘與各科皆達顯著的低度正相關($r = .251 \sim .334$, $p < .05$)；學業發展自我效能則與數學科未達顯著相關，其餘與各科皆達顯著的低度正相關($r = .231 \sim .322$, $p < .05$)。

顯示家庭社經地位與學業發展自我效能是影響學業成就的重要因素，但社經地位與作文的關係強度較低，學業發展自我效能與數學的關係強度也較薄弱。從實務的角度來看，作文能力的培養要靠豐富的閱讀、敏銳的感知能力及大量的練習等，而學校提供的閱讀資源、學校課程的教學與課堂的寫作練習，皆能彌補家庭社經地位差異產生的資源落差，達到減緩教育機會在起始點的不均等；而數學能力的培養，除了要從小累積基本的閱讀、書寫與計算能力外，還牽涉個體所秉持的內外參照架構的差異，與大魚小池效應的作用，因此可能導致與學業發展自我效能的關連強度不足。

表 3

家庭社經地位、學業發展自我效能與學業成就的相關分析

自變數 \ 依變數	學業成就						
	總積分	國文	英文	數學	社會	自然	作文
家庭社經地位	.313**	.265**	.270**	.251*	.251*	.334**	.040
學業發展自我效能	.322**	.302**	.231*	.187	.351**	.322**	.270**

* $p < .05$ 、** $p < .01$ 、*** $p < .001$

三、家庭社經地位與學業成就的描述統計與平均數差異考驗

表 4 呈現的是家庭社經地位等級與學業成就的描述統計。將總分區分成四個等級，第四等級最高，依序遞減（如附錄 2 表 4）。發現處於最高等級的家庭共有 22 人，佔全部家庭的 21.2%，而第三與第二等級的家庭皆有 22 位，佔 37.5%，第一等級的家庭有 4 人，只佔 3.8%。

顯示本研究的樣本多屬於中產階級（75%），而上層階級的人數（22 人）又高於下層階級的人數（4 人），符合研究對象就讀的學校位於傳統文教地區之印象。

表 4

家庭社經地位與學業成就的描述統計

家庭社經地位等級	個數（百分比）	平均數	標準差	對應的社會階級
第四等級	22 (21.2)	21.18	0.85	上層階級
第三等級	39 (37.5)	16.82	1.37	上層中產階級
第二等級	39 (37.5)	12.87	1.22	下層中產階級
第一等級	4 (3.8)	4.00	4.08	下層階級
總和	104(100)	15.77	4.14	-

本研究以母親的教育年數加上職業地位指數做為「家庭社經地位」的測量指標。將四個等級的家庭社經地位與學業成就的總積分進行單因子變異數分析，瞭解不同社經地位等級家庭的學業成就是否有差異。首先進行變異數同質性檢定，檢定結果顯示 Levene 統計量為 0.493，顯著性為 .688，表示各等級具有同質性，滿足變異數相等的假設，因此繼續以單因子變異數分析進行總檢定。

總檢定結果如表 5 所示，發現總檢定達顯著水準，表示四個家庭社經地位等級對學業成就產生顯著差異，因此繼續進行事後的比較，瞭解不同等級間的差異情形。

表 5

家庭社經地位與學業成就的變異數分析

	平方和	自由度	平均平方和	F	顯著性
組間	551.799	3	183.933	4.601	.005
組內	3997.754	100	39.978		
總和	4549.554	103			

本研究採用 Scheffe 法進行多重比較，主要原因是 Scheffe 法適用於各組人數不相等時，且違犯第一類型誤差（Type I error）的機率較低，另外因為報表呈現對稱性，因此表 6 以最高的第四等級為基準呈現部分比較結果。發現家庭社經地位第四等級對第二等級具有顯著差異，其餘等級間不存在差異。

從社會學的角度來說，顯示上層階級與下層中產階級的學業成就具有明顯差異，本研究的結果發現不同家庭社經地位對於學業成就，確實會產生影響並有差異，但並非全部等級皆有相同的現象，僅在部分等級間存在差異。

表 6

家庭社經地位與學業成就的多重比較表

I	J	Scheffe 法 平均差異 (I-J)	標準誤
第四等級	第三等級	3.6906	1.6859
	第二等級	5.9367*	1.6859
	第一等級	7.5636	3.4368

依變數：總積分

* $p < .05$ 、** $p < .01$ 、*** $p < .001$

四、家庭社經地位與學業發展自我效能的平均數差異考驗

接著將家庭社經地位與學業發展自我效能的總分進行單因子變異數分析，瞭解不同社經地位等級家庭的學業發展自我效能是否有差異。先進行變異數同質性檢定，檢定結果顯示 Levene 統計量為 2.817，顯著性為 .029，達顯著水準，表示各等級間並不同質，不符合變異數相等假設，接著以變異數相等的假設改用 Brown-Forsythe 進行變異數分析，並使用 Tamhane's T2 檢定法，結果顯示均等平均數的 Robust 檢定未達顯著（漸進的 F 分配 = .629，

$p = .648$)，無法進行總檢定與事後比較，此結果與過去的研究結果不同（黃建皓，2011）。

顯示不同家庭社經地位等級對學業發展自我效能並未產生等級間的差異，猜測可能是因為研究工具所測量的五個向度從教育社會學的角度分析，不同家庭社經地位的慣習、文化與社會資本皆不同，僅有「向朋友學習」這個向度與社會資本有關，「升學計畫」這個向度與慣習可能有關，其餘的向度與 Boureidu 與 Bandura 理論的關連強度較低，導致不同家庭社經地位與學業發展自我效能的差異不明顯。

五、學業發展自我效能各分量表的描述統計與相關分析

由表 7 可以發現，女生在學業發展自我效能的平均得分略高於男生 ($90.69 > 89.68$) 以獨立樣本 T 檢定檢驗男女生的平均數差異， t 值為 -3.25 ，未達統計顯著性。顯示男女生在學業發展自我效能的平均得分上無差異，結果與過去的文獻不同（Huang, 2013）。

表 7

不同性別學業發展自我效能的描述統計與平均數差異考驗

	學業發展自我效能		
	個數 (百分比)	平均數	標準差
男生	53 (51)	89.68	17.47
女生	51 (49)	90.69	13.85
總和	104 (100)	90.17	15.73

表 8 所呈現的是學業發展自我效能五個分量表的描述統計及相關分析。發現五個量表中以「做負責的學習者」的平均分數最高，達到 20.25 分，其次是「向朋友學習」的 18.09 分，「升學計畫」的

17.76 分，「時間管理」的 17.29 分，最低的是「學習和應試技巧」的 16.79 分。

顯示學生在「學習和應試技巧」的自我效能最不足，而在「時間管理」與「升學計畫」上也待加強，而一般學生在學校中雖然接受各科老師的教導，但對於上述三個分量表的自我效能，卻呈現出高需求，從實務的角度來說，一般教師很少教導學生如何學習，多是教導學生如何「考高分」，容易忽略「學習策略與方法」的教授，也很少針對如何規劃時間，確定自己的升學目標並發展升學計畫有所著墨，因此分數偏低，研究數據正好提醒學校教師除了學科知識的傳授外，掌握學習的能力與方法也是學生極缺乏待培養的層面。

從相關分析發現，分量表間的相關係數皆達顯著（ $r = .550 \sim .837, p < .001$ ），且「時間管理」與「學習和應試技巧」間的相關係數最高（ $r = .837, p < .001$ ），達顯著的高度正相關，「時間管理」對「升學計畫」與「向朋友學習」達顯著的高度正相關（ $r = .768, .749, p < .001$ ），顯示「時間管理」對於上述三個分量表的關係高度密切，但僅與「做負責任的學習者」相關性達顯著的正向中度相關（ $r = .550, p < .001$ ）。由此可知時間管理的能力與學生學業自我效能的高低息息相關，需予以重視。

表 8

學業發展自我效能各分量表的描述統計與相關分析

學業發展自我效能分量表					
	時間管理	學習和 應試技巧	向朋友學習	升學計畫	做負責任的 學習者
平均數 (標準差)	17.29 (3.70)	16.79 (3.96)	18.09 (3.62)	17.76 (3.59)	20.25 (3.08)
時間管理 相關係數	1	.837***	.749***	.768***	.550***

* $p < .05$ 、** $p < .01$ 、*** $p < .001$

六、學業發展自我效能與學業成就的描述統計與平均數差異考驗

因為學業發展自我效能每個題項最低 1 分，最高 6 分，共有 20 個題項，因此最低總分為 20 分，滿分為 120 分。本研究以 40 分為組距，將學業發展自我效能分成高效能組、中效能組與低效能組三組（如附錄 2 表 5），描述統計如表 9。

發現低效能組的人數只有 1 人，所以將中與低效能組的人數合併。其中屬於高學業發展自我效能的個數達到 85 人，佔全部人數的 81.7%，學業成就的平均為 16.92 分，中效能組的人數有 19 人，佔 18.3%，學業成就的平均為 14.67 分。以國中教育會考模擬測驗的總分為 36 分來看，高效能組學生的得分拿到一半左右的分數，中效能組則低於一半得分，顯示學業發展自我效能的高低與學業成就的相關可能不高，需藉由接下來的相關分析驗證這個可能性。

由相關分析發現，因為學業發展自我效能的分類是由人工轉換為等級變項，再合併簡化成二元變項，而學業成就為連續變項，因此兩者的相關分析可以採用點二系列相關、斯皮爾曼等級相關，或是將學業發展自我效能予以虛擬編碼，投入簡單迴歸中。結果顯示三種方法間的相關係數皆未達顯著（ $r = .13, p = .184$ ； $\rho = .137, p = .165$ ； $R = \beta = .131, p = .184$ ），證實不同程度的學業發展自我效能，與學業成就的相關強度微弱，結果與歷次 TIMSS 的評比結果發現我國中學生的學業成就與學習自信為負相關，呈現出高成就卻低自信與興趣的結果剛好相反，為何有這樣的結果，仍待後續研究的深入探討。

表 9

學業發展自我效能與學業成就的描述統計

學業發展自我效能等級	個數 (百分比)	平均數	標準差
高效能組	85 (81.7)	16.92	6.85
中效能組	19 (1.3)	14.67	5.38
總和	104 (100)	16.51	6.64

接著瞭解不同效能等級的學業成就是否有差異，以獨立樣本 T 檢定不同效能組學生的學業成就平均數差異考驗，發現高效能組的學業成就雖然高於中效能組，但統計結果顯示 t 值為 1.34， p 值為 .184，未達顯著差異。

顯示高與中效能組的學業成就差異未達統計上的顯著性，不同學業發展自我效能的等級分類，無法有效區隔出學業成就的差異，與前述的研究結果一致，即學業發展自我效能與學業成就的相關強度未達顯著性，但與過去文獻的研究結果不相符（王振宏，1999；王凱榮、辛濤、李瓊，1999；李薈、辛濤、谷生華、申繼亮，1998；徐靜，2013；Bandura, 1977, 1986; Catalina, Stanescu, & Mohorea, 2012; Galyon, Blondin, Yaw, Balls, & Williams, 2012; Weiser & Riggio, 2010; Young, Hyuck, Sunyoung, & You, 2012）。

七、以學業發展自我效能做為中介變數的中介效果分析

接著進行中介變數的檢驗，依照 Baron 與 Kenny (1986) 的檢驗步驟，首先進行家庭社經地位與學業成就的迴歸分析，以「家庭社經地位」為自變數，「學業成就」為依變數，建立迴歸模式，如表 10。發現家庭社經地位與學業成就的徑路係數達 .313 ($p < .001$)，顯示家庭社經地位對學業成就具有顯著正向的影響，總效果為 .313 (如圖 2 所示)。接著進行第二步驟的檢驗。

表 10

家庭社經地位對學業成就的迴歸分析

模式	未標準化係數	標準化係數	標準誤差	T 值
	B 之估計值	Beta 分配		
常數	8.582***		2.462	3.485
家庭社經地位	.503***	.313	.151	3.329

依變數：學業成就

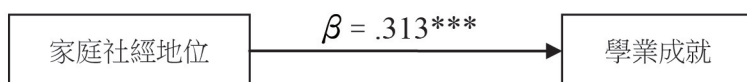
* $p < .05$ 、** $p < .01$ 、*** $p < .001$ 

圖 2 家庭社經地位與學業成就的徑路圖

以「家庭社經地位」做為自變數，將中介變數「學業發展自我效能」視為依變數，投入迴歸模式，如表 11。發現家庭社經地位的徑路係數為.013 ($p < .05$)，即一階效果為.013。顯示家庭社經地位會顯著正向的影響學業自我效能，因為自變數與中介變數的徑路係數達顯著，但本研究的樣本數未達兩百以上，因此不需要也無法進行 Sobel(1982)的間接效果檢驗，直接進入第三步驟的檢驗，如圖 3。

表 11

家庭社經地位對學業發展自我效能的迴歸分析

模式	未標準化係數	標準化係數	標準誤差	T 值
	B 之估計值	Beta 分配		
常數	89.170***		6.137	14.529
家庭社經地位	.640*	.013	.323	1.98

依變數：學業發展自我效能

* $p < .05$ 、** $p < .01$ 、*** $p < .001$

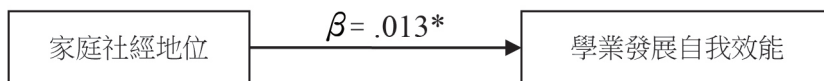


圖 3 家庭社經地位與學業發展自我效能的徑路圖

Baron 與 Kenny (1986) 中介效果檢測的最後一個步驟是控制自變數「家庭社經地位」，將自變數「家庭社經地位」與中介變數「學業發展自我效能」一起投入迴歸模式中，以「學業成就」為依變數進行預測，如表 12。發現家庭社經地位的徑路係數為.308 ($p < .001$)，而學業發展自我效能的徑路係數為.317($p < .001$)，兩者皆達顯著。顯示家庭社經地位與學業發展自我效能皆同時會對學業成就產生正向的影響，且控制家庭社經地位後的徑路係數 ($\beta = .308$) 小於未控制前的徑路係數 ($\beta = .313$)。其中家庭社經地位對學業成就的直接效果為.308，間接效果為.013 \times .317 = .00412，如圖 4 所示。

顯示家庭社經地位對學業成就的直接效果，大於學業發展自我效能間接影響學業成就的間接效果，即「家庭社經地位對學業成就的影響力高於間接透過學業發展自我效能對學業成就的影響力」，依 Baron 與 Kenny (1986) 的說法，本研究屬於「部分中介」模式，家庭社經地位會大部分正向影響學業成就，同時也會透過學業發展自我效能少部分間接正向影響學業成就。

表 12

家庭社經地位透過學業自我效能，對學業成就的迴歸分析

模式	未標準化係數	標準化係數	標準誤差	T 值
	B 之估計值	Beta 分配		
常數	-3.368		4.086	-.824
家庭社經地位	.494***	.308	.143	3.454
學業發展自我效能	.134***	.317	.038	3.561

依變數：學業成就

* $p < .05$ 、** $p < .01$ 、*** $p < .001$ 

圖 4 家庭社經地位透過學業發展自我效能影響學業成就的徑路圖

伍、研究討論與建議

一、研究討論與問題回應

(一) 不同家庭社經地位的學生，學業成就是否具有差異？

回應：研究結果發現，家庭社經地位與學業成就間存在低度的正相關 ($r = .251 \sim .334$, $p < .05$)，僅與作文未達顯著相關，且由單因子變異數分析的結果亦發現不同家庭社經地位等級的子女，與學業成就存在顯著差異，表示家庭社經地位的高低，確實會對學習成就產生影響與差異，其中第四等級與第二等級間的平均數差異甚至達到了顯

著，研究假設一獲得證實，且研究的結果與過去文獻所顯示的相符（吳慧瑛，2007；林俊瑩、黃毅志，2008；陳順利、黃毅志，2015；謝孟穎，2003；Coleman, 1966; Howie & Tjeerd, 2003; Kiamanesh, 2005; Sewell & Hauser, 1976）。

從再製理論的角度來說，家庭社經地位差異不僅僅影響經濟資本、文化資本與社會資本，父母的教養行為、家庭語言的類型、對子女學習的參與、教育期望等皆具有重要而關鍵的作用，眾多的差異讓子女進入學校前便已產生起始點的不均等，而學校與教師如何發揮教育的功能，減緩差距的擴大，賦予學生改變的可能性，是打破階級再製的關鍵因素，也是促進社會流動的重要來源。

（二）不同家庭社經地位的學生，學業發展自我效能是否具有差異？

回應：研究結果顯示家庭社經地位與學業發展自我效能間的相關性雖達顯著，但僅有微弱的正相關（ $r = .04$ ， $p < .05$ ），而單因子變異數分析的結果也指出 Levene 統計量為 2.817，顯著性為.029，達顯著水準，表示各等級間並不同質，未能滿足變異數相等的假設，以變異數相等的假設改用 Brown-Forsythe 進行變異數分析，使用 Tamhane's T2 檢定法進行檢定，發現均等平均數的 Robust 檢定未達顯著（漸進的 F 分配= .629， $p = .648$ ），無法進行總檢定及事後比較。表示不同家庭社經地位的學生，學業發展自我效能並未達到統計上的顯著差異，研究結果未能支持假設二的成立。

可能的原因在於學業發展自我效能的五個因子所測量的內容，與家庭社經地位產生的連帶影響差異過大，像是「向朋友學習」可能與文化、社會資本有關，「升學計畫」可能與慣習或父母期望有關，至於「時間管理」、「學習和應試技巧」與「做負責任的學習者」等三個向度，似乎與家庭社經地位差異的關連性較低，進而導致相關性的偏低與不顯著的結果，也與過去的研究結果不同（黃建皓，2011；Coleman & Karraker, 1997; Gecas & Schwalbe, 1983; Huang, 2013; Lin, 2003）。

但從另一個角度來說，這卻是學校教師可以施力與發揮的部分，學校教師可以嘗試在課堂中教導學生如何分配讀書時間、規劃讀書、複習與學習計劃，並在教學中加入學習與應試技巧的培養，而學校圖書館也能提供相關學習與考試技巧的書籍，增加學生的學習自信與效能，突破家庭社經地位不利的枷鎖，成功取得學業成就的進步，扭轉階級再製的命運，促成更多向上流動的機會。

(三)不同學業發展自我效能的學生，學業成就是否具有差異？

回應：研究結果顯示學業發展自我效能與學業成就(總積分)的相關性達顯著水準 ($r = .322, p < .01$)，屬於低度正相關，學業發展自我效能則與數學科未達顯著相關，其餘與各科皆達顯著的低度正相關 ($r = .231 \sim .351, p < .05$)。

研究結果似乎可以支持假設三，學業發展自我效能確實會對學業成就產生影響。但更細緻的分析卻發現學業發

展自我效能與數學科未達顯著相關，可能因為數學能力的培養需從小累積讀、寫、算能力外，還牽涉個體所秉持的內外參照架構的差異，與大魚小池效應的潛在作用，導致與學業發展自我效能的關連強度不足。

若將學業發展自我效能分組，各組與學業成就的相關皆未達顯著 ($r = .13$)，且事後比較也發現高與中效能組的學業成就未達顯著差異 ($t = 1.34, p = .184$)，而這結果似乎又否定了假設三，因此嚴格說起來，學業發展自我效能對學業成就確實有影響，但不同程度學業發展自我效能的差異，在學業成就上並無法明顯的區辨出等級上的不同，也就是不同等級的學業發展自我效能，其學業成就存在差異，但無法以統計證實其差異的存在，此結果與過去文獻所顯示的結果不盡相同（王振宏，1999；王凱榮、辛濤、李瓊，1999；李蒼、辛濤、谷生華、申繼亮，1998；徐靜，2013；Bandura, 1977, 1986; Catalina, Stanescu, & Mohorea, 2012; Galyon, Blondin, Yaw, Balls, & Williams, 2012; Weiser & Riggio, 2010; Young, Hyuck, Sunyoung, & You, 2012）。

（四）家庭社經地位是否透過學業發展自我效能，對學業成就產生影響？

回應：依據 Baron 與 Kenny (1986) 中介效果的三步驟進行分析，發現家庭社經地位會透過學業發展自我效能 ($\beta = .013$)，間接影響學業成就 ($\beta = .137$)，但家庭社經地位也會直接影響學業成就 ($\beta = .308$)，屬於部分中介模式，而家庭社經地位對學業成就的直接效果為.308，間接效

果為.00412，因此研究假設四獲得證實。結果與相關的研究相符（Weiser & Riggio, 2010）

但明顯看到直接效果遠大於間接效果（.308 > .00412），顯示家庭社經地位會大部分正向影響學業成就，同時也會透過學業發展自我效能少部分間接正向影響學業成就，雖然研究假設四獲得證實，卻發現家庭社經地位透過學業發展自我效能對學業成就的影響效果不如預期的高，後續的研究或許能深入探討可能的原因。

將上述的假設檢驗及討論結果彙整如表 13：

表 13

研究假設驗證結果

	研究假設內容	驗證結果
假設一	家庭社經地位對學業成就會產生影響	成立
假設二	家庭社經地位對學業發展自我效能會產生影響	不成立
假設三	學業發展自我效能對學業成就會產生影響	部分成立 低度相關但組間無差異
假設四	家庭社經地位會透過學業發展自我效能影響學業成就	成立 直接效果 > 間接效果

二、研究建議

依據上述的研究結果，研究的建議如下：

（一）學校教師宜教導學生學習如何學習

研究結果發現家庭社經地位與學業發展自我效能皆會影響學業成就的表現，且不同家庭社經地位的學業發展自

我效能並未呈現顯著差異。因此增進學生學習技巧可以有效抑制社經地位造成的差異，而一般在學校教學中，往往僅針對分科的課程內容進行教學，卻忽略如何讓學生學習如何學習、增加個體後設認知與自我監控的能力，導致學生花費大量的時間在學習學科知識，卻無法有效整合學科知識內與之間的關連，也無法主動進行內容的整理、歸納、比較與分析，因此學校教師可以嘗試教導學生學習的技巧，像是如何做筆記、如何使用不同顏色的文具進行重要性的區分、如何以圖像或圖表的方式將課程內容重新表徵進行再認知的程序等，藉由提升學生的學習能力，增加學生的效能感與學業成就。

(二) 學校教師宜加強學生學習的自信心與應試技巧，以增進學業成就

研究結果發現學業發展自我效能的「學習和應試技巧」分量表平均得分最低，而學業自我效能的高低，來自於個體對自己能力的評估與同儕比較的結果，因此如何提升個體自信心便十分重要。教師在教學中以正向、積極的態度鼓勵學生學習，以身教做為榜樣與示範，教導學生正確的心理歸因，另外選擇適當數量與難度的題目讓學生練習，使學生熟悉相關的考試技巧增加學生應試的經驗，改善學生的應試技巧進而提升學業成就，從而達到增加學生學習自信心的目標，讓學生在學校中能夠體驗成功並享受學習的高峰經驗。

(三) 學校宜增加關於時間規劃、學習方法與學習技巧相關書籍的購買，並鼓勵學生借閱

研究結果顯示「時間管理」與其他四個分量表的相關係數皆達顯著的高度正相關，表示時間管理的能力對於學業發展自我效能十分關鍵，但在學校中普遍缺乏對於時間管理能力、學習方法與技巧的培養與訓練，因此學校藉由讓學生閱讀大量的書籍，增進學生閱讀理解的能力，不但有助於表達能力的培養，也能增加對學科知識的理解與學習，因此相關學習方法、應試技巧與時間規劃等非認知能力的培養刻不容緩，而圖書館也應增購數量充足的相關書籍，填補學生在學習策略與技巧上的缺漏。

(四) 學校宜規劃適當的時間讓學生閱讀與學習如何學習有關的書籍

目前政府當局一直在推動閱讀能力的提升，甚至在各校設置專任的「閱讀推動教師」，給予減授鐘點的福利，目的是希望閱讀推動教師能專注在學生閱讀能力的培養，閱讀能力固然重要，但徒有閱讀能力卻不等於擁有學習的能力，因此在推動閱讀時，不但須融入各科的知識，也應讓學生閱讀與學習有關的書籍，增進學生的學習概念、技巧與考試能力，不但有助於閱讀理解能力的增加，也對學生的學習自信與效能產生正向影響。

(五) 學校與教師宜提供足夠的升學資訊並教導研擬與升學有關的各種計畫

隨著教育制度的變革，多元入學的管道讓學生擁有更多選擇的權利，但對於不同升學管道的優勢與缺點不甚清楚，有時連家長與老師也未能充分掌握，因此學校與教師宜熟悉並提供充份的升學資訊，並針對不同的需求與目標，協助擬定適合的升學計畫，並配合家長進行計畫的落實，除了有助於學習動機的提升，也能增進學習動機與態度，同時配合學習策略的培養，將有助於學業成就的提升與學業自我效能感的增長。

三、研究限制與後續的研究方向

本研究的限制與未來的研究方向，包含下列八點：

(一) 研究樣本的抽樣限制

本研究的樣本採立意取樣，非隨機抽樣，因此研究結果的適用性受限於特定條件之下，導致外在效度不足，加上少子化因素，每班的人數僅剩 25 人以下，且施測時間為會考結束後，許多受測對象請假缺席或是無意願填答，造成樣本大量流失，產生取樣上的偏差，缺乏足夠的代表性與推論性。

後續的研究考慮採取分層叢集抽樣擴大樣本數，兼具抽樣的隨機性與施測的可行性；或採取次級資料分析，使用國內外相關的大型教育資料庫（TEPS、TASA、TIMSS、PISA）所釋出的原始資料，將有助於外在效度的提升。

（二）研究變數的選擇限制

研究中所考慮的中介變數僅限「學業發展自我效能」，並無法排除其他中介變數存在的可能性，例如家庭文化資本與社會資本（Bourdieu, 1986; Bourdieu & Wacquant, 1992; Coleman, 1987, 1988）、父母親的教養行為與家長參與（李怡慧，2015；張旭，2010）、家長與教師的教育期望（楊春華，2006；蔡銘津，2014）等，因此無法確認家庭社經地位是否可能透過其他中介變數影響學業成就的適用模型。

後續的研究考慮先進行相關變數的後設分析（meta-analysis），找出過去研究中哪些變數對學業成就具有較高的效果量，事先掌握中介變數的種類與影響力，應可增加研究模型的精確度與適配性（陳瑋婷，2011；黃義良、王怡又，2016）。

（三）家庭社經地位的測量限制

關於家庭社經地位的測量，本研究原始設計採取父母親的教育程度、職業與每月收入三個測量指標，但收集資料的過程卻發現有超過半數的學生並不清楚雙親的每月收入，因此被迫排除父母親每月收入這個變項，降低在測量家庭社經地位時的內在效度。

後續的研究可採取替代的題項，例如詢問家中的藏書量、是否有獨立的房間、是否擁有電腦與相關參考書籍、過去一年內與家人出遊的次數等，做為測量每月收入的替

代指標（任春榮，2010），相信能有效提高家庭社經地位測量的精確性。

（四）家庭社經地位的指標選擇限制

對於家庭社經地位的指標認定，文獻提到不同的選擇方式，有採取雙親的資料一起合併計算、有選取雙親中數值較大的進行計算、有單獨選擇父親或母親指標，抑或先進行簡單的統計分析（相關、迴歸），找出父親或母親的指標對依變數的影響較大，再單獨進行計算等，不同的選擇方式仍存在爭議且未有定論。

後續的研究可採取競爭型模型或是對比式研究，針對不同的家庭社經地位指標分別放入研究中，單獨探討不同指標產生的影響與差異，將可取得更全面的測量結果，並探討不同指標差異的可能原因。

（五）家庭社經地位的指標運算限制

社經地位指數的計算方式，過去的研究多採取原始數值相加，或是不同指標加權和（林生傳，2005；Hollingshed, 1958）等方式，任春榮（2010）指出還有迴歸方程式法、因素分析法與項目反應理論等，哪一種指標的合併計算方法較能準確的測量與代表一個家庭的社會經濟地位，目前仍未存在共識。

後續的研究可依研究對象的屬性與特質，採取相關的指標運算方式，並在研究設計階段即納入考量。

(六) 共同方法變異的檢測限制

本研究雖有考慮並檢測共同方法變異出現的可能性，但採取的是傳統的 Harman 的單因子檢測方式，該方法簡單易用，但不夠精確，即便使用驗證性因素分析（CFA）的方式，亦無法證明測量中完全沒有共同方法變異的發生。

後續的研究若能在研究設計與測量階段，採取事先控制方式，如隔離式資料蒐集法或問卷設計法（張偉豪，2011），減少事後統計方式的使用，或是採用其他事後補救的方式，像是第三因素測試法、潛在共同方法變異變數測試法、MTMM 測試法等（周浩、龍立榮，2004；彭台光、高月慈、林鈺琴，2006），將能更精準的測量並減少共同方法變異發生的可能性，使研究設計更嚴謹。

(七) 中介效果的檢測限制

本研究所進行的中介效果檢測，是依 Barron 與 Kenny（1986）針對簡單中介模式進行的檢測步驟，但 Mackinnon, Lockwood, Hoffman, West, & Sheet（2002）的研究卻顯示在小樣本的情況下，迴歸係數會因為統計考驗力較低而未達顯著，增加違犯型二錯誤的可能性，或是使用 Sobel 檢定法時，其假設抽樣分配需常態分布，若不符合常態分配假設，統計考驗力將大幅下降。因此其他文獻建議樣本數至少需要大於兩百份；若樣本數不足或非常態分布，建議採取拔靴法（Bootstrap）推估特定母群體的信賴區間，對於中介效果的檢測將更精準（張再明、曾迎新，2014）。

後續的研究可考慮更複雜的模型，像是多重中介模型、中介性調節效果、調節性中介效果與多層次中介與調節效果等模型，或是使用結構方程模型（SEM）的統計技術，處理中介變數為潛在變數時的情形，減少測量誤差的產生（李駿，2012；陳順宇、陳譽仁，2015），亦或採用階層線性模式（HLM）的統計技術，處理多個層次間的中介或調節效果（溫福星、邱皓政，2009；蕭佳純，2015），將有助於中介變數的界定與影響程度的測量。

（八）研究推論的限制

本研究的受測樣本僅針對臺北市某完全中學國中部九年級學生，因此所得的研究結論僅限於該校與該年度的九年級學生，不宜推論至其他縣市、學校、年度或年級，避免產生過度的推論，造成外在效度的威脅。

後續的研究建議嘗試建置該校九年級學生的學業發展自我效能縱貫資料庫，收集不同時間點的資料。因為時間點是鑲嵌於學生身上，而學生則是鑲嵌於班級的層次之中，形成上下嵌套的資料結構與階層關係，適合採取更進階的潛在成長曲線模型（Latent Growth Curve Modeling，簡稱 LGCM）與多層次模型（Multilevel Modeling，簡稱 MLM）的縱貫式中介效果檢定（余民寧，2013；吳明隆，2013；邱皓政，2017），對於縱貫資料將有更精確與細緻的分析，將有助於實務上的瞭解與改善。

致謝

本文曾口頭發表於 2017 年臺師大教評中心主辦的「強化教育研究與政策暨實務連結」國際學術研討會，承蒙該場次主持人潘慧玲教授、評論人陳文彥教授的寶貴意見，及兩位匿名審查委員的審閱與指正，謹此誌謝。

參考文獻

- 方 杰、張敏強、邱皓政（2012）。中介效應的檢驗方法和效果量測量：回顧與展望。*心理發展與教育*，**1**，105-111。doi：1001-4918（2012）01-0105-111
- 王振宏（1999）。初中學生學業自我效能與學業成就關係研究。*心理發展與教育*，**1**，39-43。
- 王凱榮、辛濤、李瓊（1999）。中學生自我效能感、歸因與學習成績關係的研究。*心理發展與教育*，**4**，22-25。
- 任春榮（2010）。學生家庭社會經濟地位（SES）的測量技術。*教育學報*，**6**（5），77-82。
- 行政院主計處（1992）。*中華民國職業標準分類（第五次修正版）*。臺北市：行政院主計處編印。
- 何秋蓮（2007）。*國中生家庭社經地位、文化資本、教育期望對其學業成就影響之研究－以臺南縣國中基本學力測驗為例*（未出版之碩士論文）。國立嘉義大學教育行政與政策發展研究所，嘉義市。
- 余民寧（2013）。*縱貫性資料分析：LGM 的應用*。臺北市：心理。
- 吳羽茜、郭盛哲、呂明哲（譯）（2014）。Benokraitis, N. V 著。*社會學*。臺北市：新加坡商聖智學習。

- 吳明隆 (2013)。結構方程模式：潛在成長曲線分析。臺北市：五南圖書。
- 吳素媛 (2004)。家庭教育關連資本對國小學童數理成就影響模式之分析 (未出版之碩士論文)。國立嘉義大學科學教育研究所，嘉義市。
- 吳慧瑛 (2007)。家庭背景與教育成就：五個出生世代的比較分析。人口學刊，34，109-143。
- 宋曜廷、黃瓏瑩、郭念平、曾芬蘭 (2012)。以縱貫學業表現檢驗大魚小池效應與見賢思齊效應。中華心理學刊，54 (3)，315-330。
- 李明 (譯) (1990)。Matras, J 著。社會不平等-社會階層化與流動。臺北市：桂冠圖書。
- 李波 (2009)。論成教學生學習自我效能感的培養。繼續教育研究，5，90-92。
- 李恒 (2010)。學生自我效能感及相關概念的研究綜述。湖北第二師範學院學報，27 (7)，97-100。
- 李蒼、辛濤、谷生華、申繼亮 (1998)。中學生自我效能感、學習策略與學習成績關係的研究。教育研究與實驗，4，48-52。
- 李駿編譯 (2012)。中介作用分析 (Iacobucci, D 原著)。上海市：格致。
- 李文益 (2004)。文化資本、多元入學管道與學生學習表現－以臺東師院為例。臺東大學教育學報，15 (1)，1-32。
- 李文益、黃毅志 (2004)。文化資本、社會資本與學生成就的關連性之研究－以臺東師院為例。臺東大學教育學報，15 (2)，23-58。

- 李宗義、許雅淑編譯（2016）。**階級世代：窮小孩與富小孩的機會不平等**(Putnam, R. D.原著)。新北市：衛城。
- 李怡慧編譯（2015）。**家庭優勢**（Lareau, A 原著）。新北市：群學。
- 李華夏編譯（2007）。**有閒階級論**（Veblen, T.原著）。新北市：左岸文化。
- 李德治（2016）。**多變量分析：專題及論文常用的統計方法**。臺北市：雙葉書廊。
- 李錦旭編譯（1989）。**資本主義美國的學校教育：教育改革與經濟生活的矛盾**（Bowles, S., & Gintis, H.原著）。臺北市：桂冠。
- 沈靜、姚本先（2006）。西方學業自我概念研究的回顧與反思。**江蘇大學學報（高教研究版）**，**28**（2），22-26。
- 周浩、龍立榮（2004）。共同方法偏差的統計檢驗與控制方法。**心理科學進展**，**12**（5），942-950。
- 周新富（2008）。社會階級對子女學業成就的影響－以家庭資源為分析架構。**臺灣教育社會學研究**，**8**（1），1-43。
- 周新富、王財印（2006）。社會資本在家庭代際人力資本傳遞作用之探討。**國立臺北教育大學學報**，**19**（2），281-306。
- 周學霽、黃英哲（2007）。自我效能和集體效能：理論與研究回顧。**國北教大體育（研究類）**，**2**，67-74。
- 林生傳（2005）。**教育社會學**。臺北市：巨流圖書。
- 林俊瑩、黃毅志（2008）。影響臺灣地區學生學業成就的可能機制：結構方程模式的探究。**臺灣教育社會學研究**，**8**（1），45-88。
- 林祐聖、葉欣怡（譯）（2005）。林南著。**社會資本**。（Social Capital）。臺北市：弘智。

- 邱皓政 (2017)。多層次模式與縱貫資料分析：Mplus 8 解析應用。臺北市：五南圖書。
- 侯雅齡(2010)。科學自我概念之大魚小池效應探究：資優生教育安置方式的思考。教育科學研究期刊，55 (3)，61-87。
- 孫旻儀、蔡明學 (2007)。社經地位和學生學業成就關係之後設分析。國教學報，19，199-221。
- 徐 靜 (2013)。自我效能感在學生學習中的研究。內蒙古師範大學學報 (教育科學類)，26 (6)，65-67。
- 袁文德、艾思柏、許錦屏、梁健文、劉兆瑛、陳茂釗、佘文基與柯新豔 (2004)。學業發展自我效能量表：使用者手冊。香港：香港大學教育學院生活技能發展計畫。
- 張 旭編譯 (2010)。不平等的童年 (Lareau, A 原著)。北京市：北京大學出版社。
- 張再明、曾迎新 (2014)。中介作用的意涵及其資料分析策略之探究。家庭教育與諮商學刊，16，29-44。
- 張芳全 (2006)。社經地位、文化資本與教育期望對學業成就影響之結構方程模式檢定。測驗學刊，53 (2)，261-295。
- 張芳全 (2011)。家長教育程度、文化資本、自我抱負、學習興趣與數學成就之關係研究。臺中教育大學學報：教育類，25 (1)，29-56。
- 張芳全 (2017)。家庭社經地位對英語學習成就之影響－以幸福感為中介。學校行政，108，204-221。 doi: 10.3966/160683002017030108010.
- 張芳華 (2015)。家庭社經地位、社會資本對高中學生教育抱負影響之分析。國立臺南大學教育研究學報，49 (2)，19-40。

- 張訓譯 (2017)。大學入學甄試影響教育機會均等之分析。**育達科大學報**，45，69-90。
- 張偉豪 (2011)。**SEM 論文不求人**。高雄市：三星統計。
- 許哲耀、鄧家駒、陳郁婷 (2014)。動力還是壓力？國中生教育期望、學業自信對學習表現與身心壓力之影響。**數據分析**，9 (5)，135-146。
- 許崇憲 (2002)。家庭背景因素與子女學業成就之關係：臺灣樣本的后設分析。**中正教育研究**，1 (2)，25-62。
- 郭丁熒、許竣維 (2004)。不同社經背景小學生的數學科學業成就、文化資本、經濟暨財務資本、及社會資本關係之差異。**國立臺南大學教育學誌**，17，77-119。
- 陳順宇、陳譽仁 (2015)。**結構方程模式－中介與干擾效果**。臺北市：三民書局。
- 陳順利、黃毅志 (2015)。解除 Coleman 等人報告書的魔咒：學校中的班級因素對學業成績之影響。**教育科學研究期刊**，60 (2)，111-138。
- 陳瑋婷 (2011)。自我效能、學習策略與學業成就之關係研究。**師資培育與教師專業發展期刊**，4 (2)，83-96。
- 陸 洛編譯 (1997)。**社會階級心理學** (Argyle, M. 原著)。臺北市：巨流。
- 彭台光、高月慈、林鈺琴 (2006)。管理研究中的共同方法變異：問題本質、影響、測試和補救。**管理學報**，23 (1)，77-98。
- 曾明基 (2016)。家庭與學校差異對國中生數學成就影響之研究。**臺北市立大學學報**，47 (1)，23-40。

- 黃芳玫、吳齊殷（2010）。台灣國中學生個人特質、家庭背景與學業成績－追蹤調查資料之研究。**經濟論文叢刊**，**38**(1)，65-97。
- 黃建皓（2011）。家庭社經地位與班級經營效能對學生自我效能感之影響：階層線性模式分析。**教育經營與管理研究集刊**，**8**，107-130。
- 黃庭康（2017）。P.Willis 的<學做工>：概念方法與研究方向。**教育研究集刊**，**63**(4)，37-63。doi:10.3966/102887082017126304003
- 黃義良、王怡又（2016）。學校內部行銷、教師滿意度與學校效能關聯之探討：後設分析結合結構方程式模型之驗證。**彰化師大教育學報**，**28/29**，25-52。
- 黃毅志（1998）。臺灣地區新職業分類的建構與評估。**調查研究**，**5**，5-32。
- 黃毅志（2003）。「臺灣地區新職業聲望與社經地位量表」之建構與評估：社會科學與教育社會學研究本土化。**教育研究集刊**，**49**(4)，1-31。
- 黃毅志（2005）。教育研究中的「職業調查封閉式問卷」之信效度分析。**教育研究集刊**，**51**(4)，43-71。
- 黃毅志（2009）。國際新職業量表在台灣教育研究中的適用性：本土化與國際化的考量。**教育科學研究期刊**，**54**(3)，1-27。
- 黃毅志（2015）。職業測量方法。載於瞿海源、畢恆達、劉長萱、楊國樞（主編），**社會及行為科學研究法**（頁 433-464）。臺北市：東華。
- 黃毅志、楊雅鈞（2007）。背景因素、學業成績與成就抱負－以台東大學為例。**台東大學教育學報**，**18**(2)，31-66。

- 黃瓊儀、吳怡慧、游錦雲（2015）。國中身心障礙學生家庭社經地位、社會資本、文化資本、財務資本與學習成果之關係研究。**教育科學研究期刊**，**60**（4），129-160。
- 楊春華（2006）。教育期望中的社會階層差異：父母的社會地位和子女教育期望的關係。**清華大學教育研究**，**27**（4），71-76。
- 溫福星（2012）。**多變量分析方法與操作－SPSS 之應用**。臺北市：鼎茂圖書。
- 溫福星、邱皓政（2009）。組織研究中的多層次調節式中介效果：以組織創新氣氛、組織承諾與工作滿意的實證研究為例。**管理學報**，**26**（2），189-211。
- 詹秀雯、張芳全（2014）。影響國中生學習成就因素之研究。**臺中教育大學學報：教育類**，**28**（1），49-76。
- 詹秀雯、張芳全（2014）。影響國中生學習成就因素之研究。**臺中教育大學學報：教育類**，**28**（1）。49-76。
- 鄒浮安（1994）。家庭社經地位與學業成就之關係：後設分析。**教育研究資訊**，**2**（3），38-47。
- 趙珮晴、余民寧（2012）。自律學習策略與自我效能、學習興趣、學業成就的相關研究。**教育研究集刊**，**59**（3），1-32。
- 劉劍華、謝百亮（2015）。以徑路分析檢證五專護理科學生家庭社經地位、家庭資本、生活滿意度對其學業成就之影響。**崇仁學報**，**8**，27-46。
- 蔡政中（2015）。十二年國民基本教育學業成就與教育分流之探討。**學校行政**，**99**，226-248。
- 蔡順良（2008）。青少年多向度自我效能量表編制與驗證。**教育心理學報**，**39**，105-126。

- 蔡銘津 (2014)。教師期望形成學生自我應驗的預言之虛實探討。
樹德人文社會電子學報，7 (1)，22-31。
- 鄭英傑 (2017)。學做工還是怕做工？臺灣社會高學業成就勞動階級學生及其家長的反再製心態之分析。**教育研究集刊**，63 (4)，65-100。 doi:10.3966/102887082017126304002
- 鄭耀男、陳怡靖 (2000)。臺灣地區家庭背景對就讀公私立學校與受教育年數的影響：並檢證文化資本論、財務資本論、社會資本論之適用性。**國民教育研究學報**，6，103-140。
- 蕭佳純 (2015)。教師人格特質、復原力與創意教學關聯之研究：多層次調節式中介效果的探討。**特殊教育研究學刊**，40 (1)，85-115。 doi:10.6172/BSE.201503.4001004
- 駱明慶 (2001)。教育成就的省籍與性別差異。**經濟論文叢刊**，29 (2)，117-152。
- 駱明慶 (2002)。誰是台大學生？－性別、省籍與城鄉差異。**經濟論文叢刊**，30 (1)，113-147。
- 駱明慶 (2004)。升學機會與家庭背景。**經濟論文叢刊**，32 (4)，417-445。
- 駱明慶 (2016)。誰是台大學生？(2001-2014)－多元入學的影響。**經濟論文叢刊**，即將出版。 取自
http://www.econ.ntu.edu.tw/ter/new/index.html?n=3&m=9&l=_c。
- 謝亞恆 (2004)。族群、家庭背景與國中學業成就之研究 (未出版之碩士論文)。私立南華大學教育社會學研究所，嘉義縣。
- 謝孟穎 (2003)。家長社經背景與學生學業成就關聯性之研究。**教育研究集刊**，49 (2)，255-287。

- 簡美玲 (2006)。親子教育期望與子女學業成就、身心症狀之關聯 (未出版之碩士論文)。天主教輔仁大學兒童與家庭學系碩士班，新北市。
- 簡晉龍、任宗浩、張淑婷 (2008)。跨學科間自我概念與學業成就路徑模式之檢驗－整合模式在數學和科學領域的適用性。教育心理學報，40 (1)，107-126。
- 羅淑苑、黃毅志 (2016)。重探臺東補習教育階層化與效益的特性。教育研究學報，50 (2)，27-46。
- 龔心怡、林素卿、張馨文 (2009)。家長社經地位與數學學習動機對數學學業成就之研究－以國中基本學力測驗數學領域為例。彰化師大教育學報，15，121-142。
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1986). *Social foundation of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1989). Perceived self-efficacy in the exercise of personal agency. The British Psychological Society. *The Psychologist: Bulletin of the British Psychological Society*, 10, 411-424.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Psychology*, 51(6), 1173-1182. doi: 10.1037/0022-3514.51.6.1173

- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really?. *Educational Psychology Review*, 15(1), 1-40.
- Bourdieu, P. (1986). The forms of capital (R. Nice, Trans.). In L. C. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education*, 241-258. New York, NY: Greenwood Press.
- Bourdieu, P., & Passeron, J. (1977). *Reproduction in education, society, and culture*. London, UK: Sage.
- Bourdieu, P., & Wacquant, L. (Eds.). (1992). *An invitation to reflective sociology*. Chicago, IL: University of Chicago Press.
- Breanne, E. F., & Negel, M. L. (2013). "Where there is a will, there is an A": Examining the roles of self-efficacy and self-concept in college students' current educational attainment and career planning. *Social Psychology of Education*, 16(4), 635-650.
- Catalina, C. C., Stanescu, D. F., & Mohorea, L. (2012). Academic self-efficacy, emotional intelligence and academic achievement of romanian students. Results from an exploratory study. *Journal of Educational Science and Psychology*. II(1), 41-51.
- Coleman, J. S. (1966). *Equality of educational opportunity*. National Center for Educational Statistics, Washington, DC.
- Coleman, J. S. (1987). Families and schools. *Educational Researcher*, 16(6), 32-38.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95-120.

- Coleman, P. K., & Karraker, K. H. (1997). Self-efficacy and parenting quality: Findings and future applications. *Developmental Review, 18*, 47-85.
- Duncan, O. D. (1961). *A socio-economic index for all occupation and properties and characteristics of the socioeconomic index*. Occupation and Social Status. Glencoe, Scotland: Free Press.
- Fischer, F., Schult, J., & Hell, B. (2013). Sex differences in secondary school success: Why female studenta perform better. *European Journal of Psychology of Education, 28*, 529-543. doi: 10.1007/s10212-012-0127-4
- Galyon, C. E., Blondin, C. A., Yaw, J.S., Nalls, M. L., & Williams, R. L. (2012). The relationship of academic self-efficacy to class participation and exam performance. *Social Psychology of Education, 15*, 233-249. doi: 10.1007/s11218-011- 9175-x
- Ganzeboom, H. B., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research, 25*, 201-239.
- Gecas, V., & Schwalbe, M.I. (1983). Beyond the looking-glass self: Social structure and efficacy-baassed self-esteem. *Social Psychology Quarterly, 46*(2), 77-89.
- Harman, H. (1967). *Modern factor analysis*. Chicago, IL: University of Chicago Press.

- Hollingshed, A. B. (1958). *Social class and mental illness: A community study*. New York, NY: John Wiley and Sons.
- Howie, S. J., & Tjeerd, P. (2003, February). *Language proficiency and contextual factors influencing secondary students' performance in mathematics in south africa*. Paper presented at the American Education Research Association Annual Meeting, Chicago, IL.
- Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, 28, 1-35. doi: 10.1007/s10212-011-0097-y
- Katsillis, J., & Rubinson, R. (1990). Cultural capital, student achievement, and educational reproduction: The case of Greece. *American Sociological Review*, 55(2), 270-279.
- Kiamanesh, A. R. (2005). The role of students' characteristics and family background in Iranian students' mathematics achievement. *Prospects*, 35(2), 162-174.
- Lin. C. H. (2003). Intergenerational parallelism of self-efficacy: Moderating variables, mediating variables, and common antecedents. *Dissertation Abstracts International*, 65(7-A), 2783.
- Mackinnon, D. P., Lockwood, C. M., Hoffman, J. M, West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83-104.
- Marsh, H. W., & Parker, J. W. (1984). Determinants of students self-concept: is it better be a relatively large fish in a small pond

- even if you don't learn to swim as well?. *Journal of Personality and Social Psychology*, 47, 213-231.
- Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: its multifaceted, hierarchical structure. *Educational Psychologist*, 20, 108-125.
- Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, 79(3), 1129-1167.
- Sewell, W. H., & Hauser, R. M. (1972). *Causes and consequences of higher education: Models of the status attainment process*. Madison, WI: University of Wisconsin Press.
- Shavelson, R. J. (1976). Validation of construct interpretations. *Review of Educational Research*, 46, 407-441.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417-453.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology*, 13, 290-312.
- Song, I. S., & Hattie, J. A. (1984). Home environment, self-concept and academic achievement: A causal modeling approach. *Journal of Educational psychology*, 76, 1269-1281.
- Treiman, D. J. (1977). *Occupational prestige in comparative perspective*. New York, NY: Academic Press.

- Weiser, D. A., & Riggio, H. R. (2010). Family background and academic achievement: does self-efficacy mediate outcomes?. *Social Psychology of Education*, 13, 367-383. doi: 10.1007/s11218-010-9115-1
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461-481.
- Willis, P. (1977). *Learning to labor: How working class kids get working class jobs*. New York, NY: Columbia University Press.
- Young, J. J., Hyuck, S., Sunyoung, J., & You, K. L. (2012). The effects of academic self- efficacy, learning strategies, and perceived instructional strategies on high and low achievers' in the middle school Korean language. *Journal of Educational Policy*, 9(2), 239-257.
- Zhang, J. X., & Schwarzer, R. (1995). Measuring optimistic self-beliefs: A Chinese adaptation of the general self-efficacy Scale. *Psychologia*, 38(3), 174-181.

附錄一

同學你好：

非常感謝你協助進行這份問卷的填寫，問卷內容主要在探討「家庭的社會經濟地位對子女學習效能及學業成就的相關研究」。本研究結果將做為未來改進學校教學與效能之參考。你的用心填答，對於研究結果將產生關鍵與寶貴的貢獻。填答的所有內容，僅供學術研究之用，敬請安心回答，再次感謝你的配合，謝謝！

班級：_____ 座號：_____

一、家庭基本資料

	父 親	母 親
教育程度	<input type="checkbox"/> 博士畢業 <input type="checkbox"/> 碩士畢業 <input type="checkbox"/> 大學、科技大學或二技畢業 <input type="checkbox"/> 二專或五專畢業 <input type="checkbox"/> 高中或高職畢業 <input type="checkbox"/> 國中畢業 <input type="checkbox"/> 國小畢業 <input type="checkbox"/> 不識字	<input type="checkbox"/> 博士畢業 <input type="checkbox"/> 碩士畢業 <input type="checkbox"/> 大學、科技大學或二技畢業 <input type="checkbox"/> 二專或五專畢業 <input type="checkbox"/> 高中或高職畢業 <input type="checkbox"/> 國中畢業 <input type="checkbox"/> 國小畢業 <input type="checkbox"/> 不識字
職業	現在或退休前的工作情形 <input type="checkbox"/> 民意代表、行政、企業主管及經理人員 <input type="checkbox"/> 高層專業人員（如大專教師、醫師、律師…等） <input type="checkbox"/> 中小學、特教、幼稚園老師 <input type="checkbox"/> 一般專業人員（一般工程師、藥劑師、記者、護士…等）	現在或退休前的工作情形 <input type="checkbox"/> 民意代表、行政、企業主管及經理人員 <input type="checkbox"/> 高層專業人員（如大專教師、醫師、律師…等） <input type="checkbox"/> 中小學、特教、幼稚園老師 <input type="checkbox"/> 一般專業人員（一般工程師、藥劑師、記者、護士…等）

	<input type="checkbox"/> 技術員及助理專業人員（工程技術員、代書、藥劑生、推銷保險…等） <input type="checkbox"/> 事務工作人員（文書、打字、櫃臺、簿記、出納…等） <input type="checkbox"/> 服務及買賣工作人員（商人、廚師、理容、服務生、保母、警衛、售貨…等） <input type="checkbox"/> 農、林、魚、牧工作人員 <input type="checkbox"/> 技術工（泥水匠、麵包師父、裁縫、板金、修理電器…等） <input type="checkbox"/> 機械設備操作工及裝配工（含司機） <input type="checkbox"/> 非技術工（工友、門房、洗菜、簡單裝配、體力工…等） <input type="checkbox"/> 職業軍人（軍官、士官兵） <input type="checkbox"/> 家管 <input type="checkbox"/> 失業/待業 <input type="checkbox"/> 其他_____	<input type="checkbox"/> 技術員及助理專業人員（工程技術員、代書、藥劑生、推銷保險…等） <input type="checkbox"/> 事務工作人員（文書、打字、櫃臺、簿記、出納…等） <input type="checkbox"/> 服務及買賣工作人員（商人、廚師、理容、服務生、保母、警衛、售貨…等） <input type="checkbox"/> 農、林、魚、牧工作人員 <input type="checkbox"/> 技術工（泥水匠、麵包師父、裁縫、板金、修理電器…等） <input type="checkbox"/> 機械設備操作工及裝配工（含司機） <input type="checkbox"/> 非技術工（工友、門房、洗菜、簡單裝配、體力工…等） <input type="checkbox"/> 職業軍人（軍官、士官兵） <input type="checkbox"/> 家管 <input type="checkbox"/> 失業/待業 <input type="checkbox"/> 其他_____
每月收入	<input type="checkbox"/> 十萬元以上 <input type="checkbox"/> 八~十萬元 <input type="checkbox"/> 五~八萬元 <input type="checkbox"/> 三~五萬元 <input type="checkbox"/> 三萬元以下	<input type="checkbox"/> 十萬元以上 <input type="checkbox"/> 八~十萬元 <input type="checkbox"/> 五~八萬元 <input type="checkbox"/> 三~五萬元 <input type="checkbox"/> 三萬元以下

- 17.在測驗/考試前，有充足的準備
- 18.在升學時，諮詢朋友的意見
- 19.取得有關獎學金和經濟資助的資料
- 20.考慮自己的行為所帶來的後果

本問卷到此結束，感謝您的填答

附錄二

表 1

家庭背景問卷編碼表

	題項	編碼方式
性別	男性	1
	女性	0
家庭結構	單親	1
	雙親	2
教育程度	博士畢業	22
	碩士畢業	18
	大學、科技大學或二技畢業	16
	二專或五專畢業	14
	高中或高職畢業	12
	國中畢業	9
	國小畢業	6
	不識字	0
每月收入	十萬元以上	5
	八～十萬元	4
	五～八萬元	3
	三～五萬元	2
	三萬元以下	1

表 2

學業發展自我效能問卷編碼表

學業發展自我效能	
題項	編碼方式
非常沒有信心	1
沒有信心	2
略沒有信心	3
略有信心	4
有信心	5
非常有信心	6

表 3

國中教育會考績分換算表

科目	三等級 七標示	計分方式
國文 英文 數學 社會 自然	A++	7
	A+	6
	A	5
	B++	4
	B+	3
	B	2
作文	C	1
	六級分	1
	五級分	0.8
	四級分	0.6
	三級分	0.4
	兩級分	0.2
	一級分	0.1

表 4

家庭社經地位的分類等級

家庭社經地位指數	編碼方式	分層	對應社會階級
20 分以上	4	第四等級	上層階級
15~19 分	3	第三等級	上層中產階級
10~15 分	2	第二等級	下層中產階級
未滿 10 分	1	第二等級	下層階級

表 5

學業發展自我效能的分類等級

學業發展自我效能	編碼方式	等級名稱
80 分以上	3	高效能組
40~80 分	2	中效能組
未滿 40 分	1	低效能組

教育學誌 第四十期

2018年11月，頁121~180

測驗向度數評估方法的比較

楊彥文

敏惠醫護管理專科學校助理教授

涂柏原*

國立臺南大學教育學系副教授

摘要

本研究旨在模擬具有簡單結構以及部分複雜結構的多向度資料，透過操弄向度個數（1、2、3）、各向度能力值之間的相關（0、0.3、0.6）、各向度題數（10、20題）、以及樣本人數（250、500、1000、2000人）等條件，每個條件組合下模擬100次，以比較DETECT（Kim, 1994）、NOHARM（McDonald, 1996）、平行分析（PA; Horn, 1965）和HULL法（Ceulemans & Kiers, 2006）等四種測驗向度評估方法之表現，以各種方法正確辨認向度個數的百分比，作為比較各方法表現之標準。

研究主要發現如下：（1）模擬資料分析結果顯示，當資料是單一向度時，DETECT無法順利得到一個單向度的解；資料是二向度時，因資料結構與DETECT的理論不是十分一致，因此DETECT的表現不佳，其他三種方法皆優於DETECT；當資料是簡單結構的三向度時，DETECT的表現比在二向度的情形好。向度之間的相關

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

小於 0.3 時，PA 的表現最佳，向度之間的相關為 0.6 時，DETECT 和 NOHARM 比 PA 及 HULL 法好。如果資料是 3d1 的（即資料雖然用三向度的 MIRT 模式產生，但因為試題箭頭指向同一個方向，實際上可說是單一向的），則各種方法的表現情形與單一向度時一樣。（2）在各向度能力值之間的相關為 0.6 以下，PA 和 HULL 的表現優於 DETECT 和 NOHARM，但是當各向度能力值之間的相關變大時，DETECT 和 NOHARM 的表現優於 PA 和 HULL。（3）當試題數增加時，四種方法的正確模式判斷率也會隨之提升。（4）NOHARM、PA 和 HULL 三種方法似乎較不受樣本大小改變而影響，對 DETECT 而言，當樣本大小增加時，正確模式判斷率也會隨之增加。

關鍵字：向度數、DETECT、NOHARM、平行分析、HULL

A Performance Comparison of Test Dimensionality Assessment Methods

Yen-wen Yang

Assistant Professor,

Min-Hwei Junior College of Health Care Management

Bor-yaun Twu*

Associate Professor,

Department of Education, National University of Tainan

Abstract

The purpose of this study was to investigate the performance of four dimensionality assessment procedures, namely DETECT (Kim, 1994), NOHARM (McDonald, 1996), Parallel Analysis (Horn, 1965) and HULL method (Ceulemans & Kiers, 2006), in terms of their accuracy of identifying the numbers of dimensions given by different multidimensional data sets. With the manipulation of the number of dimensions, the correlation among the dimensions, the number of items per dimension, and the sample size, simulated responses were generated under different conditions, for each of the 100 replications per condition.

The main findings were as the following: Firstly, when the data is unidimensional, DETECT is not able to obtain the proper one-factor solution. For the two-dimensional case, DETECT's performance is not

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

good as expected due to the fact that some of the item response data was not generated to be within-item multidimensional. DETECT performs better in the three-dimensional case than in the two-dimensional case, because the item response data was generated to be between-item multidimensional one. Parallel analysis performs better than other methods when the correlation between domain abilities is less than .3, and DETECT and NOHARM outperforms Parallel analysis and HULL when the correlation becomes .6. For the so-called 3d1 data, in which the item response was generated using a M2PL model, but all items point to the same direction in the latent space, all methods give similar results as was the unidimensional case. Secondly, the PA and HULL outperformed DETECT and NOHARM when the correlation among dimensions was 0.3 or lower, and the DETECT and NOHARM outperformed PA and HULL when the correlation was 0.6 or higher. Thirdly, as the number of items increased, the accuracy of identifying the numbers of dimensions was also increased for all procedures. Finally, sample size seem did not affect NOHARM, PA and HULL, but when sample size increases, the performance of the DETECT procedure improves.

Keywords: dimensionality, DETECT, NOHARM, Parallel Analysis, HULL

壹、緒論

一、向度的意義與基本概念

許多研究者認為向度評估和理解測驗的結構是測驗發展中很重要的步驟 (Tate, 2003)，如果要根據測驗分數對受試者之表現做出有意義的推論，評估資料的向度性是必要的 (Stout, 1987; Stout, Habing, Douglas, Kim, Roussos, & Zhang, 1996)。Stone 和 Yeh (2006) 做了一個很好的結論，測驗結構的探討，讓我們“確認哪些領域或向度被測量，辨識這些領域之間的關係，以提供所假設的多向度性和測驗分數解釋的支持，並確定與所測量的構念無關的變異” (p. 211)。檢驗構念之間的關係，使我們能夠找到與所預期之構念結構一致的證據，並控制無關的構念，如果我們要保持測量的一致性和跨測驗分數的解釋，這兩者都是必要的。

向度性或向度數 (dimensionality) 一般被認為是可以充分說明受試者在測驗上答題表現所需的潛在特質之最少個數。傳統上，如果 $\theta = (\theta_1, \theta_1, \dots, \theta_m)$ ，受試者對各個試題反應組型之機率滿足下式時，則局部獨立被滿足 (Stout, 1990)：

$$P(X_1, X_2, \dots, X_J | \theta) = \prod_{j=1}^J P(X_j | \theta) \quad (1)$$

其中 X_1, X_2, \dots, X_J 分別為是考生在第 1、2、 \dots 以及第 J 題的分數， J 是測驗的題數；此時這筆測驗資料的向度數為 m 。

在實務上，局部獨立的要求是不太容易達成的，因此，McDonald (1981) 提出一個比公式 (1) 弱的假定來發展模式的估計程序，稱為弱局部獨立 (weak local independence, WLI)，指的是所有兩兩

配對試題之條件共變數的值皆為 0，而不是針對所有試題的聯合分配。對全部的 θ 向量而言，WLI 要求

$$E(\text{cov}(X_i, X_j) | \theta) = 0 \quad (2)$$

也就是只要任何試題 i 和試題 j 的給定 θ 後之共變數的期望值等於 0，在實務上只要非常接近於 0 即可，即代表局部獨立可被滿足。

Ackerman、Gierl 和 Walker (2003) 指出許多教育和心理測驗本身就是多向度的，意謂著這些測驗測量二個或更多個向度，而多向度的測驗主要可以分為兩大類型 (Adams, Wilson, & Wang, 1997)。第一種類型是每個試題只測量一種能力 (單向度試題)，而整份測驗中包含許多單向度試題子集，每一個子集中的試題都測量相同的能力，但不同子集的試題測量不同的能力，此種測驗稱為題間多向度的測驗 (between-item multidimensional test)，或是測驗具有簡單結構 (simple structure)，如圖 1 (a) 所示。這種型式最常見於心理測驗，如性向測驗、人格量表、智力測驗等都屬於這種類型，以新編多元性向測驗量表 (陳榮華、吳明雄、陳心怡, 2010) 為例，其測量的是學生在多方面之潛能，包括語文推理、數字推理、圖形推理...等八個分測驗，其結果不是一個總分，而是八個性向分數所構成的分數剖面圖 (profile)，除了可瞭解學生間之差異，亦能夠進行個人內在差異之比較。第二種類型是每一個試題可以測量到一種以上的能力，即試題作答反應同時由多個能力向度決定，此種題型所組成的測驗被稱為題內多向度的測驗 (within-item multidimensional test)，或是測驗具有複雜結構 (complex structure)。例如，數學成就測驗中經常包含計算題與應用問題兩種題型，其中計算題是屬於單向度試題，純粹是測量受試者的四則

運算能力；而應用問題則屬於二向度試題，受試者除了需要具備計算能力外，還需能夠理解問題情境，才能將答案算出來，這種多向度類型如圖 1 (b) 所示，單向度 IRT 模式不適合用在這種測驗上。如果忽略試題多向度的性質，勉強使用單向度 IRT 進行試題分析，會因為忽略了其中某幾項能力而產生偏差的結果，使測驗分數的意義變得模糊不清 (Ackerman, 1992; Reckase, 1985)。

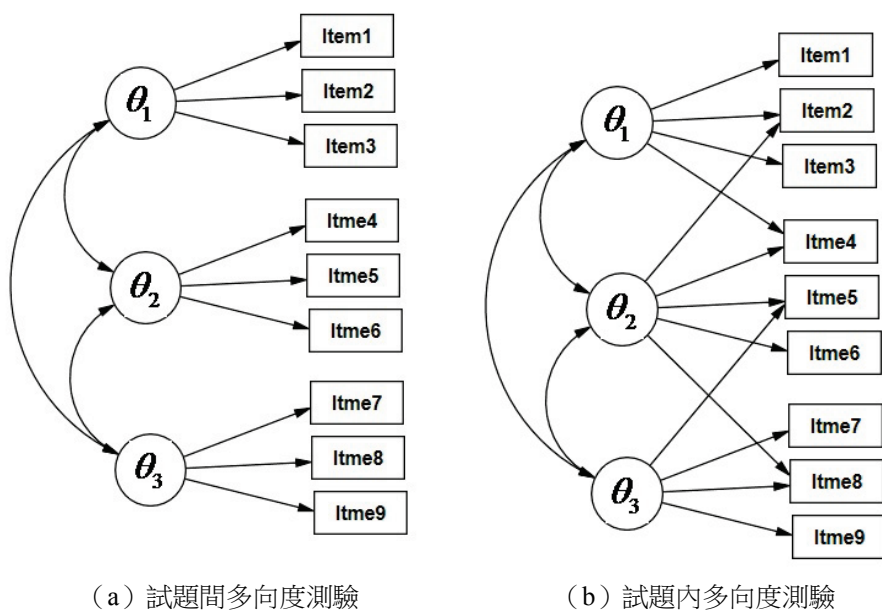


圖 1 兩種類型的多向度測驗模式

二、評估測驗向度的方法

對於 1980 年代初期以前被廣泛使用的各種評估測驗向度的方法，Hattie (1985) 曾進行了一個完整的回顧，根據所得到的結果，他認為在那個時期只有利用 McDonald 的非線性因素分析

NOHARM 模式所得到的結果是比較值得信賴的。而在 1980 年代初期，Stout 也開始從事有關向度個數檢定方法的研究，利用本質的獨立性（essential independence）以及本質的向度性（essential dimensionality）之概念，Stout（1987）提出了一個檢定偏離單一向度性的 T 統計考驗數，並將此方法發展為 DIMTEST 程式（Stout, 1987; Stout, Douglas, Junker, & Roussos, 1999）。隨後，Stout 與他的學生們又發展出 HCA/CCPROX 和 DETECT 等用來檢測資料向度數的方法與軟體（Stout, et al., 1996）。這些以條件的共變數為基礎非參數多向度評量的方法，自從被提出來之後，在測驗界中就一直深受重視。

針對這些新近的發展趨勢，Tate（2003）進行一個詳細的研究，探討這些在 Hattie 的研究發表之後被提出的方法，並提出一個向度評估方法的分類架構。Svetina 和 Levy（2014）也提出一個資料向度性評估方法分類的架構，以方便對向度評估議題感到興趣的研究者，可以根據該架構選取合適的分析方法。Svetina 和 Levy（2012）亦針對目前可用來進行向度性評估的軟體包括 Mplus、TESTFACT、NOHARM、DETECT 和 DIMTEST 等程式，提供一個簡介。這些研究說明了目前心理計量領域中的學者，對於向度性評估這個議題仍然充滿了興趣。

Muthén（1983）曾提到類別資料的因素分析（FA）與多向度試題反應理論（MIRT）模式的數學公式是等同的，Takane 和 de Leeuw（1987）提供一個 MIRT 與 FA 模式等同之正式證明，Knol 和 Berger（1991）也利用 IRT 和 FA 模式，對模擬的資料進行分析，發現以四分相關（tetrachoric correlation）矩陣進行共同因素分析的結果不會比用理論上似乎更合適的 MIRT 模式來得差。因此，探索性因素

分析 (EFA) 中常用的因數個數決定之準則，應當也適合用來決定測驗資料的向度個數。

在 EFA 的研究中，用來決定因素個數最常見的方法有四種：Cattell 的陡坡圖 (scree plot) 檢定、Horn 的平行分析 (parallel analysis, PA)、Kaiser 的特徵值大於 1 以及 Velicer 的最小平均淨相關 (minimum average partial, MAP) 檢定等。除了這些方法以外，Ceulemans 和 Kiers (2006) 新近提出一個 HULL 法，用來決定因素個數，此方法可視為是一般化的陡坡圖檢定；其主要目的是在模式適配和自由度之間取得最佳平衡，並結合統計方法來確保在因素分析的過程中，不會過度抽取因素個數。Lorenzo-Seva、Timmerman 和 Kiers (2011) 曾將 HULL 法與其他數種決定因素的方法進行比較，他們發現 HULL 法在決定主要因素個數上是表現較佳的方法。

針對目前常用向度檢測的方法中，筆者對於 DETECT、NOHARM 以及 EFA 中被認為在決定因素個數方面表現較佳的 PA 和 HULL 法感到興趣，擬針對這四種方法進行比較研究，因此以下先對這四種方法進行簡要介紹。

(一) DETECT 程序

DETECT 程序最早是由 Kim (1994) 所發展出來的，其理論基礎以及所用統計數之表現行為，則是由 Zhang 和 Stout (1999a, 1999b) 補足的。它是一個在近似簡單結構的假定之下，決定用來表徵試題分數矩陣中試題間關係向度個數之統計程序，DETECT 程序並未假定試題反應函數之型式，因此被稱為是一個非參數的程序。當研究者想知道用一個特殊的模式來將資料矩陣加以模式化所需要座標向度的個數時，DETECT 常被用來滿足此目的。

DETECT 雖然並未假定試題反應函數之型式，但對試題反應函數有數個假定，包括（1）正確反應的機率必須隨著任何一個 θ 座標值的增加而單調增加；（2）局部獨立性必須成立。DETECT 並不需要完全的局部獨立，它僅要求「成對的局部獨立」被滿足：給定 Θ 之後，第 i 題和第 J 題得分之共變數等於 0，也就是 $\text{cov}(U_i, U_j | \Theta = \theta) = 0$ ，相當於 McDonald (1996) 所謂的弱局部獨立。如果是使用一個單調遞增的 MIRT 模式能產生成對的局部獨立所需要的潛在變項之最小個數，則試題分數矩陣需要一個向度的模式來精確地表徵矩陣之間的關係，這個就是 DETECT 程序在發展過程中所使用的向度性基本定義。

基本上，DETECT 是在多向度空間中，尋找在一個與整體測驗最佳測量方向不同的某一個方向上，測驗中具有最佳測量的同質性試題群組。DETECT 的測驗最佳測量方向與 Wang (1985) 之參照組合 (reference composite) 概念是相似的，但其數學定義是不相同的 (Reckase, 2009)。主要的區別在於參照組合是應用一個單一向度 IRT 模式分析試題分數矩陣時所得到的能力向度，該能力向度實際上是多向度空間的各個座標值之加權組合方向的估計值；相反的，在 DETECT 中所使用的最佳測量方向是試題估計得到的真分數具有最大改變量 θ 點方向的平均 (Reckase, 2009)。有關這個最佳測量方向是如何得到之詳細推導過

程，請見 Zhang 和 Stout (1999a)，底下有關 DETECT 的描述摘要自 Reckase (2009) 的書。

DETECT 指數之基本概念是完全潛在空間 (complete latent space) 之座標的加權組合：

$$\theta_w = \mathbf{w}\boldsymbol{\theta}' = \sum_{k=1}^m w_k \theta_k \quad (6)$$

其中 \mathbf{w} 是在空間中 m 個座標軸之權重向量， θ_w 是座標值的組合分數，這些權重被重新量度化過，因此 θ_w 的變異數等於 1.0。

DETECT 程序決定了在 $\boldsymbol{\theta}$ 空間中組合分數所界定的方向上，將訊息函數平方根之期望值加以最大化的向量 \mathbf{w} 元素的值，如果 \mathbf{w}^* 是提供了最大的期望值之權重向量，則它的元素可以利用底下的公式估計得到：

$$w_\ell^* = c \sum_{i=1}^k E \left\{ \frac{\partial P_i(\boldsymbol{\theta})}{\partial \theta_\ell} \left[\sum_{i=1}^k P_i(\boldsymbol{\theta}) Q_i(\boldsymbol{\theta}) \right]^{\frac{1}{2}} \right\}, \ell = 1, 2, \dots, m \quad (7)$$

其中 w_ℓ^* 是 \mathbf{w}^* 的一個元素， c 是量度用的常數，以使 θ_w 的變異數為 1， k 是測驗的題數。 $P_i(\boldsymbol{\theta})$ 是試題 i 的試題反應函數， m 是模式化資料矩陣中的關係所需要的向度個數。

在執行 DETECT 程式時，一開始所有的試題得先被分割成一系列互斥的幾個子集 $\mathbf{P} = \{A_1, A_2, \dots, A_q\}$ ，分割目標在使各子集的試題測量不同的構念。這些試題子集的形成可由研究者判斷決定，或是由 DETECT 程式自動找出來。DETECT 提供了一個選項可將讓「DETECT 指標」值最大

化的試題組合列印出來。DETECT 指標 $D_w(\mathbf{P})$ 是由下式所計算得到

$$D_w(\mathbf{P}) = \frac{2}{k(k-1)} \sum_{1 \leq i < j \leq k} \delta_{ij}(\mathbf{P}) E[\text{Cov}(U_i, U_j | \theta_w)] \quad (8)$$

其中 \mathbf{P} 是測驗中試題的任何分割組合， k 是測驗中的題數， i 和 j 是任意兩題，而 U_i 和 U_j 是其得分， θ_w 是所感到興趣的組合分數，且

$$\delta_{ij}(\mathbf{P}) = \begin{cases} 1 & i, j \in A_\ell \\ -1 & \text{其他} \end{cases} \quad (9)$$

如果兩個試題是在同一個試題分割區 (partition) 中， δ 值為 1，如果它們是在不同的分割區中， δ 值則為 -1。當組合分數是依據最佳的權重得到時，這個指標記為 $D(\mathbf{P})$ ，為分割區 \mathbf{P} 「理論的 DETECT 指標」，而將此指標的值加以最大化的分割區記為 \mathbf{P}^* 。在軟體中實際上是將公式 (8) 所計算得到的值乘上 100 (Roussos & Ozbek, 2006)，以使 DETECT 指標的值在 0~5 之間，好方便解釋及應用。Kim (1994) 以及 Jang 和 Roussos (2007) 認為 DETECT 指標的值為 1 或更大時，代表試題反應資料有大的多向度性，指標的值在 0.4~1 之間則具有中度至大的多向度性；低於 0.4 的值被認為具有中度至弱的多向度性；若是 DETECT 指標的值低於 0.2，基本上是單一向度的。

除了 DETECT 指標以外，這個軟體另外提供了兩個資料向度性的指標。第一個是以 DETECT 指標的複核效率為

基礎的，資料矩陣被隨機分成兩個樣本，將公式（8）的 $D_w(\mathbf{P})$ 值最大化的分割區是由第一個樣本來決定的，該樣本所得到的 $D(\mathbf{P}^*)$ 值稱為 D_{\max} 。最大化第二個樣本之 $D(\mathbf{P}^*)$ 值的分割區也被決定，然後根據第二個樣本所得到的試題分割區，使用第一個樣本的資料來計算 $D(\mathbf{P})$ ，這個值被稱為「參照的 DETECT 指標」，以 D_{ref} 表示，這兩個值的比值 $r_{\max} = D_{\text{ref}} / D_{\max}$ 被視為是另外一個向度性的指標。如果 r_{\max} 的值接近於 1，表示由兩個樣本所發現的分割區是相同的，代表 DETECT 分析的結果是穩定的；如果 r_{\max} 的值是小的，代表二者所得到的分割區是非常的不同。Zhang 和 Stout（1999b）認為小的 r_{\max} 值代表試題分割區的得到是「因為機遇造成」的，試題反應資料可能為本質地單一向度的（essentially unidimensional）。

根據 DETECT 形成試題分割區的作法，屬於相同分割區的試題彼此間之共變數應當是正的，而不同分割區之試題彼此間的共變數是負的，因此共變數矩陣中正、負號符合期望組型之比率，可以作為另一個向度性的指標，此項指標稱為 IDN 指數。如果資料的多向度結構與一個簡單結構模式一致的話，我們所期待的共變數組型與實際觀察得到的組型是一致的，則此指數的值接近於 1。

Zhang 和 Stout（1999b）也透過兩個模擬研究展示 DETECT 的表現，他們發現當近似簡單結構成立時，在辨識主要潛在向度和估計測驗多向度性是否存在時，

DETECT 的表現相當好；即使當近似簡單結構未能成立時，他們認為 DETECT 的分析結果仍然是有幫助的，因為它仍然可以找到向度相對同質之試題群組。當近似簡單結構的假設成立時，DETECT 程序的報表會列印出非重疊群組的個數以及與組成每個群組的試題，因此，同質試題群組的個數可以視為主要向度的個數。

(二) NOHARM 程序

NOHARM (Fraser, 1988) 這個電腦程式使用非線性因素分析 (nonlinear factor analysis, NLFA) 的方法，將多向度常態肩形的補償性 IRT 模式適配到試題分數矩陣，McDonald (1996) 將這個分析方法稱為 NOHARM 模式。這個程式雖然是以一個常態肩形模式來適配試題反應矩陣，實際上它不是直接估計常態肩形模式的參數，而是以一個多項式來近似常態肩形 MIRT 補償性模式。McDonald (1996) 認為在大部份的應用中，一個三次方的多項式即具有充分的精確度。當估計模式的試題參數時，NOHARM 只使用答對每一個試題的機率 $\hat{\pi}_i$ 以及同時答對第 i 題和第 k 題之機率 $\hat{\pi}_{ik}$ 等二種訊息，試題間更高階的交互作用資訊（例如，同時答對三個試題的機率）並未被用到，使用試題分數矩陣的這兩個特色大致相當於使用試題分數的平均數以及試題配對之間的 phi 相關 (Reckase, 2009)。

以矩陣形式來表示，常態肩形 MIRT 模式之多項式近似公式為 (Reckase, 2009)

$$P(u_{ij} = 1 | \boldsymbol{\theta}_j) = \sum_{p=0}^{\infty} \gamma_{ip} h_p \left[\frac{\mathbf{a}_i \boldsymbol{\theta}'_j}{\sqrt{\mathbf{a}_i \mathbf{P} \mathbf{a}'_i}} \right] \quad (10)$$

其中 $u_{ij} = 1$ 代表第 j 個考生答對第 i 題， h_p 是常態化 Hermite - Tchebycheff 多項式的 p 次方項， γ_{ip} 是試題 i 多項式第 p 個項的係數，而 \mathbf{P} 是 $\boldsymbol{\theta}$ 的共變數矩陣。在實務上，只需要前四項（即 $p = 0, 1, 2, 3$ ），試題參數即可被精準地估計能力向度間之共變數矩陣 \mathbf{P} 出現在公式（10）之中，以允許 $\boldsymbol{\theta}$ 各個元素之間有相關存在，如果各能力向度彼此是獨立無關的，則 \mathbf{P} 是一個單元矩陣。

令 $z_j = \mathbf{a}_i \boldsymbol{\theta}'_j / \sqrt{\mathbf{a}_i \mathbf{P} \mathbf{a}'_i}$ ，則公式（10）中 $h_p(z_j)$ 前四個項的形式為 $h_0(z_j) = 1$ 、 $h_1(z_j) = z_j$ 、 $h_2(z_j) = (z_j^2 - 1) / \sqrt{2}$ 、和 $h_3(z_j) = (z_j^3 - 3z_j) / \sqrt{6}$ （Reckase, 2009）。試題 i 答對率之估計值等於公式（10）系列之第一項，也就是

$$\hat{\pi}_i = \gamma_{i0} h_0(z_j) = \gamma_{i0} = \Phi \left[\frac{d_i}{\sqrt{1 + \mathbf{a}_i \mathbf{P} \mathbf{a}'_i}} \right] \quad (11)$$

同時答對試題 i 和 k 這個配對的機率之估計值為

$$\hat{\pi}_{ik} = \sum_{p=0}^{\infty} \gamma_{ip} \gamma_{kp} \left[\frac{\mathbf{a}_i \mathbf{P} \mathbf{a}'_k}{\sqrt{\mathbf{a}_i \mathbf{P} \mathbf{a}'_i \mathbf{a}_k \mathbf{P} \mathbf{a}'_k}} \right]^p \quad (12)$$

估計試題參數時，NOHARM 使用「未加權的最小平方法」來將公式（13）中的標準最小化，

$$SS_{ik} = \sum_{i \neq k} (p_{ik} - \hat{\pi}_{ik})^2, \text{ 對所有的 } i, k \quad (13)$$

其中 p_{ik} 是實際上同時答對第 i 題和第 k 的比率。因為最小平方方法並沒有直接可用的適合度卡方檢定，McDonald (1996) 建議使用適配 m 個向度模式之後所得到的殘差共變數矩陣，各個非主對角線元素的平方和之平均

$$f_1 = 2[n(n-1)]^{-1} \sum_{j < k} (r_{jk}^*)^2 \quad (14)$$

或是各個非主對角線元素之絕對值和的平均

$$f_2 = 2[n(n-1)]^{-1} \sum_{j < k} |r_{jk}^*| \quad (15)$$

來審視模式對資料的適配情形，其中 n 為題數， r_{jk}^* 是殘差共變數矩陣中第 (j, k) 個元素。McDonald 和 Mok (1995) 以及 McDonald (1996) 建議使用驗證性因素分析中的適合度指標來協助判斷由 NOHARM 所得到的解是否令人滿意。因為 NOHARM 採用最小平方方法原則進行參數估計，唯一可用的適合度指標是 Tanaka 指標：

$$\gamma_{ULS} = 1 - \frac{Tr(\mathbf{R}^2)}{Tr(\mathbf{C}^2)} \quad (16)$$

其中 \mathbf{R} 是殘差共變數矩陣， \mathbf{C} 是樣本的共變數矩陣， $Tr(\cdot)$ 代表矩陣的跡 (trace)，這個指標類似一個決定係數，反應出變異數為某一個模式所解釋的比率 (Tanaka & Huba, 1989)。

雖然可以利用審視殘差值或 Tanaka 指標來判斷模式的適配度，缺乏正式的統計考驗數，對使用者而言，多少仍

是不便。因此 Gessaroli 和 De Champlain (1996) 提出一個近似卡方 (approximate chi-square) 統計數，用來檢定適配 m 個向度 MIRT 補償性模式後所得到的殘差共變數矩陣非主對角線之元素之值是否等於 0：

$$\chi_{G/D}^2 = (N-3) \sum_{j=1}^n \sum_{k=1}^{j-1} Z_{jk}^{2^{(r)}} \quad (17)$$

其中 N 是受試者人數， $Z_{jk}^{2^{(r)}}$ 是第 j 題與第 k 題之殘差共變數經 Fisher 的 Z 轉換後所得到的值之平方。如果 NOHARM 分析是探索性的，當虛無假設為真時，這個檢定統計數是具有自由度為 $.5n \times (n-1) - t$ 之卡方分配，其中 $t = nm + .5m(m-1)$ ， n 為題數， m 為向度數（或因素個數）。

(三) PA

在探索性因素分析各種決定因素個數的方法中，學術界一致肯定較佳的方法之一是平行分析 (Zwick & Velicer, 1986; O'Connor, 2000)。由因素分析的文獻可發現，PA 的應用已有一段很長的歷史，最早是由 Horn (1965) 所提出的，目的在解決研究者利用 Kaiser 的特徵值大於 1 和 Cattell 的陡坡圖這兩種方法來決定因素個數時所遇見的問題。PA 最適合與主成分分析 (PCA) 一起使用，若將 PA 應用在主軸因素分析 (PAF) 時，如果樣本資料需要 m 個共同因素來解釋變項之間的關係，PA 的理論認為由樣本資料之減縮相關矩陣所得到最大 m 個特徵值應當大於由具有相同變項和相同人數的隨機資料所得到的最大 m 個特徵值。一般來

說，隨機資料被重複產生多次，然後由每一筆資料所得到的第一個特徵值被排序，以計算第 95 百分位數的值，其他各個特徵值亦是如此處理。因此由樣本所得到的特徵值大於由隨機資料所得到相對應特徵值之 95 百分位數之個數，即被視為是適合解釋樣本資料的因素個數。

利用線性因素分析來處理二元計分試題，長久以來有難度因素 (difficulty factor) 的問題，自然引起讀者懷疑利用 PA 來決定二元計分資料向度數的作法是否允當。對於這個問題，Weng 和 Cheng (2005) 在不同人數、不同的因素負荷量、不同得分的比率 (二個類別人數之比率) 以及因素之間不同相關大小的條件下，利用模擬資料探討 PA 的表現。他們發現在辨認因素個數方面，整體而言，PA 有不錯的表現；無論是 Pearson 積差相關或是四分相關矩陣被使用，PA 的表現是相近的。另外，他們的研究結果亦顯示隨著變項之因素負荷量變大、樣本人數增多以及兩個類別的人數比率接近時，PA 有較好的表現。而當 95 百分位數或 99 百分位數被用來作為決定的標準時，PA 判斷的正確率比用平均數來得高。另外，Yu、Popp、DiGangi 和 Jannasch-Pennell (2007) 比較 Rasch 模式、PA 以及 TETRAD 三種方法在評估多元計分測驗資料是否具有單一向度性的表現，結果亦發現 PA 在辨認正確的因數個數方面的表現是相當不錯的。因此將 PA 應用在測驗資料的向度個數評估，應當是可行的。

在執行 PA 程序時有一些選項需考慮，比如是用 PCA、PAF 或是其他方法來抽取因素，其中比較受歡迎的是利用 PCA 來抽取因素 (Timmerman & Lorenzo-Seva, 2011)，因此，在本研究中將採用此方法。

(四) HULL 法

Ceulemans 和 Kiers (2006) 為了解決複雜的模式選擇之問題，提出了一個以數值凸包 HULL 為基礎 (convex HULL-based) 的方法，來確定 EFA 中主要共同因素的個數。HULL 法的原理是尋找一個模式適配測量值與自由度之間的最佳平衡，再結合統計的方法，以確定不會發生抽取過多的因素。

以數值凸包 HULL 為基礎的探索性程序，可以視為是一般化的陡坡圖檢定。做法是如陡坡圖一樣，畫出適配度測量值 (f) 對上自由度 (df) 的 HULL 圖 (如圖 3 所示)。在圖中的每個點代表在不同因素個數的解所獲得的適配度測量值與自由度的組合，各點旁邊的數字代表所適配的因素個數。根據此圖，較合適的因素模式，是在 HULL 圖中接近肘處的解，亦即旁邊有標註 6 的那個點所對應的，也就是 6 個共同因素。進行 HULL 法時之主要步驟如下：

- (1) 首先選擇共同因素個數的範圍，一般是採用平行分析建議的個數加上 1 作為上限；
- (2) 評估各個因素個數的解之適配度測量值；
- (3) 計算各個因素模式中適配檢定數

之自由度；(4) 在 HULL 圖中找出適配度值與自由度最佳平衡的解。

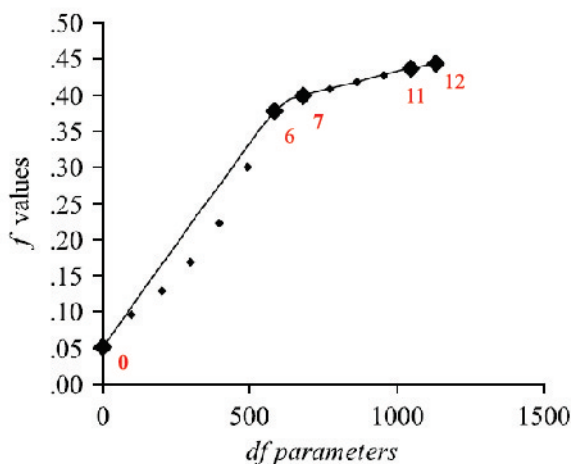


圖3 適配度值 (f) 對自由度 (df) 的 hull 圖

註：資料取自“The Hull Method for Selecting the Number of Common Factors,” by Lorenzo-Seva, U.; Timmerman, M. E.; Kiers, H. A. L. (2011). *Multivariate Behavioral Research*, 46(2), 347.

根據 Lorenzo-Seva 等人(2011)模擬研究的結果，HULL 法搭配的 CFI、CAF、RMSEA 和 SRMR 等四種適配度指標的表現都非常出色，但是 CFI 只能應用在因素是以最大概似法 (maximum likelihood, ML) 或未加權最小平方法 (unweighted least squares, ULS) 所抽取得到的情形 (稱為 HULL-CFI)，而 HULL-CAF (即根據 KMO 來進行 HULL 分析) 可以在任何因素抽取方法中使用。在實務中，Lorenzo-Seva 等人建議採用 HULL-CFI 或 HULL-CAF 來決

定因素的個數，因此，在本研究中將採用 HULL-CAF 來判斷因素個數。

三、測驗結構評估方法之相關研究

(一) NOHARM 和 DETECT 的相關研究

De Champlain 與 Gessaroli (1998) 以模擬的資料來比較 TESTFACAT 的 G^2 差異檢定 (稱為 G_{diff}^2)、LISREL8 的卡方統計量 (以 χ_{LIS8}^2 表示) 以及 NOHARM 的近似卡方 (即公式 17 的) 等三種程序的向度檢測型一錯誤率與正確拒絕虛無假設的比率 (即拒絕率)。De Champlain 與 Gessaroli 所操弄的變項包括：不同潛在特質間之相關程度 (0 和 0.7)、樣本人數 (250、500 和 1000 人)、及測驗長度 (20 和 40 題)。模擬資料用的試題參數是取自法學院入學測驗 (LSAT)，其中 a 參數的值在 1.15~0.34 之間； b 數在 -1.71~1.67 之間，而 c 參數是在 0.04~0.22 之間。LSAT 測驗包含二個向度的試題：第一個向度為演繹推理 (AR) 的試題，約佔全測驗的 25%，第二個向度是閱讀/口語推理 (LR 及 RC) 的試題，約佔全測驗的 75%。De Champlain 和 Gessaroli 在模擬研究中亦採用相同的雙向度比重來產生雙向度的試題反應資料，其中第一題至第五題只測量第一個向度，第六題至第二十題只測量第二個向度，此為試題參數結構 A。但是實徵資料中試題可能會測量一種共同能力 (比如閱讀理解)，所以他們又模擬了另一種結構的測驗，

第一題至第五題測量第一個向度，第十六題至第二十題測量第二個向度，而第六題至第十五題則同時測量兩個向度，且在兩個向度上的負荷量相同，此為試題參數結構 B。因為 NOHARM 程序中無法估計 c 參數，因此所有試題之 c 參數都固定為 0.15。

De Champlain 和 Gessaroli 的研究結果顯示，在單向度模擬資料中， χ^2_{GID} 統計量的型一錯誤率幾乎都在 0.05 以下，其中最高的為 0.07；卡方統計量 χ^2_{LIS8} 在 20 個試題 1000 個受試者的情境中，型一錯誤率膨脹至 0.68； G^2_{diff} 在所有情境中錯誤拒絕率都很高，型一錯誤率從 0.17（20 個試題 1000 個受試者）至 0.79（40 個試題 250 個受試者），這表示當資料反應實際為單向度時，採用 G^2_{diff} 可能會錯誤地拒絕虛無假設，而以為資料是多向度的。而在二向度模擬資料中，虛無假設是資料為單向度的，在 24 個模擬情境中，NOHARM 在其中 22 個的拒絕率為 100%，只有在試題參數結構 B 中有兩個情境（20 題、250 人、相關 0，及 40 題、1000 人、相關 0）的拒絕率為 99%，顯示 NOHARM 可以正確地辨別資料的多向度性。 χ^2_{LIS8} 在題數 20 人數 1000 的情境中，拒絕率為 100%，但作者並未提及在其他情境中 χ^2_{LIS8} 的表現。TESTFACT 的 G^2_{diff} 在向度間相關為 0 的 12 個情境中拒絕率皆為 100%，而在向度間相關為 0.7 時的拒絕率介於 62%（結構 B、20 題、250 個受試者）至 100%（結構 B、20 題、1000 個受試者）之間。總之，在所有情境中，

G_{diff}^2 嚴重的膨脹型一錯誤率，而 χ_{GID}^2 的結果卻接近於名義的水準 (0.05)。因此，綜合 De Champlain 和 Gessaroli 的研究結果， χ_{GID}^2 的型一錯誤率低，在二向度資料中拒絕率非常高，且不易受到樣本數、試題數、試題參數結構及潛在特質間相關的影響。雖然 NOHARM 無法估計受試者能力只能假設其為標準常態分配，而 TESTFACT 可估計每個受試者的精熟分數。但是 De Champlain 和 Tang (1993) 認為即使受試者的能力為非標準常態分配時，NOHARM 也不會受到太大的影響，所以 NOHARM 是三種方法中最適合用來作為檢定向度的方法。

Finch 和 Habing (2005) 曾探討比較使用以 NOHARM 分析結果為基礎的兩個具有近似卡方的 χ_{GID}^2 和 ALR 指標，以及 DETECT 在評估和表徵測驗資料向度的表現，其中 ALR 為近似概似比統計數 (approximate likelihood ratio statistic)。他們以某一個全州的基本技能評量和 SAT 的試題參數來模擬資料，所操弄的變項包括在潛在特質間的相關 (0、0.3、0.8 和 0.95)、題數 (15、30 和 60 題)、樣本大小 (1000 和 2000 人)、潛在特質分配的偏態 (-1.5、-0.5、0、0.5 和 1.5)、以及 c 參數的值為 0 或不為 0。藉由在簡單結構的條件下，決定潛在向度個數的準確性和試題群集的正確程度，來判斷兩個程序的表現。研究結果顯示以 NOHARM 結果為基礎的方法，在尋找向度個數的正確性與正確的試題群集與 DETECT 的表現相同。相較於

DETECT，NOHARM 通常更能夠正確判斷潛在的向度個數，但卻不太能將屬於相同群組的試題放在一起。整體而言，這兩種方法因資料的性質不同會有不同的表現；如果向度個數較少，DETECT 表現得較好，而以 NOHARM 結果為基礎的方法在向度個數較多時表現得比較好。此外，相較於 χ^2_{GID} 和 ALR 指標，DETECT 在二向度時有高估的傾向，但在六向度時有低估的現象。

Svetina (2011) 利用複雜結構的模擬資料來比較 DETECT 和 NOHARM 程序在評估測驗向度個數的表現，她操弄的因素包括：向度個數 (2 和 3 個)、資料結構類型 (試題複雜度百分比：0%、10%、30%和 50%)、向度之間的相關 (0、0.3、0.6、0.75 和 0.9)、樣本大小 (500、1000 和 2000 人)、以及每個向度的題數 (10 和 20 題)。Svetina 比較了 DETECT 法的 DETECTe (探索性)、DETECTcv (交叉驗證性) 和 NOHARM 法的 RMSR、 χ^2_{GID} 和 ALR 等五種統計數在辨識向度數的表現。在二向度和三向度補償性 MIRT 條件下，DETECT 的表現通常優於 NOHARM 的方法，尤其是當向度間的相關小於 0.6，資料複雜性小於 30%，以及有較大的樣本量時，DETECT 會得到較高比例的正確率。當複雜性增加和樣本量變小時，所有方法的表現普遍降低。DETECT 對簡單結構的試題分類是一致的，但對複雜結構試題的分類之表現則不太一致。而三個以 NOHARM 的結果為基礎的方法中， χ^2_{GID} 和 ALR 通常表現

比 RMSR 出色，當樣本人數為 500 而複雜性為 30% 或更低時， χ^2_{GID} 的表現最佳。當題數增加時，ALR 在相關 0.60，而複雜性低於 30% 時的表現會獲得改善。當使用非補償性 MIRT 模式產生資料時，以 NOHARM 為基礎的方法，特別是 χ^2_{GID} 和 ALR 是所有五個方法中表現最準確的。綜合來說，DETECT 的表現在各種複雜性水準、樣本大小、以及相關的簡單試題分類上比 NOHARM 更一致，當複雜性和相關性水準增加時，所有方法的分類率都會隨之下降。在大多數情況下，DETECT 對複雜試題的分類比以 NOHARM 結果為基礎的方法更為一致。特別是，當複雜性、試題數，和真實的向度數增加時，DETECT 比 NOHARM 表現的比較好。

(二) PA 及 HULL 法的相關研究

在探索性因素分析的研究中，因素個數的決定是關鍵的步驟，在諸多因素個數決定的方法中，平行分析 (PA) 表現相當良好。Zwick 和 Velicer (1986) 對 Kaiser 的特徵值大於 1、Bartlett 的卡方檢定、Cattell 的陡坡圖檢定、Velicer 的 MAP 法與 Horn 的平行分析等五種決定因個數的方法進行比較，操弄沒有相關存在的因素個數、樣本大小、每個因素的指標變項個數、因素負荷量大小和因素複雜性等條件來檢視上述方法在決定因素個數上的表現。在這些條件下，PA 被發現是最準確的方法。研究中還提出幾個可以改

善 PA 表現的建議，包括增加樣本量，因素負荷量，以及每個因素有關的變數個數等。

Crawford 等人 (2010) 進行了以 PA 決定因素個數的研究，研究中利用模擬母群和樣本的資料，用來比較使用主成分分析的平行分析 (PA-PCA) 與使用主軸因素分析的平行分析 (PA-PAF) 在決定潛在因素個數上的表現；同時，也對平均特徵值 (λ_M) 和第 95 百分位數特徵值 (λ_{95}) 兩個標準的精確度進行檢驗。在母群資料分析研究中，Crawford 根據一、二、四個因素之模式產生共變數矩陣，並操弄每個因素的指標變數個數、平均值和因素負荷量的變化、以及因素之間相關的大小等條件。研究結果顯示，當只有一個潛在的因素時，PA-PCA 和 PA-PAF 的表現都 100% 準確；但是當有多個潛在因素時，PA-PAF 的表現都比 PA-PCA 更準確；PA-PCA 和 PA-PAF 在平均負荷量較小和因素之間的相關較高的條件下，會出現不正確的結果。在這些情況下，不正確的結果，總是建議 1 個潛在因素。在樣本資料分析研究中，Crawford 等人所操弄的因素包括樣本量 (100 和 500 人)、因素個數 (0、1、2 和 4 個)、觀察指標變數個數 (0 個因素時，指標變數為 3、6 和 12 個；1、2 和 4 個因素時，指標變數的個數為 3 或 6 個)、因素負荷量 (0.3、0.5 和 0.7) 以及因素間的相關 (具有 2 個或 4 個因素時，因素之間的相關分別為 0、0.4、0.6、0.7 或 0.8)。分析結果顯示，當沒有潛在因素存在時，PA-PCA 和 PA-PAF 所給予的準確

度相似，兩種方法使用 λ_{95} 的表現都比 λ_M 更準確。在一個潛在因素的情形下，具有較大因素負荷量和較大樣本時的表現明顯地比較好；無論是使用 λ_M 還是 λ_{95} 規則，PA-PCA 的表現與 PA-PAF 一樣好；當每個因素有 3 個指標變數時， λ_M 比 λ_{95} 準確，然而當每個因素有 6 個指標變數時， λ_{95} 比 λ_M 來得好。在多個因素的情形下，所得到的結果與單一因素的類似，不同的平行分析在具有較大的因素負荷量和更大的樣本人數時，普遍有好的表現；當因素間沒有相關，使用 λ_{95} 規則時，PA-PAF 表現與 PA-PCA 相似；然而，當使用 λ_M 規則時，PA-PCA 表現比 PA-PAF 好。當因素間的相關程度增加時，若使用 λ_{95} ，PA-PAF 表現比 PA-PCA 更好，但這兩種方法在樣本人數為 100 且因素負荷量低時，傾向於產生不準確的結果。總之，在大多數研究的條件下，不管抽取因素的方法和特徵值規則的選擇，只要負荷量變大，樣本人數增多，每個因素測量的指標變數個數增加和因素間的相關變小，PA 準確性皆會隨之改善。

Lorenzo-Seva、Timmerman 和 Kiers (2011) 以模擬資料來比較 MAP、PA、AIC (Akaike, 1973)、BIC (Schwarz, 1978) 與 HULL 法 (分別使用了 CFI、CAF、RMSEA 和 SRMR 等四種適配度指標) 的表現，該研究操弄了底下五個條件：(a) 測量的變數個數，(b) 主要因素個數，(c) 樣本大小，(d) 因素間相關水準，(e) 共同變異數水準。研究結果顯示：HULL 法在檢測主要因素個數之表現得很好，準確率介於 85~94%，而 HULL-CFI 和 HULL-CAF 的

表現又優於 HULL-RMSEA 和 HULL-SRMR；PA 的成功率為 81%，MAP 表現得不理想，整體成功率為 51%，而 AIC 與 BIC 表現得最差，成功率不及 50%；因此他們認為 PA 是除了 HULL 法以外可以用來決定因素個數的方法。為了進一步證實 HULL 法的準確性，她們使用五因素人格量表（Five-Factor Personality Inventory, FFPI），該量表有 100 個題目，皆為 Likert 五點量尺之題目（Hendriks, Hofstee, & De Read, 1999）。FFPI 是一個平衡的問卷，亦即，測量每個因素的題目皆 20 題，其中正向題 10 題，反向題 10 題，以控制受試者填答時，不會受到「默許」（acquiescence）這個反應心向的影響。即使如此，雖然原始得分不會受到默許這個反應心向的影響，但是題目間之相關矩陣仍然會受到默許這個反應心向所造成的變異情形的影響（Bentler, 1969; 引自 Lorenzo-Seva & Rodríguez-Fornells, 2006）。FFPI 的設計方式，使得一般 EFA 分析常會得到六個因素，一個由默許的反應心向所組成的因素，加上符合大五人格理論（Big-Five model of personality）的五個因素（Lorenzo-Seva & Rodríguez-Fornells, 2006）。

Hendriks 等人在編製 FFPI 時，所使用的 2,451 人常模樣本的資料被 Lorenzo-Seva 等人（2011）拿來分析，以檢驗 HULL 法在選擇因素個數上的表現。Lorenzo-Seva 等人首先執行 PA，採用 $\lambda_{.95}$ 標準，結果 PA 建議保留 11 個因素，因此在利用 HULL 方來決定共同因素的個數時，將因素個數的範圍設為 0 到 12 個，在分別進行這 13 個因素模式的分析之後，所得到的 HULL 圖即是圖 3。介於每個點之間

畫出一條想像的直線。由圖 3 可以看到介於 2 和 5 之間的點都低於這條想像的線，因此摒棄不用。根據 HULL 法選擇最佳因素個數的規則，由圖中可以看到因素個數可能為 0、6、7、11 或 12 個，在圖 3 中分別以比較大的點呈現。FACTOR 程式為使研究者能較客觀地判斷因素個數，它計算了每個因素解的陡坡圖檢定值 st ，其中三個 st 值呈現於表 3 之中，最大的是 6 個因素解的 st 值，因此，6 個因素是最佳的選擇。

表 3

因素個數 (q)、適配度值 (f)、自由度 (df) 和陡坡圖檢定值 (st)

q		f	df	st
0	*	0.052	0	—
1		0.096	100	—
2		0.129	199	—
3		0.169	297	—
4		0.223	394	—
5		0.3	590	—
6	*	0.377	585	2.542
7	*	0.398	679	2.054
8		0.408	772	—
9		0.417	864	—
10		0.427	955	—
11	*	0.437	1,045	1.755
12	*	0.443	1,134	—

註：資料取自“The Hull Method for Selecting the Number of Common Factors,” by Lorenzo-Seva, U.; Timmerman, M. E.; Kiers, H. A. L. (2011). *Multivariate Behavioral Research*, 46(2), 358. st 為 the Scree Test value (陡坡圖檢定值)，*表示解位在凸包 HULL 上，是可能的因素個數解。

四、研究目的與問題

(一) 研究目的

根據前面的敘述，本研究擬探討 DETECT、NOHARM、PA 與 HULL 等方法，在決定向度數（或因素個數）上的表現。具體而言，本研究的主要目的在於探討向度數、向度間的相關、樣本大小及題數對 DETECT、NOHARM、PA 和 HULL 等方法在向度評估上表現的影響。

(二) 研究問題

依據上述之研究目的，本研究將試圖回答以下的研究問題：

1. 在不同向度數的資料中，DETECT、NOHARM、PA 和 HULL 等方法在向度評估正確率的表現為何？
2. 向度之間的相關對 DETECT、NOHARM、PA 和 HULL 等方法在向度評估正確率的表現有何影響？
3. 樣本大小對 DETECT、NOHARM、PA 和 HULL 等方法在向度評估方法的正確率有何影響？
4. 題數對 DETECT、NOHARM、PA 和 HULL 等方法在向度評估正確率的表現有何影響？

貳、研究方法與設計

本研究欲探討 DETECT、NOHARM、PA 和 HULL 等四種程序在測驗資料向度數檢定表現之異同，四種程序的分析性質在本研究中主要侷限在探索性部分，雖然 DETECT 和 NOHARM 皆允許研究者進行「驗證性」的分析。為要回答研究問題，擬藉著對已知結構

的模擬資料進行分析，探討這四種程序檢測向度數的能力。底下說明研究在模擬資料時所操弄的條件。

一、研究設計規劃

(一) 模擬資料之條件設定

如上面所述，本研究將透過模擬資料來探討向度數、各向度能力間的相關、各向度題數以及不同樣本人數等四個條件，對於向度檢測方法表現的影響，因此底下將描述產生模擬資料時所考慮的各項條件。

1. 資料型態：由於本研究所探討的方法中 DETECT 和 NOHARM 僅適用二元計分的資料，因此模擬的試題反應資料將侷限為二元計分的資料。
2. 向度個數：在本研究中考慮 3 種向度結構：單向度（以 1d 表示）、二向度（以 2d 表示）和三向度（包含 3d 和 3d1）等，其中 3d 表示資料是利用三個向度的 MIRT 模式產生，也必須用三個向度的 MIRT 模式才能有良好的適配度，而 3d1 的資料雖然是利用三個向度的 MIRT 模式來產生的，然而可以用一個單一向度的模式來加以適配。
3. 各向度能力值間的相關：向度間的相關包括 0、0.3、0.6 等三種，以幫助瞭解在不同強度的相關之下，這四種程序在決定向度數上的表現情形。
4. 各向度題數：為探討題數對各種方法表現的影響，本研究模擬的試題反應資料之題數為 30 及 60 題，除了 3d1 以外，2d 測驗的設計有 1/3 的試題為複雜結構 2/3 的試題為近似

簡單結構，而 3d 測驗的設計將以近似簡單結構為主，因此各個向度題數的安排如下：

1d—由於是單一向度，所以全部的試題皆測量相同的向度 30（60）題；

2d—主要測量第一個向度 10（20）題、第二向度 10（20）題、及二個向度皆測的 10（20）題；

3d—主要測量第一個向度 10（20）題、第二向度 10（20）題、及第三向度 10（20）題；

3d1—每一個試題均需要三個向度的能力，但是試題向量在空間中皆指到相同的方向，實際上可以利用單一向度模式來適配資料。因此，題數為 30（60）題。

5. 試題參數：用來產生模擬資料所用的試題參數均取自

Reckase（2009），其中，用來產生單一向度模擬資料的試題參數是取自第 153 頁的表 6.1，原先該表中的試題參數是用來產生簡單結構之三向度資料，此處是將各題所有的最大鑑別力參數那一項取出，視為是單一向度的鑑別力參數。產生 2d 和 3d 的試題參數分別取自第 204 頁的表 7.8 和第 151 頁的表 6.1。當題數為 60 題時，則將原本各向度的試題參數複製一次。而用來產生 3d1 資料之試題參數是取自第 191 頁的表 7.5。之所以採用這些試題參數，筆者是希望本研究的結果最後也可以與 Reckase 書上的結果作一個對照。

6. 樣本人數：本研究探討 250、500、1000 和 2000 人等不同的人數對於這四種程序表現的影響。

7. 模式產生的方式：除了單向度（1d）的資料是用二參數 logistic 的 IRT 模式產生以外，其餘的 2d、3d 及 3d1 的資

料將皆以多向度二參數 logistic IRT 模式(M2PL)來產生，亦即只考慮猜測參數 $c = 0$ 的情況。模擬資料的產生是以 MATLAB 程式來完成，而在每一個條件組合之下，將重複 100 次。模擬產生資料時，所操弄的各個條件摘要呈現於表 4 之中，因為單向度的測驗資料不涉及向度間相關的問題，所以模擬情境總數為 $3 \times 3 \times 2 \times 4 + 2 \times 4 = 80$ 種，其中單一向度部分共有 $2 \times 4 = 8$ 種，而其他的部分則有 $3 \times 3 \times 2 \times 4 = 72$ 種。

表 4
模擬資料分配表

操弄的條件	情境	情境總數
向度個數	1d、2d、3d 和 3d1	4
各向度能力值間的相關	0、0.3 和 0.6	3
各向度題數	10 和 20 題	2
樣本人數	250、500、1000 和 2000	4
情境總數		80

(二) 模擬步驟

本研究以 MATLAB 程式來產生模擬資料，模擬資料產生之步驟如下：

1. 以前述之試題參數作為模擬資料的真實試題參數。
2. 界定測驗的向度個數及每個向度的題數。
3. 界定模擬資料的受試者人數，受試者在各個向度能力值均由 $N(0,1)$ 抽樣得到，而產生二向度和三向度的資料時，各向度能力值之間的相關，分別設為 0、0.3 和 0.6。亦即單一向度受試者能力值是由標準常態分配抽樣得到，而二向

度和三向度的資料，受試者的能力值是由多變項常態分配抽樣得到。

4. 根據步驟 2 的試題參數與步驟 3 所產生之受試者能力值，利用 2PL 或 M2PL 模式，計算受試者答對每一個試題之機率值 $P(\theta)$ 。
5. 然後針對受試者在每一個試題上面，由齊一分配 $U(0,1)$ 隨機產生一個亂數，與步驟 4 所產生的 $P(\theta)$ 比較，若 $P(\theta)$ 大於 $U(0,1)$ 所產生之亂數值，則該題的答題反應計為答對為 1 分，反之為答錯為 0 分，藉以產生受試者每一個題目之作答反應。
6. 根據步驟 1~5 之程序，每種情境均模擬 100 次，總共有 80 種模擬情境。

二、資料分析方法

研究者利用 DETECT、NOHARM、PA 和 HULL 法等四種程序分別估計測驗資料的向度數，計算四種程序正確辨認向度個數的比率。關於這四種方法如何決定向度個數，簡要描述如下。

(一) DETECT 程序

DETECT 主要是一個探索性的技術，旨在尋求確認同質試題群組的個數，作為存在於資料中具有近似簡單結構之向度數。本研究先檢驗資料的 r_{\max} 值，在實務中 r_{\max} 大於或等於 0.8 時，所分析的資料具有近似的簡單結構 (Jang & Roussos, 2007; Stout et al., 1996)。當近似簡單結構成立時，因為 DETECT 輸出不重疊 (non-overlapping) 的試題

群組，因此，DETECT 所提供的不重疊試題之群組數等於向度數。

(二) NOHARM 程序

NOHARM 利用 ULS 來最小化其適配函數，因此缺乏正式的檢定統計數，為了彌補這個缺憾，可以使用 Gessaroli 和 De Champlain (1996) 所提出的近似卡方的統計數 $\chi^2_{G/D}$ 。本研究將使用在相鄰兩個向度模式之 $\chi^2_{G/D}$ 值的差異來檢定是否需要較多的向度，來作為決定向度數的方法。開始時以向度數較少的模式來適配資料，然後將向度個數加 1，再重新適配資料；這兩個模式 $\chi^2_{G/D}$ 之差異，具有自由度等於 1 之卡方分配，如果差異沒有達到顯著，表示向度數較少的模式是可接受的。假如向度數較高的模式提供較好的適配（即差異檢定達到顯著），該程序繼續進行，將向度數加 1，再次適配資料，以決定新的模式是否有較佳的適配。具體而言，對於本研究中單向度的資料，進行單向度與二向度的 $\chi^2_{G/D}$ 差異檢定來決定向度數，而二向度的資料，則進行單向度、二向度與三向度的 $\chi^2_{G/D}$ 差異檢定來決定向度數，以此類推。

(三) PA 法

雖然 Reckase (2009) 提到 PA 在決定向度個數上的應用時，是利用 TESTFACT 這個程式對資料進行完整訊息的試題因素分析 (full-information item factor analysis)，以得到特徵值，在本研究中，將仍採取傳統探索性因素分析的作法。本研究採用 FACTOR (Lorenzo-Seva & Ferrando,

2013) 這個程式來進行平行分析，FACTOR 的功能十分齊全，允許使用者分析資料時使用共變數矩陣、Pearson 相關矩陣、和四分相關矩陣；在決定因素個數方面，它讓使用者可以由 MAP、PA 和 HULL 中擇一；而抽取因素的方法，FACTOR 提供未加權最小平方、最大概似、極小秩因素分析 (minimum rank factor analysis, MRFA) 或 PCA 等，此外，FACTOR 程式也讓使用者自行選擇是否進行轉軸，以及設定轉軸的方法 (操作介面如圖 4 所示)。因此，本研究採用以 PCA 法作為抽取因素的方法分析的結果，以 λ_{95} 作為閾值來決定因素 (向度) 個數。

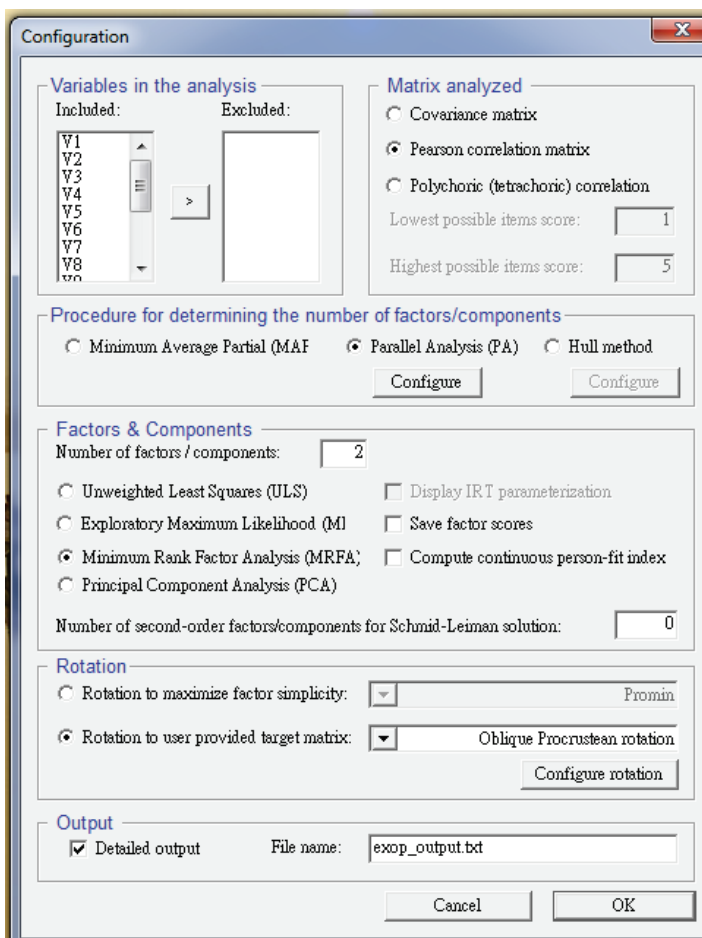


圖 4 FACTOR 程式介面

(四) HULL 法

本研究也採用 FACTOR 程式來完成 HULL 法的分析，當選擇 HULL 法來決定因素個數時，FACTOR 程式除了需要研究者決定分析的資料矩陣型態、因素抽取方法以外，FACTOR 程式也允許使用者選擇 CAF、RMSEA 或 CFI 來

作為符合其需求的適合度指標，也可由程式自動依照所使用的因素抽取方法選擇最佳的指標，後者是 FACTOR 所預設的。在這些適合度指標中，CAF 為 common part accounted for 之縮寫，亦即是以 KMO 指標為基礎的；CFI (comparative fit index) 與 RMSEA (root mean squar error of approximatim) 二者皆是結構方程模式中常見的指標。在本研究中，以 PCA 法來抽取因素，以 CAF 作為適合度指標，在 FACTOR 程式中，此組合稱為 PCA-CAF。

肆、結果與討論

本研究比較四種決定測驗向度（因素）個數的方法在判斷測驗潛在特質個數的表現之差異情形。以下先呈現在所操弄的各種條件下，四種方法在單一向度（1d）資料之表現，接著呈現二向度（2d）資料的結果，最後是三向度（包括 3d 和 3d1）資料的結果。

一、單一向度模擬資料分析結果

在本研究中 DETECT 分析無論使用探索性或交叉驗證模式的方式都無法順利得到一個單向度的解，此項結果與 Svetina（2011）的研究結果一致。因此，在單一向度的情形中，將只呈現 NOHARM、PA 和 HULL 三種程序的結果比較。由於是單一向度，所以全部的試題皆測量相同的向度。

表 5 上半部呈現的是三種程序在單一向度 30 題測驗資料中正確指出向度數的比率，PA 和 HULL 兩種程序，不論樣本大小，都可以 100% 的正確判斷出單向度的資料，而 NOHARM 程序除了在

樣本為 250 時正確率為 0.80 外，隨著樣本量增加，其正確的判斷率也隨之增加達到 0.98。表 5 下半部是單一向度 60 題資料的分析結果，由表中數據可以看到，三種不同程序不論樣本大小，幾乎都可以 100% 的正確判斷出單向度的資料。對照 30 題和 60 題的結果，可以發現隨著題數的增加，NOHARM 的 χ_{GID}^2 的正確率有增加的傾向。就整體來看，除了 30 題 250 人的情形以外， χ_{GID}^2 的表現情形與 De Champlain 和 Gessaroli (1998) 的發現類似。

表 5
各種方法對於單向度資料向度數的正確判斷率

題數	人數	方 法			
		DETECT	NOHARM	PA	HULL
30	250	—	.80	1.00	1.00
	500	—	.96	1.00	1.00
	1000	—	.97	1.00	1.00
	2000	—	.98	1.00	1.00
60	250	—	1.00	1.00	1.00
	500	—	.99	1.00	1.00
	1000	—	1.00	1.00	1.00
	2000	—	1.00	1.00	1.00

二、二向度模擬資料分析結果

在模擬二向度的資料時，整個測驗被規劃成 1/3 的試題測量第一個向度，1/3 的試題測量第二個向度，其餘 1/3 的試題同時測量了兩個向度，所以最後 1/3 的試題是文獻中所謂的複雜結構 (Gierl, Leighton, & Tan, 2006)，而其餘 2/3 的試題呈現近似結構資料。

表 6 上半部呈現的是二向度 30 題測驗資料的結果，DETECT 最初被發展出來的目的是在一個近似簡單結構的假定之下，決定用

來表徵試題分數矩陣中試題間關係所需要的向度數，Zhang 和 Stout (1999b) 提到即使當近似簡單結構未能成立時，DETECT 的分析結果可能仍然是有參考價值的。由表 6 中可以看出，對 DETECT 而言，跨所有各向度能力值間的相關和樣本大小的條件，其正確模式判斷率介於 0.15~0.66 之間(除了在相關為 0.6，樣本大小為 250，正確判斷率為 0.05 偏低外)，正確判斷率隨著樣本人數的增加而提升。對 NOHARM 而言，跨所有各向度能力值間的相關和樣本大小的條件，其正確模式判斷率介於 0.52~0.73 之間，似乎比 DETECT 程序較不受各項操弄變數所影響。PA 在各向度能力值間的相關小於或等於 0.3 時，其正確模式判斷率幾乎達 100%，當各向度能力值間的相關為 0.6 時，正確模式判斷率介於 0.16~0.99 之間，正確判斷率會隨著樣本人數的增加而提升。當各向度能力值間的相關等於 0 時，HULL 程序的正確模式判斷率介於 0.95~0.99 之間，然而當各向度能力值間有相關存在時，基本上 HULL 程序的向度判斷表現都不理想。

表 6
各種方法對於二向度資料向度數的正確判斷率

題數	相關	人數	方 法			
			DETECT	NOHARM	PA	HULL
30	0	250	.44	.53	1.00	.95
		500	.54	.52	1.00	.96
		1000	.58	.58	1.00	.96
		2000	.59	.66	1.00	.99
	.3	250	.21	.69	.97	.26
		500	.44	.62	1.00	.38
		1000	.66	.72	1.00	.23
		2000	.66	.65	1.00	.26
	.6	250	.05	.64	.16	0
		500	.15	.60	.23	.17
		1000	.51	.73	.76	0
		2000	.65	.66	.99	0
60	0	250	.39	.97	1.00	.99
		500	.48	.91	1.00	1.00
		1000	.64	.99	1.00	1.00
		2000	.80	.99	1.00	1.00
	.3	250	.22	.97	1.00	.78
		500	.48	.95	1.00	.98
		1000	.54	.99	1.00	1.00
		2000	.64	1.00	1.00	1.00
	.6	250	.12	.95	.94	.05
		500	.25	.84	.97	.26
		1000	.52	1.0	1.00	.33
		2000	.71	.98	1.00	.34

表 6 的下半部呈現的是二向度 60 題測驗資料的結果，由其中可以看出，對 DETECT 而言，當各向度能力值間的相關小於或等於 0.6 時，其正確模式判斷率介於 0.12~0.80 之間，其正確判斷率隨著樣本增加而提升，但是隨著各向度能力值間的相關變大而逐漸變小。對 NOHARM 而言，除了向度能力值之間的相關為 0.6 且樣

本人數為 500 人時，正確判斷率為 0.84 外，在其他所有的條件組合下，正確判斷率皆在 0.91~1.00 之間。這個結果與 De Champlain 和 Gessaroli (1998) 所呈現的數據有些差距，然而在 De Champlain 和 Gessaroli 的研究中所呈現的是向度數為 1 的虛無假設在二向度情形下被拒絕的次數，未必代表真正的向度數為 2，因為當進行向度數等於 2 的假設檢定時，也可能這個假設也被拒絕。而 PA 跨所有向度能力值之間的相關和樣本大小的條件，其正確模式判斷率幾乎都達 100%。HULL 程序在各向度能力值間的相關小於或等於 0.3 時，其正確模式判斷率幾乎都達 100% (除了相關為 0.3 且樣本人數 250 時，正確判斷率為 0.78 外)，當各向度能力值間的相關等於 0.6 時，基本上 HULL 程序的表現不是十分理想。

對於二向度的模擬資料，因為有 1/3 的試題同時在兩個向度上都有因素負荷量，是文獻中所謂的複雜結構，似乎影響了 DETECT 正確判斷率的表現。整體而言，在各向度能力值間的相關較低時，所有程序的正確判斷率的表現都比較好，而且有隨著樣本人數增加而提升，而隨著各向度能力值間的相關變大而減小的趨勢，基本上測驗長度增加也會明顯改善各程序的正確判斷率。整體來說，DETECT 在二向度資料的表現不如 NOHARM 的 χ^2_{GID} ，此項發現與 Svetina (2011) 的不同，她發現在二向度補償性 MIRT 的條件下，DETECT 的表現普遍優於 NOHARM。

Reckase (2009) 介紹了數個資料向度數決定之程序，其中與本研究有關的有 DETECT 和 PA。Reckase 使用一組試題參數模擬三向度的試題反應資料，該組試題包含了測量閱讀能力 10 題，測量算術計算能力 20 題，以及測量閱讀、算術計算和問題解決能力 20 題共 50 題，此種測驗組型的設計即為複雜結構。對這筆資料，

DETECT 分析結果建議為二個試題群集而非預先設計的三個群集，此結果是因為 DETECT 傾向於忽略沿著參照組合方向的試題，而且它試圖去適配一個近似簡單結構模式。DETECT 會將在某一個向度上 a 參數較大之試題群集在一起，即使這些試題中有一部份在其他的向度上 a 參數的值也是大的；DETECT 只有在具有較大的 a 參數之試題組形成互斥的試題群集時，此方法才能正確地將試題進行群集分類 (Reckase, 2009)。此結果與本研究中的二向度模擬資料的結果是一致的，因為複雜結構的試題之反應組型會影響 DETECT 的正確模式判斷率。

至於 PA 的部分，Reckase (2009) 以一個二向度的模擬資料來進行分析，結果發現 PA 可以準確地判斷出向度數，與本研究二向度資料在各個條件下 PA 的表現大都很理想的結果吻合，使用 PA 作為向度數的判斷是個好的選擇。

三、三向度模擬資料分析結果

(一) 三向度的資料之一 (3d)

三向度模擬的資料產生時，1/3 的試題主要測量第一個向度，1/3 的試題主要測量第二個向度，而另外 1/3 的試題測量第三個向度，是屬於近似簡單結構的設計。分析的結果呈現於表 7 之中。

表 7 的上半部是測驗長度為 30 題時之分析結果，由其中可以看出，當各向度能力值間的相關小於或等於 0.3 時，DETECT 的正確模式判斷率介於 0.73~1.00 之間，正確判斷率隨著樣本人數增加而提升，當向度能力值之間的相關

等於 0.6 時，基本上 DETECT 的正確判斷率會隨著人數減少而減小。對 NOHARM 而言，跨所有各向度能力值間的相關和樣本大小的條件，其正確模式判斷率介於 0.51~0.70 之間，似乎不受相關大小與人數多寡的影響。PA 在各向度能力值間的相關小於或等於 0.3 時，其正確模式判斷率幾乎達 100%，然而當各向度能力值間的相關為等於 0.6 時，PA 就幾乎無法正確做出判斷。而 HULL 程序，在各向度能力值間的相關等於 0 時，其正確模式判斷率介於 0.98~1.00 之間，當各向度能力值間有相關存在時，基本上 HULL 程序在正確模式判斷率表現都不理想。

表 7

各種方法對於三向度資料 (3d) 向度數的正確判斷率

題數	相關	人數	方 法			
			DETECT	NOHARM	PA	HULL
30	0	250	.81	.59	.99	.98
		500	.98	.67	1.00	1.00
		1000	1.00	.67	1.00	1.00
		2000	1.00	.68	1.00	1.00
	.3	250	.73	.51	.78	.21
		500	.81	.53	1.00	.26
		1000	.95	.64	1.00	.38
		2000	1.00	.68	1.00	.48
.6	250	.51	.64	.03	0	
	500	.54	.54	.02	0	
	1000	.63	.70	.37	0	
	2000	.74	.65	.35	0	

(續下頁)

(接上頁)

題數	相關	人數	方 法			
			DETECT	NOHARM	PA	HULL
60	0	250	.98	.88	.99	1.00
		500	.97	.99	1.00	1.00
		1000	1.00	.96	1.00	1.00
		2000	1.00	.99	1.00	1.00
	.3	250	.85	.93	1.00	.97
		500	.95	.95	1.00	1.00
		1000	.99	1.00	1.00	1.00
		2000	1.00	.98	1.00	1.00
	.6	250	.62	.98	.32	.17
		500	.66	.99	.42	.40
		1000	.85	.98	.61	.39
		2000	.92	.98	.65	.38

表 7 下半部呈現的是三向度 60 題的分析結果，由表中可以看出，當各向度能力值間的相關小於或等於 0.3 時，模式判斷正確率介於 0.85~1.00 之間，當各向度能力值間的相關為 0.6 時，模式判斷正確率介於 0.62~0.92 之間，判斷正確率隨著樣本人數的增加而提升。對 NOHARM 而言，跨所有各向度能力值間的相關和樣本大小的條件，其模式判斷正確率介於 0.88~1.00 之間，表現十分很好。PA 和 HULL 在各向度能力值間的相關小於或等於 0.3 時，其模式判斷正確率幾乎都達 100%，然而，當相關等於 0.6 時，兩種程序的判斷正確率都下降，表現都不理想。

整體而言，三向度（3d）的模擬資料屬於近似簡單結構，DETECT 正確判斷的比率應當會是不錯，也會比其在二向度複雜結構資料的表現好很多，與 Svetina（2011）在

三向度利用補償性 MIRT 所產生的資料中所發現的，當向度間的相關小於 0.6 且有較大的樣本時，DETECT 的表現通常優於 NOHARM 的情形類似，表 7 中 DETECT 的表現，除了在相關等於 0.6 的情形外，大致上都優於 NOHARM。NOHARM 的判斷正確率跨各向度能力值間的相關和樣本大小的條件，也都表現不錯。PA 和 HULL 在各向度能力值間的相關較低時，判斷正確率都很好，有隨著相關增加而變小的趨勢，基本上測驗長度增加也會明顯改善各種程序正確判斷率。

(二) 三向度的資料之二 (3d1)

這個部分所模擬的是受試者需要三個向度的能力始能正確回答每一個題目之試題反應資料，然而，由於所有的試題向量在潛在能力空間中皆指向相同的方向；根據 Wang (1985) 參照組合分數 (reference composite) 的概念來看，實際上是可以使用單一向度模式來適配資料的。當利用 DETECT 來進行測驗的向度數分析時，其結果與前述的單一向度的資料雷同，因此底下只呈現 NOHARM、PA 和 HULL 三種程序的結果。

表 8 上半部呈現的是三向度 (3d1) 30 題時之分析結果，由表中可以看到，PA 和 HULL 兩種程序，跨所有各向度能力值間相關和樣本大小之條件，正確的判斷的比率皆為 100%，而 NOHARM 程序的判斷正確率介於 0.56~0.83 之間。表 8 的下半部呈現的是 60 題資料的分析結果，在各

種相關和各種樣本人數中，NOHARM 正確判斷的比率在於 0.96~1.00 之間，而 PA 和 HULL 兩種程序，則在所有的條件組合下，皆有 100%的判斷正確率。

整體而言，3d1 的資料基本上是所有試題箭頭皆指向該測驗參照組合方向，可以想像成是一個單一向度的測驗，而該能力值是三個向度之參照組合的分數，亦即是用來產生資料的三個向度能力值之線性組合；因此，可以使用單向度的 IRT 模式來進行分析。就不同測驗長度的結果來看，PA 和 HULL 兩種程序模式判斷正確的比率皆為 100%，未受到題數的影響，而 NOHARM 的結果明顯受到題數的影響，當題數愈多時，判斷的正確率會愈高。

當初 Reckase (2009) 模擬 3d1 資料的目的是想要闡明關於資料向度數的三個重點：第一，向度數是資料矩陣的特性。即使資料是以多個 θ 值產生的，或者試題在沿著數個向度方向上的差異是敏感的，資料矩陣的向度性未必能夠反映資料實際複雜的程度。第二，即使測驗試題對很多向度上的差異是敏感的，如果所有試題所測量的 θ 之組合分數是相似的，則受試者與試題之交互作用的結果可能只需要用單一向度模式即可加以模式化。第三，即使試題對很多向度的差異是敏感的，如果所有受試者的能力位於多向度空間中的一條線上，所得到的作答反應資料可以用單一向度模式來解釋 (pp. 193-194)。本研究的 3d1 模擬資料之分析結果與 Reckase 所提到的第二個重點相吻合，名義

上資料是利用三個向度的 MIRT 模式產生的，實際上利用單一向度的模式即有很好的適配度，而 NOHARM、PA 和 HULL 等三種程序均有很好的模式正確判斷率。

表 8
各種方法對於三向度資料 (3d1) 向度數的正確判斷率

題數	相關	人數	方 法			
			DETECT	NOHARM	PA	HULL
30	0	250	—	.61	1.00	1.00
		500	—	.59	1.00	1.00
		1000	—	.60	1.00	1.00
		2000	—	.70	1.00	1.00
	.3	250	—	.83	1.00	1.00
		500	—	.73	1.00	1.00
		1000	—	.83	1.00	1.00
		2000	—	.74	1.00	1.00
	.6	250	—	.73	1.00	1.00
		500	—	.56	1.00	1.00
		1000	—	.73	1.00	1.00
		2000	—	.61	1.00	1.00
60	0	250	—	1.00	1.00	1.00
		500	—	.98	1.00	1.00
		1000	—	.97	1.00	1.00
		2000	—	1.00	1.00	1.00
	.3	250	—	1.00	1.00	1.00
		500	—	.98	1.00	1.00
		1000	—	.99	1.00	1.00
		2000	—	.98	1.00	1.00
	.6	250	—	1.00	1.00	1.00
		500	—	.96	1.00	1.00
		1000	—	1.00	1.00	1.00
		2000	—	.98	1.00	1.00

肆、結論與建議

本研究的目的是在探討 DETECT、NOHARM、PA 和 HULL 等常用的測驗向度數檢測的程序，對不同的條件組合下所產生的模擬資料之向度數評估的正確率。以下呈現研究的結論，以及對未來研究的建議。

一、研究結論

筆者利用四種不同的評估向度的方法，包括二種在 IRT 研究中常見的方法 (DETECT 和 NOHARM)，以及二種在因素分析中被肯定的方法 (PA 和 HULL)，來探討在不同的條件組合下所產生的模擬資料之向度數的檢測結果。研究設計所操弄的條件包括向度數、向度能力值之間的相關、各向度題數、以及樣本人數。

(一) 向度個數的影響情形

由單一向度模擬資料的分析結果可以發現，無論使用探索性或交叉驗證模式的方式，DETECT 都無法順利得到一個單向度的解，NOHARM、PA 和 HULL 這三種方法都表現很好，幾乎達 100% 的正確判斷率。二向度資料的分析結果顯示，因為模擬所產生的資料是複雜結構的，與 DETECT 的理論基礎不是十分相符合，因此影響 DETECT 的表現，NOHARM、PA 和 HULL 等三種方法的表現都優於 DETECT。三向度資料是為近似簡單結構的，因此 DETECT 的表現就比在二向度時好。整體而言，各向度能力值之間的相關小於或等於 0.3 時，PA 的表現優於其他三

種程序，但是當各向度能力值之間的相關變大(即 0.6)時，DETECT 和 NOHARM 的表現比 PA 和 HULL 好。

而另一種用三向度模式所產生的三向度資料，因在空間中所有的試題箭頭均指到同一個方向，該方向可以視為是一個參照組合；因此這筆資料是可以利用單一向度 IRT 模式來加以適配的 (Wang, 1985)。因為資料的向度數實際上是單一向度的，各個方法的表現情形也與前面的單一向度資料所看到的一致，亦即 DETECT 無法順利得到一個單向度的解，而 NOHARM、PA 和 HULL 的正確率皆幾乎達到 100%。

(二) 各向度能力值之間的相關之影響情形

綜合在各種條件組合下資料分析之結果來看，DETECT、PA 和 HULL 三種方法受到各向度能力值之間相關的影響比較明顯，隨著各向度能力值之間相關的增加，模式判斷的正確率隨之減小；NOHARM 較不受能力值間之相關的影響。整體的模式判斷正確率在能力值相關小於或等於 0.3 時，PA 的表現明顯優於其他三種方法。對於這個情形，可能的原因是因為 DETECT 假定資料是具有直交簡單結構的，PA 是以傳統的作法將相關矩陣分解為彼此無關的特徵值與特徵向量，因此所得到的向度數應亦屬直交向度的個數，而 HULL 法以 PA 所建議的因素個數加上 1 為上限，來尋求最適合的解。因此當資料的向度之間具有相關時，DETECT 和 PA 的表現會因所用的算則之假定與資料

不一致而受到影響，當 PA 所給予的向度數不正確時，HULL 法理論上也受到影響。當然，此項觀察也會受到題數的影響，值得後續研究探討清楚。

(三) 各向度試題數

DETECT、NOHARM、PA 及 HULL 等四種方法的表現，都會受到題數的影響，當題數變多時，各種模式的判斷正確率也會隨之提升。

(四) 樣本大小的影響情形

NOHARM、PA 和 HULL 三種方法似乎較不受樣本大小的影響，而當樣本人數增加時，DETECT 模式判斷正確率也會隨之增加。這種結果應該與 DETECT 在分析資料時，需要將整個樣本隨機分成兩個次樣本，以分別得到 D_{\max} 和 D_{ref} ，並計算 $r_{\max} = D_{ref} / D_{\max}$ 等之作法有關。當樣本人數少時，又分成兩個次樣本，人數會更少，所得到的結果就更不穩定。

二、研究建議

(一) 對於測驗向度數評估方法之建議

當測驗資料在各向度能力值之間的相關較低或沒有相關時，因素分析的方法（PA 和 HULL）表現的很理想，而在各向度能力值之間的相關較高時，與 IRT 較有關聯的方法（DETECT 和 NOHARM）表現相對較好。在一般的情況下，一份測驗資料的真正向度個數是不存在的，但是要精確地表徵試題反應資料間主要關係之充分的向度數是需要

的 (Reckase, 2009)。因為向度數是人與試題之交互作用所得到的資料之特性，即使同一個測驗對不同樣本施測，所得到反應資料之向度數也未必會是一樣的。由本研究所得到的結果來看，研究者在判斷向度個數時，建議參考由多個程序所得到的結果，就像目前探索性因素分析的文獻中，也建議研究者不要僅依靠單一方法來下結論 (Fabrigar, Wegener, MacCallum, & Strahan, 1999)。

(二) 對未來研究之建議

1. 考慮納入其他測驗向度數評估方法與資料型式

目前文獻中被應用來檢測向度數之方法，不只是本研究所探討的 DETECT、NOHARM、PA 和 HULL 而已，如 Tate (2003) 以及 Svetina 和 Levy (2014) 所整理歸納的還有很多，例如 TESTFACT (參數的) 與 DIMTEST (非參數的)，甚至驗證性因素分析 (confirmatory factor analysis, CFA) 等。而且在本研究中，資料的模擬乃採用二參數的試題反應理論模式來產生答對/答錯的作答反應原始資料，然而在許多模擬研究中，研究者似乎僅模擬作答反應之共變數矩陣資料，由原始作答資料進行分析理論上會比用共變數矩陣多一些誤差。因此，模擬原始資料或是模擬共變數矩陣，是否為許多已發表的研究報告之發現有差異的原因之一，未來或許是可以討論的。

2. 擴充更多模擬情境變項

就模擬資料而言，操弄的模擬條件仍有擴展的空間。本研究模擬條件主要操弄的變項包括測驗的向度個數

(1d、2d、3d、3d1)、各向度能力值間的相關(0、0.3、0.6、0.9)、各向度內試題數(10、20題)、樣本大小(250、500、1000、2000人)等四項，產生104種模擬資料來進行正確模式判斷率的比較，但模擬條件無法窮盡所有可能的情境。以往測驗向度數評估方法的比較，大多數的研究都採用模擬近似簡單結構模式的測驗資料進行比較，本研究在二向度的模擬資料中採用的是1/3的試題具有複雜結構的型態；因此，未來進行的研究除了上述的四項操弄變項以外，可以考慮再增加其他的條件，也可以考量在模擬資料時，改變試題複雜結構的比例，甚至除了二元計分的資料以外，加入多元計分的資料。

3. 納入測驗的實徵資料

本研究所使用的是模擬資料，未來可以考慮納入實徵資料的分析，例如：國中基本學力測驗、國中教育會考測驗資料、大學學科能力測驗、大學統一入學測驗，或如TIMSS和PISA等跨國性調查的實徵資料，以使研究的發現更為完整。

參考文獻

- 陳榮華、吳明雄、陳心怡(2010)。新編多元性向測驗。台北：中國行為科學社。
- Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, 29(1), 67-91.

- Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Education Measurement: Issues and Practice*, 22(3), 37-53.
- Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1-23.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *2nd International Symposium on Information Theory* (pp. 267-281). Budapest: Akademiai Kiado.
- Ceulemans, E., & Kiers, H. A. L. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex HULL based method. *British Journal of Mathematical and Statistical Psychology*, 59, 133-150.
- Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Methods*, 70, 885-901.
- De Champlain, A. F., & Tang, K. L. (1993). *The effect of nonnormal ability distributions on the assessment of dimensionality*. Paper presented at the meeting of the National Council on Measurement in Education, Atlanta, GA.
- De Champlain, A. F., & Gessaroli, M. E. (1998). Assessing the dimensionality of item response matrices with small sample sizes

- and short test lengths. *Applied Measurement in Education*, *11*, 231-253.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299.
- Finch, H., & Habing, B. (2005). Comparison of NOHARM and DETECT in item cluster recovery: Counting dimensions and allocating items. *Journal of Educational Measurement*, *42*, 149-169.
- Fraser, C. (1988). *NOHARM II: A Fortran program for fitting unidimensional and multidimensional normal ogive models of latent trait theory*. Armidale, N.S.W.: University of New England, Centre for Behavioral Studies.
- Gessaroli, M. E. & De Champlain, A. F. (1996). Using an approximate chi-square statistic to test the number of dimensions underlying the response to a set of items. *Journal of Educational Measurement*, *33*, 157-179.
- Gierl, M. J., Leighton, J. P., & Tan, X. (2006). Evaluating DETECT classification accuracy and consistency when data display complex structure. *Journal of Educational Measurement*, *43*, 265-289.
- Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement*, *9*, 139-164.

- Hendriks, A. A. J., Hofstee, W. K. B., & De Raad, B. (1999). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences, 27*, 307–325.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185.
- Jang, E. E., & Roussos, L. A. (2007). An investigation into the dimensionality of TOEFL using conditional covariance-based nonparametric approach. *Journal of Educational Measurement, 44*(1), 1-21.
- Kim, H. (1994). *New techniques for the dimensionality assessment of standard test data*. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign, IL.
- Knol, D. L., & Berger, M. P. F. (1991). Empirical comparison between factor analysis and item response models. *Multivariate Behavioral Research, 26*, 457-477.
- Lorenzo-Seva, U., & Ferrando, P. J. (2013). FACTOR 9.2: A comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models. *Applied Psychologic Measurement, 37*(6), 497-498. DOI: 10.1177/0146621613487794.
- Lorenzo-Seva, U., & Rodríguez-Fornells, A. (2006). Acquiescent responding in balanced multidimensional scales and exploratory factor analysis. *Psychometrika, 71*, 769-777.

- Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. L. (2011). The HULL method for selecting the number of common factors. *Multivariate Behavioral Research, 46*, 340-364.
- McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34*, 100-117.
- McDonald, R. P., & Mok, M. C. (1995). Goodness of fit in item response models. *Multivariate Behavioral Research, 54*, 483-495.
- McDonald, R. P. (1996). Normal ogive multidimensional model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 257-269). New York, NY: Springer
- Muthén, B. (1983). Latent variable structural equation modeling with categorical data. *Econometrics, 22*, 48-65.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*(3), 396-402.
- 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. (2009). *Multidimensional item response theory*. New York, NY: Springer.
- Roussos, L. A., Ozbek, Y. O. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement, 43*(3), 215-243.

- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461-464.
- Stone, C. A. & Yeh, C.-C. (2006). Assessing the dimensionality and factor structure of multiple-choice exams: An empirical comparison of methods using the Multistate Bar Examination. *Educational and Psychology Measurement*, 66(2), 193-214.
- Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika*, 52, 589-617.
- Stout, W. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*, 55, 293-325.
- Stout, W., Habing, B., Douglas, J., Kim, H. R., Roussos, L., & Zhang, J. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychological Measurement*, 20, 331-354.
- Stout, W., Douglas, B., Junker, B., & Roussos, L. (1999). *DIMTEST* [Computer software]. The William Stout Institute for Measurement, Champaign, IL.
- Svetina, D. (2011). *Assessing dimensionality in complex data structures: A performance comparison of DETECT and NOHARM procedures* (Doctoral dissertation, Arizona State University). Retrieved from <http://repository.asu.edu/attachments/56763/content>
- Svetina, D., & Levy, R. (2012). An overview of software for conducting dimensionality assessment in multidimensionality models. *Applied Psychological Measurement*, 36(8), 659-669.

- Svetina, D., & Levy, R. (2014). A framework for dimensionality assessment for multidimensional item response models. *Educational Assessment, 19*, 35-57.
- Takane, Y. & de Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika, 52*(3), 393-408.
- Tanaka, J. S., & Huba, G. J. (1989). A general coefficient of determination for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical & Statistical Psychology, 42*(2), 233-239.
- Tate, R. (2003). A comparison of selected empirical methods for assessing the structure of responses to test items. *Applied Psychological Measurement, 27*, 159-203.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*(2), 209-220.
- Wang, M. (1985). *Fitting a unidimensional model to multidimensional item response data: The effect of latent space misspecification on the application of IRT* (Research Report MW: 6-24-85). University of Iowa, Iowa City, IA.
- Weng, L. J. & Cheng, C. P. (2005). Parallel analysis with unidimensional binary data. *Educational and Psychological Measurement, 65*(5), 697-716.
- Yu, C. H., Popp, S. O., DiGangi, S., & Jannasch-Pennell, A. (2007). Assessing unidimensionality : A comparison of Rasch modeling,

parallel analysis, and TETRAD. *Practical Assessment, Research & Evaluation*, 12(14). Available online:

<http://pareonline.net/getvn.asp?v=12&n=14>

Zhang, J. M. & Stout, W. (1999a). Conditional covariances structure of generalized compensatory multidimensional items. *Psychometrika*, 64(2), 129–152.

Zhang, J. M. & Stout, W. (1999b). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika*, 64(2), 213–249.

Zwick, W. R. & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432-442.

教育學誌 第四十期

2018 年 11 月，頁 181~214

學校組成特色和閱讀優勢對學生閱讀成績的 跨層次影響：以台灣 2011 年 PIRLS 資料進 行多層次研究

林政達*

國立成功大學 博士候選人

莊雙喜*

國立成功大學 副教授

摘要

本研究的目的是針對學校層級如居住區域，家長對教育的投入 (PI) 和教師自我效能 (TSE) 與學生層級如學生的閱讀優勢 (PRA) 對閱讀成績的影響進行研究調查。從台灣 2011 年 PIRLS 數據庫中選出了 147 個學校 (學校層級) 和 4,217 個匹配的四年級 (學生層級) 的資料庫。我們採用階層線性建模來分析本研究中的數據。結果表明，在學生層級上，認知閱讀優勢對學生閱讀成績有積極影響。在學校層面上，家長與教師自我效能對學生閱讀成績有重要的影響。此外，本研究中，居住區域，家長投入和教師自我效能並無調節效果。

*通訊作者：林政達，聯絡方式：e-mail：johna02361@gmail.com

*通訊作者：莊雙喜，聯絡方式：e-mail：johna0236@yahoo.com.tw

關鍵字：感知閱讀優勢、居住區域、家長對教育的投入、
教師自我效能、閱讀成就、PIRLS

The Cross-Level Effects of School Characteristics and Perceived Reading Advantage on Student Reading Achievement: Multilevel Study of Taiwan PIRLS 2011

Cheng-Ta Lin

PHD candidate,

National Cheng Kung University

Shuang-Shii Chuang

Associate professor,

National Cheng Kung University

Abstract

The purpose of this study was to investigate the impacts of area, parental input (PI), and teachers' self-efficacy (TSE) according to school characteristics, as well as students' perceived reading advantage (PRA) as it relates to reading achievement. A two-level database of 147 schools (macro-level) and 4,217 matched fourth graders (micro-level) was selected from the Taiwan PIRLS 2011 database. We adopted hierarchical linear modeling to analyze the data in this study. The results indicated that at the micro-level, perceived reading advantage had a positive and significant influence on students' reading achievement. At the macro-level, parental input and teachers' self-efficacy had significant effects on student reading achievement. In

*Corresponding Author: Cheng-Ta Lin ; E-mail: johna02361@gmail.com

*Corresponding Author: Shuang-Shii Chuang ; E-mail: johna0236@yahoo.com.tw

addition, area, parental input and teachers' self-efficacy exhibited no moderator effects in this study. According to the research findings, the researchers illustrated issues pertaining to suggestions for future research.

Keywords: perceived reading advantage, area, parental input, teachers' self-efficacy, reading achievement, PIRLS

Introduction

Reading competence has been viewed as the key to knowledge acquisition. Students, who read diversified words and pictures, not only acquire messages and build up individual knowledge systems, but also improve their reading achievement. With the purpose of understanding student reading performance, many governments in the world have focused on cultivating reading competence. Countries show great interest in participating in two international reading assessments, PISA (b for International Student Assessment) and PIRLS (Progress in International Reading Literacy Study). The results of these two assessments are provided to participating countries as a reference for promoting reading literacy. According to the results of the PIRLS 2011, the average score of Taiwanese students was 553, which ranked it ninth out of forty-nine countries; however, it ranked last among a group of four Asian countries (Shanghai, Hong Kong, Singapore, and Taiwan).

In this study, the Taiwan PIRLS 2011 database was adopted with the purpose of analyzing the influence of perceived reading advantage and achievement via school characteristics. Furthermore, the interactions between perceived reading advantage and school characteristics and two different levels of achievement are investigated. Owing to interviewing results of teachers through the action research, we discovered that living area, parenting support and teachers' self-efficacy act as important components toward learning achievement. The first step of cultivating reading habits is to make students fond of

reading. Rasinski (1994) indicated that if parents have a habit of reading, their children will pay greater attention to reading activities. Furthermore, teachers' self-efficacy creates a positive learning climate for students, and it influences students learning attitude and achievement as well. Moreover, selecting a residence near high-performance schools is also one of major considerations from a parental perspective.

The motive of this study focuses on the relationship of perceived reading advantage and school characteristics toward achievement. Secondly, from the perspective of school characteristics, area, parental input, and teachers' self-efficacy are the other keys to improving achievement. Therefore, discussing the relationship between school characteristics and achievement is the second motive in this study. Third, we adopted the PIRLS database from the IEA (The International Association for the Evaluation of Educational Achievement). The specialty of the large database is a consideration of multi-level issues such as students/ schools/ areas/ countries. In addition, the structure of the database comprises nested characteristics; therefore, multi-level issues cannot be ignored.

This study includes an analysis and investigation of variables at two different levels: one is perceived reading advantage in level 1, and the other is school characteristics in level 2. Additionally, another focus of the study is whether the interaction between these variables exhibits a cross-level moderator effect. Finally, we use the Taiwan PIRLS 2011 database to examine the influence of perceived reading advantage and

school characteristics toward reading achievement from cross-level angles. We plan to provide the results and suggestions to education departments with the goal of improving reading literacy in the younger generation.

This study is aimed at investigating the influence of perceived reading advantage and reading achievement via school characteristics. The data were from the Taiwan PIRLS 2011 database. We intend to explore if school characteristics (level 2) and perceived reading advantage (level 1) show cross-level interaction related to students' reading achievement. There are three research questions as follows:

1. Does perceived reading advantage affect students' achievement?
2. Do school characteristics affect students' achievement?
3. Do perceived reading advantage and school characteristics have a mediator effect on achievement?

Literature review

Achievement

According to Merriam-Webster, achievement is defined as something that has been done or achieved through effort. Additionally, reading achievement facilitates in-depth understanding of the meaning achievements derived from learning itself. Bernard, Brauer, Abrami, and Surkes (2004) developed an effective predictive index to measure achievement. Gardner (1999) defined intelligence as one type of

achievement used to integrate knowledge toward problem solutions. Wolters and Hussain (2015) argued that grit, which has been suggested to be a higher order personality trait that is highly predictive of both success and performance can be used to predict students' academic achievement as well, and they also suggested that it is distinct from other similar traits such as conscientiousness. However, Gardner and Hatch (1989) suggested that IQ tests, which rely on paper-and-pen examinations, can also be used to evaluate competences in linguistic and logical-mathematical areas. Finally, A performance goal orientation means students' grades outperform than peers (Ames, 1992; Nicholls, Cheung, Lauer, & Patashnick, 1989). The measurements might comprise of reading for grades, the desire to be favorable students of the teacher, and the desire to outperform others in reading(Baker & Wigfield, 1999). Therefore, grades and competition might act as important driving force in their early school life to make students put more efforts in study for better learning achievement. According to the literature review, we define reading achievement in this study according to the definitions referenced in the literature and aim this study toward an understanding of overall reading competence, the purpose of literacy, the use of reading to as a source of information, the interpretation process, and straightforward process.

Perceived Reading Advantage

Winter (1973) suggested behavior to be correlated with a motivation to obtain power. Brunstein, Schultheiss, and Grässman

(1998) suggested that motive is the driving force in pursuing personal goals. Thus, both personal goals and behavior appear to have specific connections with motive. However, in the earlier ambiguous stage before constructing a motive, people only perceive the advantages they will acquire from a process. Therefore, we refer to the definition of advantage given in the Cambridge Dictionary, which is “a condition giving a greater chance of success.” Roger et al. (2005) regarded advantage as an innovative concept making one superior to others. Previous studies have indicated that advantage is the most influential agent by which to predict the associations between attitude and intention (Lean, Zailani, Ramayah, & Fernando, 2009) and acceptance (Plouffe, Hulland, & Vandenbosch, 2001). Because this study focuses on reading achievement, we argue that reading is a specific interpretation of something. Therefore, we define perceived reading advantage as when an act of reading leads to a greater opportunity for success at the early stages of forming a motive. For instance, reading animal books helps us to investigate the eating habits and habitats of animals; reading fiction broadens our imagination; reading science books helps us understand theories and engages us in the creation of prototypes, and finally, learning language leads students toward an exploration of a foreign world (Kim & Krashen, 1997). Thus, people obtain many benefits reading heterogeneous books. In sum, perceived reading advantage is an important driving force. It not only encourages students to engage in reading diversified books and cultivating heterogeneous knowledge, but also reinforces the application of learned

individual, integrated modules of knowledge into practical life in such a way as to support future success.

School Characteristics

Many students struggle with studying (Shobana & Cartledge, 2004). However, some research has indicated that school characteristics, which include area, parental input, and teachers' self-efficacy, affect students' achievement. These parameters are described in more detail as follows:

- (1) Area: Students in urban schools with a high composition of children from minority groups and economically disadvantaged households are faced with the problem of children who are failing to read (Shobana & Cartledge, 2004). However, Brasington and Haurin (2009) indicated that parents show great interest in purchasing houses in areas with high performing schools. This suggests a strong connection between better residential districts and high-performing schools. In conclusion, students living in better living districts usually study in high-performance schools, and it appears that excellent reading achievement is typically an outcome of this residential choice.
- (2) Parental input: Leibowitz (1977) illustrated that parental input, such as parental education, occupation and income, is related to children's achievement and future earnings. Flouri (2006) indicated that parental involvement is significantly

related to educational attainment. Houtenville and Conway (2008) stated that parental effort in areas including child, parent, household, and school characteristics, has a strong positive effect on achievement. Shobana and Cartledge (2004) stated that parents who are concerned about the effects of difficulties with and progress in learning as they relate to children's reading ability and comprehension are providing the kind of parental support and involvement that helps construct reading strategies and skills that serve prerequisite roles in reading success. Finally, Nussbaum, Berner, and Brady (2005) indicated that parental measures have connections with child achievement. Accordingly, parental education, occupation, household income, and parent, household, and school characteristics have been shown to be the key agents influencing achievement. However, this study takes into consideration the relationships between parental support and learning. It is posited that with parental input, it will be easier for children to improve both reading achievement and reading literacy.

- (3) Teachers' self-efficacy: Bandura (1982) indicated that efficacy influences how we think, feel, and implement projects. Thus, effective lecturers are equipped with sound, diversified teaching skills. Sharma, Forlin, and Loreman (2008) suggested that teachers should be trained to teach students with a variety of learning tools such as board games, interactive software, activities, field trips, teamwork,

brainstorming, and storytelling. Other studies have found that the more learners engage in interaction with instructors, the better learning outcomes they will have (Garrison & Cleveland-Innes, 2005; So & Brush, 2008). Kang and Im (2013) suggested that learning assistance, instructional communication, and instructional support can predict learning outcomes in online learning environments. Hall, Sabey, and McClellan (2005) argued that effective text structural strategies promote expository text comprehension. Therefore, teaching competencies and instructional programs are key factors predicting student achievement (Elam, 1971). In short, curriculum design, activities, and teaching techniques are strongly related to achievement. However, in this study, we are more concerned with teachers' job satisfaction, the degree to which they understand the goals of the curriculum, and the degree to which they believe they have met the goals of the curriculum they are teaching.

Owing to the literature review of school characteristics mentioned above, we explore connections among area, parental input, and teachers' self-efficacy with reading achievement. The items in these three areas are as follows: 1) area: district differentiation from urban to rural; 2) parental input: parental support and involvement, and 3) self-efficacy: teachers' job satisfaction, understanding of the goals of the curriculum they are teaching, and their perceptions of their ability to meet those goals.

Methodology

Research Framework

The research framework is shown as Figure 1. The variables are divided into two different levels. Reading advantage is the independent variable, and achievement is the dependent variable in Level 1. The variable of school characteristics is in level two. H1 represents the significantly direct effect between perceived reading advantage and achievement. H2 represents the significantly direct effect between school characteristics and achievement. H3 is used to investigate whether school characteristics have a moderator effect between reading advantage and achievement.

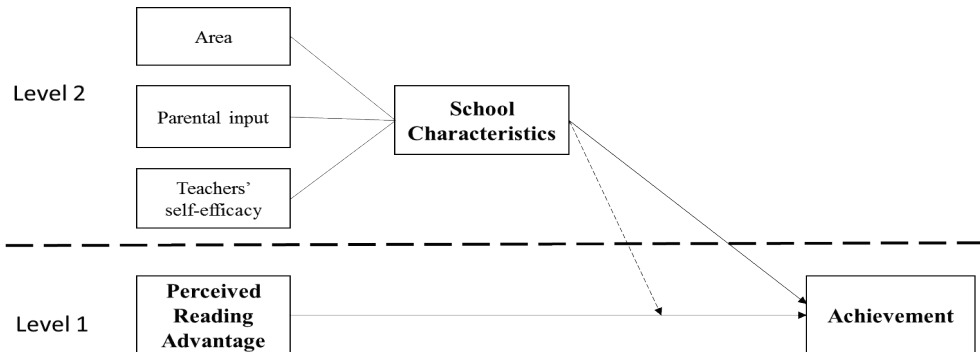


Figure 1. Research framework

The research variables used in this study were divided into in two different levels. Perceived reading advantage and achievement were in level 1, and school characteristics were in level two. Based on the

previous literature, we proposed the research framework as shown above in Figure 1 and hypotheses as follows:

- H1: If student perceive themselves to have a greater reading advantage, they will perform better on achievement tests.*
- H2a: If students live in better areas, they will perform better on achievement tests.*
- H2b: Children who receive more input from their parents will exhibit higher levels of achievement on tests.*
- H2c: If teachers are more confident when lecturing, students will perform better on tests.*
- H3a: Area has a moderator effect between reading advantage and achievement.*
- H3b: Parental input has a moderator effect between reading advantage and achievement.*
- H3c: Teachers' attitude has a moderator effect between reading advantage and achievement.*

Sampling

The research purpose was to assess the reading achievement of fourth graders in Taiwan. The data were from the Taiwan PIRLS 2011 (IEA, 2013), and the samplings were implemented in two stages through a Taiwanese population ratio. The first stage was school samplings by probability proportional to size (PPS). The rate at which schools were drawn was based on the number of students in each school. Then, in the second stage, the class was viewed as a unit to

extract a fixed number of students to participate in assessments in the selected schools. A total of 4,217 students and 147 principals from 147 different schools filled out questionnaires (see Table 1).

Table 1.

Variables and descriptive statistics

Variable	N	Mean	SD	Minimum	Maximum
Reading Advantage ^a	4217	3.42	0.68	1.00	4.00
Achievement ^a	4217	556.14	60.67	306.15	730.92
Area ^b	147	3.37	1.25	1.00	5.00
Parental input ^b	147	3.80	0.57	2.50	5.00
Teachers' self-efficacy ^b	147	4.02	0.48	2.67	5.00

^a level 1 variables; ^b level 2 variables. ◦ Level 1 (N=4217), Level 2 (N=147) ◦

Variables

Achievement is the outcome variable at level 1 in this study. The PIRLS has two types of questions: multiple choice and questions and answers. Each multiple-choice question has only one correct answer with one score. However, questions and answers will give students different scores from one to three, depending on their comprehension ability. The reliability in achievement was 0.98.

Perceived reading advantage is the only independent variable at level 1, and it contained three items using Likert four-point scales. Students could score from one to four. A higher score indicated a greater possibility that they perceived reading advantage. The items in

the questionnaire are listed as follows: like to read things that make them think, learn a lot from reading; books help to imagine other worlds. The reliability in perceived reading advantage was 0.74.

Parental input was the independent variable at level two, and it included four items using Likert four-point scales. Students could score from one to four. The items in the questionnaire are listed as follows: parental support and involvement. The reliability in parental input was 0.78.

Teachers' self-efficacy is the independent variable at level two, which comprised of three items using Likert four-point scales. Teachers could score from one to four. The items in the questionnaire are listed as follows: teaching job satisfaction, the degree to which teachers understand the goals of their curriculum, and their perception of whether or not they successfully met those goals. The reliability in teachers' self-efficacy was 0.78.

Results and Analysis

Descriptive statistics

Table 2 lists the information for the mean, standard error, and correlations in the detailed descriptive statistics. In general, the correlation between perceived reading advantage and achievement was 0.263 ($p < 0.001$). There was a good model fit based on the confirmatory factor analysis in this study, for which the values were NFI=0.99, RFI=0.98, IFI=0.99, TLI=0.99, CFI=0.99 and RMSEA=0.028. The

factor loadings were 0.67, 0.77 and 0.67 in perceived reading advantage; 0.98, 0.95, 0.98, 0.95 and 0.97 in reading achievement; 0.61, 0.67, 0.76 and 0.74 in parental input; 0.72, 0.85 and 0.70 in teachers' self-efficacy.

The convergent validity values were 0.74, 0.98, 0.79, and 0.80. The AVE values were 0.50, 0.93, 0.50, and .057. These figures were higher than those required for composite reliability and an average variance of 0.6 and 0.5 (Fornell, 1981), respectively. Hair argued that the square root of the AVE should be at least 75% higher than the correlation coefficients among the constructs (Hair, Anderson, Tatham, & Black, 1998). The square root of the AVE values 0.70, 0.96, 0.69, and 0.76 met these requirements, as shown in Table 2; thus, the constructs showed good discriminant validity.

Table 2.

Descriptive statistics

Variables	Mean	SD	PRA	ACH	PI	TSE
Perceived Reading Advantage ^a (PRA)	3.42	0.68	0.70			
Achievement ^a (ACH)	555.75	60.79	0.263***	0.96		
Parental input ^b (PI)	3.80	0.56	0.005	-0.115	0.69	
Teachers' self-efficacy ^b (TSE)	4.02	0.48	0.036	0.083	0.532***	0.76

^a level 1 variables; ^b level 2 variables; Co-relation of level 1 and 2; Level 1 (N=4217); Level 2 (N=147).

* $p < 0.05$. *** $p < 0.001$.

Hierarchical linear modeling analysis

I. Null model

In cross-level research, researchers have to examine if a cross effect exists. In an HLM study, variances in the number of dependent variables can be dismantled as the number of variances within the group and between groups. First, researchers must confirm that the variation between groups is significant and then proceed to the cross-level analysis. In order to examine the variation effect between groups in terms of achievement, we first adopted the null model. The model is as follows:

$$\text{Level 1: Achievement}_{ij} = \beta_{0j} + \gamma_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \mu_{0j}$$

Table 3 indicates that the variations in achievement ($\tau_{00}=421.101$) were significant ($\chi^2= 668.103$, $df=146$, $\rho<0.001$) among schools. It showed that the reading achievement of students was significant in different schools. In addition, the variance within the group was 3279.624. Thus, $ICC1= 421.101/(421.101+3279.624)=0.1137$. This suggests that there was an 11.37% variance among schools. According to the argument, if the ICC1 figure is greater than 0.058, the variation between groups cannot be ignored (Cohen, 1988). Therefore, when analyzing achievement, we had to consider variations between groups.

Table 3.

Null Model				
Fixed effect	Coefficient	Standard error	T-Ratio	P value
Achievement , γ_{00}	554.855	1.911	290.302	0.000
Random effect	Variance component	df	χ^2	P Value
Intercept1 , μ_{0j}	421.101	146	668.103	0.000
Level-1 , γ_{ij}	3279.624			

II. One-way ANCOVA model with random effects

When the adoption of the HLM analysis was confirmed, we had to examine the hypotheses in this study. First, we examined only one variable, perceived reading advantage, in level 1 and investigated its relationship with reading achievement. The model was as follows:

$$\text{Level 1: Achievement}_{ij} = \beta_{0j} + \beta_{1j} (\text{Perceived reading advantage}) + \gamma_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j}$$

Table 4 shows the results of perceived reading advantage ($\gamma_{10} = 21.393$, $t = 17.601$, $\rho < 0.001$). The results suggest that perceived reading advantage had a positive influence on reading achievement. The results also suggested that perceived reading advantage is an effective factor by which to predict reading achievement. That is to say, adding each unit in perceived reading advantage will increase the achievement scores by 21.393 points. Therefore, higher levels of perceived reading advantage led to better achievement on the tests. Thus, H1 was fully supported. Hence, we found that perceived reading advantage explains 7.54% $[(3279.624 - 3032.274) / 3279.624]$ of variations among students.

Also, in terms of random effect, τ_{00} was 364.417 ($\chi^2=628.633$, $\rho<0.001$), which indicated a positive significance between groups. Owing to the fact that different schools have different intercepts, we inferred that the variables in level two might have a direct influence on achievement.

Table 4.

One-way ANCOVA model with random effects

Fixed Effect	Coefficient	Standard error	T-Ratio	P value
Achievement , γ_{00}	555.932	1.794	309.740	0.000
Perceived Reading Advantage , γ_{10}	21.393	1.215	17.601	0.000
Random Effect	Variance component	df	χ^2	P value
Intercept1 , μ_{0j}	364.417	146	628.633	0.000
Level-1 , γ_{ij}	3032.274			

III. Regression model with intercepts and slopes- as-outcome

We examined the relationship of perceived reading advantage, area, parental input and teachers' self-efficacy toward students' reading achievement. The model was as follows:

$$\text{Level 1: Achievement}_{ij} = \beta_{0j} + \beta_{1j} (\text{Perceived reading advantage}) + \gamma_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Area}) + \gamma_{02} (\text{Parental input}) + \gamma_{03} (\text{Teachers' self-efficacy}) + \mu_{0j}$$

$$B_{ij} = \gamma_{10} + \mu_{1j}$$

Table 5 shows that area significantly influenced achievement ($\gamma_{01}=9.912$, $t=7.853$, $\rho<0.001$). Thus, it can be inferred that area can directly influence achievement. These results suggest that area is an effective

factor by which to predict achievement. Thus, adding each unit in area will increase the reading achievement scores by 9.912 points. Parental input significantly influences achievement as well ($\gamma_{02}= 7.509$, $t= 2.274$, $\rho<0.05$). Thus, parental input could directly influence achievement. The results suggested that parental input is an effective agent by which to predict achievement. Then, adding each unit in parental input will increase the scores in reading achievement by 7.509 points. Finally, teachers' self-efficacy significantly influenced achievement ($\gamma_{02}= -7.546$, $t= -2.030$, $\rho<0.05$). Thus, it can be inferred that teachers' self-efficacy can directly influence reading achievement. The result suggested that teachers' self-efficacy is an effective factor by which to predict reading achievement. Then, adding each unit in teachers' self-efficacy will decrease scores in reading achievement by 7.546 points; that is to say. H2a, H2b and H2c were fully supported. Additionally, in terms of random effects, we further examined the intercept variance component and found it to be significant ($\tau_{00}= 188.665$, $\chi^2=388.793$, $\rho<0.001$). In addition, area, parental input, and teachers' self-efficacy explained 55.19% $[(421.101-188.665)/421.101]$ of the variations among schools. Area, parental input, and teachers' self-efficacy had a significant influence on achievement among the 147 different schools. The results suggest that there are some variables that can be considered critical factors that affect reading achievement in level two. Future studies can explore other factors influencing achievement from a school perspective.

Table 5.

Regression model with intercepts and slopes- as-outcome				
Fixed Effect	Coefficient	Standard error	T-Ratio	P value
Achievement , γ_{00}	555.739	1.416	392.242	0.000
Area , γ_{01}	9.912	1.262	7.853	0.000
Parental input , γ_{02}	7.509	3.302	2.274	0.024
Teachers' self-efficacy , γ_{03}	-7.546	3.717	-2.030	0.044
Perceived reading advantage , γ_{10}	21.056	1.203	17.502	0.000
Random Effect	Variance component	df	χ^2	P value
Intercept1 , μ_{0j}	188.665	143	388.793	0.000
Level-1 , γ_{ij}	3032.042			

IV. Full model

We examined whether or not the cross-level interaction through the variables in level 1 and level 2 influenced reading achievement. The model was as follows:

$$\text{Level 1: Achievement}_{ij} = \beta_{0j} + \beta_{1j} (\text{Perceived reading advantage}) + \gamma_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Area}) + \gamma_{02} (\text{Parental input}) + \gamma_{03} (\text{Teachers' self-efficacy}) + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} (\text{Area}) + \gamma_{12} (\text{Parental input}) + \gamma_{13} (\text{Teachers' self-efficacy}) + \mu_{1j}$$

Table 6 indicates that the interaction effect of area, parental input and teachers' self-efficacy showed no significant influence on reading achievement ($\gamma_{11} = -0.301$, $t = -0.313$, $\rho > 0.05$); ($\gamma_{12} = 3.284$, $t = 1.119$, $\rho > 0.05$); ($\gamma_{13} = -4.795$, $t = -1.611$, $\rho > 0.05$). Hence, H3 was not supported.

Table 6.

Full model

Fixed Effect	Coefficient	Standard error	T-Ratio	P value
Achievement , γ_{00}	555.722	1.420	391.166	0.000
Area , γ_{01}	9.962	1.257	7.920	0.000
Parent input , γ_{02}	7.477	3.297	2.267	0.025
Teachers' self-efficacy , γ_{03}	-7.539	3.718	-2.027	0.044
Perceived Reading Advantage , γ_{10}	21.089	1.190	17.721	0.000
Area , γ_{11}	-0.301	0.964	-0.313	0.754
Parent input , γ_{12}	3.284	2.936	1.119	0.264
Teachers' self-efficacy , γ_{13}	-4.795	2.976	-1.611	0.107
Random Effect	Variance component	df	χ^2	P value
Intercept1 , μ_{0j}	188.637	143	388.361	0.000
Level-1 , γ_{ij}	3032.552			

Figure 2 indicates the relationships of the cross-level interaction between perceived reading advantage and parental input toward achievement. Even though there was no moderator effect (Baron & Kenny, 1986) between perceived reading advantage and achievement via parental input; too much parental input eventually resulted in a negative influence on achievement, and we could clearly observe that the intercepts of the two regression lines were different. That is to say, the more input parents offered to their children, the greater negative influence occurred between perceived reading advantage and achievement in these students on their test results.

In terms of the cross-level interaction, the regression lines shown in Figure 3 represent the relationship between perceived reading advantage and teachers' self-efficacy via achievement. Although there was no moderator effect between perceived reading advantage and achievement via teachers' self-efficacy. Arrogance on the part of

teachers eventually led to a negative influence on student achievement, and we could clearly observe the different intercepts in the two regression lines. That is to say, the more arrogant attitude that teachers exhibited while lecturing students, the more negative influence there was between perceived reading advantage and achievement in their students based on the examination results.

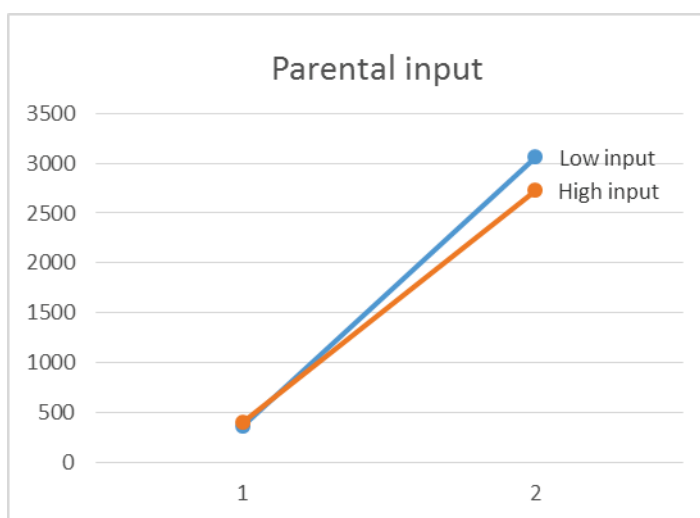


Figure 2. Cross-level interaction between perceived reading advantage and parental input and achievement



Figure 3. Cross-level interaction between perceived reading advantage and achievement via teachers' self-efficacy

Conclusions

This study investigated perceived reading advantage and school characteristics as they relate to achievement through the HLM. We used the Taiwan PIRLS 2011 database, which included 4217 fourth graders and 147 elementary schools to conduct a cross-level analysis. Owing to the fact that the hierarchical linear model involved students and school level, data integration became critically important. The following results were obtained in response to the research questions.

Perceived reading advantage and students' achievement

From the perspective of student level, the results indicated that perceived reading advantage had a positive and significant influence on reading achievement in the fourth graders under consideration; that is to say, greater reading advantage as perceived by the students led to better reading achievement based on the test scores. This meant that adding each unit in perceived reading advantage increased reading achievement scores by 21.393. From the results of student level, it could be clearly understood that there is a relationship between perceived reading advantage and achievement. If students perceive the advantages of reading, administrative staff in the schools should rearrange reading classes into the formal school schedules. Therefore, students will develop a habit of reading. This result is in line with those of previous studies (Davis, 1989; Nysveen, Pedersen, & Thorbjørnsen, 2005).

School characteristics and students' achievement

We proved that area has a positive influence on achievement; that is to say, if students live in a better residential area, they will have better reading achievement based on test scores. This meant that students studying in the urban school increased reading achievement scores by 9.912 points. The results are in line with prior study (Shobana & Cartledge, 2004). Furthermore, parental input was also verified to have a positive influence on reading achievement; that is to say, the more input parents gave to the children, the better their grades were on

reading achievement tests. This meant that adding each unit in parental input increased reading achievement scores by 7.509 points. The results are according with previous study (Flouri, 2006). Third, teachers' self-efficacy was proven to have a negative influence on achievement, which is against the hypothesis of this study; that is to say, the more confident teachers felt about their teaching techniques, the worse the children performed on reading achievement tests. This meant that adding each unit in teachers' self-efficacy decreased reading achievement scores by 7.546 points. However, the results are in line with other study (Pintrich & de Groot, 1990). However, we found no moderator effects between perceived reading advantage and achievement via area, parental input, and teachers' self-efficacy.

A sound living environment, proper parental input and good teaching attitude were all shown to be important factors influencing reading achievement. However, an environment that is too comfortable spoils children and makes them lose their desire to learn. Too much support from parents deprives children of learning opportunities. Arrogant teachers lack a positive attitude related to pursuing the latest knowledge and delivering appropriate material based on the curriculum. The teaching attitude might lead to students poor performance such as learning achievement. Table 7 indicates current Taiwan status of school characteristics and perceived reading advantage in student level.

Table 7.

The information of school characteristics and perceived reading advantage of students in Taiwan

	Issues	Results
Area	Studying in urban school	Better performance
Parental inputs	Committee participation	Finance and suggestions support
	Curriculum assistance	Expertise knowledge
	Parent-teacher communication	Differential lecturing
Teacher' self-efficacy	Arrogant	Never upgrade teaching skills
	Inertial learning attitude	Lack of latest knowledge
	Lack of evaluation system	Driving force shortage
Perceived reading advantages	Students' self-expectation	Knowledge acquisition
	Parental expectation	Pleasing parents
	Interesting curriculum	Perceived enjoyment in learning

According to the results mentioned, principals should focus on designing interesting curriculum to increase students' motivation, inviting parents to share their expertise in extra curriculum, and polishing teachers' lecturing skills to deliver students clear teaching goals on the purpose of improving students' achievement.

This study had three relevant findings. First, perceived reading advantage in level 1 had a significant influence on achievement. Second, in terms of area, parental input and teachers' self-efficacy in level 2 had significant influence on achievement. Finally, there were no moderator effects found in the study.

Contributions

- The most important finding was the effectiveness of a strategy involving collecting group and individual levels for investigating the effects in the different levels. Overall, we found that perceived reading advantage and area, parental input,

teachers' self-efficacy, and school characteristics were important agents affecting reading achievement.

Implications

Thus, the following suggestions are offered. First, teachers should inspire students and help them understand the importance of reading. For example, reading makes students think; reading provides diversified information, and reading helps students imagine other worlds and cultures. More importantly, learning helps students think and create solutions to problems. Since students learn competence through perceived reading advantage, there will be a great improvement in their grades. Second, living in a sound environment helps children maintain a peaceful learning attitude. With a quiet learning environment and clear thoughts, students can concentrate on their studies and improve their reading achievement as well. Third, proper parental input does improve children's attitude toward learning. If students have a positive learning attitude incorporating the proper levels of parental input, they can endure the pressures associated with failure. Grit is one of the major reasons individuals can overcome pressure and walk on a path to success; that is to say, proper parental input plays an important role in reading achievement. Finally, in order to provide students with an interactive curriculum and interesting activities, hardworking teachers have to update their teaching techniques and acquire the latest heterogeneous knowledge from time to time.

References

- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261-271.
- Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34*(4), 452-477.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147.
- Baron, R. M., & Kenny, D. A. (1986). Moderator-Mediator Variables Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.
- Bernard, R. M., Brauer, A., Abrami, P. C., & Surkes, M. (2004). The Development of a Questionnaire for Predicting Online Learning Achievement. *25, 1*(31-47).
- Brunstein, J. C., Schultheiss, O. C., & Grässman, R. (1998). Personal goals and emotional well-being: The moderating role of motive dispositions. *American Psychological Association, 75*(2), 494-508.
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, & User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340.
- Elam, S. (1971). Performance Based Teacher Education. What is the State of the Art? *ERIC*.

- Flouri, E. (2006). Parental interest in children's education, children's self-esteem and locus of control, and later educational attainment: Twenty-six year follow-up of the 1970 British Birth Cohort. *British Journal of Educational Psychology*, 76(1), 41-55.
- Fornell, C. A. D. F. L. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*(8), 39-50.
- Gardner, H. (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century*: Basic Books.
- Gardner, H., & Hatch, T. (1989). Multiple intelligences go to school: Educational implications of the theory of multiple intelligences. *Educational Researcher*, 18(8), 4-10.
- Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: interaction is not enough. *American Journal of Distance Education*, 19(3), 133-148.
- Hair, J. F., Anderson, R. E., Tatham, R. I., & Black, W. C. (1998). *Multivariate data analysis with readings*. NY: Paulist press.
- Hall, K. D., Sabey, B. L., & McClellan, M. (2005). EXPOSITORY TEXT COMPREHENSION: HELPING PRIMARY-GRADE TEACHERS USE EXPOSITORY TEXTS TO FULL ADVANTAGE. *Reading Psychology*, 26(3), 211-234.
- Houtenville, A. J., & Conway, K. S. (2008). Parental Effort, School Resources, and Student Achievement. *The Journal of Human Resource*, 43(2), 437-453.

- Kang, M., & Im, T. (2013). Factors of learnereinstructor interaction which predict perceived learning outcomes in online learning environment. *Journal of Computer Assisted Learning*, 29(3), 292-301.
- Kim, H. Y., & Krashen, S. (1997). Why Don't Language Acquire Take Advantage of the Power of Reading. *TESOL journal*.
- Lean, O. K., Zailani, S., Ramayah, T., & Fernando, Y. (2009). Factors influencing intention to use e-government services among citizens in Malaysia. *International Journal of Information Management*, 29(6), 458-475.
- Leibowitz, A. (1977). Parental Inputs and Children's Achievement. *The University of Wisconsin Press Journals Division*, 12(2), 242-251.
- Nicholls, J. G., Cheung, P., Lauer, J., & Patashnick, M. (1989). Individual differences in academic motivation: Perceived ability, goals, beliefs, and values. *Learning and Individual Differences*, 1, 63-84.
- Nussbaum, B., Berner, R., & Brady, D. (2005). The creative future. *Business review weekly*, 27(32), 58-62.
- Nysveen, H., Pedersen, P. E., & Thorbjørnsen, H. (2005). Intentions to Use Mobile services: Antecedents & Cross-Service Comparisons. *Journal of the Academy of Marketing Science*, 33(3), 330-346.
- Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40.

- Plouffe, C. R., Hulland, J. S., & Vandenbosch, M. (2001). Research Report: Richness Versus Parsimony in Modeling Technology Adoption Decisions—Understanding Merchant Adoption of a Smart Card-Based Payment System. *Information Systems Research, 12*(12), 208–222.
- Rasinski, T. V. (1994). *Fast Start: A parental involvement reading program for primary grade students*. New Orleans, LA.
- Roger, Y., Price, S., Randell, C., Fraser, D. S., Weal, M., & Fitzpatrick, G. (2005). Ubi-learning integrates indoor and outdoor experiences. *Communication of the ACM, 48*(1), 55-59.
- Sharma, U., Forlin, C., & Loreman, T. (2008). Impact of training on pre-service teachers' attitudes and concerns about inclusive education and sentiments about persons with disabilities. *Disability & Society, 23*(7), 773-785.
- Shobana, M. R., & Cartledge, G. (2004). Making Home an Advantage in the Prevention of Reading Failure: Strategies for Collaborating With Parents in Urban Schools. *Preventing School Failure: Alternative Education for Children and Youth, 48*(4), 15-21.
- So, H. J., & Brush, T. A. (2008). Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: relationships and critical factors. *Computers & Education, 51*(1), 318-336.
- Winter, D. G. (1973). The power motive. *American Psychological Association*.

Wolters, C. A., & Hussain, M. (2015). Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacognition Learning, 10*, 293-311.

國立南大學教育學系「教育學誌」徵稿要點

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
2018.09.26 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）、參考文獻、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、本要點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」

