

主題論文 Special Issue

教育心理、輔導與測評 Educational Psychology, Counseling, Testing & Assessment

- 1 進行多層次建模最小可行的樣本數建議：貝氏模擬取向
曾明基

Sample Size Requirements of Using Multilevel Models: Bayesian Simulation Study
Ming-Chi Tseng

- 27 量表長度簡化研究：「簡式中小學教師主觀幸福感量表」修訂
余民寧、陳柏霖、陳玉樺

Study of Scale-Items Reduction: The Reconstruction of Subjective Well-Being Scale
Min-Ning Yu / Po-Lin Chen / Yu-Hua Chen

- 57 國小學童正向情緒與創造力關聯之研究：以創造力傾向及創意自我效能為中介變項
蕭佳純

The Correlation between Elementary Pupils' Positive Emotion and Creativity with
Creativity Tendency and Creative Self-efficacy as The Mediators
Chia-Chun Hsiao

一般論文 Article

- 85 Teachers' Cognitive Processes from the Perspective of Distributed
Cognition: An Analysis Focused on Teachers-Made Assessments
Chin-Tu Liang

分布認知觀點的教師認知歷程：以教師自編評量為分析焦點
梁金都



2017年12月出版
季刊



定價250元 GPN 2009405238



教育研究與發展期刊

第十三卷第四期

Journal of Educational Research and Development Vol.13 No.4

國家教育研究院

教育研究與發展期刊

Journal of Educational
Research and Development

本期主題 | 教育心理、輔導與測評

2017年12月出版 Educational Psychology, Counseling, Testing & Assessment

【2017年臺灣社會科學核心期刊TSSCI第一級】

國家教育研究院

NATIONAL ACADEMY for EDUCATIONAL RESEARCH

目錄 第十三卷第四期

i	編輯委員
v	執行主編的話
	主題論文：教育心理、輔導與測評
1	進行多層次建模最小可行的樣本數建議：貝氏模擬取向 曾明基
27	量表長度簡化研究：「簡式中小學教師主觀幸福感量表」修訂 余民寧 / 陳柏霖 / 陳玉樺
57	國小學童正向情緒與創造力關聯之研究：以創造力傾向及創意自我效能為中介 變項 蕭佳純
	一般論文
85	分布認知觀點的教師認知歷程：以教師自編評量為分析焦點 梁金都
119	徵稿啟事
123	審稿辦法
125	《教育研究與發展期刊》投稿者基本資料表
126	授權書

Contents Vol.13 No.4

iii	Journal of Educational Research and Development
v	Words from the Executive Editor in Chief
	Special Issue: Educational Psychology, Counseling, Testing & Assessment
1	Sample Size Requirements of Using Multilevel Models: Bayesian Simulation Study Ming-Chi Tseng
27	Study of Scale-Items Reduction: The Reconstruction of Subjective Well-Being Scale Min-Ning Yu / Po-Lin Chen / Yu-Hua Chen
57	The Correlation between Elementary Pupils' Positive Emotion and Creativity with Creativity Tendency and Creative Self-efficacy as The Mediators Chia-Chun Hsiao
	Article
85	Teachers' Cognitive Processes from the Perspective of Distributed Cognition: An Analysis Focused on Teachers-Made Assessments Chin-Tu Liang
121	Call for Papers
125	Journal of Educational Research and Development Submission Form
127	Transfer of Copyright Agreement

教育研究與發展期刊 第十三卷第四期

- 發行人 許添明（國家教育研究院院長）
- 總編輯 許添明（國家教育研究院院長）
- 副總編輯 郭工賓（國家教育研究院副院長）
- 本期主編 林世華（國立臺灣師範大學教育心理與輔導學系兼任教授）

「教育心理、輔導與測評」品質促進小組

王文中（香港教育大學教育心理學系講座教授）
邱皓政（國立臺灣師範大學管理學院教授）
翁麗禎（國立臺灣大學心理系教授）
陳學志（國立臺灣師範大學教育學院院長）
陳秉華（國立臺灣師範大學教育心理與輔導學系兼任教授）
溫福星（東吳大學國際經營與貿易學系教授）

編輯委員

一、課程與教學類

卯靜儒（國立臺灣師範大學教育學系教授）
李隆盛（中臺科技大學校長）
范信賢（國家教育研究院課程及教學研究中心副研究員）
黃政傑（靜宜大學教育研究所終身榮譽教授）
黃炳煌（國立政治大學教育學系名譽教授）

二、教育政策與行政類

林子斌（國立臺灣師範大學教育學系副教授）
秦夢群（國立政治大學教育行政與政策研究所教授）
張鈿富（淡江大學教育政策與領導研究所教授兼教育學院院長）
曾大千（國家教育研究院教育制度及政策研究中心研究員）
楊國賜（亞洲大學經營管理學系講座教授）

三、教育心理、輔導與測評類

任宗浩（國家教育研究院測驗及評量研究中心副研究員）
余民寧（國立政治大學教育學系特聘教授）
林世華（國立臺灣師範大學教育心理與輔導學系兼任教授）
洪儷瑜（國立臺灣師範大學特殊教育系教授）
張郁雯（國立臺北教育大學教育學系教授）

四、教師培育與專業發展類

周愚文（國立臺灣師範大學教育學系教授）
林佩蓉（臺北市立大學幼兒教育學系副教授）
張德銳（輔仁大學師資培育中心教授）
陳竑濬（國立政治大學教育學院教授）
陳麗華（淡江大學課程與教學研究所教授）

Journal of Educational Research and Development

Vol.13 No.4 December 31, 2017

- **Publisher**

National Academy for Educational Research

- **Editor in Chief**

Tian-Ming Sheu President, National Academy for Educational Research

- **Vice Editor in Chief**

Kung-Bin Guo Vice President, National Academy for Educational Research

- **Executive Editor in Chief**

Sieh-Hwa Lin Adjunct Associate Professor, Department of Educational Psychology & Counseling, National Taiwan Normal University

Quality Improvement Team

Wen-Chung Wang	Chair Professor, Department of Psychological Studies, The Education University of Hong Kong
Haw-Jeng Chiou	Professor, Department of Business Administration College of Management, National Taiwan Normal University
Li-Jen Weng	Professor, Department of Psychology, National Taiwan University
Hsueh-Chih Chen	Professor & Dean, College of Education, National Taiwan Normal University
Ping-Hwa Chen	Adjunct Professor, Department of Educational Psychology & Counseling, National Taiwan Normal University
Fur-Hsing Wen	Professor, Department of International Business, Soochow University

Editorial Board

Derray Chang	Professor, Center of Teacher Education, Fu Jen Catholic University
Dian-Fu Chang	Professor & Dean of College of Education, Graduate Institute of Educational Policy and Leadership, Tamkang University
Yu-Wen Chang	Professor, Department of Education, National Taipei University of Education
Andy Chen	Professor, College of Education, National Chengchi University
Li-Hua Chen	Professor, Graduate Institute of Curriculum and Instruction, Tamkang University
Joseph M. Chin	Professor, Graduate Institute of Educational Administration and Policy, National Chengchi University
Yu-Wen Chou	Professor, Department of Education, National Taiwan Normal University
Hsin-Hsien Fan	Associate Researcher, Research Center for Curriculum & Instruction, National Academy of Educational Research
Ping-Huang Huang	Professor Emeritus, Department of Education, National Chengchi University

Li-Yu Hung	Professor, Department of Special Education, National Taiwan Normal University
Jenq-Jye Hwang	Lifetime Emeritus Professor, Graduate Institute of Education, Providence University
Lung-Sheng Steven Lee	President, Central Taiwan University of Science and Technology
Pei-Jung Lin	Associate Professor, Department of Early Childhood Education, University of Taipei
Sieh-Hwa Lin	Adjunct Associate Professor, Department of Educational Psychology & Counseling, National Taiwan Normal University
Tzu-Bin Lin	Associate Professor, Department of Education, National Taiwan Normal University
Chin-Ju Mao	Professor, Department of Education, National Taiwan Normal University
Tsung-Hau Ren	Associate Researcher, Research Center for Testing & Assessment, National Academy of Educational Research
Ta-Chien Tseng	Researcher, Research Center for Educational System & Policy, National Academy of Educational Research
Kuo-Shih Yang	Chair Professor, Department of Business Administration, Asia University
Min-Ning Yu	Distinguished Professor, Department of Education, National Chengchi University

執行主編的話

國立臺灣師範大學教育心理與輔導學系兼任教授 林世華

首先，恭賀國家教育研究院教育研究與發展期刊榮獲 2017 年 TSSCI 教育學門第一級期刊的殊榮。感謝本刊編輯委員與同仁的辛勤，投稿者和讀者的支持。希望舊雨新知能繼續支持和愛護本刊！

其次，「教育研究與發展期刊」13 卷 4 期主題為「教育心理，輔導與測評」，投稿者的論文經由嚴謹的審查過程與評閱，本期共收錄頗具研究參考價值的四篇論文。簡要說明如下：

第一篇為曾明基的〈進行多層次建模最小可行的樣本數建議：貝氏模擬取向〉。本文之研究是採取模擬研究法來比較貝氏方法和 ML 估計法在多層次模型以及成長模型建構時，所需最小可行的分析樣本單位數。研究結果顯示貝氏方法進行多層次模型以及成長模型建構，所需的樣本數較小。本論文的刊出將可帶動進一步的專業對話與更多確認研究的出現。

第二篇為余民寧、陳柏霖及陳玉樺的〈量表長度簡化研究：「簡式中小學教師主觀幸福感量表」修訂〉。本文之研究是將「主觀幸福感量表」修訂為「簡式中小學教師主觀幸福感量表」。研究結果顯示：簡式量表具有不錯的外在效標關聯效度。本文指出簡式量表發展的方向：如何在量表簡化過程中考慮與原量表的一致性，極小化簡化所造成的損失。


第三篇為蕭佳純的〈國小學童正向情緒與創造力關聯之研究：以創造力傾向及創意自我效能為中介變項〉。本文之研究是採取結構方程模式的方法，來了解創造力傾向與創意自我效能中介於正向情緒與創造力之間的間接影響效果。此類中介效果模式是教育與心理學研究正夯的重要課題之一，本期出版也正好趕上時尚的研究課題！

第四篇為梁金都的〈分布認知觀點的教師認知歷程：以教師自編評量為分析焦點〉。本文之研究是以某國小為研究個案，以學校實施「雙向細目表」的教師

自編評量為中介物，以了解教師的認知歷程。研究結果顯示：教師認知階段依次呈現質疑、反思、轉向、創新等，認知系統大致呈現朝向上升方向發展。本文以教學現場脈絡，兼具教學與評量雙主題，值得與現場教師分享。

最後，本期四篇論文得以出版，感謝論文投稿者認真撰稿，審查委員、品質促進小組委員及編委會編輯委員審慎評閱，以確保本期論文的最佳品質。同時也要感謝編輯行政團隊之全心全力付出與協助，使得本期得以順利出刊。

臺灣師大教育心理與輔導系退休



謹識

2017 年 12 月 10 日

進行多層次建模最小可行的樣本數建議： 貝氏模擬取向

曾明基 國立東華大學師資培育中心兼任助理教授

摘 要

本研究經由模擬研究的方式同時比較貝氏方法和 ML 估計法在多層次模型以及成長模型建構時，最小可行的分析樣本單位數，並同時考慮存在隨機遺漏下，在多層次模型以及成長模型建構所需的樣本數調整。研究發現，使用貝氏方法進行多層次模型以及成長模型建構，所需的樣本數較小且可以獲得穩定的參數覆蓋率以及統計考驗力，值得加以推廣。

關鍵詞：貝氏方法、多層次模型、成長模型



Sample Size Requirements of Using Multilevel Models: Bayesian Simulation Study

Ming-Chi Tseng

Adjunct Assistant Professor, Center for Teacher Education, National Dong Hwa University

Abstract

This paper shows practical guidelines of sample size requirements when results are analyzed by multilevel models. The study found that when Bayesian method is used for multilevel model, stable parameters and power are attained through fewer samples.

Keywords: Bayesian method, multilevel model, growth model



壹、緒論

多層次研究由於涉及到不同層次的分析，因此要考慮到複雜的抽樣問題，若以兩層次模型為例，在橫斷性研究中，多層次模型的兩個不同層次主要可區分為個體層次與總體層次，而在縱貫性研究中，多層次模型一般以成長模型表示，成長模型的兩個不同層次主要可區分為個體內層次與個體間層次。雖然多層次研究在橫斷性與縱貫性的名稱不同，但是概念一致。以教育學門為例，在橫斷性研究的兩層結構中就必須同時考慮從個體（學生）層次和總體（學校）層次抽多少樣本數，而在縱貫性研究的兩層結構中就必須同時考慮個體內（時間）層次和個體間（學生）層次的樣本數。

在橫斷性研究中，多層次模型可以下式表示：

$$\begin{aligned} y_{ij} &= B_{0j} + r_{ij} \\ B_{0j} &= r_{00} + u_{0j} \end{aligned} \quad (1)$$

其中 y_{ij} 代表個體（學生）層次 i 巢套在總體（學校）層次 j ， B_{0j} 是隨機截距， r_{ij} 和 u_{0j} 分別是個體層次和總體層次的誤差。

在縱貫性研究中，多層次模型可以下式表示：

$$\begin{aligned} y_{ti} &= B_{0i} + r_{ti} \\ B_{0i} &= r_{00} + u_{0i} \end{aligned} \quad (2)$$

其中 y_{ti} 代表個體內（時間）層次 t 巢套在個體間（學生）層次 i ， B_{0i} 是隨機截距， r_{ti} 和 u_{0i} 分別是個體內層次和個體間層次的誤差。由公式（1）和公式（2）可知，多層次模型在橫斷性研究和縱貫性研究雖然所探討的議題不同，但在公式的表答上沒有差異。

在多層次研究中，由於資料分析牽涉到跨層級的數據，因此樣本規模的決定相較於單一層次的分析更為複雜，畢竟樣本大小影響到統計推論的有效性。對於多層次模型的兩層結構而言，Kreft 與 Leeuw（1998）曾建議採用 30/30 準則，也就是總體層次不少於 30 組（如：30 間學校數），且每組不少於 30 人（如：30 個學生）來決定樣本規模。基本上 Kreft 等人所提出的 30 / 30 準則並不適用於縱貫性研究，畢竟在進行縱貫性研究時，研究者要重複蒐集每位受試者 30 次的觀測資料是相當困難的事，因此，前述原則基本上僅適用於橫斷性研究的多層次模型。此外，一般在大型資料庫中，可以蒐集 3 波到 4 波的資料已經相當不容易，因此

這個準則並不適用在縱貫性研究，在縱貫性研究中，由於個體內層次觀察的時間點不多，考慮到檢定力和參數估計穩定性，一般會將個體間層次的樣本數增大，也就是蒐集的学生數要多。可見，不論是進行多層次模型或是成長模型，合理的分析樣本數是需要加以考量的。

在多層次模型或是成長模型建構時，最大概似法（Maximum Likelihood, ML）是經常使用的方法，在充分大的樣本數時，ML 估計可以產生漸進有效性和一致性的估計值（Stoel & Garre, 2011）。然而在實證上，受限於研究經費及研究場域或樣本的特殊性，在橫斷性的多層次模型或縱貫性的成長模型建構上，研究者在進行多層次分析時常常無法收集到足夠的樣本數進行跨層級的模型建構。

有別於 ML 估計法所需要的大樣本數要求，近來，貝氏估計法極受矚目，研究人員若有先前的文獻或既得的研究結果可供參閱，在貝氏架構下可以將這些資料透過先驗分配使其後驗分配更符合實際的研究情況，一般認為這樣的估計方式可以在較小樣本數的多層次模型或是成長模型建構中獲得穩定的參數估計（Hox, van de Schoo, & Matthijsse, 2012; Muthén & Asparouhov, 2012）。

基於此，本研究旨在探討進行多層次模型以及成長模型建構時，所需的樣本數與檢定力的關係，並同時比較 ML 估計法和貝氏估計法在模型建構上所需合理樣本數的差異，供實證研究者進行多層次模型以及成長模型建構時合理樣本數的參考。

貳、文獻探討

本研究在多層次分析的樣本數決定上，主要聚焦於兩個議題，其中一個是參數估計正確性（accuracy）的考量，也就是需要多大的樣本數才可以估計出不偏的參數，而另一個為達成研究者所期望的檢定力時，需要多大的樣本數（Heck & Thomas, 2009），因此，對於多層次分析各層所需的樣本數如何配置，其重要性不言而喻。目前在國內、外多層次分析的實證以及模擬研究探討上（溫福星、邱皓政，2011；楊志堅、劉心筠、楊志強，2004；Hox, 2010; Maas & Hox, 2004, 2005; Kreft & De leeuw, 1998; Raudenbush & Bryk, 2002; Snijders, 2005），已有相關研究探討進行模型建構時，合理或可行的分析樣本單位數為何，供實證研究者參閱。然而，上述的模擬以及實證研究皆使用 ML 估計法進行說明，忽略了進行多層次

分析時，研究者可能在實證上會有樣本數不足的問題，而導致參數估計偏誤，模式無法收斂等。有別於 ML 估計法，本研究在後續也將介紹貝氏估計法供實證研究者參閱，並同時比較兩種不同估計法在進行多層次分析時，所需樣本數的差異。

一、多層次分析的合理樣本數決定

對於使用多層次模型進行橫斷性或是縱貫性研究，各層所需的樣本數如何配置，向為研究的重大挑戰。有關不同樣本規模對參數估計正確性的影響，在模擬研究部分，Maas 與 Hox（2005）在一系列條件下探討對於兩層結構的多層次模型參數估計正確性的影響，其在模擬研究中操弄總體層次個數（ $N = 30、50、100$ ）、個體層次大小（ $N = 5、30、50$ ）以及組內相關（Intraclass correlation coefficient, ICC）（ $ICC = .1, .2, .3$ ），模擬結果顯示固定參數估計值（截距與斜率）的偏誤微不足道。其次，就總體層次樣本數對標準誤的影響而言，Maas 與 Hox 發現總體層次樣本數對於固定效果的迴歸係數的標準誤的影響是小的，當總體層次樣本數為 30 時，斜率的未覆蓋率（Coverage）為 6.0%，而截距的未覆蓋率為 6.4%，與常用的未覆蓋率 5% 相差不大，這樣的差異是可以接受的。此外，個體層次大小在三種不同條件下的 95% 信賴區間覆蓋率都有良好的表現。當個體層次樣本數變大時，對於參數覆蓋率會有所改善，但增加總體層次樣本數改善幅度較佳，而在三種不同 ICC 條件下，固定參數的覆蓋率明顯優於變異數的覆蓋率。此外，Maas 與 Hox（2004）也指出，當總體層次常態分配的誤差假設違反時，非常態的殘差會對固定效果的參數估計產生影響，當總體層次的群組個數越多時，則會有較穩定的參數估計，但若增加個體層次的樣本數並不能改善估計結果，為了得到不偏的固定參數估計值，合理的總體層次樣本數應為 50 組。Hox（2010）進一步指出，在多層次分析中針對總體層次以及跨層級的交互效果進行分析時，統計檢定力主要依賴於總體層次樣本數而非總樣本數。此外，Snijders（2005）也指出對於固定參數的估計檢驗，使用較小的個體層次樣本並不會帶來負面的影響。Raudenbush 與 Bryk（2002）也認為無論個體層次樣本規模的大小為何，固定效果的估計是不偏的。

而在實証研究方面，溫福星、邱皓政（2011）使用 1988 年美國教育追蹤統計資料庫（NELS88）進行多層次模型建構，指出當總體層次學校數分成 1003、23 及 10 間時，固定效果的參數估計穩定，但隨機效果的估計受到總體層次樣本數影

響頗大，估計標準誤的顯著性在總體層次樣本數為 1003 間學校時顯著，其餘的不顯著。Mok (1995) 也使用真實教育資料探討不同樣本規模對多層次模型固定參數、變異數以及共變數的多種參數估計值的影響，其研究發現在總樣本數固定的前提下，使用較多學校搭配每校內較少的學生會比使用較少的學校搭配較多的學生較有統計檢定力，且參數估計偏誤較小。

從上述模擬以及實証研究結果可知，影響多層次模型固定參數估計正確性以及統計檢定力的主要因素為總體層次樣本數的多寡，在估計模型時，增加總體層次樣本數比增加個體層次樣本數較有統計檢定力，因此在後續的模擬分析時，將固定個體層次的樣本數，僅著重在操作總體層次樣本數的大小。

在多層次分析中，除了固定參數的估計外，對於總體層次以及個體層次的隨機效果變異數估計也是很重要的議題 (Cohen, 1998)。在 Maas 與 Hox (2005) 的模擬研究中，許多實驗情況下的隨機參數估計值的偏誤很小，而最大偏誤是發生在最小樣本數與最大 ICC 時。而且總體層次的樣本數對於變異數估計值的標準誤影響很大，當總體層次樣本數為 30 時，總體層次截距變異以及斜率變異的參數未覆蓋率為 8.9% 以及 8.8%。儘管當總體層次樣本數增加為 50 時有所改善，但在三種不同總體層次樣本數下，變異數標準誤的估計正確性皆相對固定參數標準誤差。此外，溫福星、邱皓政 (2011) 的實証研究也指出隨機效果的估計受到總體層次樣本數影響頗大，估計標準誤的顯著性在不同總體層次樣本中出現相異的結果，跨層級的交互作用在大樣本上顯著但在中、小樣本卻未顯著，得到相反的結論，在隨機部分有明顯的變化，顯示總體層次樣本數所影響的是標準誤，連帶影響統計檢定力，使得顯著性考驗結果受到影響。而 Kreft 與 De Leeuw (1998) 認為，一般來說為了獲得研究者所需的檢定力，觀察值要多，除非研究者所探究的議題有非常強且容易被偵測到的效果，當總體層次樣本數很少時，隨機成分會被低估，或有較大的標準誤出現，對於跨層級效果要有足夠檢定力，總體層次樣本數不能太少，需大於 20。而 Maas 與 Hox (2005) 也指出為了得到固定效果標準誤的估計值不偏，總體層次樣本數至少要 50 組。此外，Hox (2010) 也提出多層次模型樣本數的配置經驗法則，當研究者關注跨層級交互作用時，總體層次樣本數要大一些，至少要有 50 個總體層次樣本，而每個總體的個體層次樣本數為 20 人，但若關注的焦點是隨機效果參數，如變異數、共變數以及標準誤，則總體層次樣本需改為 100 個，而每個總體層次的個體層次樣本數為 10 人。

值得注意的是，上述多層次模型的模擬以及實證研究結果，都是在既有的設定下以模擬或是實證的方式所進行的結果推論，都是在不同研究目的下，例如固定效果、隨機效果或是標準誤不偏情況的檢定力要求，必須有樣本數限制，因此並沒有統一的標準。且過去有關多層次模型不同層次樣本數配置的研究，均建立在橫斷性的多層次模型上，無論是總體層次或是個體層次的樣本數規劃，均難以應用到二階層的縱貫性成長模型研究設計上，畢竟成長模型的個體內層次（相當於多層次模型的個體層次）為時間點，在教育學門縱貫性研究設計上很難收集超過 5 個波次的測量。雖然楊志堅等人（2004）針對成長模型進行模擬研究，探討樣本數與檢定力的關係，並指出當樣本數為 300 個以上有效樣本時，檢定力幾乎可以得到 .8 以上合適的結果，而當樣本數增加至 600 個以上時，即便在比較嚴苛的情況下，成長模型平均成長率的檢定力也可以得到合適的結果。此外，巫博瀚（2012）的模擬研究發現，無論是無條件成長模式或條件成長模式，當第二層分析單位數達 100 或 100 人以上時，第一層的分析單位數（波數）為三波、四波、五波或六波，對於迴歸係數的估計都是不偏的。但當第二層的樣本規模（人數）較小時，隨機效果的估計會有嚴重的偏誤，而當提高第二層分析單位的樣本規模時，則隨機效果參數的估計將會愈正確。如果研究者只關心模式中的固定效果時，使用小規模樣本（100 人）即能獲得良好的估計，當關注模式中的隨機效果時，必須使用較大規模的樣本，方能滿足參數估計正確性的要求。

然而，實證上個體間層次的樣本數在多波次測量時，常面臨樣本數流失的問題，很難與上述多層次模型與成長模型的模擬與實証研究的完全平衡資料相提並論，可知過去多層次模型與成長模型的研究結果仍有其應用上的限制。基於此，本研究後續的模擬研究設計除了考量多層次模型與成長模型外，也將進一步考量包含隨機遺漏值時，在多層次模型與成長模型分析合理的樣本數調整，並探討樣本數的差異對於參數估計正確性與統計考驗力（power）的影響。

二、貝氏估計法在多層次分析的應用

傳統的多層次分析主要使用 ML 進行估計，然而 ML 方法在符合大樣本的前提下，參數估計具有一致性，並且具有漸進不偏性和有效性。但在實證的多層次模型或是成長模型建構時，多層次模型總體層次或是成長模型個體內以及個體間層次的樣本數常常無法符合大樣本數的要求。此外，多層次模型個體層次觀察樣

本數也常呈現不平衡狀況（例如各校內的學生數可能不相同）或是成長模型的個體內層次蒐集波次過少，此時不平衡的程度以及測量波次過少的問題會影響到 ML 估計的正確性，而貝氏方法恰好能夠補足其缺點（Kaplan & Depaoli, 2012; Lee, 2007; Muthén & Asparouhov, 2012）。

貝氏方法與 ML 頻率統計取向差異頗大，其建立在局部限制的情境下，透過不斷重複抽取樣本以形成參數的後驗分配，並進一步估計此部分的平均數與標準差，做為參數的點估計值與分散程度的指標，在每一個參數自成一個維度估計的情況下，無需經由複雜的積分計算，便能獲得準確的參數估計值，因此，貝氏方法特別適合應用在多層次模型或是成長模型這類複雜模型的參數估計上，可有效降低模型估計時樣本數的要求（Hox, van de Schoot, & Matthijsse, 2012; Muthén & Asparouhov, 2012）。

Hox、van de Schoot 與 Matthijsse（2012）將貝氏估計法使用在多層次結構方程模型的總體層次樣本數估計上，並透過模擬研究的方式與 Meuleman 與 Billiet（2009）使用 ML 估計法進行多層次結構方程模型的模擬研究比較，相較於 Meuleman 與 Billiet 使用 ML 的模擬研究指出進行多層次結構方程模型的總體層次合理樣本數應為 50 到 100 個樣本，Hox 等人發現在相同的估計參數情況下，使用貝氏方法所需要的樣本數僅為 20 個便可以達到穩定與有效的參數估計值，可見貝氏方法可有效的降低估計複雜模型時所需的樣本數要求。惟目前在多層次模型或是成長模型的模擬與實證研究上，並未有相關的研究同時比較貝氏方法與 ML 法在模型估計時，所需合理樣本數的差異，此一不足正是本研究極欲補足之處，企盼後續分析結果可供實證研究者進行多層次模型或是成長模型建構時決定樣本數的參考。

貝氏方法相較於傳統頻率學派的 ML 估計法在本質上有很大的差異，因此，以下簡介貝氏方法的特點（Kaplan & Depaoli, 2012; Lee, 2007; Muthén & Asparouhov, 2012）供讀者參考。首先，貝氏學派認為任何一個未知變量都具有不確定性，既然未知變量可以看作隨機變量，那麼將未知參數看作隨機變量也是合理的，用一個機率分布來描述。因此，貝氏方法視未知參數為隨機變量，有自己的分布，而傳統的頻率學派僅把未知參數看作普通的未知變量而不是隨機變量，因此在進行參數估計時，ML 估計結果僅會得到一個固定參數，而貝氏方法除了固定參數外，還會增加估計參數的一個後驗機率區間。

其次，貝氏方法除了同頻率學派一樣利用樣本訊息進行參數估計外，也會利用先驗訊息，利用先驗訊息使其數量化以形成先驗分布並加入到統計推斷中，進而提高統計推斷的品質。若忽視先驗訊息，有時會得出不合理的結論，但若研究者對於先驗訊息瞭解不多，在貝氏方法下可以使用無訊息的先驗分布進行參數估計。

貝氏方法主要利用樣本訊息和先驗訊息導出未知參數的後驗分布並進行各種統計推斷，後驗分布的機率融合了樣本訊息和先驗訊息，比傳統的抽樣理論有著更佳的合理性和靈活性。若以貝氏公式表示，可得 $P(\theta|y) = P(y|\theta)P(\theta)/P(y)$ ，其中 $P(\theta)$ 為參數 θ 的先驗分布， $P(y|\theta)$ 就是傳統頻率學派所求得的概似函數， $P(y)$ 是 y 的邊際分布， $P(\theta|y)$ 稱為 θ 的後驗分布，從貝氏公式可以清楚知道後驗分布綜合了數據的樣本訊息和先驗訊息，也就是融合了全部的訊息在裡面，比傳統頻率學派僅考慮樣本訊息更為合理，充分利用了所有已知的訊息，避免了訊息浪費，由此可知貝氏方法的研究主要利基於這個後驗分布 $P(\theta|y)$ 。貝氏方法分析流程可以圖 1 表示。

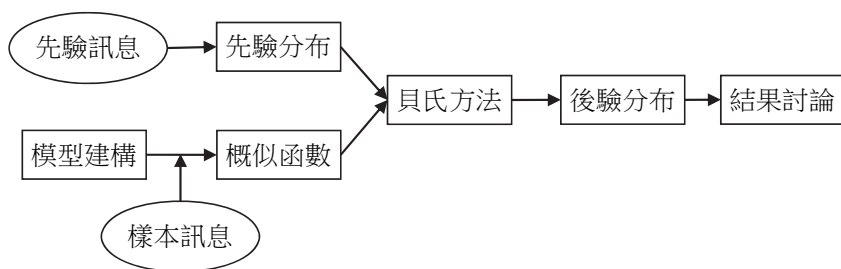


圖 1 貝氏方法分析流程

參、研究方法

本研究目的主要經由模擬研究方式同時比較貝氏方法和 ML 估計法在多層次模型以及成長模型建構時，最小可行的分析樣本單位數，並同時考慮存在隨機遺漏下，貝氏方法和 ML 估計法在多層次模型以及成長模型建構所需的樣本數調整。依據研究設計，利用蒙地卡羅法產生本研究所需的模擬資料，進一步透過貝氏方

法以及 ML 法估計參數，並透過 90% 信賴區間覆蓋率以及 80% 的統計考驗力比較兩種估計方法在多層次模型以及成長模型建構時，所需最小可行樣本數的差異。

一、模擬研究設計

本研究所假設的理論模式僅聚焦於多層次模型以及成長模型的非條件化模型，並未考量非線性模型，以下分別就多層次模型以及成長模型的模擬設計進行說明。

（一）多層次模型

從上述模擬以及實証研究結果可知，影響多層次模型固定參數、變異數估計正確性以及統計檢定力的主要因素為總體層次樣本數的多寡，在估計多層次模型時，增加總體層次樣本數比增加個體層次樣本數更具有統計檢定力，因此在後續的模擬分析時，將固定個體層次的樣本數為 20，僅操作總體層次樣本數的大小。在國內教育學門，無論在國小、國中、高中以及大學，在同一間學校裡收集到 20 個樣本數應是可行的。

此外，在多層次模型中主要透過檢視零模型的 ICC 來判斷是否必須以多層次的統計技術來分析多層次的資料。Cohen（1988）認為 ICC 在不同的研究領域差異很大，因此在不同的研究領域有不同的 ICC 判斷標準。不過，Cohen 提出了 3 種可供參閱的 ICC 標準，這 3 種數字反應出 ICC 的效果大小，他認為當 ICC 小於 .059 時，算是相當小的，其效果可以略而不計，相當於虛無假設的零相關無法被拒絕，而 ICC 介於 .059 到 .138 時，這樣的大小算是中度相關，至於 ICC 大於 .138 時則為高度組內相關，Cohen 認為 ICC 在中度相關以上時就不能忽略相似性的存在，其對迴歸係數估計標準誤與統計檢定力的影響就不能被忽視，因此當 ICC 大於 .059 時，則必須考慮多層次模型的統計分析。

基於此，本研究在 ICC 的參數設定上主要參照 Cohen（1988）的 ICC 標準，將模式的 ICC 設定為 .059 以及 .138 兩種進行後續的多層次模型模擬比較。

在模型設定上，設定隨機截距平均數為 0，組內層次變異數為 10，當組間層次變異數設定為 .627 時，ICC 為 .059 ($.627 / (10 + .627) = .059$)，當組間層次變異數設定為 1.601 時，ICC 為 .138 ($1.601 / (10 + 1.601) = .138$)。設定個體層次的樣本數均為 20 人，僅改變組間層次樣本數的大小。

此外，目前國內外在多層次模擬研究設計時並沒有考慮隨機遺漏機率，因此

本研究在後續模擬比較上，除了進行完整資料的模擬比較外，有別於過往的模擬研究，也將同時考量隨機遺漏機率为 25%、35% 以及 50% 時進行多層次模型建構所需合理樣本數的調整。

（二）成長模型

國內外在探討多層次模型的合理樣本數需求上，無論是模擬或實證研究，主要集中於橫斷性的多層次模型，相較於縱貫性的成長模型合理樣本數需求，較少探討。Muthén 與 Muthén（2002）歸納過往的模擬以及實證研究，認為在非條件化的成長模型架構下，可將成長模型的截距設定為 0，斜率設定為 .2，截距與斜率的相關設定為 0，截距變異數設定為 .5，斜率變異數設定為 .1，各題項的殘差設定為 .5，以進行後續成長模型的樣本數或是效果值差異分析。基於此，本研究在後續成長模型的模擬研究母群參數設定上，主要參考 Muthén 與 Muthén（2002）的模型設定。

在成長模型的模擬研究設定上，主要依據 Muthén 與 Muthén（2002）的模型設定測量時間點 4 個，此外，也加入測量時間點為 3 個的成長模型，並比較其在合理樣本數需求上的差異。而在成長模型的樣本數隨機遺漏設定上，主要參閱 Muthén 與 Muthén（2002）以及 Wang 與 Wang（2012）的成長模型模擬研究設計，將 4 個測量波次的樣本遺漏機率分別設定為 0%、10%、20%、30%。

二、貝氏方法

在貝氏方法中，先驗訊息的設定至關重要，若以人為的方式操弄先驗訊息，將可能使不適配的模型適配，或是得出完全相反的結果（Kaplan & Depaoli, 2012），因此在先驗訊息的設定上需要謹慎處理。先驗訊息主要可以區分為兩種，一是有訊息先驗，一是無訊息先驗（Lee, 2007）。如果所選擇的先驗分布使得後驗分布和他是同一族分布，這種先驗稱為共軛先驗，這種先驗分布常被採用，主要是因為計算簡單，但在實證研究中需考量先驗分布的合理性。若研究者對模型的參數知道很少或是一無所知，在這種情況下則使用無訊息先驗。在多層次模型架構下，一般常將變異數的無訊息先驗設定為變異數的倒數服從 Gamma (.001, .001) 或是 Gamma (0, 0)（Browne & Draper, 2006; Gelman, 2006; Kaplan & Depaoli, 2012），這樣的設定並不會對參數估計的結果產生重大影響。基於此，本研究在後續的貝氏方法模擬研究上，在多層次模型主要設定變異數的倒數服從

Gamma (.001, .001)，而設定成長模型截距以及斜率變異數的倒數服從 Gamma (0, 0)，無論設定哪一種，並不會對後續與 ML 方法的比較分析結果產生差異。

此外，貝氏方法的推論主要基於後驗分布，因此，如何得到未知參數複雜的後驗分布對貝氏方法的估計至關重要，在一般情況下後驗分布常常沒有明顯的表達式，也不是常見的分佈，這些複雜的非標準形式的後驗分布估計維度數常常很高，因此在估算上是一大問題，在本研究的貝氏方法模擬估算上主要使用馬可夫鍊蒙地卡羅法 (Markov chain Monte Carlo, MCMC) 的方式，結合資料本身的資訊與參數先驗分布，構成聯合後驗分布來處理高維、複雜的積分運算問題，MCMC 方法將複雜的抽樣問題轉化為一系列簡單的抽樣，而不是直接從複雜的後驗分布抽取樣本，一般常見的 MCMC 方法為 Metropolis-Hastings 取樣法 (M-H; Chib & Greenberg, 1995) 與 Gibbs 取樣法 (Gelfand & Smith, 1990; Geman & Geman, 1984; Tanner & Wong, 1987)，本研究主要使用 Gibbs 取樣法。Gibbs 取樣法以各參數形成的限制後驗分布來逼近聯合後驗分布，透過不斷的取樣以及更新參數，以其最後的結果能逼近真正的參數值 (Chib & Greenberg, 1995; Gelman, Carlin, Stern, & Rubin, 2003)。Gibbs 取樣法具體步驟如下，假設 y 是數據， θ 是未知參數，把參數向量 θ 分塊為 $\theta = (\theta_{(1)}, \dots, \theta_{(k)})$ 。選取參數 θ 的一個初始值 $\theta^{(0)}$ ，假定第 i 次疊代開始時參數 θ 的值是 $\theta^{(i-1)}$ ，則第 i 次疊代為：

第一步：從滿足條件分布的 $p(\theta_{(1)} | y, \theta_{(2)}^{(i-1)}, \dots, \theta_{(k)}^{(i-1)})$ 中抽取一個樣本 $\theta_{(1)}^{(i)}$ 。

第二步：從滿足條件分布的 $p(\theta_{(2)} | y, \theta_{(1)}^{(i)}, \theta_{(3)}^{(i-1)}, \dots, \theta_{(k)}^{(i-1)})$ 中抽取一個樣本 $\theta_{(2)}^{(i)}$ 。

.....

第 K 步：從滿足條件分布的 $p(\theta_{(k)} | y, \theta_{(1)}^{(i)}, \theta_{(2)}^{(i)}, \dots, \theta_{(k-1)}^{(i)})$ 中抽取一個樣本 $\theta_{(k)}^{(i)}$ 。

對 $i = 1, \dots, n$ 重複以上各步驟，從而得到樣本 $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n)}$ 。有了後驗分布的分配後，便可以進行統計推斷了。貝氏方法透過不斷重複抽取樣本（如上述第一步到第 K 步的 Gibbs 取樣法）以形成參數的後驗分布，再估計此分布的平均數與標準差，作為參數的點估計值與離散程度的指標。

本研究後續在多層次模型以及成長模型的合理樣本數探討上，將利用貝氏方法進行參數估計並與 ML 方法比較，期望透過貝氏方法在處理複雜模型估計上的優點，有效降低模型建構時所需的樣本數要求，供實證研究者參考。

三、研究因子

本模擬研究主要同時比較貝氏方法和 ML 估計法在多層次模型以及成長模型建構時，最小可行的分析樣本單位數，並同時考慮存在隨機遺漏下，貝氏方法和 ML 估計法在多層次模型以及成長模型建構所需的樣本數調整，以下分別說明在多層次模型和成長模型中的研究因子。

無論是貝氏方法或是 ML 法在多層次模型以及成長模型的建構，均使用 MPLUS 軟體進行分析，每個研究均模擬 10,000 次，藉以獲得穩定的估計（Muthén & Muthén, 2002）。此外，受限於研究資源以及時間，在貝氏方法的 10,000 次模擬中均設定每一個模擬的 MCMC 迭代次數為 1,000 次，而前 500 次的抽樣捨棄不用，以最後 500 次的後驗分布期望做為參數估計的結果。

（一）多層次模型

在多層次模型中主要設定 ICC 為兩種，分別為 .059 與 .138，代表組內相關係數的差異（Cohen, 1988）。在模型設定上，設定隨機截距平均數為 0，組內層次變異數為 10，當組間層次變異數設定為 .627 時，ICC 為 .059，當組間層次變異數設定為 1.601 時，ICC 為 .138。此外，遺漏值設定為 0%、25%、35%、50%，代表實證樣本收集上存在的隨機遺漏。

（二）成長模型

在成長模型中主要設定截距為 0，斜率為 .2，截距與斜率的相關設定為 0，截距變異數為 .5，斜率變異數為 .1，各題項的殘差設定為 .5。在測量波次上設定為 3、4 兩種（Muthén & Muthén, 2002）。遺漏值在四個不同時間點分別設定為 0%、10%、20%、30% 四種（Muthén & Muthén, 2002; Wang & Wang, 2012）。

有關貝氏方法和 ML 估計法在多層次模型以及成長模型參數估計正確性的判斷，主要是以 90% 的參數覆蓋率（Wang & Wang, 2012）進行檢核標準，本研究所使用的 MPLUS 軟體在進行模擬研究的參數覆蓋率檢定自動以 95% 的參數覆蓋率作為分析，當 MPLUS 軟體參數覆蓋率分析結果達到 90% 以上即符合可接受的標準，兩種不同估計法的統計考驗力標準均為 .8（Muthén & Muthén, 2002）。此外，在本研究引註關於樣本數以及統計考驗力的文獻中，Muthén 與 Muthén（2002）呈現統計考驗力和參數覆蓋率，Hox、van de Schoot 與 Matthijsse（2012）呈現估計參數偏誤與參數覆蓋率，Meuleman 與 Billie（2009）呈現估計參數偏誤、標準誤偏誤、

統計考驗力和參數覆蓋率。上列研究所使用指標將一併列出在本研究中，但主要著重在統計考驗力和參數覆蓋率的說明上。

肆、研究結果與討論

本研究同時比較貝氏方法以及 ML 估計法合理樣本數的需求，屬於探索性質，在總體層次樣本數的設定上，雖然參考過去相關研究（溫福星、邱皓政，2011；楊志堅、劉心筠、楊志強，2004；Hox, 2010; Kreft & De leeuw, 1998; Maas & Hox, 2004, 2005; Raudenbush & Bryk, 2002; Snijders, 2005）進行樣本數設定，但為了更細緻比較貝氏方法以及 ML 估計法合理樣本數的差異，因此在樣本數的間距設定上，在多層次模型主要設定以 20 人為單位，在成長模型主要設定以 10 人為單位。

一、多層次模型合理樣本數差異：貝氏 VS ML

首先進行貝氏方法以及 ML 估計法在橫斷性的多層次模型建構時，當 ICC 大於 .059 以上時，合理的樣本數比較。

由表 1 至表 2 可知，需使用多層次模型進行建構時，無論模型是否存在隨機遺漏，當個體層次樣本數為 20 人情況下，貝氏方法在總體層次需要 31 個樣本進行多層次模型建構以及分析，此時模型的參數估計以及統計考驗力才穩定，當總體層次低於 31 個樣本時，參數覆蓋率因總體層次樣本數過小而逐漸低於 90% 的參照值。當 ICC 增加至 .138 時，使用貝氏方法在總體層次樣本僅需要 10 個，即可以獲得穩定的參數估計結果以及統計考驗力。

但若使用傳統的 ML 方法進行模型建構，當 ICC 接近 .059 時，總體層次最少需要 42 個樣本才可獲得相對穩定的參數估計以及統計考驗力，與 Maas 與 Hox（2004）所建議的合理的總體層次樣本數為 50 組相近。此外，最大偏誤是發生在最小樣本數與最大 ICC 時，也與 Maas 與 Hox（2005）的模擬研究相似。即便 ICC 增加至 .138，總體層次樣本依舊需要 31 個，與貝氏方法相比，ML 方法所需要的樣本數相對較大，對於實證研究者而言，需考量研究經費以及樣本數的增加。

表 1 多層次模型下 ICC = .059 時隨機截距變異數的模擬參數估計

貝氏					ML				
ICC = .059	Coverage	Power	Parameter bias	Standard error bias		Coverage	Power	Parameter bias	Standard error bias
no missing									
620 (31)	.905	1.000	.113	.022	820 (41)	.916	.802	.042	.004
600 (30)	.897	1.000	.114	.028	800 (40)	.918	.791	.042	.004
missing = 25%									
620 (31)	.906	1.000	.122	.031	840 (42)	.906	.808	.050	.017
600 (30)	.893	1.000	.124	.046	820 (41)	.909	.790	.052	.012
missing = 35%									
620 (31)	.907	1.000	.106	.012	840 (42)	.913	.804	.045	.008
600 (30)	.900	1.000	.112	.037	820 (41)	.915	.792	.045	.008
missing = 50%									
620 (31)	.901	1.000	.118	.028	840 (42)	.913	.815	.042	.008
600 (30)	.899	1.000	.124	.028	820 (41)	.911	.799	.042	.008

註：1. 當總樣本數為 100 (5) 時，100 表示總樣本數，(5) 表示組間層次樣本數。若遺漏值 = 25% 則最後的估計總樣本數為 75 (5)，若遺漏值 = 35% 則最後的估計總樣本數為 65 (5)，若遺漏值 = 50% 則最後的估計總樣本數為 50 (5)，其餘依此類推。

2. 隨機截距變異數的模擬參數估計結果與 ICC 的參數估計結果相近，在此僅呈現隨機截距變異數的模擬參數估計結果。

表 2 多層次模型下 ICC = .138 時隨機截距變異數的模擬參數估計

貝氏					ML				
ICC = .138	Coverage	Power	Parameter bias	Standard error bias		Coverage	Power	Parameter bias	Standard error bias
no missing									
200 (10)	.902	1.000	.097	.770	560 (28)	.902	.979	.121	.013
180 (90)	.897	1.000	.115	.749	540 (27)	.896	.972	.125	.014
missing = 25%									
200 (10)	.904	1.000	.091	.807	620 (31)	.900	.987	.130	.006
180 (9)	.897	1.000	.121	.766	600 (30)	.896	.984	.134	.019

(續下頁)

表2 多層次模型下 ICC = .138 時隨機截距變異數的模擬參數估計(續)

貝氏					ML				
missing = 35%									
200 (10)	.903	1.000	.105	.795	600 (30)	.901	.989	.123	.008
180 (9)	.899	1.000	.130	.753	580 (29)	.895	.985	.132	.019
missing = 50%									
200 (10)	.902	1.000	.105	.782	580 (29)	.902	.982	.113	.004
180 (9)	.898	1.000	.103	.751	560 (28)	.899	.979	.115	.015

二、成長模型合理樣本數差異：貝氏 VS ML

進一步在縱貫性的成長模型架構下，比較貝氏方法以及 ML 估計法合理樣本數的差異。由表3及表4可知，當測量時間點為3個，無論模型是否存在隨機遺漏，貝氏方法在個體間層次僅需要90個樣本即可進行成長模型建構以及分析，且模型的參數估計以及統計考驗力穩定，但若使用傳統的ML方法進行成長模型建構，當測量時間點為3個時，個體間層次最少需要310個樣本才可獲得相對穩定的參數估計以及統計考驗力，與楊志堅等人（2004）針對成長模型進行的模擬研究相似，當樣本數為300個以上有效樣本時，檢定力幾乎可以得到.8以上合適的結果。

當測量時間點進一步增加至4個時，貝氏方法與ML方法相比所需要的樣本數差異不大，貝氏方法需要60個，而ML方法需要70個，可見在成長模型架構下，當研究者無法收集較大的樣本數時，或許可考慮增加測量波次以獲得相對穩定的參數估計以及統計考驗力。但若受限於樣本特殊性、研究經費以及測量波次的侷限時，可以考慮使用貝氏方法藉以有效降低合理樣本數的需求。

表3 成長模型，3個時間點

no missing	Mean				Slope				Var Mean				Var Slope			
Sample Size	Coverage	Type I error	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias
貝氏																
80	.948	.052	.000	.011	.950	.830	.003	.016	.952	1.000	.045	.064	.959	1.000	.154	.095

(續下頁)

表 3 成長模型，3 個時間點（續）

no missing	Mean				Slope				Var Mean				Var Slope			
70	.948	.052	.000	.014	.952	.775	.002	.026	.949	1.000	.050	.072	.946	1.000	.174	.125
ML																
250	.949	.051	.002	.003	.950	1.000	.003	.000	.943	1.000	.005	.006	.945	.810	.005	.019
240	.950	.050	.002	.003	.949	1.000	.003	.000	.943	1.000	.006	.009	.946	.794	.006	.016
missing	Mean				Slope				Var Mean				Var Slope			
Sample Size	Coverage	Type I error	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias
貝氏																
90	.948	.052	.000	.010	.949	.821	.004	.007	.948	1.000	.041	.044	.963	1.000	.148	.100
80	.949	.051	.000	.012	.950	.772	.005	.020	.949	1.000	.046	.059	.962	1.000	.169	.119
ML																
310	.950	.050	.001	.002	.949	1.000	.003	.006	.947	1.000	.005	.000	.946	.810	.005	.011
300	.946	.054	.001	.005	.953	1.000	.003	.011	.946	1.000	.004	.002	.948	.799	.005	.005

註：1. Muthén 與 Muthén（2002）指出，使用 MPLUS 軟體進行 Monte Carlo 模擬研究時，研究者所設定的母群參數將假定為真並作為虛無假設，而 MPLUS 所檢定的假設為當母群參數顯著不為 0 的對立假設。因此，當所設定的參數不為 0 時，軟體所呈現的為統計檢定力，當所設定的參數為 0 時，軟體所呈現的是第一類型錯誤率。本研究假定成長模型的截距平均為 0，所呈現的結果為第一類型錯誤率，而其他參數所呈現的為統計考驗力。

2. P bias 為 Parameter bias 縮寫，Se bias 為 Standard error bias 縮寫。

表 4 成長模型，4 個時間點

no missing	Mean				Slope				Var Mean				Var Slope			
Sample Size	Coverage	Type I error	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias
貝氏																
50	.958	.042	.002	.026	.956	.851	.003	.038	.947	1.000	.082	.088	.948	1.000	.099	.096
40	.956	.044	.003	.050	.956	.749	.004	.054	.949	1.000	.099	.116	.955	1.000	.121	.134

（續下頁）

表 4 成長模型，4 個時間點（續）

no missing	Mean				Slope				Var Mean				Var Slope			
ML																
50	.940	.060	.003	.028	.944	.886	.003	.018	.923	.978	.028	.029	.926	.810	.021	.029
40	.940	.060	.003	.028	.937	.806	.003	.029	.912	.935	.034	.036	.920	.693	.024	.038
missing	Mean				Slope				Var Mean				Var Slope			
Sample Size	Coverage	Type I error	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias	Coverage	Power	P bias	Se bias
貝氏																
60	.955	.045	.001	.032	.955	.857	.002	.038	.947	1.000	.068	.075	.952	1.000	.083	.092
50	.957	.043	.001	.037	.953	.775	.000	.037	.946	1.000	.082	.091	.958	1.000	.099	.119
ML																
70	.942	.058	.003	.015	.942	.926	.003	.023	.927	.994	.022	.028	.932	.808	.026	.027
60	.943	.057	.003	.015	.939	.882	.004	.029	.926	.987	.025	.026	.930	.736	.030	.030

茲將本研究在多層次模型以及成長模型中貝氏方法與 ML 法模擬合理樣本數的最小需求摘錄如下，供實證研究者參閱：

表 5 多層次模型以及成長模型在貝氏方法與 ML 法最小可行樣本數摘要

貝氏方法			ML 方法	
多層次模型	no missing	missing	no missing	missing
ICC =.059	620 (31)	620 (31)	820 (41)	840 (42)
ICC =.138	200 (10)	200 (10)	560 (28)	620 (31)
成長模型	no missing	missing	no missing	missing
Time = 3	80	90	250	310
Time = 4	50	60	50	70

註：當總樣本數為 620 (31) 時，620 表示總樣本數，(31) 表示組間層次樣本數。

伍、建議

本研究主要經由模擬研究方式，同時比較貝氏方法和 ML 估計法在多層次模型以及成長模型建構時，最小可行的分析樣本單位數，並同時考慮存在隨機遺漏下，貝氏方法和 ML 估計法在多層次模型以及成長模型建構所需的樣本數調整。研究發現，無論是否存在隨機遺漏，使用貝氏方法進行多層次模型以及成長模型建構時所需的合理樣本數較小，且可以獲得穩定的參數覆蓋率以及統計考驗力，值得加以推廣。

此外，貝氏方法除了可有效降低複雜模型樣本數的估計需求外，也被進一步應用在處理高維的複雜數據，如含二分或有序的潛變量模型、多層次模型、非線性模型以及混合模型等，不僅提供了應用貝氏方法的理論依據，還具有顯著的實用價值。

雖然本研究企圖透過模擬研究，探討貝氏方法相較於 ML 估計法，在多層次模型以及成長模型建構時，使用較小的樣本數下可獲得穩定的參數估計。但受限於研究者能力，並無法以更深入淺出的方式說明多層次模型、成長模型建構，以及貝氏方法的估計過程，因此提供本研究的語法於附錄供實證研究者參閱。新近國內已有專書（邱皓政，2017）詳細探討多層次模型、成長模型的建構，可讀性佳，實證研究者在詳閱該書後，若欲建構多層次模型或成長模型，可配合文末語法帶入所估計實證模型的參數，進一步做為修訂樣本數的參考。

謝誌

感謝 3 位審稿者細緻且嚴謹的審查，明基要在此致上最高的敬意與謝意！雖然文章刊登，但並不表示現階段有能力可以完整回答以及修訂審稿者的提問和要求。期許自己不斷精進專業，10 年後，能提出更令大家滿意的答案！

也感謝期刊小編棄而不捨的來電，和適時打氣，讓這篇研究經歷了兩年，可以重見天日！千言萬語，除了感謝，還是感謝！

參考文獻

- 巫博瀚(2012)。成長曲線模式之樣本單位數決定研究：蒙地卡羅模擬（未出版之博士論文）。國立成功大學教育研究所，臺南市。
- 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。
- 溫福星、邱皓政(2011)。多層次模型方法論：階層線性模式的關鍵問題與試解。臺北市： $\alpha\beta\gamma$ 實驗室。
- 楊志堅、劉心筠、楊志強(2004)。縱貫研究以潛在成長模式分析之樣本數與檢定力研究。《教育與心理研究》，27(3)，603-626。
- Browne, W. J., & Draper, D. (2006). A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis*, 1, 473-514.
- Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85, 398-409.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2003). *Bayesian data analysis*. New York, NY: Chapman & Hall.
- Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6, 721-741.
- Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *American Statistician*, 49, 327-335.
- Cohen, J. (1998). Determining sample sizes for surveys with data analyzed by hierarchical linear models. *Journal of Official Statistics*, 14, 267-275.
- Gelman A. (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1, 515-533.
- Heck, R. H., & Thomas, S. L. (2009). *An introduction to multilevel modeling techniques* (2nd ed.). New York, NY: Routledge.
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.
- Hox, J. J., van de Schoot, R., & Matthijsse, S. (2012). How few countries will do? Comparative survey analysis from a Bayesian perspective. *Survey Research*

- Methods*, 6, 87-93.
- Kaplan, D., & Depaoli, S. (2012). Bayesian structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 650-673). New York, NY: Guilford.
- Kreft, I., & Leeuw, J. D. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.
- Lee, S.Y. (2007). *Structural equation modeling: A Bayesian approach*. Chichester: John Wiley & Sons.
- Maas, C. J. M., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, 58, 127-137.
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1, 86-92.
- Meuleman, B., & Billiet, J. (2009). A Monte Carlo sample size study: How many countries are needed for accurate multilevel SEM? *Survey Research Methods*, 3, 45-58.
- Mok, M. (1995). Sample size requirements for 2-level designs in educational research. *Multilevel Modelling Newsletter*, 7, 11-15.
- Muthén, B. O., & Asparouhov, T. (2012). Bayesian SEM: A more flexible representation of substantive theory. *Psychological Methods*, 17, 313-335.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 4, 599-620.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Snijders, T. A. (2005). Power and sample size in multilevel linear models. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Sciences* (Vol. 3, pp. 1570-1573). Chichester, England: Wiley.
- Stoel, R. D., & Garre, F. G. (2011). Growth curve analysis using multilevel regression and structural equation modeling. In J. J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 97-111). New York, NY: Routledge.

- Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association*, 82, 528-550.
- Wang, J., & Wang, X. (2012). *Structural Equation Modeling: Applications using Mplus*. Hoboken, NJ: John Wiley & Sons.

附錄一：

title: HLM, ML, ICC=.138, N=560, Missing=25%

montecarlo:

```
names are y;  
nobs = 560;  
seed = 3454367;  
nrep = 10000;  
ncsizes = 1;  
csizes = 28 (20);  
missing = y;
```

analysis:

```
type = twolevel;  
estimator = ML;  
process = 2;
```

model population:

```
%within%  
y*10;  
%between%  
y*1.601; !icc = 1.601/11.601 = .138
```

model missing:

```
[y*-1.099]; !missing = 25%
```

model:

```
%within%  
y*10 (w);  
%between%  
y*1.601 (b);
```

model constraint:

```
new(icc*.138);  
icc = b/(w+b);
```

output:

```
tech9;
```

附錄二：

title: HLM, Bayes, ICC=.138, N=100, Missing=50%

montecarlo:

```
names are y;  
nobs = 100;  
seed = 3454367;  
nrep = 10000;  
ncsizes = 1;  
csizes = 5 (20);  
missing = y;
```

analysis:

```
type = twolevel;  
estimator = BAYES;  
process = 2;  
fbiter = 1000;
```

model population:

```
%within%  
y*10;  
%between%  
y*1.601;
```

model missing:

```
[y*0];
```

model:

```
%within%  
y*10 (w);  
%between%  
y*1.601 (b);
```

model priors:

```
b~IG (.001,.001);
```

model constraint:

```
new(icc*.138);  
icc = b/(w+b);
```

output:

tech9;

附錄三：

title: GLM, ML, N=70, Missing = 0, 10, 20, 30%

montecarlo:

names are y1-y4;

nobs = 70;

seed = 3454367;

nrep = 10000;

missing = y1-y4;

analysis:

estimator = ML;

process = 2;

model population:

i s | y1@0 y2@1 y3@2 y4@3;

[i*0 s*.2];

i*.5;

s*.1;

i with s@0;

y1-y4*.5;

model missing:

[y1*-15 y2*-2.20 y3*-1.39 y4*-.85]; !missing = 0,10,20,30%

model:

i s | y1@0 y2@1 y3@2 y4@3;

[i*0 s*.2];

i*.5;

s*.1;

i with s@0;

y1-y4*.5;

output:

tech9;

附錄四：

title: GLM, Bayes, N=80, Missing = 0,10,20%

montecarlo:

names are y1-y3;

nobs = 80;

seed = 3454367;

nrep = 10000;

missing = y1-y3;

analysis:

estimator = BAYES;

process = 2;

fbiter = 1000;

model population:

i s | y1@0 y2@1 y3@2;

[i*0 s*.2];

i*.5;

s*.1;

i with s@0;

y1-y3*.5;

model missing:

[y1*-15 y2*-2.20 y3*-1.39];

model:

i s | y1@0 y2@1 y3@2;

[i*0 s*.2];

i*.5;

s*.1;

i with s@0;

y1-y3*.5;

output:

tech9;

量表長度簡化研究： 「簡式中小學教師主觀幸福感量表」修訂

余民寧 國立政治大學教育學系教授

陳柏霖 玄奘大學應用心理學系副教授

陳玉樺 國立政治大學教育學系博士候選人（通訊作者）

摘 要

本研究旨在發展一份適合華人使用，經濟、有效之評估主觀幸福感之工具。本研究將過去所編制之「主觀幸福感量表」修訂為「簡式中小學教師主觀幸福感量表」，以符合臨床應用與實務研究上之經濟性與效率性的需求。依據全國各縣市學校數比率進行抽樣後，共得有效選題樣本 1,046 人，在原量表情緒幸福感、心理幸福感與社會幸福感三向度的架構下，根據評定量尺模式與試題差異功能分析結果，選出 15 題的「簡式中小學教師主觀幸福感量表」。然後，重新以分層隨機抽樣取得 1,180 名中小學教師為有效樣本，執行主觀幸福感驗證性因素分析（CFA），發現整體模式具備相當不錯之適配度。本研究發現：以「臺灣憂鬱情緒量表」作為外在效標，進行效標關聯效度分析，發現心理、社會、情緒幸福感與憂鬱變項間，均有顯著負相關，顯示本簡式量表具有不錯的外在效標關聯效度。針對上述研究結果並提出相關建議，以供未來研究與實務工作者之參考。

關鍵詞：主觀幸福感、短題本、試題反應理論



Study of Scale-Items Reduction: The Reconstruction of Subjective Well-Being Scale

Min-Ning Yu

Professor, Department of Education, National Chengchi University

Po-Lin Chen

Associate Professor, Department of Applied Psychology, Hsuan Chuang University

Yu-Hua Chen

Ph. D. Candidate, Department of Education, National Chengchi University (Corresponding Author)

Abstract

The multidimensional Subjective Well-being Scale (SWS) had become one of the most widely-used inventories for measuring the subjective well-being construct including psychological well-being, social well-being, and emotional well-being. For economic efficiency and widespread use, we developed a short form of the Subjective Well-being Scale, the SWS-SF. There were two sets of sample: development sample and validation sample, in which 1,046 and 1,180 teachers randomly selected according to the ratio of the number of schools from cities around Taiwan. Through two-stage process of item selection and model confirmation, the 39-item full scale was reduced to 15-item brief version. The items in the brief version of Subjective Well-being Scale were selected according to the rating scale model (RSM) and differential item functioning (DIF) of item response theory. The structure of the brief version was tested by the confirmation factor analysis (CFA). Results show that: The processes of parameter estimation, item selection, cross-validation, and comparisons with the full scale and other construct (Taiwan Depression Scale) are presented and discussed. The results show that the three-factor model performed adequate goodness-of-fit in brief version, and the psychological, social, and emotional well-beings all had significantly negative relationships with depression. This indicates that the factor structure of the brief version of Subjective Well-being Scale was the same as the full scale and also there was a good criterion-related validity in teachers' subjective well-being model. Finally, some suggestions for practice usage and future research are proposed.

Keywords: subjective well-being, short form version, item response theory



壹、緒論

隨著現代文明高速發展，超乎人們想像的進步速度帶來了蓬勃發展與富裕繁榮，也帶來繁忙的生活步調與壓力負荷，導致現代人罹患憂鬱、焦慮與恐慌等身心症的比例越來越高，心理健康成為普羅大眾與學術研究所共同關心的議題（陳柏霖、余民寧，2015）。自正向心理學崛起後，學者們紛紛認為心理健康的意涵不單指沒有心理疾病而已，還應該包含正向情緒的發展、正向心理與社會機能的呈現（Keyes, 2002; Ryff & Singer, 1998）。

對於心理健康的探討，大多集中在幸福感的研究。心理學家們將主觀幸福感做為反應個體心理健康的重要指標，發展心理健康意義上的主觀幸福感測量。Keyes（2002, 2003, 2005a, 2005b）進一步提出心理健康的內涵包括了情緒、心理與社會幸福感，並以主觀幸福感做為心理健康的代表症狀。其中，情緒幸福感強調個體對於自身生活滿意程度的主觀認知與正負向情緒的感受強度，當個體感受到較多的正向情緒以及較少的負向情緒，且對生活達到某種滿意程度時，即感受到幸福（Diener, 1984; Diener, Suh, Lucas, & Smith, 1999; Ryan & Deci, 2001）；心理幸福感則是關注個體對於自身的意義與自我實現，認為幸福感是個體全心全意投入一項活動，且充分發揮自身潛能後所達到的心理滿足感受（Ryff, 1989, 1995; Ryff & Keyes, 1995）；社會幸福感則是重視外在社會環境對個體的影響，認為個體成長過程中，同時面對內在及外在的壓力，需從中整合以取得平衡，才能夠獲得幸福感（Keyes, 1998, 2002, 2005a; Keyes & Waterman, 2003）。對此，本研究採取理論建構取向，將主觀幸福感（以下簡稱 SWB）定義為：「個體不受職業類別，在當下、平時，亦或是某一期間等情況下，個人以其情感狀態、心理及社會機能觀點，來對其生活進行主觀的覺知和評鑑的一種狀態」（余民寧、謝進昌、林士郁、陳柏霖、曾筱婕，2011），該定義涵蓋心理、社會與情緒幸福感三個成分。

面對主觀幸福感此一龐大複雜之構念，如何進行科學化的測量，是研究者所面臨的一大挑戰。就以其中的心理幸福感為例來說，Ryff 最初在 1989 年曾編制因素長達 120 題的心理幸福感量表，但後來為了方便研究者進行研究，後續也發展出各因素（分量表）為 14 題、9 題與 3 題的短題本量表（Ryff & Keyes, 1995; Ryff, Lee, Essex, & Schmutte, 1994）。國內陸洛（1998）提出中國人幸福感量表完整版 48 題，但後來亦以選取因素負荷量較高之題目作為篩選方式，陸續提出簡短版 20

題、極短版 10 題等短題本測量工具 (Lu & Lin, 1998; Lu, 2006)。上述這些選題做法是植基於古典測驗理論，它的可能缺點即因選題適切性不足或使用非典型樣本（例如：病患等）而影響選題效能，而且古典測驗理論是以信度（reliability）與測量標準誤觀點來驗證測量精確度，然測量標準誤係由信度計算得來，而信度的估算又同時受到試題參數與其他試題間的關聯程度兩者所影響，因而造成研究者在選題過程中，無法得知某一試題的加入與否對整份量表信度的影響程度，亦即在建構短題本測量工具時無法立即性地掌握其測量精確度（余民寧，2009）。

本研究目的旨在發展一份適合華人使用，經濟、有效之評估主觀幸福感之工具。本研究擬以過去余民寧等人（2011）所編制之 39 題「主觀幸福感量表」（Subjective Well-Being Scale, SWBS）為對象，精簡修訂成為一份具有 15 題的「簡式主觀幸福感量表」（Subjective Well-Being Scale Short Form, SWBS-SF），以符合臨床實務應用上之經濟性與效率性的需求。因此，本研究擬提出一套新的量表修訂策略，並簡述其執行架構如下：首先，在原量表情緒幸福感、心理幸福感與社會幸福感三向度的架構下，以分層隨機抽樣取得 1046 名中小學教師有效樣本，採用評定量尺模式（Rating Scale Model, RSM）分別從試題難度（item difficulty）、適配度檢驗、階段難度（step difficulty）與試題差異功能分析（differential item functioning, DIF）等估計角度，選出 15 題「簡式中小學教師主觀幸福感量表」；其次，將工具精簡為 Likert 四點量表計分（原量表為五點量表計分）、進行題目文字修正後，重新再以分層隨機抽樣方式取得另一筆 1180 名中小學教師的有效樣本，接著採用驗證性因素分析方法檢定短題本的因素結構之型態與因素負荷量，並以 RSM 方法進行試題難度估計與資料適配度之檢驗；最後，採用「臺灣憂鬱情緒量表」作為外在效標，以進行效標關聯效度分析，並進行檢定短題本工具之信、效度等心理計量特質指標。

貳、文獻探討

一、主觀幸福感的定義與測量

主觀幸福感的研究源自於西方兩種傳統的觀點，其一為快樂主義取向的幸福觀，其二則為完善論取向的幸福觀（Ryan & Deci, 2001; Waterman, 1993）。所謂

的主觀幸福感內涵包括：正向情緒、負向情緒、與生活滿意度。正向與負向情緒可視為主觀幸福感中的情緒層面評估，生活滿意度可視為主觀幸福感中的認知層面評估。當個人覺得整體生活達到某種滿意程度，且感受到正向的情感平衡，亦即較常體驗到愉快情緒、較少體驗到不愉快或痛苦情緒，此種主觀感受即為幸福（Andrews & Withey, 1976; Diener, Emmons, Larsen, & Griffin, 1985）。

相對於前者，完善論取向的幸福觀則強調生活目的與意義對個人的影響，認為幸福感不單只是由愉悅和快樂的情緒所構成，而應該是個人全身心投入活動，所獲得生理或心理需求的滿足，進而達到自我實現（Waterman, 1993）。Ryff（1989）認為個人經歷生活挑戰，充分發揮自我潛能達到完美體驗即能感受到幸福，又稱為「心理幸福感」。

近年來，Keyes（1998）另外提出「社會幸福感」的概念，有別於過去將幸福感界定為個人對生活主觀評估的滿意度和影響的情緒知覺（情緒與認知取向）（Compbell, Converse, & Rodgers, 1976; Diener, 1984），或是個體的心理功能（心理幸福感）（Ryff, 1989; Ryff & Keyes, 1995），進一步探究個人與社會關係的適應、社會任務的實踐等社會領域，評估個人與他人、群體與社會互動的幸福感。

Keyes（2002, 2005a）、Keyes 和 Waterman（2003）以因素分析方式找出心理健康的組成包括情緒、心理與社會幸福感，且這三種組成間彼此相關、但又可分別存在之個別因素。因此，作為個人心理健康的測量指標。若從操作型定義的觀點來看，「心理健康是一種正向情感與正向生活機能症狀的併發症（syndrome）」。其中，正向情感部分，Keyes 等人認為「情緒幸福感」是個人對於生活中自己的情緒狀態覺知與評估，且概念上與 Andrews 和 Withey（1976）、Diener 等人（1999）所認為對生活的滿意度、以及所感受到的正負向情緒強度相同；而正向機能部分，則包括了心理與社會生活機能，前者即是 Ryff（1989, 1995）、Ryff 和 Keyes（1995）所稱之「心理幸福感」，而後者則是 Keyes（1998）延伸擴充的「社會幸福感」。綜合上述，可形成三向度「主觀幸福感」（subjective well-being）構念。

二、試題反應理論與量表長度簡化

在臨床與實務應用領域中，常因考量到節省時間成本、快速篩檢目的或施測對象作答能力，故選擇短題本作為研究工具。但在短題本工具的發展過程中，簡化量表可採用的方法雖然有很多，但大多出現使用非典型樣本與效度過度高估這

兩類方法學的問題 (Levy, 1968)。從古典測驗理論觀點來看，短題本難以避免被控訴的，即是自動套用全題本信效度及降低效度要求等兩項一般化原罪 (Smith, McCarthy, & Anderson, 2000)。Smith 等人 (2000) 更進一步指出，長題本本身的效度不佳、短題本各因素的內容涵蓋範圍縮小與信度降低、以及兩者間的重疊變異量過低、因素結構或階層改變等情形，都是量表長度簡化所可能遭遇的問題。

然而，試題反應理論從個別試題的觀點出發，以試題特徵曲線來解釋受試者能力特質、試題特性與測驗反應之間的關係 (余民寧, 2009)。Rasch 模式也是客觀的測量，受試者的能力估計值與試題估計值是相互獨立的 (王文中, 2004; 陳柏熹、王文中, 1999)。因此，當資料符合 Rasch 模式時，代表對受試者能力的估計不會受到試題難度影響，具有「試題獨立」(test-free) 特性；同理，對試題難度的估計也不會受到受試者能力高低的影響，具有「樣本獨立」(sample-free) 特性 (趙小瑩、王文中、葉寶專, 2007)。

鑒於本研究所使用的主觀幸福感量表 (長題本) 採用李克特氏量尺模式的五點計分方法，因而選擇 Andrich (1978) 所提出之 Rasch 測量家族中的評等量尺模式 (Rating Scale Model, RSM) 進行分析。RSM 適用於試題反應資料之間具有次序大小關係的情況，例如：非常同意 (5 分)、同意 (4 分)、沒意見 (3 分)、不同意 (2 分)、非常不同意 (1 分)，且量尺中各點數之間的閾值相同，意即受試者在回答所有題目時，對各選項間的心理距離是一樣的，也因此適用於所有題目皆使用相同評等量尺的資料 (施慶麟、王文中, 2006)，在 RSM 中，受試者 n 在第 i 題中得到 j 分比 $j-1$ 分的勝率 (odds) 的對數為：

$$\log\left(\frac{P_{nij}}{P_{ni(j-1)}}\right) = \theta_n - (\delta_i + \tau_j)$$

P_{nij} 和 $P_{ni(j-1)}$ 是受試者在第 i 題得 j 分和 $j-1$ 分的機率， δ_i 是第 i 題的整體難度 (overall difficulty)，或簡稱為「難度」， τ_j 是第 i 題得 j 分的「閾參數」(threshold parameter) 或稱為「階段難度」(step difficulty)。RSM 假設每道試題均有一個整體難度 δ_i ，且要求所有試題都必須共享一套閾參數，且為了解決量尺的不確定性，RSM 將整份測驗的難度平均設定為 0，階段難度的總和亦設定為 0。因此，在 RSM 模式中， τ_j 是固定效果，與受試者或試題均沒有交互作用，亦即受試者在回答量表試題時，對於各選項之間的差異都具有相同之判斷。

Rasch 模式亦可用於檢視量表的信、效度。使用 Rasch 多點計分模式時（以 Likert 五點量表為例），計分類別應符合等級反應（graded response）假定，順序類別應具有互斥、單一意義與詳盡等特性（Guilford, 1965; Stone & Wright, 1994），且在所有的試題都是潛在變項的前提下，適當的等級反應類別應符合：(1) 每一個類別至少有 10 筆觀察值，(2) 觀察分配是有規律的，(3) 類別的平均測量值是單調遞升的，(4) 偏離反應適合度統計（Outfit）均方小於 2.0，(5) 閾值之間至少相差 1.4 個 logits 以上、並小於 5 個 logits（Linacre, 2002）。但實際研究中卻常出現違反等級順序，低類別產生高難度的閾值失序（disordered thresholds）現象。Bond 和 Fox（2007）建議：「針對失序的類別考慮和相鄰的類別合併」，透過 Rasch 模式的階段難度分析可調整評分量尺的順序類別，例如將 5 個順序類別中無效的估計類別加以合併成為 Likert 四點量表，進而提升量表的信度。

在效度分析部分，Wright 與 Masters（1982）出將資料的適配統計，推論作為測驗效度的做法。有兩種效度的證明方法：當試題適配模式時，表示試題的校準是有效的；當受試者表現的適配統計被接受時，表示受試者的測量是有效的。因此，適配度（model-data fit）檢驗，可被用來檢視模式與資料觀察值之間是否合適，使用加權（weighted）及未加權（unweighted）兩種適配指標 MNSQ（mean squares）值，並藉由 Wilson-Hilferty 轉換法將 MNSQ 轉換為近似常態化的 t 分配值，以便利檢視。

隨著量表研究的發展，雖然在心理測驗上建議使用長式、多維度的量表較為精準，但在實際應用上，往往是短式、簡化的版本更能符合研究者的需求。因此，如何能讓節省下的時間成本值得所損耗的測驗效度，便成為量表長度簡化的一個重要議題（Smith et. al., 2000）。Levy（1968）和 Smith 等人（2000）認為，在進行短題本的信效度檢核時，需注意以下幾點：透過實證方式，檢核量表簡化所耗損的效度是否符合時間成本效益；短題本不可直接套用全題本的信效度，而應獨立取樣建立其信效度資料；確認短題本符合全題本的因素或階層結構與所涵蓋的內容，以及長短題本的重疊變異量；確認短題本的每個因素都達到合理的信度水準；與比較長短題本的分類正確性（hit ratio）是否差異過大。

為了嚴謹地建構本研究的「簡式中小學教師主觀幸福感量表」，內在效度部分係指量表內容的代表性或取樣適切性，本研究擬採取獨立與嵌入兩種模式，亦即長、短題本使用不同樣本分開施測，並以驗證性因素分析進行交叉驗證。效標

關聯效度部分係指量表與外在效標間的關係，本研究擬以臺灣憂鬱情緒量表作為外在效標，以驗證主觀幸福感的三因素結構與憂鬱變項間之效標關聯模式的適配程度。

參、研究方法

一、研究參與者

為了嚴謹地建構「簡式中小學教師主觀幸福感量表」，在內在效度上，量表施測採取嵌入與獨立兩種模式，先進行長題本施測篩選出短題本資料，並進行交叉驗證；而後重新抽樣進行短題本施測。因此，本研究共包含兩群研究參與對象，第一個選題樣本（item selection sample）作為「主觀幸福感量表（長題本）」的選題與試題反應分析，第二個交叉驗證樣本（cross-validation sample）目的則在建立「簡式中小學教師主觀幸福感量表（短題本）」之信、效度資料。茲分述如下。

（一）選題樣本

本研究的原始量表係依據教育部公布 96 學年度全國各層級學校數（教育部全球資訊網，2007），以分層隨機抽樣方式按照各層級學校所占比率，抽取 500 所學校。其次，依據全國 25 縣市各級學校之比率，決定各縣市所應分配之學校數目。最後，兼顧性別考量，每校共抽取 8 位教師（男女教師各 4 位），使抽樣的預期樣本數達 4000 人。透過此抽樣過程，確保所抽樣之樣本能依原始母群體之學校數多寡均衡分配，使樣本能具有母群體之代表性。經實際抽樣後的資料整理，本研究共得有效的教師樣本數為 1,046 人，有效樣本回收率為 26.2%，在沒有催收的情況下，尚符合一般問卷調查的回收率情況，一般大約僅有 20% 到 30% 之間而已（Connelly, Brown, & Decker, 2003; Newell, Rosenfeld, Harris, & Hindelang, 2004; Sheehan, 2002）。

本研究有效選題樣本共計 1,046 名教師，排除遺漏值後，性別分布為女性受試者 547 人（52.3%），男性受試者 494 人（47.2%）；教育程度分布為大學學歷者 689 人（65.9%），碩士（含以上）者為共 339 人（32.4%）。

（二）交叉驗證樣本

本研究在修訂短題本時，其樣本抽樣程序與前次相同，係依據教育部公布 102

學年度全國各層級學校之學校數（教育部全球資訊網，2013），以分層隨機抽樣方式按照各層級學校所占比率，抽取國小 348 所、國中 121 所、高中 45 所、高職 20 所，共 534 所學校，每學校抽樣 4 位教師，共計抽取 2,136 名教師，回收之有效樣本數為 1,180 人，回收率 55.2%。排除遺漏值後，有效交叉驗證樣本共計 1180 名教師，女性受試者 645 人（54.7%），男性受試者 534 人（45.3%），教育程度方面以大學學歷者最多，為 611 人（51.8%），其次為碩士（含以上）者，共 551 人（46.7%）。

二、研究工具

（一）主觀幸福感量表

本研究所使用的「主觀幸福感量表」，係採用余民寧等人（2011）所編制的量表。其中「心理幸福感」分量表主要用以探討個人內在的自我心理調適與對生活的宏觀知覺，「社會幸福感」分量表係以探討公眾與社會準則來評量自我的生活機能，而在「情緒幸福感」分量表則是探討個人對生活中自我情緒狀態的覺知與評估。

該量表包括 39 題，依受試者填寫反應自身情況的程度，從極不同意至非常同意等，分別給與 1 ~ 5 分方式計分，經將反向題反向計分後（反向題共 16 題），使得分愈高者代表受試者在該向度之主觀幸福感傾向愈強。經本研究選題樣本分析後，各分量表的 Cronbach' s α 分別為：.72、.70、.87，皆達 .70 以上，且總量表內部一致性信度值為 .84，顯示本量表在主觀幸福感及其三個子向度測量上，均具有相當不錯的穩定性與一致性。

（二）臺灣憂鬱情緒量表

本研究採用余民寧、劉育如與李仁豪（2008）所編制之本土化「臺灣憂鬱情緒量表」（Taiwan Depression Scale, TDS），作為測量受試者心理疾病程度之依據。依受試者反應自身情況的程度，分別給與 0 至 3 分不等，表示「從不如此」至「總是如此」不同程度的作答反應，分數愈高代表憂鬱情緒傾向愈嚴重。

本量表內容包含四個向度因素，分別為認知、情緒、身體、與人際關係等。本量表以教師樣本（即交叉驗證樣本）進行分析後，四個向度之內部一致性信度係數值分別為 .82、.82、.81 與 .82，總量表內部一致性信度係數值則為 .93，顯示本量表在憂鬱情緒及其四個子因素測量上，均具有相當不錯的穩定性與一致性。

在本研究中，本量表將用來提供簡式中小學教師主觀幸福感量表的效標關效度之證據。

三、研究程序與資料分析

本研究包含兩次修訂，第一次修訂為選題與 DIF 分析。為達到文字的一致性，本研究進行第二次修訂，並進行驗證性因素效度考驗。在資料分析部分有兩個重點，其一是依據試題反應分析結果進行選題，其二則為短題本的信、效度考驗。本研究使用 ConQuest、Amos 等統計軟體，來進行相關的資料統計分析。

（一）第一次修訂：選題與差異試題功能分析

本研究以第一群參與者進行選題與差異試題功能分析，運用 Rasch 模式的評定量尺模式（即 RSM）分別從試題難度估計、適配度檢驗與階段難度進行分析。RSM 模式假定量表中各評定點數之閾值（threshold parameter）相同，且對於每位受試者而言，其效果是固定的，因此適用於分析所有題目皆使用相同評等量尺的資料（施慶麟、王文中，2006）。

首先，針對心理、社會與情緒幸福感等三個向度進行模式的試題難度估計、適配度檢驗與階段難度分析。其分析程序與判讀方法如下所述：

1. 試題難度估計：題目難度值應介於 -3 ~ 3 範圍，數值愈大代表該試題愈困難，數值愈小代表該試題愈簡單（余民寧，2009）。
2. 適配度（model-data fit）檢驗：適配度考驗係以檢視模式與原始資料觀察值之間是否合適，以作為刪題的參考指標。在適配指標上，ConQuest 軟體提供加權（weighted）及未加權（unweighted）兩種適配指標 MNSQ（mean squares）值，根據 Linacre（2006）對 MNSQ 合理範圍的建議， $MNSQ > 2$ 表示該題將扭曲或破壞測量系統； $MNSQ$ 在 1.5 ~ 2 之間表示該題對測量的建構雖不具生產性，但也不具破壞性； $MNSQ$ 在 0.5 ~ 1.5 之間表示該題對測量具生產性； $MNSQ < 0.5$ 表示該題對測量有較少生產性。此外，也可以 t 值判斷題目是否符合模式選取標準，同樣具有加權與未加權 t 值，其理想值應介於 ± 2 之間。
3. 階段難度分析：階段難度值應該由小到大呈現單調遞增排列，數值愈大表示受試者愈不容易達到該階段的標準，且閾值之間至少相差 1.4 個 logits 以上、並小於 5 個 logits（余民寧，2009；Linacre，2004）。

其次，刪除難度過於困難或容易、適配度不佳試題後，進行性別的差異試題

功能分析（DIF），當題目難度值在兩個群體間的差距達 0.5logits 時，此時的 DIF 效果已不容忽視（Wang，2008），將予以刪除該題。

（二）第二次修訂

在不改變題意的前提下，修訂短題本中的部分文字，例如：將社會幸福感分量表中，擴大「社區」一詞概念改為「社區（社會、社群）」等。此外，為了檢驗短題本在不同樣本的穩定性，重新進行分層隨機抽樣，針對心理、社會與情緒幸福感三個向度，以 CFA 檢核短題本的因素結構與模式適配度，並以 RSM 進行試題難度估計與資料適配度檢驗，採 EAP/PV 信度來考驗量表的信度，當係數高於 .70 時，代表量表之內部一致性信度良好。最後，以臺灣憂鬱情緒量表進行效標關聯效度分析，並對於以短題本作為篩選工具的效益進行評估。

在模式適配指標的選取與參數估計上，本研究是以 Amos 軟體並使用最大概似估計法進行各項參數值的估計，同時採納 Hu 和 Bentler（1998）的建議，除選取傳統卡方值外，至少在各類別（即絕對適配指標、相對適配指標等）中再選取 1 個或以上的適配指標。而在判定模式適配的準則上，本研究參考 Hu 與 Bentler（1999）、Vandenberg 和 Lance（2000）的建議，以 GFI、CFI、NFI、NNFI（以上指標需大於 .90）及 RMSEA（需小於 .08）等四項指標，做為模型適配度之判準依據。

參、結果與討論

一、長題本施測與選題

（一）難度估計與試題適配度檢驗

初步分析結果，長題本試題難度分布略成負偏態，39 題中的難度估計值介於 -1.000 ~ 0.722 之間，尚介於 -3 ~ 3 的合理範圍，且難度值負數約占半數（19 題），表示對受試者而言，長題本試題的難易適中，而試題難度估計誤差範圍從最小的 0.024 至最大的 0.107 皆很微小，代表以三向度模式來估計主觀幸福感量表（長題本）尚稱精確，如表 1 所示。

在適配度指標方面，第 15、17 題之 $MNSQ > 1.5$ ，表示該試題可能測到受試者其他面向的特質，第 34、35、36 題之 $MNSQ < 0.5$ ，則表示這 3 題與其他試題

相似性高，以上 5 題對測量受試者的主觀幸福感貢獻較少。大多數 MNSQ 都介於 0.5 ~ 1.5 之間，適配情形良好。不過由 T 值來看，出現許多絕對值過高的 T 值，一般而言，T 值的絕對值在 2 或 3 以內是較佳的，然而根據 Linacre (2006) 的建議，如果 MNSQ 值是可接受的話，則 T 值可以被忽略。

表 1 主觀幸福感量表（長題本）量表的試題反應分析結果摘要

向度	測量 指標	試題摘要內容	難度	標準誤	Outfit		Infit	
					MNSQ	T	MNSQ	T
心理 幸福 感	獨立 自主	01. 被他人強烈意見影響 (-)	0.722	0.025	1.11	2.4	1.10	2.6
		02. 對自己的看法很有信心，與一般的 輿論相反 *	0.245	0.025	0.91	-2.1	0.88	-3.0
		03. 依據自己的見解，而不是別人的價 值觀	-0.274	0.026	0.91	-2.0	0.89	-2.4
	環境 掌控	04. 我覺得我能掌控我生活中的一切情 況	-0.479	0.027	0.75	-6.4	0.75	-5.8
		05. 日常生活中的瑣事常讓我感到沮喪 (-)*	0.626	0.025	0.99	-0.2	0.99	-0.3
		06. 我能妥善處理日常生活中所需要負 擔的責任	-0.719	0.027	0.66	-8.9	0.65	-8.1
	生活 目標	07. 我現在活在當下，不想去思考未來 (-)	0.581	0.025	1.36	7.4	1.32	7.7
		08. 有些人會覺得生活沒有目標，但我 不會 *	-0.501	0.027	0.97	-0.7	0.98	-0.4
		09. 有時候我會覺得自己好像已經完成 人生中所有該做的事 (-)	0.493	0.025	1.37	7.5	1.36	8.5
	自我 接納	10. 當回顧我的一生時，我會很欣然接 受	-0.352	0.026	1.06	1.3	1.05	1.2
		11. 我喜歡我人格中存在的多個面向	-0.347	0.026	0.94	-1.4	0.93	-1.5
		12. 在很多方面，我會對自己的成就感 到失望 (-)*	0.641	0.025	1.06	1.4	1.06	1.5
	與他人 建立積 極關係	13. 與他人維持親密關係是一件感到困 難的事 (-)	0.588	0.025	0.17	3.7	0.17	4.2
		14. 願意花時間幫助別人，別人認為肯 付出	-0.470	0.027	1.92	-1.9	1.90	-2.1
		15. 沒有體驗過與他人維持溫暖又信任 的關係 (-)*	0.374	0.025	1.81	14.9	1.83	17.1

(續下頁)

表 1 主觀幸福感量表（長題本）量表的試題反應分析結果摘要（續）

向度	測量 指標	試題摘要內容	難度	標準誤	Outfit		Infit	
					MNSQ	T	MNSQ	T
心理 幸福 感	個人 成長	16. 擁有新經驗，進而挑戰自己對世界的想法	-0.876	0.028	0.98	-0.5	0.97	-0.6
		17. 人生是一種不斷學習、改變的持續性過程	-0.888	0.028	2.69	26.8	2.38	20.2
		18. 放棄嘗試大幅改變我的生活方式很久 (-)*	0.636	0.107	1.17	3.6	1.15	3.8
社會 幸福 感	社會 統整	19. 不屬於社區中的一份子 (-)*	-0.066	0.025	1.18	3.8	1.18	4.5
		20. 我覺得我與社區中的每個人都很親近	0.236	0.024	0.78	-5.5	0.77	-6.8
		21. 我的社區是一個令人感覺舒適的地方	-0.406	0.025	0.80	-5.0	0.79	-5.6
	社會 接納	22. 我覺得一般人對幫助他人是不求回報的	-0.141	0.025	1.04	1.0	1.01	0.3
		23. 我認為一般人對他人的問題是漠不關心的 (-)	0.204	0.024	0.87	-3.0	0.86	-3.9
	社會 貢獻	24. 一般人是善良的 *	-1.000	0.026	0.79	-5.3	0.78	-4.9
		25. 有寶貴的東西值得留給這個世上 *	-0.654	0.026	0.94	-1.5	0.94	-1.3
		26. 日常行為舉止無法對社區產生任何貢獻 (-)	0.040	0.024	0.86	-3.5	0.85	-4.1
	社會 實現	27. 我沒有任何重要的東西可以貢獻給社會 (-)	-0.055	0.025	1.25	5.4	1.25	6.1
		28. 對每個人而言，這個世界已經變得比以前更好	0.588	0.024	1.18	3.9	1.17	4.3
		29. 覺得社會已經停止再進步了 (-)*	0.181	0.024	1.01	0.3	1.00	0.1
	社會 一致性	30. 對我來說，社會並沒有改善許多 (-)	0.380	0.024	1.05	1.1	1.04	1.1
		31. 對我來說，這個世界太複雜了 (-)	0.289	0.024	1.22	4.7	1.20	5.2
		32. 無法理解世界上到底發生什麼事 (-)*	-0.111	0.025	1.24	5.1	1.25	6.1
	情緒 幸福 感	33. 我覺得很容易去預測社會即將發生什麼事	0.515	0.093	0.92	-2.0	0.90	-2.7
覺察 生活 滿意	公認 快樂	34. 覺得自己時時充滿喜悅 *	0.086	0.027	0.43	-17.0	0.41	-18.1
		35. 我感覺到自己神采飛揚	0.196	0.027	0.43	-16.8	0.41	-18.7
		36. 認為自己是個極快樂的人 *	0.326	0.027	0.49	-14.5	0.47	-16.2
	覺察生 活滿意	37. 我的生活非常沈穩祥和	-0.368	0.028	0.55	-12.3	0.56	-11.8
		38. 很滿意目前自己的生活 *	-0.167	0.028	0.65	-9.1	0.64	-9.6
		39. 覺得自己的生活很豐盛 *	-0.073	0.061	0.57	-11.5	0.57	-11.8

註：(-) 代表反向題、* 代表選入短題本

三向度主觀幸福感模式的 4 個階段難度估計值，係採升幂方式排列，依序為：-2.025（step 1）、-0.363（step 2）、-0.023（step 3）、2.410（step 4）。根據 Linacre（2004）對等級量表閾值的建議：閾值需呈現單調遞增，且閾值之間至少相差 1.4 個 logits 以上、並小於 5 個 logits。而本研究資料顯示，各階段閾值均呈單調遞升，且估計誤差小，表示估計結果精準，但第二、三階段閾值相差小於 1.4 個 logits，且原始與調整適合度之 T 值大於 2，表示階段閾值的適合度較差，如表 2 所示。因此，在短題本的階段參數設定上，將考慮減少成為 3 個階段難度參數即可。

表 2 主觀幸福感量表（長題本）的試題閾值分析結果摘要

	閾值	標準誤	Outfit		Infit	
			MNSQ	T	MNSQ	T
0			3.32	22.9	2.20	10.6
1	-2.025	0.021	1.59	7.8	1.43	7.2
2	-0.363	0.017	1.97	11.8	2.09	18.0
3	-0.023	0.015	2.60	17.5	2.61	25.4
4	2.410*		2.78	18.9	2.83	24.8

（二）性別的差異試題功能分析（DIF analysis）

在性別分析上，在長題本 39 題中沒有超過 0.5 個 logits 以上的題目存在，此即表示並無明顯的 DIF 現象存在，如表 3 所示。最後，綜合難度估計、適配度檢驗與 DIF 分析結果，參酌文獻與長題本的因素結構，修改第 25 題為「我對這個社區（社會、社群）做一些有意義的事」，形成短題本的中小學教師主觀幸福感量表，並改採用 Likert 四點量表（非常同意、同意、不同意、非常不同意），以再進行第二次施測與信效度分析。

表 3 長題本量表的性別差異試題功能分析摘要

題號	估計值（男性）	估計值（女性）	DIF	檢測結果
01	-0.019	0.019	0.038	無
02	-0.110	0.110	0.220	無

（續下頁）

表 3 長題本量表的性別差異試題功能分析摘要（續）

題號	估計值（男性）	估計值（女性）	DIF	檢測結果
03	-0.092	0.092	0.184	無
04	-0.054	0.054	0.108	無
05	-0.025	0.025	0.050	無
06	-0.012	0.012	0.024	無
07	0.049	-0.049	0.098	無
08	-0.024	0.024	0.048	無
09	0.002	-0.002	0.004	無
10	0.079	-0.079	0.158	無
11	0.064	-0.064	0.128	無
12	0.052	-0.052	0.104	無
13	-0.009	0.009	0.018	無
14	0.033	-0.033	0.033	無
15	-0.006	0.006	0.012	無
16	0.060	-0.060	0.120	無
17	0.136	-0.136	0.272	無
18	0.064	-0.064	0.128	無
19	-0.011	0.011	0.022	無
20	-0.089	0.089	0.178	無
21	-0.013	0.013	0.026	無
22	0.027	0.027	0.054	無
23	-0.059	-0.059	0.118	無
24	-0.037	-0.037	0.074	無
25	-0.006	0.006	0.012	無
26	0.006	-0.006	0.012	無
27	0.008	-0.008	0.016	無
28	-0.055	0.055	0.110	無
29	-0.031	0.031	0.062	無
30	0.031	-0.031	0.062	無
31	0.030	-0.030	0.060	無

（續下頁）

表 3 長題本量表的性別差異試題功能分析摘要（續）

題號	估計值（男性）	估計值（女性）	DIF	檢測結果
32	-0.006	0.006	0.012	無
33	-0.112	0.112	0.224	無
34	-0.004	0.004	0.008	無
35	-0.007	0.007	0.014	無
36	-0.004	0.004	0.008	無
37	-0.017	0.017	0.034	無
38	0.019	-0.019	0.038	無
39	0.004	-0.004	0.008	無

（三）選題與修訂

雖然主觀幸福感量表（長題本）39 題的自陳量表作答時間在 5-8 分鐘以內，但適配度檢驗發現，部分題目可能測到跨面向特質（如 15、17 題），部分題目與其他相似性過高（如 34、35、36 題）。此外，長題本採用五點量尺計分，其中第二、三階段閾值相差小於 1.4 個 logits，階段閾值的適合度較差，因此研究者著手進行主觀幸福感量表（長題本）之簡化與修訂。

過去建構這種具有階層性因素結構的短題本時，研究者有時候會只追求維持二階因素，挑選對二階因素貢獻量最大的一階因素，導致一階因素減少進而降低二階因素涵蓋範圍，影響量表效度（Costa & McCrae, 1992）。為了確保短題本與全題本有相同的因素結構，Smith、McCarthy 與 Anderson（2009）建議將選題的抽取單位降到題項，分別以一階因素為群組去抽選，來改善代表性不足之情況。其中，在「與他人建立積極關係中」該群組中，第 14 題 MNSQ 介於 1.5 ~ 2.0 之間不具生產性捨棄，比較第 13 題與 15 題之後，研究者認為雖然第 13 題有少量生產性效果，但試題文字描述「與他人維持親密關係是一件感到困難的事（反）」容易誤導受試者將與他人建立的正向積極關係限定為親密關係，因此從題意內涵上考量，選取第 15 題「沒有體驗過與他人維持溫暖又信任的關係（反）」較為適切。此外，考量簡式中小學教師主觀幸福感三向度的題數盡可能均等，因此情緒幸福感的兩個群組皆增加選取（2 題），而在「公認快樂」該群組中，因為第 35 題的 MNSQ<0.5 少量生產性效果，但試題文字描述「我感覺到自己神采飛揚」受試者較不易理解，因此捨棄該題，選入其他 34 與 36 題。

因此，本研究以余民寧等人（2011）所提出之主觀幸福感三向度的測量指標作為一階因素之群組，綜合試題難度估計、適配度檢驗、階段難度與差異試題功能分析的結果，在原始二階「主觀幸福感量表」三因素模式（即含心理、社會、情緒幸福感三因素）（余民寧等人，2011）下，審視各因素的內涵後選入 15 題，作為「簡式主觀幸福感量表」的題目，如表 1 所示。

二、簡式中小學教師主觀幸福感量表（短題本）施測結果

（一）驗證性因素分析

以驗證性因素分析檢定「簡式中小學教師主觀幸福感量表」的模式適配度，結果如下表 4 所示。比較一階三因素模式與二階三因素模式的結果發現，二階三因素模式的適配表現雖可接受，但其中可能在某些測量結構面向還是不夠適配，一階三因素模式則相對較佳，除卡方值達顯著外，SRMR、RMSEA 都能分別達到 Hu 和 Bentler（1999）建議的 .08、.06 的適配水平，而相對適配指標 NNFI = .875 與 CFI = .897 接近 .90 水準。

本研究進一步檢視各測量指標的標準化因素負荷量發現，各數值皆已達到顯著水準，各分量表因素負荷量數值分布情形如下：在「心理幸福感」分量表中，除 item1 因素負荷量為 .33 較低外，其餘均分布在 .49~.67 之間；在「社會幸福感」分量表中，從最小的 .45~ 最大的 .69；在「情緒幸福感」分量表中，從最小的 .72~ 最大的 .83，皆具有不錯之解釋水準（參見圖 1 所示）。此外，心理、社會與情緒幸福感三者間之潛在相關係數分別為： $\phi_{12}=.91$ 、 $\phi_{13}=.81$ 、 $\phi_{23}=.65$ ，皆達顯著正相關。短題本驗證性因素分析的結果大致與 Keys、Shmotkin 與 Ryff（2002）以及 Keys（2005）所提出的因素結構一致，也與長題本的模式相符（余民寧等人，2011）。

表 4 簡式中小學教師主觀幸福感量表的模式適配度摘要

模式	df	χ^2	SRMR	RMSEA	NNFI	CFI
一階三因素模式	87	583.517	.0506	.076	.875	.897
二階三因素模式	87	817.882	.0638	.086	.840	.867

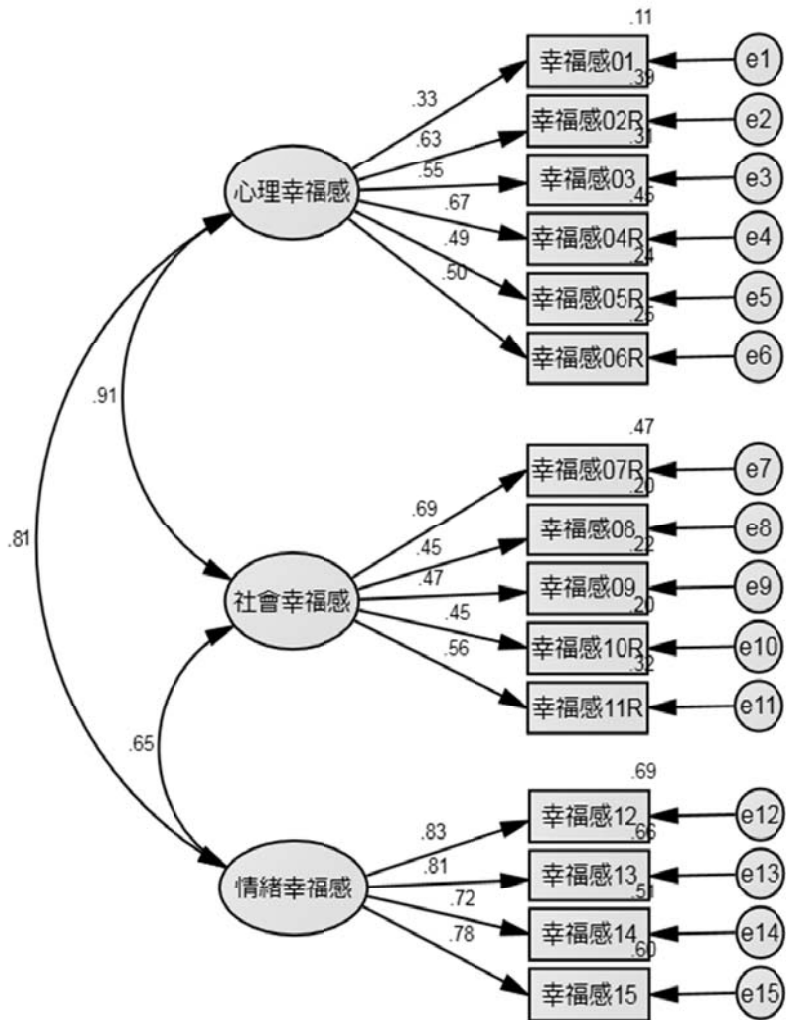


圖 1 簡式中小學教師主觀幸福感三因素模式圖

(二) 難度估計與試題適配度檢驗

短題本試題難度分布略成負偏態，15 題中的難度估計值介於 -1.543 ~ 0.842 之間，尚介於 -3 ~ 3 的合理範圍，且難度值負數約占半數（8 題），表示對受試者而言，短題本試題的難易適中，而試題難度估計誤差範圍從最小的 0.038 至最大的 0.086 皆很微小，表示以三向度模式估計短題本試題難度尚稱精確，如表 5 所示。

在適配度指標方面，短題本所有試題 MNSQ 都介於 0.59 ~ 1.46 之間，適配情形良好，對主觀幸福感測量有生產性效果。不過由 T 值來看，仍有部分絕對值過高的 T 值，根據 Linacre (2006) 的建議，如果 MNSQ 值是可接受的話，則 T 值可以被忽略。

在階段難度估計值方面，三個階段難度閾值呈現單調遞增排列，依序為：-2.775 (step 1)、-0.575 (step 2)、3.350 (step 3)，閾值之間相差至少 1.4 個 logits 以上、並小於 5 個 logits，表示短題本改採 Likert 四點量表計分是頗為合理的選擇。階段難度估計誤差小，表示短題本的階段難度估計結果頗為精準（下表 6）。

表 5 簡式中小學教師幸福感量表（短題本）量表的試題反應分析結果摘要

題號	試題摘要內容	三向度模式					
		難度	標準誤	Outfit		Infit	
				MNSQ	T	MNSQ	T
01	對自己的看法很有信心，即使與一般輿論相反	0.456	0.038	1.28	6.4	1.29	6.1
02	日常生活中的瑣事常讓我感到沮喪（-）	0.771	0.037	1.08	1.9	1.07	1.7
03	有些人會覺得生活沒有目標，但我不會	-0.391	0.039	1.09	2.2	1.10	2.3
04	在很多方面，我會對自己的成就感到失望（-）	-0.009	0.038	1.04	1.1	1.07	1.6
05	從來沒有體驗過與他人維持溫暖又信任關係（-）	-1.543	0.039	1.32	7.3	1.36	8.6
06	已經放棄嘗試大幅改變我的生活方式很久（-）	0.715*	0.086	1.46	10.1	1.42	8.8
07	覺得不屬於社區（社會、社群）中的一份子（-）	-0.814	0.040	1.12	2.8	1.17	4.0
08	相信一般人是善良的	-0.201	0.040	1.05	1.3	1.07	1.7
09	對這個社區（社會、社群）做一些有意義的事	-0.151	0.040	0.78	-5.8	0.80	-4.9
10	覺得社會已經停止再進步了（-）	0.842	0.039	1.29	6.7	1.29	6.3
11	無法理解世界上到底發生什麼事（-）	0.324*	0.080	1.02	0.4	1.03	0.8
12	自己時時充滿喜悅	0.018	0.038	0.59	-12.0	0.61	-10.6
13	自己是個快樂的人	0.669	0.038	0.67	-9.2	0.68	-8.9
14	滿意目前的生活	-0.377	0.039	0.70	-8.4	0.74	-6.8
15	自己生活很豐盛	-0.310*	0.066	0.67	-9.3	0.70	-7.9

註：（-：反向題）。

表 6 簡式中小學教師幸福感量表（短題本）的試題閾值分析結果摘要

閾值	標準誤	Outfit		Infit	
		MNSQ	T	MNSQ	T
0		45.63	192.1	3.28	17.4
1	-2.775	1.54	11.6	1.33	6.5
2	-0.575	2.10	21.0	2.11	21.4
3	3.350*	2.25	23.2	1.86	15.1

（三）題目與受試者適合度之分布

由圖 2 所示可知，由左至右分別呈現「心理」、「社會」、與「情緒幸福感」三個向度的受試者與題目的對應分布。由圖顯示，受試者的能力分布在心理幸福感、社會幸福感與情緒幸福感等三個向度上皆大於試題難度值的分布，顯示受試者的主觀幸福感程度相對於試題難度而言是偏高的。本短題本「簡式中小學教師主觀幸福感量表」適合應用於多數具中等能力程度的受試者，其試題難度參數估計值約座落於 ± 1 個 logits 範圍之間。而本圖顯示受試者的能力分布遠大於試題難度值的分布，此乃這批受試者是全國中小學教師之故，教師一般都被認為是一群相對較為幸福的工作族群。

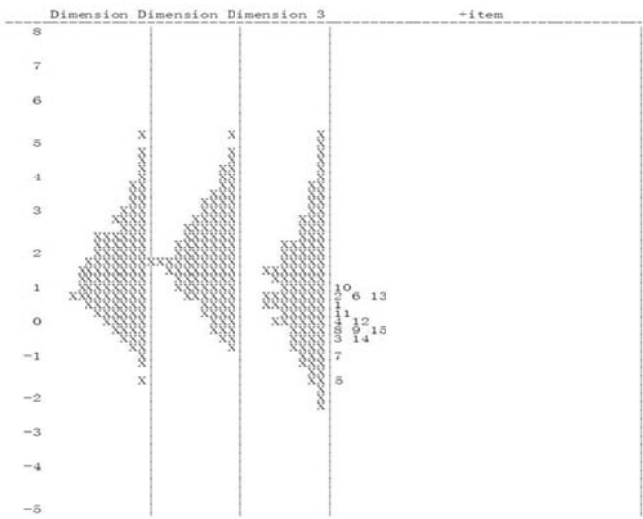


圖 2 受試者與題目的對應分布

（四）分量表信度及相關

以多向度 **RSM** 分析主觀幸福感的三個分量表，如表 7 結果顯示，心理幸福感、社會幸福感和情緒幸福感三個分量表的信度分別為 0.82、0.77、0.78，EAP/PV 信度係數皆高於 .70，表示各分量表之內部一致性信度良好。此外，心理與社會幸福感的相關為 .86，心理與情緒幸福感的相關為 .77，社會與情緒幸福感的相關為 .61，三個分量表之間為中高度相關。

表 7 三個分量表的信度係數與相關係數

分量表	1	2	3
1 心理	(.82)		
2 社會	.86	(0.77)	
3 情緒	.77	0.61	(0.78)

註：（）內為各分量表的信度係數。

三、簡式中小學教師主觀幸福感量表與臺灣憂鬱情緒量表之效標關聯效度

正向心理學指出，幸福感是個人重要的內在心理能量之一，並影響個體心理健康（Sin & Lyubomirsky, 2009; Snyder & Lopez, 2002, 2007）。心理健康的個體所主觀知覺到的幸福感，不僅僅是一種快樂、滿足的正向感受，也包含了個體在面對內、外在不同環境所發展出的高度正向機能。亦即，對內，個體展現正面的心理機能，具有自我導向的自主性、能接納現在與過去的自我、以及強調積極的個人成長等；對外，個體發揮正面的社會機能，認知自我是屬於社會的一分子，能接受社會的複雜性與個人對社會的貢獻與價值。因此，就積極面來看，高度的主觀幸福感是具有生產力的（productive），可調節憂鬱傾向、改善憂鬱症狀（余民寧、許嘉家、陳柏霖，2010；Sin & Lyubomirsky, 2009）；而就消極面來看，高度幸福感則與自殺、憂鬱等心理疾病有顯著的負相關（Keyes & Waterman, 2003）。因此，本研究以余民寧等人（2008）所編制之「臺灣憂鬱情緒量表」為效標，提供簡式中小學教師主觀幸福感量表的外在效標關聯效度之證據。

以簡式中小學教師主觀幸福感量表的三因素幸福感模式，進行與憂鬱變項間

的關聯分析，模式驗證結果如表 8 所示。SRMR 與 RMSEA 皆能符合 Vandenberg 和 Lance（2000）所建議之最低上限標準 .10 和 .08，但相對適配指標 NNFI = .807 與 CFI = .836 卻偏離可接受之底線值 .90，整體看來，三因素幸福感與憂鬱間之效標關聯模式的適配表現尚可接受。進一步檢視此三因素幸福感與憂鬱變項間之潛在相關係數分別為 -.74、-.65 與 -.69，皆達負向顯著水準（參見圖 3 所示），顯示以中小學教師為研究樣本進行主觀幸福感測量，與憂鬱變項間有顯著的負相關，亦即「主觀幸福感愈高的教師，其憂鬱程度愈低」，這與 Ryff 和 Keyes（1995）、Keyes（2005a）的研究分析結果一致，亦對本研究簡式中小學教師主觀幸福感量表提供了有效的效標關聯效度之證據。

表 8 憂鬱情緒量表與簡式中小學教師主觀幸福感量表間之關聯模式分析摘要

模式	df	χ^2	SRMR	RMSEA	NNFI	CFI
一階三因素主觀幸福感與憂鬱量表之效標關聯模式	102	1011.200	.0979	.095	.807	.836

肆、結論與建議

本研究之「簡式中小學教師主觀幸福感量表」發展，係採理論建構取向（theoretical approach），參考 Keyes（2002, 2005a）、Keyes 和 Waterman（2003）的三因素幸福感（心理、社會、情緒）模式，將過去余民寧、謝進昌、林士郁、陳柏霖、曾筱婕（2011）所編制之 39 題「主觀幸福感量表」進行文字修訂、並精簡成為 15 題的「簡式中小學教師主觀幸福感量表」，以符合臨床應用與實務研究上之經濟性與效率性的需求。本量表包含「心理幸福感」（6 題）、「社會幸福感」（5 題）與「情緒幸福感」（4 題）三個分量表，採 Likert 四點評量作答。以評定量表模式（RSM）進行試題分析，短題本試題難度估計值介於 -1.543 ~ 0.842 之間，難易適中。在適配度指標方面，所有試題 MNSQ 都介於 0.59 ~ 1.46 之間，適配情形良好，對主觀幸福感測量具有生產性效果。在階段難度估計值方面，三個階段難度閾值均呈現單調遞增排列，依序為：-2.775（step 1）、-0.575（step

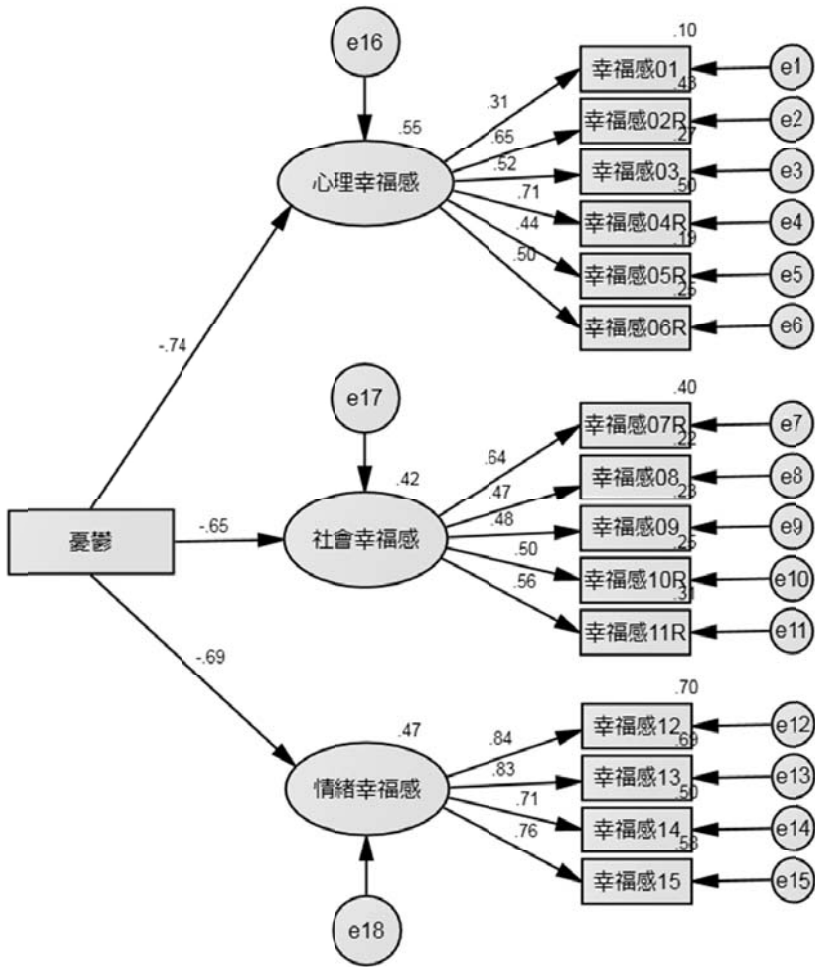


圖 3 憂鬱情緒與簡式中小學教師主觀幸福感間之關聯模式圖

2)、3.350 (step 3)，閾值之間相差至少 1.4 個 logits 以上、並小於 5 個 logits，表示短題本採用 Likert 四點量表計分是合理的選擇。在信度部分，三個分量表的信度係數分別為：0.823、0.770、0.777，EAP/PV 信度係數皆高於 .70，表示各分量表之內部一致性信度良好。在效度部分，驗證性因素分析發現，相對適配指標 NNFI = .875 與 CFI = .897 接近 .90 水準，SRMR = .0506 ($< .08$) 與 RMSEA = .076 ($< .08$)，顯示此模式具有可接受之模式適配度 (Hu & Bentler, 1999)，

符合 Keyes、Shmotkin 與 Ryff (2002) 以及 Keys (2005) 所提出之主觀幸福感的三因素模式，提供本研究簡式主觀幸福感內在效度之證據。此外，以「臺灣憂鬱情緒量表」作為外在效標，進行三因素模式與憂鬱變項間的關聯分析發現，相對適配指標 NNFI = .807 與 CFI = .836，SRMR = .0979 ($< .10$) 與 RMSEA = .095 ($< .10$)，顯示三因素幸福感與憂鬱間之效標關聯模式的適配表現尚可接受 (Vandenberg & Lance, 2000)，提供本研究簡式主觀幸福感量表外在效度之證據。

以下茲針對各項研究結果，提出幾項結論與建議，供未來研究與實務工作者之參考。

一、主觀幸福感之理論建構取向

本研究在驗證「主觀幸福感量表」具三因素模式上，檢驗心理、社會與情緒幸福感之測量結構與品質，發現整體測量模式適配度，符合 Hu 和 Bentler (1999) 提出的建議，模式具有相當不錯的適配度表現。僅在 item1「我對自己的看法很有信心，即使它與一般的輿論相反」(獨立自主)的因素負荷量較低，其餘測量指標皆具備不錯的構念解釋權重。整體而言，本研究「簡式中小學教師主觀幸福感量表」之因素結構，與 Keyes (2002) 以及 Keyes (2005) 的研究結論相當一致，也與長題本的因素結構相同。整體而言，主觀幸福感為實徵上有關聯，但概念上不同之三種構念所組合成的量表。

二、量表長度簡化的信、效度分析

量表簡化大致上可採用分量表(分測驗)抽樣、分層抽樣、題項抽樣、因素抽樣、受試者特質抽樣等不同方法 (Levy, 1968)，但如何避免因長度簡化所可能伴隨產生的因素內容涵蓋範圍縮小、因素結構或階層改變、以及信度降低與效度不佳等問題，便成為發展短題本方法學上的重要議題。本研究採取試題反應理論的觀點，運用 RSM 與 CFA 來檢視短題本的信、效度。信度部分，採用 EAP/PV 來考驗量表的內部一致性信度。效度部分，長、短題本使用不同樣本分開施測，並以驗證性因素分析進行因素結構驗證，提供內在效度證據；並以臺灣憂鬱情緒量表作為效標，驗證主觀幸福感的三因素結構與憂鬱變項間之效標關聯模式的適配程度，提供外在效度證據。綜合上述研究結果顯示，簡式中小學教師主觀幸福感量表具有良好之信、效度，可作為評估主觀幸福感之良好工具。

三、未來研究的建議

修訂後之「簡式中小學教師主觀幸福感量表」共計 15 題，內容涵蓋心理、社會與情緒幸福感三個分量表，不僅在因素結構上與原量表相符，且具有良好之信效度，可作為省時、符合經濟效益之良好評估工具。因此，未來研究可進一步以評等量尺模式進行「主觀幸福感」決斷分數研究，對於幸福感低落的個體，可建議其尋求諮商員或心理醫生的輔導與協助。此外，本研究係以教師為研究樣本，未來研究如可擴及一般社會大眾、高齡長者、心理疾病患者等不同族群，並發展常模，以促使本量表之應用價值大為提升。

謝誌

本文作者擬感謝科技部補助本研究案經費，補助編號為：NSC 102-2410-H-004-191-SS2。關於本論文，作者要感謝匿名評審提供寶貴的審查建議。

參考文獻

- 王文中(2004)。Rasch 測量理論與其在教育和心理之應用。**教育與心理研究**，**27**，637-694。
- 余民寧(2009)。**試題反應理論(IRT)及其應用**。臺北市：心理。
- 余民寧、許嘉家、陳柏霖(2010)。中小學教師工作時數與憂鬱間的關係——主觀幸福感的觀點。**教育心理學報**，**42**(2)，229-252。
- 余民寧、陳柏霖、許嘉家(2010)。教師憂鬱傾向的影響因素之研究。**輔導與諮商學報**，**32**(2)，73-97。
- 余民寧、劉育如、李仁豪(2008)。臺灣憂鬱症量表的實用決斷分數編制報告。**教育研究與發展期刊**，**4**(4)，231-257。
- 余民寧、謝進昌、林士郁、陳柏霖、曾筱婕(2011)。教師主觀幸福感模式建構與驗證之研究。**測驗學刊**，**58**，55-85。
- 施慶麟、王文中(2006)。評等量尺的 Rasch 分析。**教育與心理研究**，**29**，399-421。
- 教育部全球資訊網(2007)。**各級學校概況表**。取自 http://140.111.34.54/statistics/content.aspx?site_content_sn=8869
- 教育部(2013)。**各級學校概況表**。取自 http://140.111.34.54/statistics/content.aspx?site_content_sn=8869
- 陳柏熹、王文中(1999)。生活品質量表的發展。**測驗年刊**，**46**(1)，57-74。
- 陳柏霖、余民寧(2015)。從幽谷邁向巔峰：臺灣民眾心理健康的分布。**成人及終身教育學刊**，**24**，35-65。
- 陸洛(1998)。中國人幸福感之內涵、測量及相關因素探討。**人文及社會科學**，**8**，115-137。
- 趙小瑩、王文中、葉寶專(2007)。正負向情感及身體激起量表之發展與試題反應分析：焦慮症與憂鬱症三角模式的再探討。**測驗學刊**，**54**，223-257。
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, **43**, 561-573.
- Andrews, F. M., & Withey, S. B. (1976). *Social indications of well-being*. New York, NY: Plenum.

- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluation and satisfaction*. New York, NY: Russell Sage Foundation.
- Connelly, N. A., Brown, T. L., & Decker, D. J. (2003). Factors affecting response rates to natural resource-focused mail surveys: Empirical evidence of declining rates over time. *Society & Natural Resources*, 16, 541-549.
- Costa, P. T. Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Diener, E. D. (1984). Subjective well-being. *Psychological Bulletin*, 95, 542-575.
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49, 71-75.
- Diener, E. D., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125, 276-302.
- Guilford, J. P. (1965). *Fundamental statistics in psychology and education* (4th ed.). New York, NY: McGraw-Hill.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to under parameterized model misspecification. *Psychological Methods*, 3, 424-453.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly*, 61, 121-140.
- Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, 43 (2), 207-222.
- Keyes, C. L. M. (2003). Complete mental health: An agenda for the 21st century. In C. L. M. Keyes, & J. Haidt (Eds.), *Flourishing: Positive psychology and the life well-lived* (pp. 293-312). Washington, DC: American Psychological Association.
- Keyes, C. L. M. (2005a). Mental illness and/or mental health? Investigating axioms of

- the complete state model of health. *Journal of Consulting and Clinical Psychology*, 73(3), 539-548.
- Keyes, C. L. M. (2005b). Chronic physical conditions and aging: Is mental health a potential protective factor? *Ageing International*, 30(1), 88-104.
- Keyes, C. L. M. (2007). Promoting and protecting mental health as flourishing: A complementary strategy for improving national mental health. *American Psychologist*, 62(2), 95-108.
- Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology*, 82, 1007-1022. <http://dx.doi.org/10.1037/0022-3514.82.6.1007>
- Keyes, C. L. M., & Waterman, M. B. (2003). Dimensions of well-being and mental health in adulthood. In M. H. Bornstein, L. Davidson, C. L. M. Keyes & Moore, A. Kristin (Eds.), *Well-being: Positive development across the life course* (pp. 477-497). Mahwah, NJ: Lawrence Erlbaum Associates.
- Levy, P. (1968). Short-form tests: A methodological review. *Psychological Bulletin*, 69, 410-416.
- Linacre, J. M. (2002). Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.
- Linacre, J. M. (2004). Optimizing rating scale category effectiveness. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement* (pp. 258-278). Maple Grove, MN: JAM.
- Linacre, J. M. (2006). *Winsteps: Rasch model statistical software*. Chicago, IL: MESA.
- Lissitz, R. W., & Green, S. B. (1975). Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology*, 60, 10-13.
- Lu, L. (2006). Cultural fit: Individual and societal discrepancies in values, beliefs and SWB. *Journal of Social Psychology*, 146, 203-221.
- Lu, L., & Lin, Y. Y. (1998). Family roles and happiness in adulthood. *Personality and Individual Differences*, 25, 195-207.
- Newell, C. E., Rosenfeld, P., Harris, R. N., & Hindelang, R. L. (2004). Reasons for nonresponse on U.S. navy surveys: A closer look. *Military Psychology*, 16(4), 265-

276.

- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78.
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52, 141-166.
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57, 1069-1081.
- Ryff, C. D. (1995). Psychological well-being in adult life. *Current Direction in Psychological Science*, 4, 99-104.
- Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69, 719-727.
- Ryff, C. D., Lee, Y. H., Essex, M. J., & Schmutte, P. S. (1994). My children and me: Mid-life evaluations of grown children and of self. *Psychology and Aging*, 9, 195-205.
- Ryff, C. D., & Singer, B. H. (1998). Human health: New directions for the next millennium. *Psychological Inquiry*, 9, 69-85.
- Sheehan, K. B. (2002). Online research methodology: Reflections and speculations. *Journal of Interactive Advertising*, 3(1), 24-38.
- Sin, N. L., & Lyubomirsky, S. (2009). Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. *Journal of Clinical Psychology*, 65(6), 467-487.
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12, 102-111.
- Snyder, C. R., & Lopez, S. J. (Eds.) (2002). *Handbook of positive psychology*. New York, NY: Oxford University.
- Snyder, C. R., & Lopez, S. J. (2007). *Positive psychology: The scientific and practical explorations of human strengths*. Thousand Oaks, CA: Sage.
- Stone, M. H., & Wright, B. D. (1994). Maximizing rating scale information. *Rasch*

Measurement Transactions, 8, 386.

Vandenberg, R. J., & Lance, C. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practice, and recommendations for organizational research. *Organizational Research Method*, 3, 4-70.

Wang, W. C. (2008). A critique of Raju and Oshima's prophecy formulas for assessing the reliability of item response theory-based ability estimates. *Applied Psychological Measurement*, 32, 261-266.

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, 64(4), 678-691.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA.

國小學童正向情緒與創造力關聯之 研究：以創造力傾向及創意自我效能 為中介變項

蕭佳純 國立臺南大學教育學系教授

摘 要

本研究以 368 名國小五年級的學童為研究樣本進行分析，以了解正向情緒、創造力傾向、以及創意自我效能等因素對創造力的直接影響，以及正向情緒透過創造力傾向的中介、創意自我效能的中介對創造力所造成的間接影響，以及正向情緒透過創意自我效能再透過創造力傾向的中介對創造力所造成的間接影響，並以結構方程模式加以檢驗。經由結構方程模式統計檢定後，整體模式所獲得的指數顯示模式尚可被接受。對整體效果的分析顯示，創造力傾向對創造力具直接影響；正向情緒以及創意自我效能雖然對創造力不具有直接影響，但是正向情緒透過創造力傾向對創造力產生間接影響；以及正向情緒可以透過創意自我效能再透過創造力傾向對創造力產生間接影響。最後，針對分析結果，本研究將提出相關的討論與建議。

關鍵詞：正向情緒、創造力傾向、創造力、創意自我效能



The Correlation between Elementary Pupils' Positive Emotion and Creativity with Creativity Tendency and Creative Self-efficacy as The Mediators

Chia-Chun Hsiao

Professor, Department of Education, National University of Tainan

Abstract

Total 368 G5 pupils were sampled for analyzing the direct effects of positive emotion, creativity tendency, and creative self-efficacy on creativity, the indirect effects of positive emotion on creativity with the mediating effects of creativity tendency and creative self-efficacy, as well as the indirect effects of positive emotion on creativity through the mediating effects of creative self-efficacy and then creativity tendency in this study. Structural Equation Model was further utilized for the verification. After the statistical test through Structural Equation Model, the acquired overall model index shows the model being acceptable. The analysis of the overall effect reveals direct effects of creativity tendency on creativity, no direct effect of positive emotion and creative self-efficacy on creativity, but indirect effects of positive emotion on creativity through creativity tendency, as well as indirect effects of positive emotion on creativity through creative self-efficacy and then creativity tendency. Finally, relevant discussions and suggestions aiming at the analyses are proposed in this study.

Keywords: positive emotion, creativity tendency, creativity, creative self-efficacy



壹、前言

因應創意經濟時代需要，培育學生創造力，已成為美國、新加坡、香港等世界各國厚植人力及提升國家競爭力的教育發展方向（Chan & Yuen, 2014; Chang, Chuang & Bennington, 2011），雖然創造力過去被認為是奇怪的、古怪的，並無法控制的，但近幾十年來已有所轉變，並將其視為是未來教育課程的核心，可以有效面對變遷的重要能力（Puccio, Murdock & Mance, 2007）。Gibson（2010）強調，「創造力」已被認為是解決 21 世紀龐大社會、政治、經濟問題的一種手段，這都顯示「創造力」已是各先進國家所重視的目標，因此 21 世紀也已進入了所謂的「創造力時代」（Florida, 2002）。哪些因素會影響創造力因素呢？在過去的研究中可知，學生個人以及家庭環境層面的影響仍為最多研究者所討論的（蕭佳純，2011）。另一方面，葉玉珠（2000）從生態系統模式的觀點也提出：個人特質對個體的創意有最直接的影響，所以，本研究以個人層面的因素作為探討影響國小學童創造力的因素，所關切的個人層面因素包括正向情緒、創意自我效能及創造力傾向。

近代正向心理學的發展，所強調的重點包括強調樂觀、正向情緒以及正向意義（常雅真、毛國楠，2006），過去研究指出，正向情緒會促進創意問題解決的可能性（Gasper, 2004; Kaufmann & Vosburg, 2002）。且依 Fredrickson（1998）的擴展 - 建立理論（broaden-built theory）觀點，正向情緒可以消除原有的負向情緒，進一步促進創意問題解決。在創造力發展的過程中，必然會遭遇許多壓力與挫折，具正向情緒的學童可能會較抱持著正向的期待，面對壓力時較能自我調適，所以可能較有高創造力，所以本研究目的一為，了解具有正向情緒的學生，其創造力是否也會較高。而除了正向情緒之外，綜合多位學者的看法（胡夢蕾，2006；葉玉珠、吳靜吉、鄭英耀，2000；Amabile, 1996; Oldham & Cummings, 1996; Runco & Sakamoto, 1999; Runco & Walberg, 1998）表示，高創造力者大致上具有某些創造力的個人特質，例如高度的工作熱誠、精力旺盛、自信、好奇、大膽、樂觀、理性主義等等，這些特質在相關研究中也曾以創造力傾向來討論（胡夢蕾，2006），由此可知，在學生個人特質上，除了正向情緒，創造力傾向可能也是一個值得切入討論的重點。目前創造力的研究大多偏向認知的部分，但是情意部分的創造力傾向研究則相對較少（李偉清，2012），尤其國小學童創造力傾

向的實證性研究更是不多見（洪文東，2002；蔡笑岳、朱雨潔，2007；Hwang, Chen, Dung, & Tang, 2007; Magoun, Eaton, & Owens, 2002）。所以，討論具有創造力傾向的國小學童是否也會有高創造力，此為本研究之目的二。

除了正向情緒、創造力傾向的影響之外，Bandura（1997）強調，個體行為產生係在動態的歷程下形成，因此彙整這些可能影響個體創意行為產生的要素並了解其間的關係，實屬必要。社會認知理論強調個體的認知歷程對其行為的影響，尤其是個體自我效能更是預測個人行為的重要指標（洪素蘋、黃宏宇、林珊如，2008）。因此，本研究欲了解，創意自我效能較高的學生，則他的創造力表現是否也會較高，此為本研究之目的三。除了正向情緒、創造力傾向及創意自我效能的直接影響之外，本研究還想討論創意自我效能及創造力傾向的中介效果。在創造力的過程中，學生可能會經歷許多挫折、困境，擁有正向情緒的學生，較可能擁有高創造力傾向的特質，如具有冒險精神、毅力較強等，則較易產生高創造力，或者是較易具有高度的創意自我效能，對於創造力較有高度信念、信心，而產生較高的創造力，表示正向情緒可能透過創造力傾向的中介而對創造力產生間接效果，也可能透過創意自我效能的中介而對創造力產生間接效果，此為本研究之目的四、五。當個人認為自己具有製造創意成品的能力之信念較高時，也就是說，創意自我效能較高時，則他的創造力傾向也會較高，再加上本研究的目的五，本研究合理推論正向情緒可能透過創意自我效能，再透過創造力傾向的中介而對創造力產生間接影響，此為本研究之目的六。相較過去研究可發現，有鑑於國內目前有關於正向情緒的討論尚在起步，而進一步討論正向情緒、創造力自我效能、創造力傾向與創造力間關聯的研究更是付之闕如，本研究的進行正可彌補此一缺口。除此之外，在討論正向情緒與創造力的國外相關研究多以成人為主（Gasper, 2004; Kaufmann & Vosburg, 2002），而國內的相關研究更少，因此本研究以國小生為對象，考量學童填寫問卷時的耐性與成熟度，選擇五年級學童為對象，據上所述可凸顯本研究的價值。

貳、文獻探討與假設發展

一、正向情緒與創造力的關聯

在創造力心理學發展的早期，「創造力」(creativity)牽涉到四個P，即「個人」(person)、「歷程」(process)、「產品」(product)與環境(place / press)等四要素(葉玉珠，2006)。葉玉珠、吳靜吉、鄭英耀(2000)認為創造力乃個體在特定領域中，產生一個所處的社會文化脈絡中具有「原創性」與「價值性」產品的歷程；也就是說創造性產品是「意向」(含傾向、動機、態度、承諾)、「技巧/策略」、「個人的知識」(含經驗)與「環境」互動的結果，簡言之，創造力即是一種人格特質、能力與歷程的整體表現，因此本研究以個人層面切入，討論各因素對創造力的影響。Seligman 與 Csikszentmihalyi 於 2000 年提出正向心理學(positive psychology)概念，他們認為，個人心理之發展及潛能之發揮，必須要能擴展個人正向情緒經驗(Seligman & Csikszentmihalyi, 2000)。Seligman (2002)主張，正向情緒包括對過去、現在與對未來的三個面向，也就是對過去感到滿意，對現在感到快樂，對未來感到樂觀。Fredrickson (2006)也認為，正向情緒需要認知的評價與意義的建立。綜合上述，正向情緒的相關意義雖有不同，但大抵上，正向情緒可以分為感官上的快感與意義上的滿足兩大類，追求理想、實踐自我、充滿意義與滿足、需要認知評價的快樂被視為是真實的快樂，有助於個體發展。

Fredrickson (2002)認為正向情緒的內涵包含：歡愉、知足以及自信；Seligman (2002)所提出的現在層面的正向情緒包含愉悅與滿足感的心流，前者偏重感官刺激的接收，產生的情緒較為短暫；後者強調當下投入的享受，產生的正向情緒較為持久，因其動用到長處的發揮，進而增進個體的自我效能與成就感，故較為學者所提倡。而常雅珍、毛國楠(2006)除了上述的構面之外，更加入了同理心的內涵。本研究統整 Seligman 的時間軸分類和常雅珍、毛國楠(2006)的看法，把正向情緒的內涵歸納為愉悅、知足、自信以及同理心四個構面。

至於在正向情緒與創造力的相關研究方面，正向情緒有助於創造力是如何產生呢？一般皆採 Fredrickson (1998)的「擴展—建立理論」之觀點，此理論認為正向情緒如滿足、自信等，會使個體擴展思考、行動技能，另外，它會擴展我們的注意力、思考速度、運用社會資源之能力，最重要的能消除負向情緒所帶來之

負作用，使個體較能有彈性或創造性的解決問題。不少探討情緒與創造力關係之研究（Ashby, Isen, & Turken, 1999; Isen, 2000; Schwarz, 2000）可發現，正向情緒有助於創造力表現。具正向情緒者產生的新奇點子較具有負向情緒者更多（Shapiro, Weisberg, & Alloy, 2000），例如 Grawith, Munz 及 Kramer（2003）就發現正向情緒在創造力表現上高於負向與中性情緒組，但是負向情緒卻不會降低創造力的表現。由以上的討論可知，正向情緒能促進創造力的表現，是因為情緒在幻想或情緒狀態中，主體可以引發出較寬廣的聯想網路，而這個被擴大的聯想網路就可以促使個體在問題解決時，增進擴散性思考和變通能力的表現。綜合以上可知，個體在正向情緒狀態下，會更有意願探索新奇的事物，增加創意解決問題的可能性（Gasper, 2004; Kaufmann & Vosburg, 2002），也就是說，情緒越正向，則創造力應會越高。

Fredrickson（2003）特別舉出 Isen（2001）二十多年的情緒研究發現，當人們有正向情緒時人們的想法較有創意、整合和對資訊更開放。Baumann 和 Kuhl（2005）透過回憶引發正負向情緒的實驗研究，發現啟動正向情緒之後，個體可以克服先前的優勢化反應抑制，迅速對局部訊息作出正確反應。Forgas, Vargas 和 Laham（2005）的研究發現，正向情緒有利於建設性問題解決的表現。Fredrickson 和 Branigan（2005）研究發現，激發受試者的正向情緒之後，有助於認知廣度的擴展。Isen（2001）研究發現，接受正向情緒操弄的受試者，其問題解決的表現優於未接受正向情緒操弄的受試者。根據邱發忠、陳學志、徐芝君、吳相儀與卓淑玲（2008）的整理，認為正向情緒以兩種機制來影響創造力的表現，其一是認知彈性的提升，也就是促使個體認知到兩個想法可以聯結在一起的能力；其二，正向情緒可以提升認知處理的效能，降低無關訊息的處理。過去研究多是從情緒層面著手，以狀態（state）論來強調正向情緒對於創造力的影響，較少研究論及正向情緒的特質。本研究從正向情緒的特質（trait）論著手，討論正向情緒對於創造力的影響，因為對國小學童而言，若學童自己多數時間能維持正向情緒，則他們面臨創造力激發過程所可能伴隨的壓力、挫折、挑戰等等，可能較易面對或克服，換言之，較正向情緒的國小學童較可能有高創造力。綜合以上理論與相關研究，本研究發展假設一為：學生的正向情緒對創造力具有正向影響。

二、創造力傾向與創造力之關聯

提出創造力「投資理論」的 Sternberg 與 Lubart (1999) 強調，創造力需要的六大資源之一就是「人格特質」，林幸台 (2002) 進一步在他的研究中表示，性格與意志是能否成為創意人的重要關鍵。而在人格特質中，最廣被研究者認同對創造力影響最大的，莫過於創造力傾向 (蕭佳純，2015)。高創造力者具有哪些傾向？葉玉珠 (2000) 的研究發現，影響創造力表現的傾向共有 16 項，例如：具有冒險精神、願意成長、喜歡與人互動、喜歡嘗試、不斷求進步、興趣廣泛等等。洪文東 (2002) 的研究後發現，創造型兒童可歸納出 20 項情緒特徵，包含有：喜歡求證事物、有想像力、觀察力敏銳、毅力堅強、超人的記憶、善用各種符號、喜歡排列組合等等。而類似的結果也在以成人為對象的研究如葉玉珠、吳靜吉與鄭英耀 (2000)、洪久賢、溫秀玲、蔡長艷與宋慧娟 (2003)、林碧芳與邱皓政 (2008) 的研究中可發現。據此，本研究推導假設二：創造力傾向對學生的創造力有正向影響。至於為何國內有關創造力傾向與創造力直接關聯的研究並不多，乃是因為國內研究者在評量創造力的程度時，大都採用 Williams (1980) 的創造力傾向量表的現象，換言之，將創造力傾向的高低視為創造力高低的謬誤現象。本研究將這兩部分視為不同變項，有別於過去研究的測量方式分開衡量，這也是本研究另一個價值所在。

除了正向情緒、創造力傾向對於創造力具有直接影響外，本研究欲討論的另一個重點為：學童的正向情緒是否可透過創造力傾向的影響而進一步對創造力產生間接效果。Williams (1980) 發展的創造力思考與創造力傾向測驗指出，創造力傾向屬於「情意」領域，而創造力思考屬於「認知」領域，所以本研究將創造力傾向視為情意特徵。據此，研究者將創造力傾向詮釋為，對創造性活動較為積極的情意態度。一般而言，當學童認知到有興趣或重要性時，可能會慢慢地培養出情意，也就是說，學童的高正向情緒在若干時日後，應可漸漸表現出高創意人的情意特徵，進而提高創造力。因此，學童若具有高正向情緒，是否能夠提升屬於情意領域的創造力傾向，爾後進一步提升創造力，目前並無相關研究，有待本研究加以檢證。據此，本研究發展假設三：學生的正向情緒會透過創造力傾向的中介效果而間接對學童的創造力產生影響。

三、創意自我效能與創造力之關聯

根據 Bandura (1977) 的社會認知理論，自我效能可能是調控動機與行為的一個重要機制。但是，Bandura (1977) 也強調，自我效能具有情境與領域的特定性，在不同的學科學習將產生不同的自我效能信念。所以，Tierney 跟 Farmer (2002) 將自我效能理論與 Amabile (1988) 的創造力理論加以結合，提出「創意自我效能」(creative self-efficacy) 的概念，指的是自我效能在特殊領域的應用。正如 Chen, Gully 與 Eden (2001) 所指，傳統對一般性自我效能的測量無法類推至不同領域內自我效能的測量。所以，本研究將以創意自我效能為變項，意指「自己認為有製造創意成品能力之信念」。

自我效能係屬特定領域與情境的概念 (Bandura, 1977; Linnenbrink & Pintrich, 2002)，所以在推估特定學習活動時，往往具有高度的預測能力。諸多有關於創造力的研究發現，創意自我效能會對創意表現或創造性產品產生正向的預測效果 (汪美香、黃炳憲, 2012; 林碧芳、邱皓政, 2008; 陳玉樹、郭銘茜, 2013; Gong, Huang, & Farh, 2009; Jaussi, Randel, & Dionne, 2007; Tierney & Farmer, 2002)。由此可知，創意自我效能對創造力的表現，扮演著關鍵性的角色 (陳玉樹、郭銘茜, 2013)。由於學生在從事創造力的過程中可能遭遇挫折與阻礙，需要長時間的投入，如果沒有強烈的自信心，很難維持創造力的展現，可見創意自我效能對於創造力行為的影響。所以，本研究合理推論，當學生創意自我效能愈高時，對於自己的創造力行為會愈具信心，也會更願意接受挑戰，以藉此展現出更具有創造力的行為。據此，本研究提出發展假設四為：創意自我效能對於創造力具有正向影響。此外，Morris 於 1989 年所提出的「認知調整模式」(cognitive tuning model) 表示，情緒是個體評估外在環境安全與否的指標，同時個體也會依當時的情緒來調整認知系統的運作 (Fredrickson, 2004)。若個體處於正向的情緒中，感受到情緒是安全的，會比較願意冒險，比較富有創意 (Schwarz, 2000)。綜合以上可知，本研究認為，創意自我效能除了創造力具有直接影響之外，正向情緒也可能透過創意自我效能的中介而對創造力產生間接影響，因為具有正向情緒的學生他可能對創造力產生較高的信心，亦即創意自我效能也會較高，所以創造力也才會隨之較高。所以本研究想討論學生是否會因為當時的正向情緒，而使得擴展了思考、注意力，對創意自我效能更高，而更有點子、創意解決問題，進而提升

了學生的創造力。據此，本研究發展假設五為：學生的正向情緒可以透過創意自我效能的中介而對創造力產生間接影響。除了假設三及假設五的發展之外，從國小學童的場域現象來看，當學生的創意自我效能較高時，他對於創意成品的信心會較高，較能使用創意思考策略，則他可能也會有較高的創造力傾向，因為學童若是對於創造力有較高的自信心，則他可能會進一步有較高的挑戰性、開放性等創造力傾向，所以，綜合假設三以及假設五的推論，推導出假設六為，學生的正向情緒可以透過創意自我效能，再透過創造力傾向的中介，而對創造力產生間接效果。

參、研究設計

一、研究架構

本研究架構如圖 1 所示

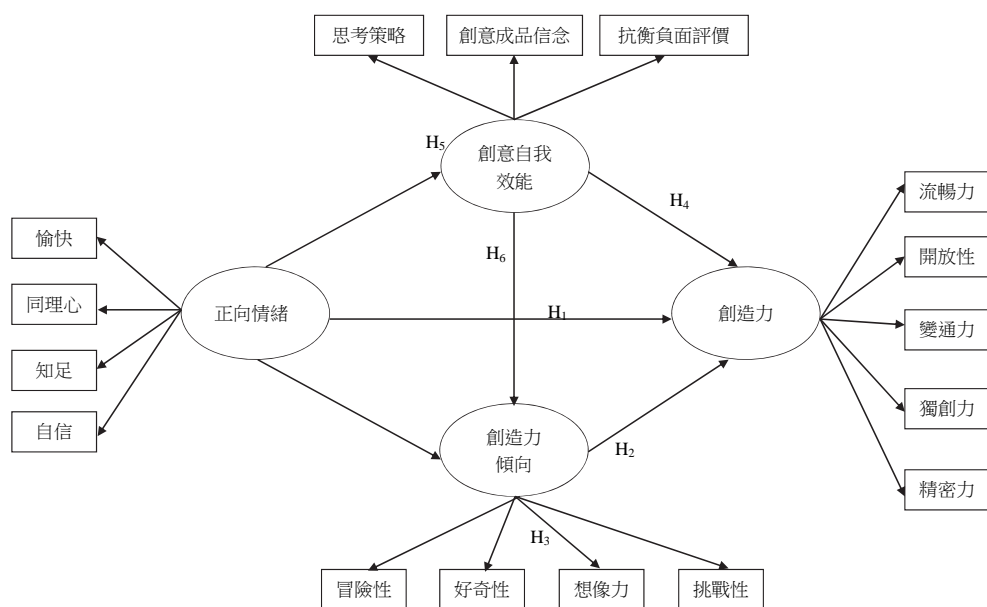


圖 1 研究架構圖

假設一為：學生的正向情緒對創造力具有正向影響。

假設二為：創造力傾向對學生的創造力有正向影響。

假設三為：學生的正向情緒會透過創造力傾向的中介效果而間接對學童的創造力產生影響。

假設四為：創意自我效能對於創造力具有正向影響。

假設五為：學生的正向情緒可以透過創意自我效能的中介而對創造力產生間接影響。

假設六為：學生的正向情緒可以透過創意自我效能，再透過創造力傾向的中介，而對創造力產生間接效果。

二、研究對象

本研究對象分為預試及正式施測兩部分，預試時採便利抽樣，抽取臺南地區的國小五年級學生共 200 位。在正式樣本的母群體為全臺灣的國民小學，取樣方式採用分層比例抽樣，依據教育部（2010）所彙整之資料，係將母群體分為四個區域（北、中、南、東），再依各區域比例抽取學校。最後共選取 50 所國民小學，以五年級學生為對象，每所學校選取一個班級，並從班級中抽取 10 位學生進行施測。而每個班級抽取 10 位學生的原因乃是因為本研究所使用的創造作業乃是委由導師於課後時請學生留下來填答，所以考量學生的配合度以及本研究評審所需批改的份數，決定一個班級抽取 10 位學生。若小規模學校的班級人數不到 10 人，則全班都協助填答。因此問卷總共發出 500 份學生問卷，經催收後，一共有 368 位學生參與施測，而男生有 186 位，女生有 182 位，問卷回收率為 72%。必須注意的是，這 10 位學生是教師考量配合度後所隨機抽取的學生，也就是徵詢學生願意於課後留下來接受評量後，再從中選取 10 位學生，但是否會因配合度較高而對研究結果產生偏誤，這是需要留意的。

三、研究工具

（一）創造力測驗

本研究所使用的創造力測驗乃使用林幸台、王木榮（1999）修訂自 Williams（1980）的創造力測驗，此創造性思考活動要受試者利用題本上印妥的線條畫圖，並為畫好的圖取個名字。本測驗一共有 12 幅未完成的圖，請學生利用格內已有的

線條完成有意義的圖畫（必須使用上這些線條），受試時間為 20 分鐘。本測驗一共評得五種分數，流暢力、開放性、變通力、獨創力以及精密性。其中，流暢力指學生每畫一個最基本的圖形，即可有 1 分，最高為 12 分。開放性按照指導手冊的給分標準有四種情形，分別給予 0~3 分，最高可得 36 分。變通力係將圖畫分類後，再計算一共出現幾類圖畫，即得幾分，最高可得 12 分。獨創力是以圖形所引發的反應為主，依下列比例算出幾分標準，5% 以上的反應給 0 分；2~4.99% 以上的反應給 1 分；2% 以下的反應給 2 分；其他具有想像與創造力而在指導手冊中查不到的反應給 3 分。精密力係以圖畫是否對稱為計分標準，計分標準介於 0~3 分。評分方式乃是依照指導手冊，但為了避免單一評分者的偏誤，所以在內部一致性信度方面，本研究邀請 5 位曾擔任 powertech 的評審教師，針對 368 位學生的創意作品進行評量，評量時間為期三個月。至於評分方式部分，Amabile（1996）建議應以「相對標準」而非「絕對標準」對作品進行評分，亦即 5 位評分者要從這 368 個學生的作品判斷相對上的創意高低，而給予評分，相信此一評分將對評分者造成極大的認知負荷，可能造成效度上的偏誤，因此除了給予三個月的評分時期外，本研究輔以考量評分者的組內評分者一致性，本研究採納劉昆夏、鄭英耀、王文中（2010）的建議，依斯布公式進行校正，求得校正後的評分者一致性分別為 .71。

（二）正向情緒量表

本研究所使用的正向情緒量表為自編量表，乃是研究者參考 Fredrickson（2002）以及常雅珍、毛國楠（2006）的量表為依據來進行編制，為 Likert 六點量表，預計一共分為愉悅、自信、知足及同理心四個構面。作者初編量表的題目如：「我開心的過每一天」、「我的心情通常是愉快的」、「我看到別人受委屈時，我會去安慰他」、「我擁有的已經夠多了，應該感到滿足」、「我對自己感到自豪」，一共編制有 46 題。經 200 位學生的預試樣本及探索性因素分析後，最後共抽取 4 個因素 32 題，累積解釋變異量為 66.49%。四個因素分別命名為愉悅（共 12 題）、同理心（共 9 題）、知足（共 4 題）、以及自信（共 7 題）。信度 Cronbach's α 值分別為 .86、.83、.88、.86，總量表的 Cronbach's α 值為 .89。而正式樣本經二階驗證性因素分析後，卡方值為 1640.31， $p < .05$ 達顯著水準，RMSEA、GFI、AGFI、CFI、IFI 及 SRMR 分別為 .07、.94、.91、.96、.96、.04，組成信度分別為 .81、.80、.82、.83，結果顯示本量表的整體適配度尚佳。

（三）創意自我效能量表

本研究所使用的創意自我效能量表為自編量表，乃是研究者參考林碧芳與邱皓政（2008）所編制的「創意自我效能量表」，並根據國小學童的特性加以修訂，來測量國小學生的創意自我效能程度，為 Likert 六點量表。此量表一共有 12 題，分別有「當我面對新問題時，我相信我能很快聯想到很多個解決方案」、「當我在作報告時，我相信我能做出令人耳目一新的作品」、「當老師不接受我的創意成品時，我想我仍會堅持自己的理想」等等。經 200 位學生的預試樣本以及探索性因素分析顯示的 KMO 值是 0.81，共得到三個因素，分別命名為「創意思考策略」，共包含四題；「創意成品信念」，共包含三題；「抗衡負面評價」，共包含三題，共可解釋 71.29% 的變異量，信度 Cronbach's α 值分別為 .89、.86、.87，總量表的 Cronbach's α 值為 .88。正式樣本經二階驗證性因素分析後，卡方值為 411.38， $p < .05$ 達顯著水準，RMSEA、GFI、AGFI、CFI、IFI 及 SRMR 分別為 .04、.91、.91、.93、.93、.05，組成信度分別為 .82、.84 以及 .87，顯示本量表的整體適配度尚佳。

（四）創造力傾向量表

本研究所使用的創造力傾向量表乃使用林幸台、王木榮（1999）修訂自 Williams（1980）的創造力傾向測驗，此測驗共包含四個分量表：好奇心（11 題）、想像力（8 題）、冒險性（6 題）、以及挑戰性（8 題），所以此量表共 33 題。本研究中使用 Likert 六點量表，分數越高代表學生之創造力傾向越高。正式樣本經二階驗證性因素分析後，卡方值為 422.38， $p < .05$ 達顯著水準，RMSEA、GFI、AGFI、CFI、IFI 及 SRMR 分別為 .04、.91、.91、.92、.94、.05，組成信度分別為 .80、.88、.86 以及 .87，顯示本量表的整體適配度尚佳。

肆、研究分析與結果討論

本研究在變項的處理上，是將各題項加總後，以組合分數做為觀察指標。由表 1 看出，大部分的相關係數皆達 0.05 的顯著水準，顯示 16 項指標間有相關性，且各觀察指標之間相關的絕對值皆無太過接近 1，符合基本適配指標的理想標準。本研究在進行適配度考驗之前，先了解多變項常態分配假設是否遭到違背，分析發現偏態絕對值介於 0.45~2.06 之間，絕對值均小於 3；峰度介於 0.074~3.08 之間，

絕對值亦小於 10。所以各觀察變項的偏態係數與峰度係數屬於可接受範圍，使用最大概似法進行模式的估計（陳正昌、程炳林、陳新豐、劉子鍵，2009）。以下，本研究依陳正昌等人（2009）的建議，進行基本適配度、整體模式適配度、內在結構適配度的分析。

表 1 觀察指標間的相關係數矩陣

變項	構面	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
正向情緒	1. 愉快	1															
	2. 同理心	.71*	1														
	3. 知足	-.15*	-.02	1													
	4. 自信	.52*	.50*	.09	1												
創意自我效能	5. 思考策略	.61*	.61*	-.05	.53*	1											
	6. 創意成品信念	.51*	.53*	.03	.44*	.72*	1										
	7. 抗衡負面評價	.52*	.50*	.02	.41*	.62*	.62*	1									
創造力傾向	8. 冒險性	.49*	.48*	-.06*	.49*	.48*	-.06*	.49*	1								
	9. 好奇心	.51*	.54*	.04	.50*	.54*	.04	.50*	.73*	1							
	10. 想像力	.42*	.44*	.14*	.47*	.44*	.14*	.47*	.65*		1						
	11. 挑戰性	.43*	.43*	-.31*	.37*	.43*	-.31*	.37*	.62*	.59*	.49*	1					
創造力	12. 流暢力	.14*	.10	-.08	.11*	.17*	.18*	.07	.27*	.09	.27*	.25*	1				
	13. 開放性	.18*	.17*	.17*	.21*	.02	.04	.03	.03*	.06	.16	.05	.58*	1			
	14. 變通力	.19*	.18*	-.02	.12*	.17*	.19*	.07	.07	.12	.06*	.07	.33*	.20*	1		
	15. 獨創力	.12*	.11*	.03	.07*	.09*	.07	.09*	.04	.11*	.04*	.08*	.29*	.31*	.09	1	
	16. 精密力	.12*	.12*	.00	.11*	.07*	.03	.03	.05	.09*	.11*	.06	.38*	.76*	.08	.30*	1

** $p < .001$ * $p < .05$

一、結構模式適配度評鑑

（一）基本適配度分析

從表 2 得知 X 變項測量誤差變異數（ $\delta_1 \sim \delta_4$ ）、Y 變項測量誤差變異數（ $\varepsilon_1 \sim \varepsilon_{12}$ ）和潛在依變項殘餘誤差（ $\zeta_1 \sim \zeta_3$ ）為正值，且達顯著水準。誤差變異 t 值介於

4.541~12.549 皆大於 1.96 且達顯著水準，潛在變項與觀察變項之間的標準化因素負荷量量 ($\lambda^x_{11} \sim \lambda^y_{53}$) 介於 .30~.91，僅少數低於 .5 或高於 .9 的情形；估計參數標準誤介於 .038~.668，沒有過大的標準誤。上述結果顯示本研究所提出模式符合基本適配標準，可進一步檢驗整體模式適配度及內在結構適配度。

表 2 結構模式估計參數顯著性考驗及係數摘要表（非標準化解）

參數	參數估計值	標準誤	t 值	參數	參數估計值	標準誤	t 值
λ^x_{11}	0.760	.045	16.899*	β_3	0.228	.075	3.028*
λ^x_{21}	0.755	.044	17.177*	δ_1	0.272	.032	8.497*
λ^x_{31}	0.731	.082	8.914*	δ_2	0.250	.031	8.187*
λ^x_{41}	0.659	.056	11.846*	δ_3	0.870	.149	5.839*
λ^y_{11}	0.470	-	-	δ_4	0.648	.057	11.328*
λ^y_{21}	0.681	.038	17.877*	ε_1	0.118	.012	9.785*
λ^y_{31}	0.601	.038	15.964*	ε_2	0.097	.017	5.888*
λ^y_{12}	0.357	-	-	ε_3	0.187	.019	9.696*
λ^y_{22}	0.801	.086	9.274*	ε_4	0.356	.030	11.974*
λ^y_{32}	0.761	.085	8.953*	ε_5	0.225	.029	7.772*
λ^y_{42}	0.759	.088	8.667*	ε_6	0.357	.036	9.841*
λ^y_{13}	1.032	-	-	ε_7	0.494	.046	10.629*
λ^y_{23}	6.819	.668	10.206*	ε_8	2.135	.180	11.847*
λ^y_{33}	0.319	.094	3.400*	ε_9	2.915	.642	4.541*
λ^y_{43}	1.347	.258	5.212*	ε_{10}	2.632	.210	12.554*
λ^y_{53}	4.737	.438	10.808*	ε_{11}	17.827	1.422	12.538*
γ_1	0.707	.063	11.153*	ε_{12}	16.742	1.843	9.085*
γ_2	-0.132	.154	- .853	ζ_1	0.500	.070	7.163*
γ_3	0.679	.101	6.745*	ζ_2	0.269	.068	3.953*
β_1	0.092	.097	0.941	ζ_3	0.691	.186	3.710*
β_2	0.465	.151	3.081*				

註：- 表該參數作為對應觀察變項之參照指標，無需估計。

* $p < .05$

（二）整體模式適配度分析

根據學者們（余民寧，2006；邱皓政，2003）的觀點，整體適配度考驗包含絕對適配度、增值適配度和精簡適配度三方面評鑑。整體模式適配度考驗顯示本研究的 $\chi^2_{(98)}=298.744$ ， $p<0.05$ ，但因 χ^2 值常會隨著樣本人數波動，一旦樣本人數很大時，幾乎所有的模式都可能被拒絕。因此本研究主要參酌其他適配度指數來評鑑模式與觀察資料的適配程度。

在絕對適配度考驗指標方面，RMSEA = .08、SRMR = .068，略大於 .05 的良好適配標準，且 GFI = .90、AGFI = .85，其中 AGFI 未大於 .90 的標準，但亦相當接近。嚴格來說，這並不能算是良好的適配，不過，依據 Doll 與 Lyon（1998）的建議，GFI 與 AGFI 介於 0.8 至 0.89 之間就代表各模式已有合理適配，因此，本研究模式與所蒐集到的樣本資料間的適配程度應是還可接受的。增值適配度考驗指標方面，NFI = .93、CFI = .95、IFI = .95、RFI = .91，此四個指標皆大於 .90 的標準。而在精簡適配度考驗指標方面，PNFI = .76、PGFI = .64，此兩項指標皆大於 .50 的標準，且 AIC = 374.74 比獨立模式之 4322.874 小，此符合理論模式的 AIC 必須小於獨立模式的 AIC 之標準，顯示本模式能以精簡的變項數有效反映變項間的關係（余民寧，2006）。如上述顯示本研究模式與觀察資料的整體模式適配度達理想標準，且為一精簡模式。

（三）內在結構適配度分析

評估理論模式內在品質可以從個別項目信度、潛在變項組合信度、潛在變項平均變異抽取量、估計參數顯著性考驗等方面來進行（陳正昌等人，2009）。如表 3 在個別項目信度方面，顯示 16 個測量指標個別項目信度介於 .23~.83 之間，仍有 5 個指標未達 .5 以上標準，以創造力占多數。潛在變項組合信度方面介於 .35~.71，其中創造力傾向、正向情緒雖未達到 .6 以上標準，但亦相當接近，但是創造力的潛在變項組成信度偏低。潛在變項平均抽取變異量介於 .35~.72，多符合 .5 以上標準，但創造力未符合標準。由上可知，本研究所建構的模式其內在結構適配度雖然有 5 個個別項目信度及 2 個潛在變項組成信度未達標準稍嫌不理想，但因已相當接近適配值，且基本適配標準評鑑與整體模式適配度評鑑皆達到標準，說明本研究模式對於觀察資料還是有一定解釋力。需注意的是，創造力量表不論在組成信度、平均變異抽取量均未達標準且偏低，表示此研究所使用的創造力測驗用以評量國小學童的適切性有其商榷之處。究其原因發現，本研究所使用的創

造力測驗僅囊括圖型測驗，而國小學童多接觸語文材料，此點或許可做為研究顯示創造力測驗較不符合統計檢定標準的解釋之一。

表 3 模式內在品質適配度

變項	標準化參數估計值	個別信度	潛在變項組合信度	平均抽取變異量
創意自我效能				
思考策略	0.81	0.66	0.88	0.71
創意成品信念	0.91	0.83		
抗衡負面評價	0.81	0.66		
創造力傾向				
冒險性	0.51	0.23	0.82	0.53
好奇心	0.86	0.74		
想像力	0.79	0.62		
挑戰性	0.73	0.53		
正向情緒				
愉快	0.82	0.67	0.84	0.57
同理心	0.83	0.69		
知足	0.73	0.53		
自信	0.63	0.40		
創造力				
流暢力	0.58	0.33	0.46	0.35
開放性	0.71	0.51		
變通力	0.49	0.24		
獨創力	0.30	0.09		
精密力	0.76	0.57		

註：- 表該參數作為對應觀察變項之參照指標，無需估計。

二、模式各潛在變項間效果

圖 2 為模式中各變項間的直接效果，即模式中所估計的參數，而各項效果值則整理於表 4。

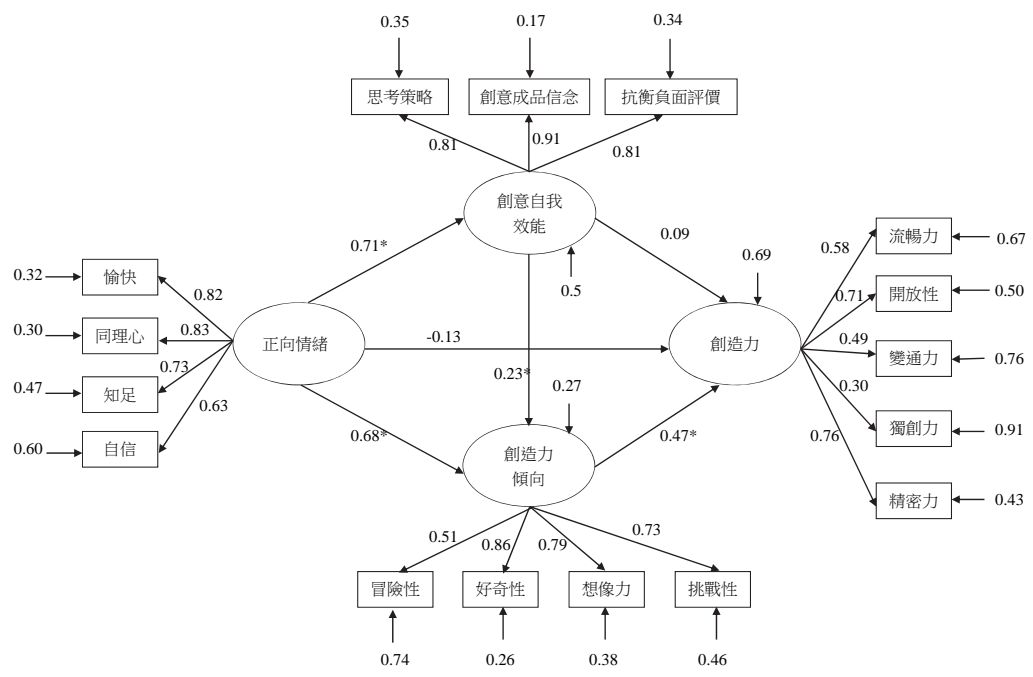


圖 2 研究結果

表 4 模式效果整理

潛在變項→潛在變項	直接效果	間接效果	全體效果
正向情緒 → 創造力	-	0.40*	0.40*
創意自我效能 → 創造力	-	0.11*	0.11*
創造性傾向 → 創造力	0.47*	-	0.47*
創意自我效能 → 創造力傾向	0.23*		
正向情緒 → 創意自我效能	0.71*		

其中正向情緒對創造力的間接效果乃是正向情緒透過創造力傾向對創造力的間接影響（ 0.68×0.47 ），再加上正向情緒透過創意自我效能、創造力傾向對創造力的間接影響（ $0.71 \times 0.23 \times 0.47$ ），兩者相加為 0.40，而創意自我效能對創造力的間接效果則是 $0.23 \times 0.47=0.11$ 所求得。本研究進一步經由 AMOS 的

Bootstrapping 方法，進行中介效果之考驗，發現正向情緒對創造力的間接效果值為 0.397 (95% CI: .152 ~ .416)，因為 95% 的信賴區間顯著異於 0，所以間接效果顯著。另外，創意自我效能對創造力的間接效果值為 0.112 (95% CI: .097 ~ .121)，因為 95% 的信賴區間顯著異於 0，所以間接效果也達到顯著。

(一) 創造力之結構模式

由圖 2. 及表 4 可知正向情緒對創造力的影響未達顯著水準 ($\lambda = -.013$, $p = 0.12$)，此與過去研究 (邱發忠等, 2008; Isen, 2000; Schwarz, 2000; Grawith, Munz & Kramer, 2003) 結果大不相同，表示當學生的正向情緒愈高，卻對創造力並不會直接產生影響，因此本研究假設一未獲得成立。若與過去研究對應來看，過去研究在衡量創造力表現時，多是將創造力的態度、認知等視為是創造力表現的衡量，更多是討論創意問題解決的能力或認知，較少是像本研究直接測量創造力的高低程度。但這是否是造成本研究此部分結果與過去研究不同之處不得而知，但卻是值得思考的一點，如果真是如此，表示正向情緒對於創造力的影響歷程應不是直接影響那麼的簡單，而本研究後續的分析或許正可以說明這點。另一個可能的解釋在於，本研究將正向情緒視為一種特質，而非如過去研究一樣視為一種狀態，這是否也可能是正向情緒未對創造力產生直接影響的因素之一，值得更多研究的討論。但是必須注意的是，本研究的正向情緒構面僅限於愉快、同理心、知足、自信四個構面，但是，「正向情緒」範圍很廣，還包括意義感、成就感、安全感、信賴感、希望感等等，本研究僅限於探討此四個構面，因此在推論時應只限於此四個構面的討論，但這也是本研究重要的研究限制。再者，創意自我效能對創造力的影響也未達顯著水準 ($\lambda = .009$, $p = 0.22$)，此與過去研究 (汪美香、黃炳憲, 2012; 林碧芳、邱皓政, 2008; 陳玉樹、郭銘茜, 2013; Gong, Huang, & Farh, 2009; Jaussi, Randel, & Dionne, 2007; Tierney & Farmer, 2002) 結果也大不相同，表示當學生的創意自我效能愈高，對創造力卻並不會直接產生影響，因此本研究假設四也未獲得成立。而其原因是是否如正向情緒的檢驗一樣，是因為本研究的創造力評量並非過去研究的創造力態度、認知或情意等，不得而知。除此之外，創造力傾向對創造力的影響達到顯著，且直接效果值為 .47，此研究結果和既有文獻 (洪文東, 2002; 洪榮昭、康鳳梅、林展立, 2003) 的結論相同，表示當國小學童的創造力傾向愈高，則創造力表現亦愈佳，因此本研究假設二獲得成立。而與過去的研究比較來看，大多數的創造力傾向與創造力的關聯研究是以國小學童

為對象，而本研究也是以國小學童為對象。換言之，國小學童創造力傾向高則創造力越佳的情形可說是得以支持，但在其他教育階段如高中、大學是否也是如此，則有待更多研究的投入。

此外，正向情緒透過創造力傾向對創造力的影響亦達顯著水準，其間接效果值為 $0.32 (0.68 \times 0.47)$ ；而創意自我效能透過創造力傾向的間接效果值為 $0.11 (0.23 \times 0.47)$ ；正向情緒透過創意自我效能再透過創造力傾向對創造力的間接影響則為 $0.08 (0.71 \times 0.23 \times 0.47)$ 。這也表示，正向情緒雖然對創造力的直接效果並不顯著，卻是可以透過創意自我效能與創造力傾向的中介而對創造力產生間接影響；而創意自我效能雖然對創造力的直接效果並不顯著，卻是可以透過創造力傾向的中介而對創造力產生間接影響，這表示創造力傾向為一個相當重要的中介變項。但過去並未對此中介影響有所著墨，而是將正向情緒、創造力傾向、創意自我效能對創造力的影響分成各自的關係、觀點作討論，甚少綜合上述變項一併討論其與創造力之間的關聯。於此經本研究檢驗發現，國小學童的正向情緒必須經由創造力傾向的中介才可對創造力產生影響，因此本研究的假設五獲得支持。除此之外，正向情緒也可以經由創意自我效能，再透過創造力傾向的中介才可對創造力產生影響，因此本研究的假設六獲得支持。而中介影響雖然未大於直接效果，但由創造力傾向的中介影響來看，實在不容小覷創造力傾向的中介效果，但因為創意自我效能對創造力的影響未達顯著，因此本研究的假設四未獲得支持。這表示，過去文獻若單獨來看創意自我效能與創造力的關係，多數發現創意自我效能是對創造力有正向影響的，但是在本研究中的結構圖中，同時納入了正向情緒以及創造力傾向，此時創意自我效能的效果卻變得不顯著，且中介的效果也不顯著。這是否意味著創意自我效能不重要？恐怕不是如此，應該說創意自我效能的效果在與創造力傾向一起比較時，顯得較無影響力了，所以直接效果以及間接效果都不顯著。以此部分的研究結果來看，國內目前有關於正向情緒的討論尚在起步，而進一步討論正向情緒、創造力自我效能、創造力傾向與創造力間關聯的研究更是付之闕如，本研究的進行正可以彌補此一研究缺口，尤其本研究與國內外的實證研究結果有些許不同，可提供未來研究作為再進一步深入探討的基礎與起點。

而在潛在變項殘差變異量部分，創意自我效能殘差變異量（ ζ_1 ）為 0.5 ，亦即創意自我效能可被正向情緒解釋的總變異量為 50% 。創造力傾向殘差變異量（ ζ_2 ）

為 0.27，即創造力傾向可被正向情緒解釋的總變異量為 73%。創造力殘差變異量 (ζ_3) 為 0.69，即創造力可被正向情緒、創造力傾向、創意自我效能解釋的總變異量為 31%。對創造力的直接效果值而言，僅有創造力傾向具有直接影響效果（效果值為 0.47）；而間接效果中也以正向情緒透過創造力傾向對創造力的影響較大（效果值為 0.32）。統整以上的分析結果可知，學生個人因素的直接影響效果中僅有創造力傾向，而在間接效果中也是以創造力傾向的效果較大。由此可知，對於國小學童創造力的影響而言，創造力傾向的重要性更甚於正向情緒及創意自我效能，於此，再次強調創造力傾向的重要性。

伍、結論與建議

於文獻回顧中，本研究獲得正向情緒、創造力傾向以及創意自我效能等因素與創造力間相互影響的概念模式，此概念模式經由假設的設立及模式界定過程中，形成本研究之假設模式，並經檢定後得整體模式可被接受。以下將逐一討論各假設之驗證結果：

一、創造力傾向對創造力具直接影響

本研究同時討論創意自我效能、正向情緒以及創造力傾向對於創造力的直接影響，研究結果發現，創造力傾向對於創造力具有正向的直接影響，但是正向情緒以及創意自我效能卻無直接影響。也就是說，對於國小五年級學童來說，如果創造力傾向程度越高，創造力表現也會越好。但是與過去研究不一樣的是，正向情緒、創意自我效能並不如預期的對創造力產生直接影響，本研究推敲原因可能是，過去討論正向情緒、創意自我效能對於創造力的影響多是討論對於創造力的情意或態度，較少真正討論對於創造力表現的影響。換言之，正向情緒或是創意自我效能都只是學童「認為」自己會有高度的創意表現，但是不論是正向情緒或是創意自我效能，到真正創造力表現的歷程恐怕還有許多因素須充分討論，而非僅是直接影響那麼簡單，但是否真是如此則有待未來研究進一步檢證。此外，也有可能是因為本研究的對象是五年級學童，對這時期的學生來說，或許他們自己並未感受到創意自我效能的程度差異，或是正向情緒對於創造力的可能影響，也就是說，即使學童具有高度的正向情緒或創意自我效能，他們也不知道要如何將

它轉化以影響創造力，也就是說，這中間的影響過程需要更多的中介效果討論，以下的結論或許可證實這個面向的討論。

二、正向情緒透過創造力傾向對創造力產生間接影響

本研究正向情緒對創造力不會產生正向的直接影響，但它仍可透過創造力傾向的中介而間接對創造力產生影響。表示國小學童若具有高正向情緒對創造力是沒有效用的，而是必須再進一步透過學生的創造力傾向，如此才會有高創造力的表現。所以為了提升學童的創造力，正向情緒需輔以高度的創造力傾向，才能對創造力產生相輔相成的正向間接效果。

三、正向情緒透過創意自我效能再透過創造力傾向對創造力產生間接影響

正向情緒除了可以透過創造力傾向的中介而間接對創造力產生影響之外，本研究還發現，正向情緒可透過創意自我效能，再透過創造力傾向的中介而間接對創造力產生影響。需注意的是，正向情緒透過創意自我效能，再透過創造力傾向，以及透過創造力傾向的中介的間接效果值總和為 0.397，相較於創造力傾向對於創造力的直接效果值 0.47 相當接近，表示正向情緒藉由創意自我效能、創造力傾向的中介效果雖然未大於創造力傾向的直接效應，但也不可忽視此中介效應。換句話說，若要大幅提升學童的創造力，除了直接提升學童的創造力傾向之外，藉由正向情緒以提升創意自我效能，再透過創造力傾向的提升也是一個好方法。

四、創造力的測量模式大致良好

本研究在各項指標負荷量和相關係數皆達顯著水準，表示這些構念用以衡量創造力模式得以作合理推論。但是還要注意的是，本研究的創造力測量，在內在結構模式的測量時，不論是潛在變項組成信度或是平均變異抽取量有未符合標準的情形，另外，創造力傾向此一變項中的一個測量指標：冒險性，以及正向情緒中的自信的個別項目信度也未達適配標準，而獨創力、流暢力、變通力等的信度也不足 0.4，雖然在 SEM 中已考慮測量誤差的問題，但測量信度偏低將導致受測者的真實創造力變異程度也偏低，可能影響創造力與正向情緒或自我效能的相關性被削弱，導致這兩個變項與創造力的結構關係不符預期，這也有可能是造成創

意自我效能以及正向情緒對於創造力的影響不顯著的可能原因之一。

五、對國小學童創造力發展之具體建議

(一) 教師應多加提升學生的創造力傾向

本研究發現創造力傾向對於創造力具直接且正向的影響效果，而且創造力傾向也是一個重要的中介變項，所以，提升學生的創造力傾向應該是創造力教育中的重要關鍵因素。而從創造力傾向的測量指標來看，較為重要的影響指標是好奇心以及想像力，如果能夠從提升學生的好奇心以及想像力應是不錯的做法。在教學上，教師可以設計一些與生活實際情境相結合的例子，吸引學生的注意之外，也讓他們能夠有思考的空間，發現問題之外，誘發他們的好奇心進而想要解決問題。另外，教學中可以多加善用聚斂式教學法以及發散式教學法交替運用，讓學生針對某一議題去討論，設計情境讓他們多加聯想，增進學生的想像力。也就是說，如果學生能夠於學校教育中多多提升自己的好奇心與想像力，應該可以具體提升學生的創造力。

(二) 維持學童的正向情緒並輔以創意自我效能及創造力傾向，以提升創造力表現

本研究發現，雖然正向情緒不會對創造力產生直接影響，但是創造力的表現可以藉由正向情緒、創意自我效能、創造力傾向的關聯以促進，所以，正向情緒仍然是有其重要性。儘管在本研究中是將正向情緒視為一種長期特質而非短期狀態，但是無可否認的，正向情緒仍然是會受到身邊的人事物而有一些短暫的波動，所以維持一個穩定的正向情緒還是相當重要的。以國小五年級的學童來說，正向情緒的來源可能多來自於與週遭他人的互動，也就是與教師、同儕、家人的互動，如果能夠讓學童生活在一個正向的環境中，營造和樂的氛圍，多鼓勵、少責罵，多肯定，少指責，應該可以培養學生穩定的正向情緒特質。而正向情緒的培養除了與他人的互動之外，也可以透過活動的設計加以培養，因為正向情緒包含同理心、知足、自信等等，如果學校的教學活動可以有一些讓學童培養同理心、知足的機會，例如服務學習、志工服務等等，或許也可以培養學童的正向情緒。但是僅有正向情緒是不夠的，必須輔以創意自我效能以及創造力傾向才能促進創造力。所以學效的教學活動設計除了提升學童的正向情緒之外，也應該在其中融入創意自我效能，逐步的設計一些難度，讓學童可以逐步克服之後，建立他們的成就感，

進而增進學童創造力的自信心。甚至可以設計具有挑戰性、冒險性的教案，以提升學童的想像力、好奇心，以增進學童的創造力。此外，雖然正向情緒可以透過創意自我效能、創造力傾向的關聯而促進創造力，卻不代表否定負向情緒的影響，因為「負向情緒/情感」一詞的包含面向也很廣，包含了挫折、不平、生氣、憤怒、苦悶、悲憫、憂傷、悲傷、失望、絕望、無助、嫉妒、討厭、焦慮等等，其中有些可能不利於創造力，有些可能有助於創造力。所以，情緒對於創造力的影響恐怕需要更多的研究投入以探討之間的關聯。

六、對未來研究之建議

就本研究的研究變項來看，影響學童創造力表現的因素相當複雜，但是本研究僅以正向情緒、創意自我效能與創造力傾向作為主要探討的因子，較缺乏家庭層面、學校層面及社會層面的探討，未來研究可納入上述四大層面中其他重要的因素。另一方面，本研究的創造力以及創造力傾向的冒險性以及正向情緒的知足不符合信度適配標準，建議後續研究者可以本研究的量表為範本，續編制符合臺灣本土學童的創造力傾向量表以及正向情緒量表，以供未來研究者使用。

參考文獻

- 余民寧(2006)。潛在變項模式－SIMPLIS的應用。臺北市：高等教育。
- 李偉清(2012)。國小資優生創造傾向量表之編製研究。**特殊教育研究學刊**，37(1)，79-102。
- 汪美香、黃炳憲(2012)。專業技能與創造力的關聯：創意自我效能中介角色。**南台學報**，37(4)，1-14。
- 林幸台(2002)。青少年的創造力發展。**應用心理研究**，15，63-67。
- 林幸台、王木榮(1999)。威廉斯創造力測驗。臺北市：心理。
- 林碧芳、邱皓政(2008)。創意教學自我效能感量表之編製與相關研究。**教育研究與發展期刊**，4(1)，141-170。
- 邱發忠、陳學志、徐芝君、吳相儀、卓淑玲(2008)。內隱與外顯因素對創造作業表現的影響。**中華心理學刊**，50(2)，125-145。
- 邱皓政(2003)。結構方程模式：LISREL的理論技術與應用。臺北市：雙葉。
- 洪久賢、溫秀玲、蔡長艷、宋慧娟(2003)。綜合活動領域教師創意教學成效之提昇：影響因素與策略運用。行政院國科會補助專題研究計劃成果報告，未出版。
- 洪文東(2002)。創造型兒童之思考特性與科學創造力的關聯性。**屏東師院學報**，16，355-394。
- 洪素蘋、黃宏宇、林珊如(2008)。重要他人回饋影響創意生活經驗？以模式競爭方式檢驗創意自我效能與創意動機的中介效果。**教育心理學報**，40(2)，303-322。
- 洪榮昭、康鳳梅、林展立(2003)。傑出科技創作學童創造特質分析—以機器人競賽為例。**師大學報：科學教育類**，48(2)，239-254。
- 胡夢蕾(2006)。我國創造力與人格特質研究之回顧與探析。**教育學刊**，26，215-239。
- 常雅珍、毛國楠(2006)。以正向心理學建構情意教育之行動研究。**師大學報：教育類**，51(2)，121-146。
- 陳正昌、程炳林、陳新豐、劉子鍵(2009)。多變量分析方法統計軟體應用(第五版)。臺北市：五南。

- 陳玉樹、郭銘茜（2013）。四向度成就目標對教師創意教學表現之影響：創意自我效能的中介效果與團隊學習行為的跨層級調節效果檢定。**教育科學研究期刊**，**58**（3），85-120。
- 葉玉珠（2000）。創造力發展的生態系統模式及其應用於科技與資訊領域之內涵分析。**教育心理學報**，**32**（1），95-121。
- 葉玉珠（2006）。**創造力教學 - 過去、現在與未來**。臺北市：心理。
- 葉玉珠、吳靜吉、鄭英耀（2000）。影響科技與資訊產業人員創意發展的因素之量表編製。**師大學報：科學教育類**，**45**（2），39-63。
- 蔡笑岳、朱雨傑（2007）。中小學生創造性傾向、智力及學業成績的相關研究。**心理發展與教育**，**2**，36-41。
- 蕭佳純（2011）。學校創新氣氛、教師內在動機與教師創意教學表現關連之研究：多層次調節式中介效果之探討。**當代教育研究季刊**，**19**（4），85-125。
- 蕭佳純（2015）。國小學童科學創造力成長歷程之縱貫性分析。**科學教育學刊**，**23**（1），23-51。
- 劉昆夏、鄭英耀、王文中（2010）。創意產品共識評量之信、效度析論。**測驗學刊**，**57**（1），59-84。
- Amabile, T. M. (1996). *Creativity in the context: Update to the social psychology of creativity*. Oxford, England: Westview Press.
- Amabile, T. M. (1988). A model of creativity and innovation in organizations. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 10, pp. 123-167). Greenwich, CT: JAI.
- Ashby, F. G., Isen, A. M., & Turken, U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, *106*(3), 529-550.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191-215.
- Bandura, A. (1997). *Self efficacy: The exercise of control*. New York, NY: Freeman.
- Baumann, N., & Kuhl, J. (2005). Positive affect and flexibility: Overcoming the precedence of global over local processing of visual information. *Motivation and Emotion*, *29*(2), 123-134.
- Chan, S., & Yuen, M. (2014). Personal and environmental factors affecting teachers'

creativity-fostering practices in Hong Kong. *Thinking Skills and Creativity*, 12, 69-77.

Chang, C. P., Chuang, H. W., & Bennington, L. (2011). Organizational climate for innovation and creative teaching in urban and rural schools. *Quality Quantity*, 45, 935-951.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, 4, 62-83.

Doll, B., & Lyon, M. A. (1998). Risk and resilience : Implications for delivery of educational and mental health services in schools. *School Psychology Review*, 27(3), 348-363.

Florida, R. (2002). *The rise of the creative class and how it's transforming work, leisure, community, & everyday life*. New York, NY: Basic Books.

Forgas, J. P., Vargas, P., & Laham, S. (2005). Mood effects on eyewitness memory: Affective influences on susceptibility to misinformation. *Journal of Experimental Social Psychology*, 41, 574-588.

Fredrickson, B. L.(1998). What good are positive emotions? *Review of General Psychology*, 2(3),300-319.

Fredrickson, B. L. (2002). Positive emotions. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 12-134). New York, NY: Oxford University Press.

Fredrickson, B. L.(2003). The value of positive emotions: The emerging science of positive psychology is coming to understand why it's good to feel good. *American Scientist*,91(4), 330-335.

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotion. *The RoyalSociety*, 359, 1367-1377.

Fredrickson, B. L. (2006). The broaden-and-build theoryof positive emotions. In M. Csikszentmihalyi & I. S.Csikszentmihalyi (Eds). *A life worth living: Contributions to positive psychology* (pp. 85-103). New York, NY: Oxford University Press.

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and through-action repertoires. *Cognition and Emotion*, 19(3), 313-332.

- Gasper, K. (2004). Permission to freely? The effect of happy and sad moods on generating old and new ideas. *Creativity Research Journal*, 16(2&3), 215-229.
- Gibson, R. (2010). The “Art” of creative teaching: Implications for higher education. *Teaching in Higher Education*, 15(5), 607-613.
- Gong, Y., Huang, J. C., & Farh, J. L. (2009). Employee learning orientation, transformational leadership, and employee creativity: The mediating role of employee creative self-efficacy. *Academy of Management Journal*, 52(4), 765-778.
- Grawith, M. J., Munz, D. C., & Kramer, T. J. (2003). Effects of member mood states on creative performance in temporary workgroups. *Group Dynamics: Theory, Research, and Practice*, 1, 41-54.
- Hwang, W. Y., Chen, N. S., Dung, J. J., & Yang, Y. L. (2007). Multiple representation skills and creativity effects on mathematical problem solving using a multimedia whiteboard system. *Educational Technology & Society*, 10(2), 191-212.
- Isen, A. M. (2000). Positive affect & decision making. In M. Lewis and J. Haviland-Jones (Eds.), *Handbook of Cognition and Emotion* (2nd ed.) (pp. 417-435). New York, NY: Guilford.
- Isen, A. M. (2001). An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. *Journal of Consumer Psychology*, 11(2), 75-85.
- Jaussi, K. S., Randel, A. E., & Dionne, S. D. (2007). I am, I think I can, and I do: The role of personal identity, self-efficacy, and cross-application of experiences in creativity at work. *Creativity Research Journal*, 19(2&3), 247-258.
- Kaufmann, G., & Vosburg, S. K. (2002). The effects of mood on early and late idea production. *Creativity Research Journal*, 14(3-4), 317-330.
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, 31(3), 313-327.
- Magoun, D., Eaton, V., & Owens, C. (2002). *IT and the attitudes of middle school girls: A follow-up study*. Paper presented at the meeting of National Educational Computing Conference, San Antonio, TX.
- Oldham, G. R., & Cummings, A. (1996). Employee creativity: Personal and contextual

- factors at work. *Academy of Management Journal*. 39(3), 607-634.
- Puccio, G. J., Murdock, M. C., & Mance, M. (2007). *Creative leadership*. London, England: Sage.
- Runco, M. A., & Sakamoto, S. O. (1999). Experimental studies of creativities. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp.273-296). New York, NY: Cambridge.
- Runco, M. A., & Walberg, H. J. (1998). Personal explicit theories of creativity. *The Journal of Creative Behavior*, 32(1), 1-17.
- Schwarz, N.(2000). Emotion, cognition and decision making. *Cognition and Emotion*, 14, 433-440.
- Seligman, M. E. P. (2002). *Authentic happiness. Using the new positive psychology to realize your potential for lasting fulfillment*. New York, NY: The Free Press.
- Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology-An introduction. *American Psychologist*, 55(1), 5-14.
- Shapiro, P. J., Weisberg, R. W., & Alloy, L. B.(2000). *Creativity and bipolarity: Affective patterns predict trait creativity*. Paper presented at the meeting of Convention of the American Psychological Society, Miami, FL.
- Sternberg, R.J. & Lubart, T.I.(1999). The concept of creativity: Prospects and paradigms. In R.J. Sternberg (Ed.), *Handbook of creativity* (pp. 3-15). New York, NY: Cambridge University.
- Tierney, P., & Farmer, S. M. (2002). Creative self-efficacy: its potential antecedents and relationship to creative performance. *Academy of Management Journal*, 45(6), 1137- 1148.
- Williams, F. E. (1980). *Creativity assessment packet (CAP): Manual*. Buffalo, NY: D.O.K.

Teachers' Cognitive Processes from the Perspective of Distributed Cognition: An Analysis Focused on Teachers-Made Assessments

Chin-Tu Liang

Principal, Yu-Sin Elementary School, Changhua County

Adjunct Professor, Graduate Institute of Education, National Chung Cheng University

Abstract

Few studies deal with how the gap between artifacts and task in terms of members' cognition distributed (Halverson & Clifford, 2006) in the field of teaching practice. The present study explored how the teachers' cognition evolved with a close look at the artifact, and the processes adopted by "Two-Way Specification Table" (TWST) for teachers-made assessments as an artifact designed for nine students' tests in "SiSi" Elementary School. Participatory observation, interviews, and documents were employed. The following conclusions were drawn: Firstly, after the implementation of TWST in SiSi Elementary School, teachers had produced 13 kinds of cognitive systems respectively. Secondly, during the process of implementing TWST, teachers gradually presented different stages of cognitive systems such as questioning, reflection, transformation, and creativity. Moreover, continuous discussions and sharing, the need for experts' assistance, developing the technological systems, and the positive attitude of principal were key elements in the functioning of the TWST. Lastly, after the implementation of TWST in SiSi Elementary School, it was found that teachers' cognitive systems were slightly improved.

Keywords: artifact, cognitive process, distributed cognition, teachers-made assessment, Two-Way Specification Table



分布認知觀點的教師認知歷程： 以教師自編評量為分析焦點

梁金都 彰化縣育新國小校長、國立中正大學教育學研究所兼任助理教授

摘 要

本研究主要是採取個案研究法，並選取「西西」國小（化名）為研究個案，且以該校實施「雙向細目表」的教師自編評量為中介物，透過了觀察和訪談與文件分析等方法，蒐集 11 位研究參與者與該表持續交互影響的實地資料，以了解教師們的認知歷程。最後本研究獲致了以下結論：一、西西國小實施雙向細目表後，教師們分別產生了十三種認知系統。二、雙向細目表的歷程，教師們漸次呈現質疑、反思、轉向、創新等不同的認知階段。三、持續地討論與分享、需要專家協助、科技系統引入、校長積極推動的態度等，乃是發揮雙向細目表功能的關鍵因素。四、西西國小實施雙向細目表，教師們的認知系統大致呈現朝向上升方向發展。

關鍵詞：人造物、認知過程、分布認知、教師自編評量、雙向細目表



Introduction

The transform and development of teachers' thinking, especially about their teaching activities, is closely related to the students' learning outcomes in the day-to-day running of school (Firestone & Riehl, 2005). Liang and Hung (2011) point out that teachers' thinking and behaviors are transformed or evolved and even make more consistent with each other by way of continuous sharing and discussing about teaching issues among them in the school. With it, when teachers converse and discuss with other colleagues about teaching practice, they can not only collaborate with each other but also achieve the teaching tasks in school (Allison, 1996; Catano & Stronge, 2006). From the perspectives of school leadership, for upgrading the qualities of students' learning, principals should understand various cognitions about teaching issues among teachers and provide more and more opportunities for their professional learning activities in teachers' professional learning communities through continuous interactive processes for teaching skills and professional knowledge (Liang & Lin, 2013).

Individuals' cognitions are distributed in situations and contexts of the organization from the perspective of distributed cognition. Such this, we are able to analyze the cognitive systems of actors (i.e., the thoughts and actions of actors) by way of the interactive processes between them and school working contexts as "artifacts" (e.g., checklists, teaching program) (Halverson & Clifford, 2006, p. 583-586). Furthermore, we have to consider the possibility that, while cognition can be distributed, they need a few "sources" for this distribution such that can operate conjointly. And we also have to consider the possibilities that each of these so-called sources, or intellectual partners (actors), can also grow such that each subsequent joining of partners will become more intelligent (Salomon, 1993, p. 111). As Perkins (1993, p. 90) also notes, the surrounding-the immediate physical and social resources outside the persons, which participates in cognition, not just act as a source of input and a receiver of output, but as a vehicle of thinking. In other words, the theory of distributed cognition maintains that teaching practices wherein communicative artifacts are put can naturally create the interaction between teachers and artifacts, and the teachers' cognition could be analyzed and

understood further. And then it even can transform teachers' cognitions and thus become more intelligent. In brief, the theory of distributed cognition can analyze teachers' cognitions and their developmental processes when principal or instructional leader adopts school working contexts as artifacts (i.e., checklist or teaching program) and interacts with teachers over and over.

Most research studies investigating distributed cognition issues focused on school and instructional leadership (e. g., Halverson & Clifford, 2006; Liang, 2011; Liang & Hung, 2011). Relatively few studies deal with how the gap between artifacts and tasks in terms of members' cognition distributed (Halverson & Clifford, 2006) in the field of teaching practice, especially the analysis of the developmental processes of teachers' cognitions through the checklist for "Two-Way Specification Table" (TWST) on teachers-made assessments as an artifact. Consequently, the findings of the present case study, which examined and analyzed teachers' cognitive systems by means of the interaction between teachers and TWST in contexts of "SiSi Elementary School", should contribute to the development of future innovative strategies for teachers' teaching and their professional knowledge.

Literature Review

The design of the present study was based on a review of the existing literature on the analytical approach for distributed cognition, the features of "Two-Way Specification Table", and the relationship between them.

The Analytical Approach for Distributed Cognition

The theory of distributed cognition stresses that individuals' cognitions are distributed in the social situation (Perkins, 1993). It could analyze the individuals' cognitions by way of the interactive processes between actors and the artifacts (e.g., teaching program, checklist) (Halverson & Clifford, 2006). Distributed cognition also suggests that owing to the interactive processes among members, situation, and the artifact it could bring up the transform and development on individual cognitions (Liang

& Hung, 2011). Moreover, Salomon (1993) further argues that individuals' inputs could affect the nature of the joint, distributed system through their collaborative activities (i.e., artifacts), so as to affects their cognitions, then their subsequent participation is altered, leading to subsequent altered joint performances and products. As a result, based on the perspective of such theory it could analyze the individuals' cognitions and understand how they acquire the new knowledge and abilities through collaborative activities as artifacts, which causes the interactive processes among individuals' cognitions, situation, and artifacts (Gan & Zhu, 2007; Hands, 2010; Hutchins, 1990).

With respect to analyzing the dimensions on distributed cognition theory (see Figure 1), Halverson and Clifford (2006) point out that it analyzes the development and practice of principals' leadership and distinguishes the policies from instruments (i.e., artifacts) based on the perspective of distributed cognition. Thus, the analyses on distributed cognition theory are guided by three key questions: 1) What is task? 2) What are the relevant artifacts? 3) How do tasks and artifacts link in the cognitive system? Briefly, the artifacts are adopted as analytical objects in terms of the traits of school tasks (i.e., macro tasks or micro tasks), and second, cognitive systems of members (i.e., principal's cognitive systems and teachers' cognitive systems) are understood by way of continuous interactive processes between artifact and such members (Liang & Hung, 2011). Three key questions are as follows:

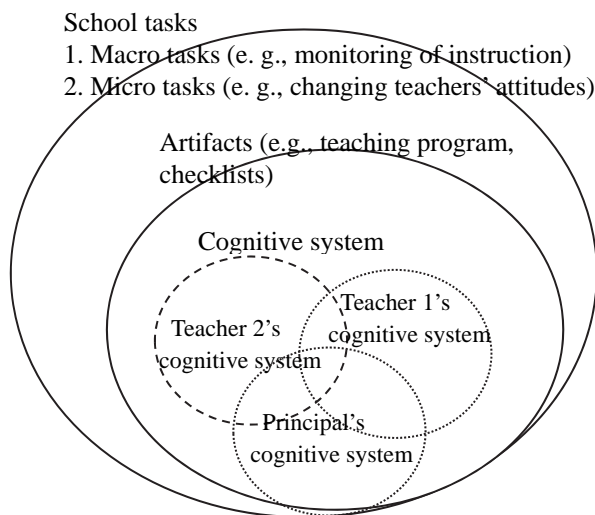


Figure 1 The Analyzing of Dimensions on Distributed Cognition Theory

Note 1. The circular dashed line means the actors' cognitive system and will change as the actors interact with the artifact continuously.

Note 2. Adapted from "Analysis of the cognitive process of a principal's instructional leadership from distributed cognition perspectives," by C. T. Liang, & C. C. Hung, 2011, *Educational Research Quarterly*, 19(2), p. 51.

School tasks

From the perspective of distributed cognition, the traits of the actors' cognitive system are understood from what tasks (i.e., macro or micro tasks) they select in school contexts, where macro tasks involve descriptions at the more abstract tasks or the large-scale organizational level and micro tasks describe the specific behaviors derived from the macro tasks (Spillane, Halverson, & Diamond, 2001). For example, from a school leadership perspective, macro task such as "monitoring of instruction" is composed of micro task such as "changing teachers' attitudes" (Halverson & Clifford, 2006, p. 583). van der Meij and Boersma (2002: 198) also argue that the choice of tasks should be as clear as possible in order to facilitate analysis the traits of the cognitive system. Such these, from the perspective of distributed cognition, for completing the macro tasks and leadership purpose the choice of the micro tasks have to be as clear as it could be in order

to analyze the actors' cognitive systems.

Artifacts

Individual cognitions as an artifact are distributed in substance and culture (Spillane et al., 2004). And an artifact is regarded as the vehicles of thinking on staff members (Perkins, 1993) and it also has social functions for communicate understanding (Hutchins, 1995). Furthermore, the artifact is penetrated within leading activities so as to complete the setting tasks (Pea, 1993). And its design includes the features that reflect the intentions of a designer (i.e., principal) on suggested uses or effects, and it is also practiced by users (i.e., teachers) (Halverson & Clifford, 2006). Namely, the actors' (i.e., principal or teachers) thoughts and actions are understood by way of the interaction between artifact and them in the leading activities or tasks. In particular, the artifact may be a language, technology, or physical space as observed cultural level on organization (Schein, 1985), formal structure and rule, as well as all kinds of agreements (Spillane et al., 2004), such as the policies or programs (Spillane et al., 2004), the checklists (Halverson & Clifford, 2006), and the logs or buildings (Gagliardi, 1990).

Cognitive system

Cognitive system involves the artifacts (i.e., macro task or micro task) and actors (i.e., principal or teachers) who are needed to complete micro and macro tasks. In other words, the actors' cognitive systems are presented in "Interaction between actors and artifacts" and "The choice of actors for the features of the artifacts" for completing micro and macro tasks (Halverson & Clifford, 2006, p. 586). In particular, how actors interact with artifacts as well as what actors select the features of artifacts represent the actors' thoughts and actions (i.e., cognitive systems) (Spillane et al, 2004, p. 23). Specifically, principal's cognitive systems are appeared when they select and interact with artifacts to complete the tasks, and similarly teachers interact with such artifacts represent their cognitive systems (Halverson & Clifford, 2006, p. 585-586).

The Features of "Two-Way Specification Table"

Concerning the features of TWST, we can firstly introduce the structure of Bloom's revised taxonomy, and then interpret its implementations in Taiwan.

The structure of Bloom’s revised taxonomy

Generally speaking, to complete the teaching tasks TWST as a checklist (i.e., an artifact) that include the knowledge dimension or teaching materials contents as well as cognitive process dimension, and we use it within the teachers-made assessments on students’ learning tests (Lee, 2011). Concerning the TWST, the structure of Bloom’s revised taxonomy (see Table 1) called “Bloom 2001” is the most representative version that is refinement and extension of original work named “Bloom 1956” which is an often utilized tool for classifying six educational objects such as knowledge, comprehension, application, analysis, synthesis, and evaluation based on what teachers expect their students to learn (Spindler & Tech, 2015). In particular, the original taxonomy consisted of only a cognitive process dimension; the revised version incorporates a knowledge dimension as well. And second change for cognitive process dimension in terminology, form in noun form (knowledge, comprehension, etc.) were written in verb form (remember, understand, etc.) that emphasizes the active cognitive behavior desired from a student (Seaman, 2011).

Table 1 *The Structure of Bloom’s Revised Taxonomy*

knowledge dimension	Cognitive process dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual						
Conceptual						
Procedural						
Meta-cognitive						

Note. Adapted from“Taxonomy for Learning, Teaching and Accessing: A Revision of Blooms’ Educational Objectives (p. 29),” by W. Anderson, & D. R. Krathwohl, 2001, New York, NY: Longman; “A revision of Bloom’s Taxonomy: An overview,” by D. R. Krathwohl, 2002, *Theory into Practice*, 41(4), p. 216.

Regarding the “Bloom 2001”, it includes two parts. On the one hand, there are four categories and eleven items in knowledge dimension. The four categories of knowledge are factual, conceptual, procedural, and meta-cognitive. Specifically, factual knowledge is considered to be knowledge of terminology and fact (i. e., knowledge of specific details,

and knowledge of terminology). Conceptual knowledge entails a deeper understanding of contents (i. e., knowledge of classifications, knowledge of principles, and knowledge of theories). Procedural knowledge maintains the knowledge of how to do something (i. e., knowledge of subject-specific skills, and knowledge of criteria for determining the timing to use appropriate procedure). Metacognitive knowledge signifies knowledge of one own cognition (i. e., strategic knowledge, knowledge about cognitive tasks, and self-knowledge) (Anderson & Krathwohl, 2001; Spindler & Tech, 2015).

On the other, six categories and nineteen items are classified in the cognitive process dimension listing additional verbs within each of six levels which more clearly delineate their nature. They are as follows: Titled remember means recalling or remembering the information (recognizing, and recalling), titled understand stresses explaining ideas (translating, interpreting, exemplifying, classifying, summarizing, comparing, and explaining), titled apply entails using the information in a new way (executing, and implementing), titled analyze maintains distinguishing between the different parts (differentiating, organizing, and attributing), titled evaluate entails justifying a stand or decision (checking, and critiquing), and titled create signifies creating a new viewpoint (generating, planning, and producing) (Anderson & Krathwohl, 2001; Spindler & Tech, 2015).

Concerning above-mentioned, TWST not only defines the categories for such two dimensions on teachers-made assessment for students' learning tests but also helps teachers know how the criteria of these categories differ. Moreover, it is important for teachers to integrate the curriculum, teaching, and assessment at such tests for promoting their teaching effectiveness (Yen & Lin, 2003).

The practice of "Two-Way Specification Table" in Taiwan

Concerning TWST on teachers-made assessments for students' learning tests in Taiwan (see Table 2). It integrates the curriculum, teaching, and assessment designed by teachers, which usually uses extracting various knowledge categories (e. g., factual, conceptual, etc.) and different cognitive levels (e. g., remember, understand, etc.) from the units of teaching materials (e. g., first unit, second unit, etc.) on teachers-made assessment for students' learning tests. Due to the design of TWST, especially how one

defines cognitive levels and what extracts knowledge categories based on materials units at tests for upgrading students’ learning outcomes. They are very difficult for teachers (Wang, Wang, Wang, & Huang, 2003). Thus, they try to use an alternative method as TWST which adopts different types of questions (e.g., true or false, multiple choice items, filling in the blanks, essay, etc.) standing for different knowledge categories and various cognitive levels in the checklist for understanding their learning outcomes.

Besides, in this table “Subject” refers to mathematics, social science, etc., with “Grade” referring to grade 1, grade 2, etc.; “Teacher” referring to Teacher A, Teacher B, etc.; and with “Total” referring to total number of the questions. To summarize, after the students’ learning tests teachers can schematically understand what knowledge categories of teaching material units are needed and know which cognitive levels that the tested students have achieved.

Table 2 “Two- Way Specification Table” of Teachers-Made Assessment for Students’ Test

Subject		Grade		Teacher				
Knowledge dimension		Cognitive process dimension						Total
		Remember	Understand	Apply	Analyze	Evaluate	Create	
First unit (knowledge of factual, conceptual, procedural, and meta-cognitive)	True							
	Choice							
	Filling							
	Essay							
second unit (knowledge of factual, conceptual, procedural, and meta-cognitive)	True							
	Choice							
	Filling							
	Essay							

(Continued on the next page)

Table 2 “Two- Way Specification Table” of Teachers-Made Assessment for Students’ Test (continued)

Subject		Grade		Teacher				
Knowledge dimension		Cognitive process dimension						Total
		Remember	Understand	Apply	Analyze	Evaluate	Create	
Third unit	True							
(knowledge								
of factual,	Choice							
conceptual,								
procedural,	Filling							
and meta-								
cognitive)	Essay							
Total								

Note 1. The true refers to true and false, with choice referring to multiple choice items, and filling referring to filling in the blanks. Such true and false, etc. as questions teacher chooses standing for different knowledge categories and various cognitive levels from units of materials on teachers-made assessments.

Note 2. Adapted from “The Establishment of Chinese General Proficiency Indicator,” by Y. L. Cheng, 2006, *Journal of Liberal Arts and Social Sciences*, 2, p. 124-126; “Two-Way Specification Table as mathematics,” by I. J. Lu, & G. I. Wu, 2011, *In-service Education Bulletin*, 28(5), p. 98.

Observe that, in Table 2 regarding the students’ tests teachers usually firstly analyze and choose knowledge categories (i. e., factual or conceptual knowledge) of the units of materials (e. g., first or second unit of material of Social Science subject), and secondly edit the questions as the true or false, multiple choice items, as well as filling in the blanks, which stand for simple cognitive levels (i. e., remember, understand, and apply), respectively. In contrast, for achieving higher cognitive levels (i. e., analyze, evaluate, and create), teachers might adopt the procedural or meta-cognitive knowledge of units of materials (e. g., second or third unit of material of Social Science subject) and use as types of essay for editing questions that makes the students try to interpret and explain at tests (Chien, 2009).

Overall, it is an important process for teachers to conduct the remedial teaching plans and improve the teaching activities when they understand the knowledge categories of teaching materials’ units and the students' cognitive process during such period

(Lu & Wu, 2011). Therefore, to enhance teaching effectiveness and completing their teaching tasks, teachers should make their way to integrate the curriculum, teaching, and evaluation through designing the teachers-made assessments for students' learning test based on TWST.

The Relation between Distributed Cognition and “Two- Way Specification Table”

According to the above-mentioned reasons, researcher can analyze teachers' cognitive systems through the interaction between teachers and TWST based on the theory of distributed cognition. In other words, teachers' cognitive systems and their developmental processes are understood when they extract what knowledge categories of materials' units are needed and define how various cognitive levels are reached for students by proposing the types of questions at students' tests in the light of TWST. That is, as an artifact, TWST is adopted as an analytical object in terms of micro school task of changing teachers' attitudes and macro school task of instruction monitoring. In particular, teachers can complete the teaching tasks effectively and upgrade the students' learning outcomes based on the TWST at students' tests, and further improve their professional abilities in the teaching activities or the remedial teaching strategies.

However, at the same time teachers should face the difficulties and challenges including editing the types of questions that can extract the knowledge categories of teaching materials' units and define the cognitive levels for students' learning. Therefore, it becomes a very important issue to edit teachers-made assessments at students' learning tests for upgrading the learning outcomes of students.

Study Approach

The design of this study was based on the discussion presented in the introduction and literature review section.

Study Framework and Methods

Case studies investigated “why” and “how” in real-life contexts (Yin, 2002). For distributed cognition theory, the present study employed the artifact as “TWST” on teachers-made assessments for students’ tests, which was chosen as a micro task for changing teachers’ attitudes derived from the macro task for monitoring of instruction in school contexts. In particular, the TWST was interacted with teachers to analyze their thoughts and actions (i.e., cognitive systems) at the “SiSi” Elementary School. The “how” was provided by identifying teachers’ cognitive systems when they interacted with TWST. And the “why” was provided by analyzing the teachers’ cognitive systems and their possible developmental processes at their school (see Figure 2).

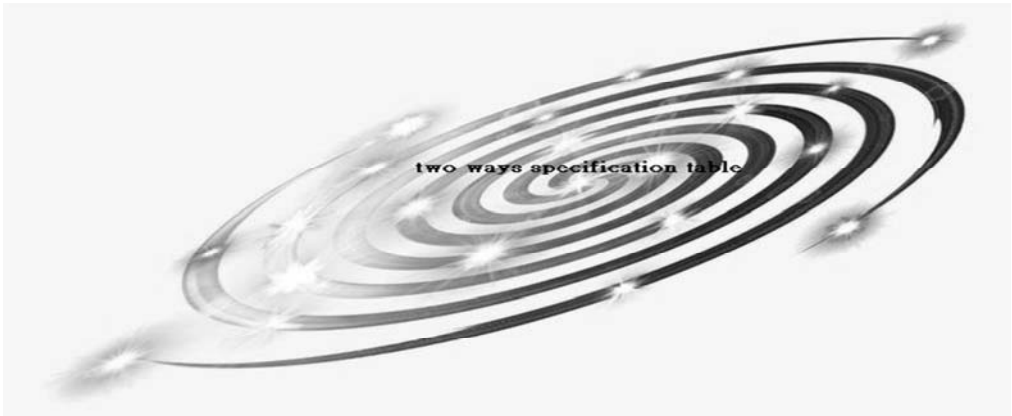


Figure 2 Teachers’ Cognitive Systems and Their Developmental Processes at SiSi Elementary School

Note 1. The light spots and their shining lights entailed the teachers’ cognitive systems and their possible developmental processes, respectively, when they interacted with TWST in the school contexts over and over.

Note 2. Adapted from: Author drew

Participants and Contexts

In present case study, the data were obtained from “Principal P” and ten teachers at “SiSi” Elementary School. In 2014, principal P selected TWST as an artifact asked policy by “W” county Government in Taiwan to interact with teachers in order to promote their teaching effectiveness. In SiSi Elementary School, ten teachers were recruited as

participants to practice TWST. In this study, they exhibited differences traits in gender, levels of seniority, and individual characteristics (see Table 3).

Table 3 *Characteristics of the Eleven Participants at “SiSi” Elementary School*

Name	Sex	Seniority	Members' characteristics and commitment to the school
Principal P	M	28	Enjoys good interpersonal relationships, exhibits concern for member needs as well as a focus on student learning, motivating teacher enthusiasm, and improving teacher professional abilities
Director A	F	20	Plans the curriculum and teaching, emphasizes the spirit of the law, and exhibits a conscientious attitude toward work
Teacher A	F	15	Team leader for teaching affairs, and responsible for TWST
Teacher B	F	24	6th grade teacher who emphasizes professional autonomy and continually expresses contrary opinions on administrative matters
Teacher C	F	15	6th grade teacher with a straightforward personality who is good at helping students visualize information in the textbook
Teacher D	F	16	5th grade teacher who has excellent teaching skills and a conscientious attitude toward teaching
Teacher E	F	7	5th grade teacher with excellent class management skills who emphasizes teachers' professional autonomy
Teacher F	M	19	4th grade teacher who exhibits proficient teaching skill but rarely interacts with colleagues
Teacher G	M	7	3th grade who exhibits few interactions with staff, unstable emotions, and mediocre teaching performance
Teacher H	F	4	2nd grade teacher who supports the school administration and enjoys good interpersonal relationships with colleagues
Teacher I	F	21	1st grade teacher who focuses on classroom teaching and values professional development and others' opinions

The recruited eleven participants were mainly considered in line with the ecological distribution of the school. We try our best to look for the members with different

backgrounds to meet their job categories (i. e., school leader, executive administrators, and teachers), gender (i. e., three male and seven female staff members V. S. twelve male staff and twenty- eight female staff members), age of service (10 years or less, 11 to 20 years, and 21 to 30 years), teaching field (from first to sixth grade teachers), and the degree (personal characteristics and commitment to the school) so as to facilitate the collection of information on the interaction between teachers and TWST. In brief, the characteristics of the recruited eleven participants at “SiSi” Elementary School were chosen such that they could stand for the whole school.

“SiSi” Elementary School is located in the countryside of the “W” County Government, Taiwan. There are twenty classes and approximately forty teachers in this school. “Principal P” had been serving as principal for fifteen years, and had good interpersonal relationships with teachers, exhibiting concerns for teachers’ needs, setting up the professional learning communities, and constantly successful experiences at the school. In particular, he focused on creating a positive culture that had emphasized student learning, promoting teacher enthusiasm, and enhancing teacher professional expertise, etc.

Although the backgrounds of most students were from working-class families in this school, most teachers in this school possess excellent teaching skills and enthusiasm. The teachers can have sufficient support of equipment (e. g., transparency projectors and DVDs) when they need in their teaching environment. In practical, the teachers constantly discussed the teaching issues with other teachers; e.g., on how teaching strategies could be used and which kind of teaching skill should be applied. Occasionally, they also discussed about teachers-made assessments at students’ tests, which involved teachers’ cognitions and teaching technology ... etc.

Methods

Participatory observations, interviews, and documents were employed in the present case study to analyze and interpret the interactions between TWST and teachers. The study was conducted within the duration of four periods (i. e., commencing stage, initial stage, medial-term stage, and final stage) from January 2014 to April 2015 (see Figure 3).

Months Periods	2014												2015			
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4
commencement	First and second tests															
initial stage	Third and fourth tests															
medial-term	Fifth and sixth tests															
final stage	Seventh to ninth tests															

Figure 3 The Timing of Data Collection for the Four Periods from January 2014 to April 2015

Participatory observations primarily occurred in the offices of school staff members and classrooms. The beginning of the first test of the entry, the time was about 1-2 hours; mid-term (4-12 months) observed the frequency was more frequent and longer (about 2-3 hours), and as far as possible in the distribution the different time of the week to obtain a comprehensive and diverse information; the number of the final stage gradually reduced, and slowly out of the field, the main collection of previously not received the main information (see Table 4). The data coding were identified with respect to the date of observation (“Observation Year-Date”, e.g., Ob14-1205, Ob15-0103).

Table 4 *Timing of Participatory Observations*

Month (during 2014-2015)	Times	Time (hours)	Total (hours)
01	4	1~2	6
02	6	1~2	8
03	6	1~2	10
04	9	2~3	18
05	11	2~3	20
06	10	2~3	24
07	7	2~3	15
08	9	2~3	20
09	11	2~3	22

(Continued on the next page)

Table 4 *Timing of Participatory Observations (continued)*

Month (during 2014-2015)	Times	Time (hours)	Total (hours)
10	8	2~3	11
11	8	2~3	11
12	7	2~3	10
01	7	1~2	10
02	5	1~2	8
03	4	1~2	6
04	2	1~2	5
Total			204

Moreover, interviews were primarily based on interview outlines that were reviewed from the professional opinions of three professors and two elementary school principals (see Table 5). And the coding that were identified with respect to the participant and date of the interview (“Interview participant Year-Date”, e.g., Ipp14-0103, Ita15-1212, with “pp” referring to “principal p” , “with “ta” referring to “teacher a”). The interviews took a semi-structured approach, allowing respondents to express their views more freely and to delineate the activities and their meanings in various contexts. For unclear interviews, researcher tried to seek more explicit and in-depth interviews with questioning. For example, in the interview, one of the teachers pointed out that “it was not necessary to think about these problems in the past, now I find the TWST bring a lot of teaching changes so that I would have noticed the ability of students to analyze and evaluate [higher cognitive level]” (Itb14-0410), she argued that “ compared to the TWST, although teachers’ experience will lead to changes in teaching activities, it will not pay attention to the analysis and evaluation of higher-level ability” (Itb14-0410). This means that the TWST than the teaching experience, can bring the teaching changes faster.

In addition to active data collection, the contents of document as official documents or meeting documents involved in school daily life and were identified with respect to the document and orders (“Document Order”, e.g., D-05).

Table 5 *Interview Outlines*

1. Do you like using the TWST for teachers-made assessment at students' tests and why?
2. What are your opinions for implementing TWST for teachers-made assessment at students' tests and why?
3. How does TWST affect your teaching activities?
4. Could you tell me what knowledge categories of materials' units you choose based on TWST and how use the types of questions (i. e., true or false, multiple choice items, filling in the blanks, and essay) to reach such categories?
5. What knowledge categories of materials' units you choose based on the features of TWST with the colleagues? How do you do if you face the difficult problems?
6. Could you tell me how you define the various cognitive levels such as titled for remember, understand, apply, analyze, and create and how use the types of questions (i. e., true or false, multiple choice items, filling in the blanks, and essay) to reach such cognitive levels?
7. What problems of cognitive levels you define with the colleagues? How do you do if you face the difficult problems?
8. How you or your colleagues take the right things for teachers-made assessment at students' tests and what?
9. Could you implement TWST for teachers-made assessment at students' tests again and again after pushing forward this plan and why?
10. Could you think the implementation of TWST is better for learning outcomes of students and improving teachers' expertise and why?

Research Credibility

In the field data obtained by the study, the researcher will correct the reliability of the information by participatory observations, interview and document as collection methods. Secondly, interviews with different respondents for the same theme, as well as the collection of positive and negative views on them, to benefit the authenticity of the gathering information. For example: “for teachers, it [implementing TWST] is not just an obligation for teaching, but also to pay attention to whether we care about our rights and interests, such as enriching teaching equipment, etc. (Itb14-0315)”; “Only for pushing forward the TWST for teachers-made assessment at students' tests and not to know what teachers need, this action is not fair. (Itg14-0313)” The above statements confirm that the principal push forward the practice of TWST is very difficult work due to teachers doubt it.

In addition, the original data and coding information collected by the researcher further allowed the participant to review and submit to the research peer (Teacher Z) for checking to minimize the subjectivity of the researcher and improve the reliability of data analysis.

Findings

Based on the data for implementing TWST from four periods at “SiSi” Elementary School (e. g., commencing stage, initial stage, medial-term, and final stage), we obtained the following findings regarding the teachers' cognitive systems and their possible development processes.

Commencement of Implementation for TWST: First and Second Tests

The teachers-made assessments for students' tests needed to be reviewed by TWST for ensuring their professional qualities due to the teachers in elementary schools were asked by “W” County Government in this semester (i.e., February 2014-July 2014). This requirement was to learn more teaching techniques (i. e., teachers-made assessments) for promoting teachers' professional abilities by way of delivering and sharing with artifact (TWST) in organization (Cook & Yanow, 1993). Relatively, the teachers at SiSi Elementary School did not believe that this could promote their professional skills, because such a requirement was a top-down bureaucratic system. Namely, the communications concerning teaching techniques and activities between school administration (i.e., principal) and teachers might not happen (Lieberman & Grolnick, 1996). In other words, it was a boring policy and bought up the negative efforts in school. As were expressed by one teacher and an observed note at the first test separately: “I thought I just made such assessments based on common senses as my teaching experiences in the past,, so now the implementation of this plan [TWST] is boring for us”. (Itg14-0103) “For completing such requirement, some teachers casually filled the checklist [TWST] on their made assessments....., so it is a superficiality behavior”. (Ob14-0116)

These negative voices were spread everywhere in the school at the commencement of implementation for TWST. However, the actors' (teachers') various cognitive systems were understood when they constantly interacted with artifact (TWST) which was derived from the macro-tasks (i.e., monitoring of instruction) and micro-tasks (i.e., changing teachers' attitudes) (Liang & Hung, 2011). For example, relative to such negative voices one teacher had been identified as a professional teacher vocalized a positive view: "The tests [TWST] for improving teachers' teaching planning can help teachers to understand the actual situations of students' learning outcomes (Itd14-0212)". Moreover, other teachers, after the first test, started to review the problems and find the reminder function by TWST for their made assessments, although they endured the pressure from authority asking such policy. It also created a more impartial opinion to them, as views were explained by two teachers: "We could easily hold the focal point about the teaching materials through TWST checking to promote the effectiveness of teaching activities". (Ida14-0305) "Principal always checks it [TWST] again and again so that I should make it carefully on such assessments; it is a reminder which is important thing for me, especially for a novice". (Ith14-0313)

Furthermore, at second test the participative teachers shared and discussed with each other about their made assessments which were reviewed by TWST once again. It could gradually adjust the cognitive systems of actors (participative teachers) when they continually interacted with artifact (TWST) (Halverson & Clifford, 2006). Meanwhile, some teachers, professional attitudes, appeared different thoughts and behaviors. For examples: "For teachers teachers-made assessment based on TWST is a professional affair, but it, asking by school administration, is not a proper thing". (Itb14-0331) "These [teachers-made assessments in the past] are unfair for students and need to be corrected by way of reviewing with TWST....., so such policy is needed". (Iti14-0408)

Some [teachers] proposition teachers-made assessments directly take sampling questions from DVD [question bank for students' tests] on computer. It is unfair to low social status students because they have few chances to practice such questions [relative to some students who high social status have many

times to practice such questions over and over due to their parents may print them from DVD]. (Itf14-0402)

Consequently, it [TWST] contributed the sharing and exchange among members [teachers] (Salomon, 1993). The negative opinions of some teachers appeared on campus when TWST prior to the implementation, and other teachers turned around gradually to the fair and positive opinions after the first test. Moreover, after the second test teachers started thinking about their professional attitudes which created the different thoughts and behaviors. That is to say, during the commencement of implementation for TWST, teachers presented cognitive systems such as negative opinions, fair and positive opinions, and thinking about their professional attitudes.

Initial Stage for TWST: Third and Fourth Tests

Although the foregoing self-professional issues most of the teachers faced, a few teachers adopted the perfunctory attitudes that written assignments they casually filled the checklist [TWST] in reaction to this policy in school. At the same situation, a few teachers thought such reviewing TWST for teachers-made assessments as a bad thing which violated the professional autonomy of teachers. And they insisted they had rich professional competences for completing the teaching tasks and thus didn't like to be suspected and challenged about their professional abilities (Hargreaves, 2001), as expressed by one teacher at third test: "I believe I have excellent teaching abilities, but for me such requirement [filled checklist of TWST] from school is a suspect behavior....., for us the perfunctory filling checklist means a resistant attitude to such school policy". (Itg14-0515)

In this stage a resistant attitude the teachers showed and their professional abilities were questioned. However, after reviewing the made TWST, some teachers found the situations that students' performances in tests centered on the lower cognitive level as remembering level and understanding level and only even in the single level (e.g., in remembering or understanding level). These highlighted the problems and myths about teachers' teaching strategies and processes. By such doing, teachers examined and

reflected such problems by employing the opportunities of sharing and discussing with others, and gradually found the more available teaching activities or strategies to resolve them. Namely, members (teachers) were able to create more intelligent, reflective, and applicative actions and outcomes when they deeply understood the related issues and contents (teachers-made assessments) (Meyer, 2007), as was explained by two teachers: “We find out our problems [teaching activities and focuses] in favor in level of remembering and understanding or only in remembering level after checking the checklist [TWST]” (Itc14-0513)

If we are able to discuss the details and content of TWST more often, we trust we could design more flexible cooperative learning activities for students and more suitable teaching materials for upgrading the students’ learning outcomes, such as upgrading the cognitive level from remembering to the analysis by the strategies of cooperating learning. (Itc14-0603)

Since practicing TWST at SiSi Elementary School many times, some teachers were gradually familiar with its connotations and significances. For example, several teachers could make distinctions for their assessments with each other on the basis of TWST and thus attract interesting and discussing in it for other teachers. Such artifact (i.e., TWST) as a vehicle for communicating understanding could advance the opinions for sharing and exchange among the members (i.e., teachers) so that they grew more intelligent (Hutchins, 1995; Salomon, 1993). As were explained by one teacher and the principal: “The all-inclusive contents of teachers- made assessments are showed according to the checklist (TWST), which include the all teaching units, all knowledge categories, and multiple cognitive levels”. (Ipp14-0626) “The opinions of sharing and discussing on TWST among teachers can distinguish the outcomes among students and provide a little strategies of remedial teaching for them”. (Ite14-0630)

Based on such situations, some teachers still adopted the attitudes of resistance and doubt in initial stage for TWST. And then some started to reflect the teaching questions at the same time after constant interacting and contacting with it, and even rose up the

more teachers to make differentiated for assessments. In other words, they could create more new knowledge due to the facilitation of the reciprocities and cooperation among teachers when the artifacts intervened (Liang, 2011). In sum, the attitudes of resistance and doubt, reflecting their teaching problems, and make differentiated for assessments teachers presented such cognitive systems in initial stage for TWST.

The Medial Stage for TWST: Fifth and Sixth Tests

Teachers had been regarded TWST as a part of teachers-made assessments after practicing the fourth tests, and they started to mind its connotations such as how material contents or units distributed, what types of questions adopted, and even how cognitive levels defined or distinguished. In fact, such difficulties of skills on TWST got more passive for some teachers, as was expressed by one teacher:

We [same grade teachers] have to discuss and compare the connotations and the types of questions about teachers-made assessments so that can understand the differences about cognitive levels on them.....; this is a routine in every student tests, but it is difficult for us. For other teachers the proportion of each material unit on the teachers-made assessments is emphasized than the cognitive levels. (Itc14-0919)

However, most teachers gradually considered and discussed the cognitive levels and the teaching materials on TWST. Namely, artifacts could promote their influence due to the functions of sharing and discussing about them among members (Louis & Marks, 1998). More specially, evaluating how cognitive levels of questions were diagnosed functions and adjusting how teaching activities practiced were needed when some teachers designed the teachers-made assessments for students' tests at school. And even a few teachers as same grade or same subject teachers got together to complete on each teachers-made assessment and shared and discussed the qualities of such assessments. As was said by one teacher:

I think I have been considered the various types and contents of questions on teachers-made assessments, for example, after filling repeat the table [TWST] type of essay as evaluation cognitive level and conceptual knowledge of third unit of mathematics material we show many times on assessments, especially through talking and discussing with other teachers..... (Ite14-1112)

Such this, the practice of TWST embodied the traits of “teachers as professional leaders” (Lima, 2008) over time, as was explained one meeting. “At a faculty meeting, Teacher F expresses he is gradually able to take into account the teaching actions to reach various cognitive levels (e.g., understanding and analysis) on teachers-made assessments” (D-04). In other words, it naturally stimulated the several teaching strategies for teachers. For example:

Some teachers feel more related to the understanding level from the perspectives of teaching subjects (e.g., calculation questions in mathematics), and relatively few stress the higher cognitive levels such as evaluation and creation deriving from adopting the posted teaching method....., we decide to change the flexible teaching strategies” .(Iti14-1219)

Thus, teachers started to adjust their teaching strategies and available tried to fill the gaps between their teaching and students’ learning due to TWST facilitating practice attitudes. Based on above-mentioned points, in the medial phase of implementation for TWST, the teachers of SiSi Elementary School firstly encountered the skills difficulties that they couldn’t define the cognitive levels about teachers-made assessments. Moreover, for promoting students’ learning outcomes the teachers shared and discussed with each other about the teachers-made assessments in order to start up the opportunities for facilitating the teaching strategies. In other words, teachers showed their agencies in the routine activities [TWST] by making relevant work programs and injecting new ideas (Chen, 2007). In brief, encountering skills difficulties, sharing and discussing with each other about assessments, and facilitating their teaching strategies were three cognitive

systems in this stage.

Final Stage for TWST: Seventh, Eighth, and Ninth Tests

For teachers TWST launched a chance for facilitating their teaching strategies due to gradually understanding the traits about various cognitive levels in terms of teachers-made assessments. Furthermore, in this case, it was necessary for expert teachers to help other teachers to get more external resources, during the transformational processes of their teaching strategies in order to understand how cognitive levels on such assessments were identified. And the teachers' teaching experiences also was the key factors for promoting such skills. Namely, the interaction among individual cognitions, individual abilities, collective culture, and collective resources in organization were also able to upgrade the organizational competitiveness (Küpers, 2007). As was explained by one observation note:

In a discussing meeting, Teacher E expresses the action on TWST to discuss with other teachers that can review the students' outcomes and understand the students' learning problems after the tests.....; she also suggests teachers adopt it [TWST] and further use the remedial teaching system [collective resources] in school to supplement the relevant knowledge categories or materials units for students..... (Ob15-0106)

In addition, during the implementation of TWST, experts' participating and giving professional opinions could contribute and refine the teachers' professional abilities and further promote the collective effectiveness in the organization (Kolikant, McKenna, & Yalvac, 2006). And constant discussing and sharing in teachers' professional learning communities also were to bring up their mutual understanding and develop the common knowledge (Cook & Yanow, 1993), as was expressed by one teacher and one document:

I use the TWST on assessments by way of discussing with other teachers, especially expert teacher [Teacher D] concerning the 'reading comprehension',

in ‘Curriculum Learning Community’ formed from the professors’ opinions, expert teachers’ views, and actual situations....., for me, the use of TWST is a professional development processes” (It15-0312).

Hiring professors to explain the importance of the TWST on the assessment for teachers, including the content and types of the questions, not only explain the meaning of each level, but also take one by one for example....., especially teachers how to edit the various types of questions. (D-7)

General speaking, most of the teachers had invested time in the investigation and development about their teaching practice with TWST; Relatively other teachers who had no more time to interact with it were also influenced. Especially, the pushing teachers’ professional development attitudes principals intended actively, which could further strengthen the function of artifact in order to contribute teachers’ actual teaching outcomes (Liang & Hung, 2011). Two teachers pinpointed his opinions: “Most of the teachers always discuss with same grade teachers about the problems on teachers-made assessments regarding TWST that can strengthen the [professional] abilities.....; in other words, this will put a lot of pressure on me”. (It15-0325) “The principal actively push TWST and continue discusses with teachers over and over....., that should help us have a clearer concept on it” (It15-0326).

Namely, organizational learning could deliver the artifact to staff members so as to they obtained collectively the technologies and abilities about it (Cook & Yanow, 1993). That is, in the past teachers only discussed the contents of TWST during the tests period, and recently they continually triangulated their teaching thoughts and actions with teaching practice and assessments of TWST as an artifact and kept on discussing and sharing with their team or communities members so that they promoted their teaching outcomes after the year and a half after implementation of TWST.

Surprisingly, the school had been researched and developed the information system to TWST that could give teaching reminder messages when the user [teachers] signed students’ scores of tests in this system, and even suggested the effective teaching topics

and strategies to teachers such as what the proportions of teaching material were provided and how cognitive levels were needed. Scilicet, the use of information technology could avoid the abuse of human and material resources, and made teachers-made assessments based on standard process improved their teaching and students' learning outcomes (Wang et al., 2003), as was explained by two teacher. "From system-reminder messages in this test I know the several problems on understanding [cognitive] level, so I decide to adopt the multiple instructional media to improve the students' understanding abilities". (It15-0410)

I found that the questions on assessments I edited that inclined to analyze level, but the students seem to be unable to get high scores. Therefore, I will especially emphasize adopting the teaching strategy for actual comparison, such as the use of experimental way to understand the different factors, there will be different effects, and let the students themselves to see. This may enhance the ability of students to analyze. (It15-0407)

For these reasons, the school integrated the human and material resources and adopted the expert opinions to break through the existing boundaries and limitations during the final stage of implementation of TWST. And the principal's pushing attitudes and teachers' triangulation used on their thoughts with teaching practice and assessments of TWST could help the teachers adopt more effective teaching activities. In sum, the integration of human and material resources, the needs of experts' opinions, principal's positive attitudes, and the triangulation used on teachers' thoughts with teaching practice and assessments of TWST were four cognitive systems in this stage.

Briefly, teachers' cognitive systems slightly inclined toward the rising trend in the year and a half after implementation of TWST (see Figure 4).

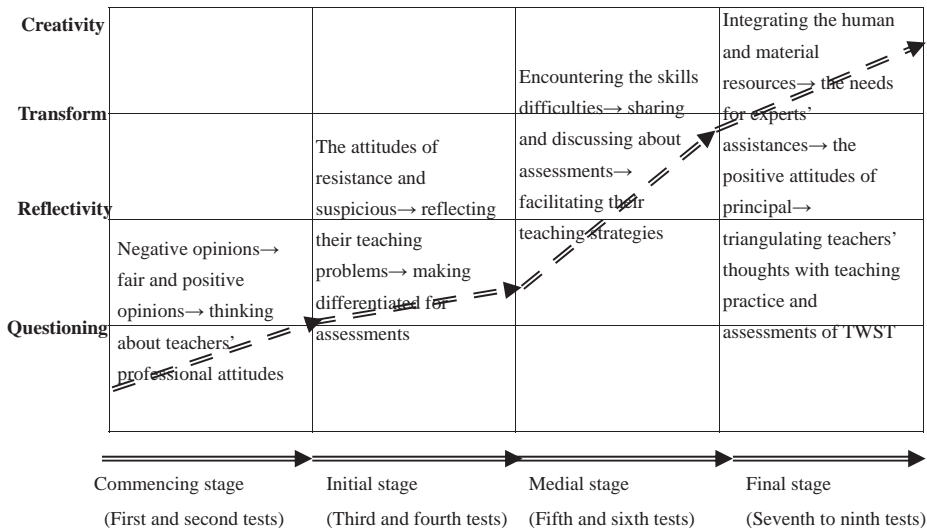


Figure 4. The Processes of Teachers’ Cognitive Systems for Implementation of TWST at SiSi Elementary School

Note. The arrows indicate the development direction of teachers’ cognitive systems at four stages.

Conclusions and Suggestions

The present study explored how the teachers’ cognition evolved with a close look at the artifact, the processes adopted by TWST on teachers-made assessments as an artifact at nine students’ tests in SiSi Elementary School. The following conclusions and suggestions were drawn:

After the Implementation of TWST in SiSi Elementary School, at Least Thirteen Types of Teachers’ Cognitive Systems were Identified

Based on the data at nine tests, after the implementation of TWST in SiSi Elementary School, thirteen types of teachers’ cognitive systems were found. Firstly, during the commencement of implementation for TWST, teachers presented cognitive systems such as negative opinions, fair and positive opinions, and thinking about their

professional attitudes. Secondly, the attitudes of resistance and suspicious, reflecting their teaching problems, and making differentiated for assessments teachers presented such cognitive systems in the initial stage on TWST. Thirdly, we analyzed the encountering the skills difficulties, sharing and discussing about assessments, and facilitating teachers' teaching strategies, which were three cognitive systems in the medial stage. Lastly, integrating the human and material resources, the needs for experts' assistances, developing the technological systems, and the positive attitudes of principals, and triangulating teachers' thoughts with teaching practice and assessments of TWST teachers presented in the final stage.

During the Process of Implementing TWST, Teachers Gradually Presented Different Stages of Cognitive Systems Such As Questioning, Reflection, Transforming, and Creativities

After the implementation of TWST, negative opinions, fair and positive opinions, and thinking about teachers' professional attitudes leaned slightly a stage of questioning. In the initial stage, the reflective significance was created from the attitudes of resistance and suspicion, reflecting their teaching problems, and making differentiated for assessments. Moreover, the transformational period in the medial stage represented teachers appeared encountering the skills difficulties, sharing and discussing about assessments, and facilitating their teaching strategies. Lastly, the creativity period was found at the final stage, which appeared from the integrating the human and material resources into triangulating teachers' thoughts with teaching practice and assessments of TWST.

Continuous Discussions and Sharing, the Needs for Experts' Assistances, Developing the Technological Systems, and the Positive Attitudes of Principal Were Key Elements in the Functioning of the TWST

In the light of the findings, teachers presented negative opinions or attitudes of resistance and suspicion in commencement and initial stages, and thus facilitating

teachers' teaching strategies or triangulating teachers' thoughts with teaching practice and assessments of TWST by way of the actions of continuous discussing and sharing when they contacted and encountered the skills difficulties concerning proportion of materials contents and topics of cognitive levels. Moreover, teachers understood the features about TWST, how the cognitive levels were discriminated and how students' outcomes were defined, due to experts' assistances. In addition, the technological systems, remedial teaching system and reminder- messages, were pulled into TWST for students' tests identified the various cognitive levels and knowledge categories of materials' units to stir up teachers adopting the available teaching strategies. Finally, principal's positive attitudes strengthened the functions of TWST and teachers' professional abilities in order to promote the teachers' teaching outcomes. In sum, the four factors were the critical factors for exerting the functions of TWST.

After the Implementation of TWST in SiSi Elementary School, Teachers' Cognitive Systems Were Slightly Moving Toward the Rising Trend

After implementation of TWST, owing to constantly interaction with TWST teachers gradually developed the thirteen cognitive systems and experienced the periods of questioning, reflectivity, transformation, and creativity, especially the critical factors of continuous discussions and sharing, experts' assistances, developing the technological systems, and principal's positive attitudes. In brief, such these cognitive systems were slightly moving toward the rising trend after implementation of TWST.

In result, we could point out the following suggestions: 1) The schools could use the existing artifacts such as checklists, teaching plans, remedial teaching system to facilitate the discussing and sharing among teachers in order to create more and more cognitive systems development for them. 2) Principals can promote the teachers' professional development by way of encouraging the discussing and sharing among teachers, grasping the experts' assistances, pulling into technological systems, and actively pushing attitudes. 3) Continuously pushing relevant artifacts based on the educational values are required that could extend and promote the educational or teaching outcomes.

Reference

- Allison, D. J. (1996). Problem finding, classification and interpretation: In search of a theory of administrative problem processing. In K. A. Leithwood, J. Chapman, D. Corson, P. Hallinger & A. Hart (Eds.), *International handbook of educational leadership and administration* (pp. 477-549). Dordrecht, Netherlands: Kluwer.
- Anderson, W., & Krathwohl, D. R. (2001). *Taxonomy for learning, teaching and assessing: A revision of Blooms' educational objectives*. New York, NY: Longman.
- Catano, N., & Stronge, J. H. (2006). What are principals expected to do? Congruence between principal evaluation and performance standard. *NASSP Bulletin*, 9(3), 221-237.
- Chen, H. J. (2007). The dialectical relationship between organizational structure and teachers' agency: A case study in a Taiwanese primary school. *Changhua University of Education: Journal of Education*, 12, 103-128.
- Chen, P. Y., & Chiao, C. C. (2009). Distributed leadership and professional learning community: A case study of innovative teaching in a high school. *Journal of Research in Education science*, 54(1), 55-86.
- Cheng, Y. L. (2006). The Establishment of Chinese General Proficiency Indicator. *Journal of Liberal Arts and Social Sciences*, 2, 115-136.
- Chien, F. M. (2009). A study on teaching assessment develops and designs for social studies in elementary school. *Social Studies of Education*, 14, 129-149.
- Cook, S. D., & Yanow, D. (1993). Culture and organizational learning. *Journal of Management Inquiry*, 2(4), 430-459.
- Firestone, W. A., & Riehl, C. (2005). Conclusion. In W. A. Firestone & C. Riehl (Eds.), *A new agenda for research in educational leadership* (pp. 171-184). New York, NY: Teachers College.
- Gagliardi, P. (1990). *Symbols and artifacts: Views of the corporate landscape*. New York, NY: Aldine de Gruyter.
- Gan, Y., & Zhu, Z. (2007). A learning framework for knowledge building and collective wisdom advancement in virtual learning communities. *Educational Technology &*

Society, 10(1), 206-226.

Hands, C. M. (2010). Why collaborate? The differing reasons for secondary school educations' establishment of school-community partnerships. *School Effectiveness and School Improvement*, 21(2), 189-207.

Halverson, R. R., & Clifford, M. A. (2006). Evaluation in the wild: A distributed cognition perspective on teacher assessment. *Educational Administration Quarterly*, 42(4), 578-619.

Hargreaves, A. (2001). The emotional geographies of teachers' relations with colleagues. *International Journal of Educational Research*, 35, 503-527.

Hutchins, E. (1990). The social organization of distributed cognition. In L. B. Resnick, J. M. Levine & S. D. Teasey (Eds.), *Perspectives on socially shared cognition* (pp. 283-307). Washington, DC: American Psychological Association.

Hutchins, E. (1995). How a cockpit remembers its speed. *Cognitive Science*, 19, 265-288.

Kolikant, Y. B., McKenna, A., & Yalvac, B. (2006). The emergence of a community of practice in engineer education. *New Directions for Teaching and Learning*, 108, 7-16.

Küpers, W. M. (2007). Phenomenology and integral pheno-practice of wisdom in leadership and organization. *Social Epistemology*, 21(2), 169-193.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212-218.

Lee, K. C. (2011). The designs and examples on "Two-Way Specification Table" of paper tests. *Journal of Education Research*, 208, 95-109.

Liang, C. T. (2011). The development and formation of a school's collective wisdom in an elementary school: From the perspective of distributed cognition. *Chung Cheng Educational studies*, 10(2), 97-139.

Liang, C. T., & Hung, C. C. (2011). Analysis of cognitive process of a principal's instructional leadership from a distributed cognition perspective. *Contemporary Educational Research Quarterly*, 19(2), 41-80.

Liang, C. T., & Lin, M. D. (2013). An analysis of the developmental process of school's

- collective wisdom from the perspective of organization learning: Taking “the plan of reading brand certification” as an investigative focal point. *Educational policy Forum*, 16(1), 103-137. DOI:10.3966/156082982013021601004
- Lieberman, A., & Grolnick, M. (1996). Networks and reform in American education. *Teachers College Record*, 98(1), 7-45.
- Lima, J. A. (2008). Department networks and distributed leadership in schools. *School Leadership and Management*, 28(2), 159-187.
- Lin, M. D., & Liang, C. T. (2014). An analysis of distributed leadership practice in an elementary school. *Educational Review*, 43, 1-43. DOI: 10.3966/156335272014120043001
- Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 106(4), 532-575.
- Lu, I. J., & Wu, G. I. (2011). “Two-Way Specification Table” as mathematics. *In-service Education Bulletin*, 28(5), 97-104.
- Meyer, A. D. (2007). Strategic epistemology: Innovation and organizational wisdom. In E. H. Kessler & J. R. Bailey (Eds.), *Handbook of organizational and managerial wisdom* (pp. 357-374). Los Angeles, CA: Sage.
- Pea, R. D. (1993). Distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognition: Psychological and educational considerations* (pp. 47-87). New York, NY: Cambridge University.
- Perkins, D. N. (1993). Person-plus: A distributed view of thinking and learning. In G. Salomon (Ed.), *Distributed cognition: Psychological and educational considerations* (pp. 88-110). New York, NY: Cambridge University.
- Salomon, G. (1993). No distributed without individuals' cognition. In G. Salomon (Ed.), *Distributed cognition: Psychological and educational considerations* (pp.111-138). New York, NY: Cambridge University.
- Schein, E. H. (1985). *Organizational culture and leadership*. San Francisco, CA: Jossey-Bass.
- Seaman, M. (2011). Bloom's taxonomy: Its evolution, revision, and use in the field of

- education. *Curriculum and Teaching Dialogue*, 13(1 & 2), 29-43.
- Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher*, 30(3), 23-28.
- Spillane, J. P., Halverson, R., & Diamond, J. B. (2004). Towards a theory of leadership practice: A distributed perspective. *Journal of Curriculum Studies*, 30(1), 3-34.
- Spindler, M., & Tech, V. (2015). Collaborative analysis and revision of learning objects. *North American Colleges and Teachers Agriculture Journal*, 59(2), 111-115.
- van der Meij, H., & Boersma, K. (2002). Email use in elementary school: An analysis of exchange patterns and content. *British Journal of Educational Technology*, 33(2), 189-200.
- Wang, T. H., Wang, W. L., Wang, K. H., & Huang, S. C. (2003). Study a web-based assessment system as teacher's training program evaluation technique. *Journal of Internet Technology*, 4(2), 71-78.
- Yen, L. C., & Lin, S. P. (2003). The explore on revised edition of Blooms' cognitive domain categories. *Journal of Education Research*, 105, 94-106.
- Yin, R. K. (2002). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

徵稿啟事

2005.1.18 編輯委員會會議通過
 2006.2.13 編輯委員會會議修正通過
 2006.5.15 編輯委員會會議修正通過
 2006.8.11 編輯委員會會議修正通過
 2007.7.13 編輯委員會會議通過
 2009.8.10 編輯委員會會議修正通過
 2011.2.10 編輯委員會會議修正通過
 2012.8.16 編輯委員會會議修正通過
 2016.3.21 編輯委員會會議修正通過

《教育研究與發展期刊》(Journal of Educational Research and Development) 為國家教育研究院發行之教育學術刊物，著重在教育領域內之研究與發展 (R&D) 方面相關議題，並推廣國內外教育學術研究與交流。所有稿件之徵、審稿辦法比照科技部人文社會科學研究中心「臺灣社會科學引文索引」資料庫 (TSSCI) 之相關規定辦理，歡迎踴躍賜稿。

● 徵稿事項

1. 本刊為季刊，全年徵稿，徵稿主題如下：

徵稿主題	出刊日期
教師培育與專業發展	3 月 31 日
課程與教學	6 月 30 日
教育政策與制度 (含教育行政、學校行政等)	9 月 30 日
教育心理、輔導與測評	12 月 31 日

2. 所有稿件皆隨到隨審，原則上，編輯部於收稿後五個月內處理完成並告知作者「刊登與否」。

● 投稿原則

1. 請以電腦打字，中英文不拘，中文撰稿文長以 15,000 字為原則，至多為 20,000 字（含中英文摘要、註釋、參考書目、附錄、圖表等），經審查通過之修正文稿字數不得超過 22,000 字。中文摘要請勿超過 500 字，英文摘要不超過 300 字，並列出中英文關鍵字至少各 3~5 個。如以英文撰稿，請勿超過 12,000 字。

2. 來稿請使用線上投稿系統，請有意投稿者至網址 <http://journal.naer.edu.tw> 註冊並上傳中英文摘要及全文電子檔（Word 或 PDF 格式），需詳細填列共同作者、服務機關、最高學歷、專長領域、聯絡電話、電子郵件等資訊，俾利編輯部確認身份。
3. 投稿正文及中英文摘要中，請勿出現任何個人資料。
4. 來稿文件之註釋（採當頁註方式）及參考書目，請用 APA 格式【第六版】。詳細規範可至本期刊網站下載說明文件，或來函（請附回郵信封）至編輯部索取。
5. 未依本刊所要求之格式來稿，本刊將逕予退稿。
6. 來稿如未獲採用，本刊將致函作者審查結果，但不寄還稿件，請投稿者自留原稿。
7. 本刊因編輯需要，保有必要之文字刪修權。
8. 兩名作者以上之稿件，應分別列明各人之貢獻。
9. 單一作者單期投稿並不限制 1 篇，但如當期該作者已通過審查之作品 2 篇以上，由編輯委員會議決定擇優刊登、刊登篇數及錄用期數；除當期選錄刊登作品外，告知作者通過作品預定刊登之期數並徵得其同意，刊登原則為一年之內。

● 著作財產權事宜

1. 為維學術倫理，請勿一稿多投，如有抄襲，改寫等侵犯他人著作權之情況者，由作者自負相關法律責任。
2. 本刊授權方式為非專屬授權（Non-exclusive License）予出版單位，來稿一經刊登，需於期限內簽署著作授權利用書掛號回寄編輯部，本刊將敬贈作者當期刊物 6 冊，不另支稿酬。

● 稿件審查

1. 本刊所有稿件採雙向匿名內外審查制度，由本刊編輯委員聘請相關領域學者專家二人審查之。經審查委員要求修改之文章，於作者修改之後再由編輯委員決定是否刊登。
2. 凡本刊接受刊登之稿件，得視編輯需要，經編輯委員會同意後，擇期刊登。

● 稿件交寄

來稿請使用「線上投稿系統」（<http://journal.naer.edu.tw>）註冊作者詳細資料並上傳作品電子檔案。

Call for Papers

Journal of Educational Research and Development (JERD) is the official periodical of National Academy for Educational Research, the national level for educational research institute in Taiwan. JERD provides a forum to discuss the issues regarding perspectives of educational research and development (R&D) to convey a brand new worldwide vision.

Publishing Schedule (2017~2018)

Subjects	Date of publication
Teacher Education and Empowerment	31, March
Curriculum and Instruction	30, June
Educational Policy and Administration	30, September
Educational Psychology, Counseling, Testing & Assessment	31, December

Information for authors

1. Manuscripts must be word processed and double spaced on A4 (210 x 297 mm) size paper with margins at least 1inch on all sides.
2. The suggested maximum length of manuscripts is 12,000 words including an 300-word abstract, texts, tables, footnotes, appendixes, and references. 3-5 keywords or key phrases are required.
3. The manuscript should be in MS Word Format or Portable Document Format (PDF), and the file size is limited to 5MB. Please go to http://journal.naer.edu.tw/contribution_login.asp to register to the Online Submission and Peer Review system and upload your manuscript.
4. All submissions should follow standard APA style (6th Edition).
5. The manuscript should be original, and has not been published previously. Do not submit material that is currently being considered by another journal.
6. Authors of the articles being accepted are required to sign the Transfer of Copyright Agreement form.
7. Author of the manuscript should fill a submission form.
8. Author will receive 6 copies of the issue of the journal containing their article.

If more details regarding JERD are needed, please contact:
jerd@mail.naer.edu.tw

- **Review Procedures**

1. The manuscripts not meeting the requirement will not be reviewed, or accepted and will be returned for modification.
2. The author should not place his/her name on any of the manuscript pages to ensure anonymity during the review procedure.
3. All paper will be submitted for anonymous peer review by domain experts.
4. Acceptance, revision, or rejection letter will be mailed within 5 months after the manuscript is received.

Enclose three copies of the manuscript. Please send all manuscripts, the compact disk of the text files and the submission form to the following address by registered mail:

Editorial Board, Journal of Educational Research and Development
National Academy for Educational Research
No.2, Sanshu Rd., Sanxia Dist., New Taipei City 23703, Taiwan (R.O.C)

審稿辦法

2005.1.18 編輯委員會會議通過

2006.4.17 編輯委員會會議通過

2010.2.8 編輯委員會會議修正通過

2011.2.10 編輯委員會會議修正通過

2012.8.16 編輯委員會會議修正通過

2014.8.25 編輯委員會會議修正通過

● 壹、審稿流程

本刊之審查包括預審、初審、複審。

一、預審

主編就來稿之性質、格式、體例及嚴謹程度進行審查。

二、初審

1. 通過預審之文章由編輯委員會聘請兩位審查人進行匿名審查。

2. 初審意見分為五類：

- (1) 極力推薦採用 (90 分以上)、(2) 推薦採用 (80-89 分)、(3) 修正後不必再送原審者審查 (75-79 分)、(4) 修正後再送原審者審查 (70-74 分)、(5) 不予採用 (69 分以下)。

分數達 75 分以上者列為候選刊登名單，並提經編輯委員會會議議決刊登。

3. 若兩位審查人分數相差過大時，且其中一位分數達 75 分以上者，應送第三位審查人審查，本刊將依據第三位審查人之意見決定是否刊登。

4. 兩位審查人分數皆 74 分以下者，予以退稿。

三、複審

1. 若審查人建議為「修正後不必再送原審查者審查」及「修正後再送原審查者審查」之文章，本刊將請作者修改，作者須於二周內寄回，並隨文附上「修改、答辯相關說明」，本刊將把修改之稿件及此說明文件交由主編或原審查人進行複審；本刊將根據複審意見提經編輯委員會會議議決刊登。

2. 所有通過複審之論文列為候選名單，並交由編輯委員會會議討論收錄之期數。

四、審查迴避

1. 本刊之編輯委員會成員及國家教育研究院現職人員以作者身分投稿，不得參與審稿流程（預審、初審及複審）。

2. 本刊當期主編以作者身分投稿，由總編輯指定代理人進行預審。

3. 投稿作者未經編輯委員會主席同意，不得參與擬收錄文稿之討論事項。

4. 編輯委員會聘請審查人時，應考量專長之符合性及研究表現優良者，宜避免審查人與作者有下列關係：
 - (1) 近三年曾任職同一系、所、科或單位。
 - (2) 近三年曾有指導博士、碩士論文之師生關係。
 已獲聘請之審查人，如自行發現與該文作者有以下利害關係、宜予迴避，請速與本刊編輯人員聯絡：
 - (1) 近二年發表論文或研究成果之共同作者。
 - (2) 審查論文時有共同執行研究計畫。
 - (3) 配偶或三親等以內之血親或姻親。
 - (4) 與該文有利益衝突之可能。

審查迴避事項如有疑義，由編輯委員會議決。

四、其他事項

1. 當期主題收錄以不超過 8 篇為原則，另收錄其它主題至少 2 篇，以平衡部分作者等待刊登時間略長之困擾。
2. 等待刊登時間超過半年之論文，列為優先刊登名單。
3. 「審查迴避名單」可由作者提出 2~3 名作為主編預審推薦參考。
4. 同一作者（包含共同作者），其作品以一年刊登一次為原則；特殊情況則於編輯委員會議上提案討論。
5. 於正式出版前：
 - (1) 如發現違反學術倫理情況，由編輯委員會召開會議共同商議處理方式；
 - (2) 如論文存有限期內無法改善之問題，授權由該期主編決定處理方式。
6. 為確保作者權益，編輯部聯絡審查人時明確告知：審查回件期限以一個月回函為原則；若回件時間逾一周以上，即報告主編並另行推薦審查人。

● 貳、稿件刊登

經審查為「極力推薦採用」、「推薦採用」、「修正後不必再送原審查者審查」及原審查人複審通過之稿件，將提請編輯委員會進行刊登確認，通過後本刊將寄發「接受刊登證明」及「出版同意授權書」，以利文章刊登出版。

● 參、撤稿作業原則

- 一、作者應以書面掛號方式，提出撤稿申請。
- 二、對凡已進入初審階段之稿件，若作者提出撤稿申請，本刊一年內不接受投稿。

《教育研究與發展期刊》投稿者基本資料表

Journal of Educational Research and Development Submission Form

姓名 Author (s)	中文： 英文：	投稿日期 Submission date	
投稿篇名 Title	中文： 英文：		
擬投稿之主題 Category of submission	<input type="checkbox"/> 教師培育與專業發展 (Teacher Education and Empowerment) <input type="checkbox"/> 課程與教學 (Curriculum and Instruction) <input type="checkbox"/> 教育政策制度 (Educational Policy and Institution) <input type="checkbox"/> 教育心理、輔導與測評 (Educational Psychology, Counseling, Testing & Assessment) <input type="checkbox"/> 其他 (Others) _____		
稿件字數 Word count	全文共_____字 (含中英文摘要、正文、參考書目、附錄、圖表等) (Abstract, text, references, appendixes, tables are included)		
服務單位 及職稱 Affiliation & Position			
最高學歷 Highest Degree			
專長領域 Specialization			
通訊地址 Address			
聯絡電話 Telephone	(O) (H) (M)		
電子郵件 E-Mail			

國家教育研究院期刊雜誌著作授權利用書

作者（即撰稿人）於《教育研究與發展期刊》第 卷第 期所發表之

論文：_____

本人同意下列所載事項：

- 一、擔保本著作有授權利用之權利，且本篇論文保證未曾發表或出版，
本著作並無不法侵害他人著作權或其他權利之情事。
- 二、同意全部內容授權國家教育研究院作無期限、地域、方式、性質、
次數等限制之利用，並得授權第三人利用。
- 三、國家教育研究院得於不破壞著作原意之範圍內修改稿件。
- 四、同意對國家教育研究院不行使著作人格權。

立授權書人（作者）：_____（簽名）

身分證字號：_____

戶籍地址：_____

聯絡電話：_____

Email：_____

年 月 日

國家教育研究院

Transfer of Copyright Agreement

The copyright of this article is transferred to National Academy for Educational Research (NAER, Taiwan, R.O.C.), effective if and when the article is accepted for publication in the *Journal of Educational Research and Development* sponsored by NAER.

Title of the article: _____

Author(s): _____

1. The copyright transfer covers the exclusive right to reproduce and distribute, including reprints, translations, photocopies, electronic reproductions, and other forms of electronic publication.
2. The signed Agreement ensures that the NAER has the author's permission to modify and publish the article.
3. The Author warrants that the article is original and has not been published before.
4. The Author ensures that co-authors are informed of the terms of the copyright transfer and signs on their behalf.
5. The Author retains the rights to use all or part of this article for his/her own works, such as books, lectures, websites or teaching materials without the need to obtain further permission.

Signed: _____ Name printed: _____

Date: _____

教育研究與發展期刊

第十三卷・第四期 2017年12月31日出刊／創刊日期：2005年6月

出版者：國家教育研究院

總編輯：許添明

副總主編：郭工賓

本期執行主編：林世華

本期品質促進小組委員：王文中、邱皓政、翁儷禎、陳秉華、陳學志、溫福星

編輯委員：任宗浩、余民寧、洪儷瑜、張郁雯

執行編輯：張雲龍、朱麒華

助理編輯：陳璵哲

地址：23703 新北市三峽區三樹路 2 號

電話：(02)7740-7187

傳真：(02)7740-7184

網址：<http://www.naer.edu.tw/>；<http://journal.naer.edu.tw/>

本期刊載文章電子檔案同步於 <http://journal.naer.edu.tw/> 下載

排版印刷：財政部印刷廠

電話：(04)2495-3126

展售：政府出版品展售中心

五南文化廣場：臺中市中山路 6 號

TEL: 04-22260330；FAX: 04-22258234

國家書店松江門市：臺北市松江路 209 號 1 樓

TEL: 02-25180207；FAX: 02-25180778

網址：<http://www.govbooks.com.tw/>

定價：每期新臺幣 250 元

季刊：每年 3 月、6 月、9 月、12 月出版

版權所有・翻印必究

Journal of Educational Research and Development

Vol.13 No.4 December 31, 2017

Date Founded: June 30, 2005

Publisher : National Academy for Educational Research

Editor in Chief : Tian-Ming Sheu

Vice Editor in Chief : Kung-Bin Guo

Executive Editor in Chief : Sieh-Hwa Lin

Quality Improvement Team : Wen-Chung Wang / Haw-Jeng Chiou / Li-Jen Weng /
Hsueh-Chih Chen / Ping-Hwa Chen / Fur-Hsing Wen

Editorial Board : Tsung-Hau Ren / Min-Ning Yu / Li-Yu Hung / Yu-Wen Chang

Executive Editor : Weng-Long Chang / Chyi-Hwa Chu

Assistant Editor : Nancy Chen

Address : No.2, Sanshu Rd., Sanxia Dist., New Taipei City 23703, Taiwan (R.O.C)

Price : NT\$250 (for each copy)

Copyright@2017

National Academy for Educational Research



GNP 2009405238