

國立政治大學教育學系
教育心理與輔導組博士論文

指導教授：余民寧 博士

中介效果的時序設計與因果推論

研究生：洪兆祥

中華民國一〇七年七月

謝辭

七年前的仲夏，剛退伍的我總算在二度挑戰時成功，拿到政大教育系博士班的入場券，懷著忐忑不安又有些驕傲的心情步入井塘樓。而今，這裡是除了家裡之外，最感熟悉的地方。

這本論文，於我不僅是智識上的挑戰，更是認識自己的最佳途徑。在研究主題上，我是個慢熱的人，花了近五年的時間，在各個課程、研究及教學的過程中，才逐漸浮出一個清晰的輪廓。為此，特別要感謝諸位造就我的師長。感謝指導教授余民寧老師，在我詞不達意的報告中，總能率先看出這份論文的價值，並支持我持續向前。感謝口試委員游森期老師、吳俊育老師給予的寶貴意見，指出本文的不足及修正方向。感謝胡悅倫老師，在這幾年的合作中，不只給予各種機會與訓練，更以身教展示做人處世之道。感謝邱皓政老師，從踏入研究所至今的十年，在近人生 1/3 的時光中，無論研究或教學，讓我持續前進至此的都是在您課堂中打下的基礎。感謝詹志禹老師，那一個學期的科學哲學，徹底打開了我狹隘的視野，引發了我對於科學本質更深層的興趣。感謝徐聯恩老師在寫作上的指導之外，能在創新中心一展長才也是重要的體驗。

研究與教學的路上，偶爾有風有雨，好在我有一群很神的隊友相挺。最明白我的柏霖，總是給我各種機會以及最大的容忍，督促我不要太廢。學長姊小葉、Greg、菁芬、March、金香，一起畢業的革命同志玉樺、忠蔚，研究室開心夥伴佩欣、師瑤、靜軒，好同學經維、智修，ABC lab 的同學師長們，陳嘉成院長、曾敏珍主任、林怡弟老師、王素芸老師、陳婉真老師，政大的好山好水，以及所有陪伴我的親友們，感謝您們的支持與鼓勵。

七年的馬拉松終於要抵達終點，在背後支持我的最大力量，是始終相信我、以我為傲的爸媽，體貼的伴侶欣怡，讓我能無後顧之憂地全心投入學術工作。這本論文獻給在天上看顧著我的爺爺奶奶們，我做到了！

洪兆祥

2018 大暑

中介效果的時序設計與因果推論

摘要

中介模型的本質在於解釋研究變數之間的因果傳遞機制，乃是當代社會科學領域中最常被使用的研究架構。然而大多數研究使用的橫斷中介設計，缺乏時間因素將難以進行因果推論的判斷，嚴重危害研究結果的正確性。本研究旨在探討中介模型如何進行完備的因果推論。其研究目的之一：探究中介模型的時序設計與因果推論的關係。其研究目的之二；探討現存的中介模型在因果推論上的完備條件。其研究目的之三：提出增強因果推論完備條件的可能策略為何。

本研究共有兩個子研究：研究一以 443 位國三學生為對象，採用五次模擬考成績為效標變數，並以「數學焦慮量表」、「數學學習動機量表」為工具，進行為期一年的追蹤測量。經描述統計分析、相關分析、路徑分析與中介檢定，結果顯示：

1. 在橫斷中介模型當中，數學焦慮為原因的因果推論較具完備條件。
2. 在序列中介模型當中，數學焦慮為原因的因果推論較具完備條件。
3. 在縱貫中介模型當中，數學焦慮為原因的因果推論，與數學學習動機為原因的因果推論，兩者的完備條件相當。

研究二以台灣青少年成長計畫資料庫當中的 1529 位國一學生為對象，分別選取親子互動、自尊與憂鬱做為研究變數，在八年之間時進行三次追蹤測量。經描述統計分析、相關分析、路徑分析與中介檢定，結果顯示：

1. 在橫斷中介當中，自尊的中介效果達顯著水準。
2. 在序列中介當中，自尊的中介效果達顯著水準。
3. 在縱貫中介當中，自尊的中介效果未達顯著水準。

本研究根據資料分析結果進行討論，並對未來研究提出建議。

關鍵詞：中介模型、時序設計、因果推論

Sequence Design and Causal Inference of Mediation Effect

Abstract

The purpose of mediation model is to explain the causal transfer mechanism among variables and it is a commonly used research framework in social science. However, most of the studies adapted cross sectional mediation analysis, which lacked sufficient time in making causal inferences. Such a practice may lead to a severe inaccuracy in data interpretation.

Therefore, this study aims to investigate how mediation models can achieve comprehensive causal inferences. The first research purpose is to examine the relationship between sequence design and causal inference of mediation models. The second research purpose is to investigate the comprehensiveness of the causal inferences in the existing mediation models. The third research purpose is to discuss possible strategies that can be used to enhance causal inference.

There are two research conducted in this study. For the first research, the participants were 443 ninth graders underwent a one year panel measurement. Five mock exam grades were used as the criterion variable, and the anxiety of mathematic exam and the motivation of learning mathematic were the research variables. With descriptive analysis, correlational analysis, path analysis, and mediation test, the results are as follow:

1. In the cross-sectional mediation model, it is more adequate to assume the anxiety of mathematic exam as the reason from causal inference. 2. In the sequential mediation mode, it is more adequate to assume the anxiety of mathematic exam as the reason from causal inference. 3. In the longitudinal mediation model, the anxiety of mathematic exam and the motivation of learning mathematic as the reason from causal inference are equally adequate.

The second research included 1529 seventh graders from TYP databased. Parent-child interaction, self-esteem and depression were used as the variables, and three panel measurements were taken over eight years. With descriptive analysis, correlational analysis, path analysis, and mediation analysis, the results are as follow:

1. In the cross-sectional mediation model, the mediation effect of self-esteem was significant. 2. In the sequential mediation mode, the mediation effect of self-esteem was significant. 3. In the longitudinal mediation model, the mediation effect of self-esteem was not significant.

The results and findings are discussed and implications for future research are also suggested.

Keywords: mediation effect, sequence design, causal inference

目次

第一章 緒論	1
第一節 研究動機.....	1
第二節 研究目的.....	3
第三節 名詞解釋.....	3
第二章 文獻探討	5
第一節 因果推論的邏輯與影響因素.....	5
第二節 中介模型的原理及其應用.....	12
第三節 縱貫設計的中介模型.....	18
第四節 自我迴歸取向的中介模型.....	30
第五節 強化中介模型因果推論的可能策略.....	34
第三章 研究方法	37
第一節 研究一資料說明.....	37
第二節 研究二資料說明.....	39
第三節 分析模型.....	41
第四章 結果與討論	45
第一節 研究一：國三數學科模擬考資料分析結果.....	45
第二節 研究二：臺灣青少年成長歷程資料庫分析結果.....	65
第三節 綜合討論.....	74
第五章 結論與建議	77
第一節 結論.....	77
第二節 研究限制.....	79
第三節 研究建議.....	80
參考書目	83

表目次

表 2-1-1 Mill 所提五種歸納法的整理摘要表.....	7
表 2-3-1 第三變項 Z 影響方式示意圖.....	14
表 2-4-1 潛在成長模式之截距與斜率變數在中介模型中的角色.....	29
表 4-1-1 全體樣本在研究變項上之描述統計 (N=987).....	47
表 4-1-2 低數學成就樣本在研究變項上之描述統計 (N=443).....	48
表 4-1-3 研究一各波次橫斷模型 A 中介分析.....	49
表 4-1-4 研究一各波次橫斷模型 B 中介分析.....	51
表 4-1-5 研究一橫斷中介的因果推論完備條件檢核表.....	52
表 4-1-6 研究一各啟始點序列模型 A 中介分析.....	53
表 4-1-7 研究一各啟始點序列模型 B 中介分析.....	54
表 4-1-8 研究一序列中介的因果推論完備條件檢核表.....	56
表 4-1-9 研究一縱貫模型 A 中介分析.....	60
表 4-1-10 研究一縱貫模型 B 中介分析.....	63
表 4-1-11 研究一縱貫中介的因果推論完備條件檢核表.....	64
表 4-2-1 臺灣青少年計畫研究變項之描述統計 (N=1529).....	66
表 4-2-1 各波次橫斷模型中介分析.....	67
表 4-2-3 研究二橫斷中介的因果推論完備條件檢核表.....	67
表 4-2-4 研究二序列模型中介分析.....	68
表 4-2-5 序列中介的因果推論完備條件檢核表.....	70
表 4-2-6 研究二縱貫模型中介分析.....	72
表 4-2-7 研究二縱貫中介的因果推論完備條件檢核表.....	74

圖目次

圖 2-3-1 三波追蹤資料之序列中介概念圖	21
圖 2-3-2 四波重複測量的自我相關概念圖	22
圖 2-3-3 四波重複測量的自我迴歸概念圖	23
圖 2-3-4 四波重複測量的交叉延宕模式概念圖	24
圖 2-3-5 五波重複測量的潛在成長模式概念圖	26
圖 2-4-1 二波追蹤資料之自我迴歸取向縱貫中介概念圖	31
圖 2-4-2 三波追蹤資料之自我迴歸取向縱貫中介概念圖	32
圖 2-5-1 五波追蹤資料之自我迴歸取向縱貫中介概念圖	34
圖 3-3-1 單向 ARMe5 模型 (僅有縱貫中介)	41
圖 4-1-1 研究二縱貫模型 A 中介架構概念圖	57
圖 4-1-2 研究一縱貫模型 B 中介架構概念圖	58
圖 4-1-3 研究一縱貫模型 A 中介分析結果圖	60
圖 4-1-4 研究一縱貫模型 B 中介分析結果圖	62
圖 4-2-1 研究二縱貫模型中介架構概念圖	71
圖 4-2-2 研究二縱貫模型中介分析結果圖	73

第一章 緒論

本章從研究動機、研究目的及名詞解釋等三部分進行陳述，旨在說明本研究探討縱貫中介模型與因果推論的動機，並透過感興趣議題與方向擬定研究目的，最後針對重要名詞涵義進行解釋。

第一節 研究動機

因果關係 (causality) 的探求乃科學典範的終極追求，千百年來無數聰慧的心智貢獻一生於此 (Pearl, 2000)。在科學典範越來越重視統計方法學 (statistical methodology) 的趨勢之下，大量的數據分析結果被用以進行重要的決策並影響著世界運行與人們的日常生活。然而良好的決策品質是建立在正確的分析架構及高品質的測量資料的基礎之上。科學典範的進步也體現於這些條件的不斷精進，此乃每一位學人應念茲在茲之事。

縱貫面設計 (longitudinal design) 在方法學當中一直被視為修正橫斷面設計 (cross-section design) 諸多不足的最佳解決方案。雖然二者間各有優缺點，但學界已經開始重視增加縱貫面設計的應用以促進科學發展。是故，許多處理縱貫資料的高階方法在近年來蓬勃發展，在心理學、教育學及管理學領域中，尤以結構方程式模型 (structural equation modeling, SEM) 與多層次模型 (multilevel modeling, MLM) 為討論重點 (邱皓政, 2017)。在這二種取向中，最重要的交集點莫過於中介模型的探討。

中介模型 (mediation model) 在社會科學領域中是應用度極高的一種研究設計，旨在探討預測變項 X、中介變項 M 及依變項 Y 三者間的單向關係，其本質與目的乃因果推論 (causality inference)。然而，在因果推論的過程中，許多使用者經常誤以為中介變數本身乃理論的核心。探究其本質而言，「時間效果」才是最重要的核心。 $X \rightarrow M \rightarrow Y$ 這樣的結構是一個普通的中介，本身並不包含時間因素在其中，這樣的橫斷面方法是缺乏正確性的。目前，絕大多數的中介模型應用皆以橫斷面設計進行，導致眾多學者對其因果推論有所詬病。為增強因果推論，

最根本的改進方式乃置入時間向度，也就是從橫斷面設計改為縱貫面設計。在處理縱貫中介的資料上有許多種方法足以勝任，包含序列設計（sequential design）、潛在成長模型（latent growth model, LGM）、自我迴歸模型（autoregressive model, AR）等方法。中介的概念是揭露社會科學動態過程較佳的一種徑路。中介變項本身扮演著回答不可見層次問題的重要角色。換言之，中介模型不僅描述現象，更做到解釋的層次。然而，在解釋的層次上，倘若係基於錯誤的理解或應用，則越多的研究也將帶來更多的誤解。故探討橫斷面設計與縱貫面設計二者在中介架構及因果推論上的差異，乃本研究動機之一。

在縱貫中介的各種理論模型當中，從統計控制與因果推論的角度觀之，目前學界以基於自我迴歸取向，涵蓋三個變項、三波測量的完整縱貫設計（又稱九宮格設計）為最佳模型。由於完整縱貫設計較全面地考慮測量時點、延宕期數與自我迴歸的控制，在估計特定時間點效果上具有較佳的統計推論程度。然而，Mitchell 與 Maxwell（2013, pp. 329）在文章的最後提到「下三角設計可做為完整縱貫設計的一種替代方案用以檢驗縱貫中介效果」，這段話激起研究者的一個想法：其一，為了估計出一條橫跨時間的中介路徑，所需要與不需要的資訊是什麼？其二，吾人發現一條顯著的縱貫中介的路徑，這就足以進行因果的推論嗎？故探討縱貫中介模型當中的因果推論是否完備、如何更完備，乃本研究動機之二。

第二節 研究目的

根據研究動機，本研究之目的可以盧列如下：

一、探究中介模型的時序設計與因果推論的關係

因果關係是人類解釋事物運作機制的一種想像，其存在與否是一個可否證的假設。因果關係的確立，應是一系列基於邏輯、證據的細緻推論過程。本研究將探討因果推論的方法，以及中介模型的時序設計應用在因果推論上的現況。

二、探討現存的中介模型在因果推論上的完備條件

中介模型在研究設計上，除了最常被採用的橫斷中介之外，納入時間概念的縱貫中介，更是有許多種不同的設計。本研究將基於因果推論的重要因素做為完備條件檢核表，並逐一檢視現有的中介設計在應用於因果推論上的限制。

三、提出增強因果推論完備條件的可能策略為何

本研究在檢驗各種中介模型的因果推論完備條件後，將提出強化因果推論的改善策略，並進一步說明相應的統計檢驗程序。最後，搭配實際資料的分析演示與歸納研究結果，做為相關領域學者制定研究分析架構之參考，並對未來相關研究提出建議。

第三節 名詞解釋

為利於本研究之討論與分析，茲詮釋重要概念之意涵如下：

一、時序設計（sequence design）

指研究變項在理論模型及實際測量當中的順序安排，涉及到邏輯上的順序、測量時間點的順序及測量波數的議題。在本研究中可分為橫斷設計、縱貫設計二大類。

橫斷設計是指，所有研究變數都在同一個時間點進行測量。而縱貫設計是指，同一群個體在不同時間內的行為變化情況的設計架構，類似受試者內設計，每一

個時間點都巢套 (nest) 在個體之下。主要特徵是分不同時間點來收集資料。縱貫設計有許多種類型，例如：趨勢研究 (trend study) 與世代研究 (cohort study) 這類針對母體推論而不強調相同樣本追蹤的方法。本研究所指稱的縱貫設計，皆為固定樣本覆訪研究 (panel study)，是在不同時間點檢視同一批受訪者，乃心理學及相關領域最常使用的追蹤設計方式。

二、中介模型 (mediation model)

中介模型在自然科學及社會科學中有著悠久的歷史 (Fisher, 1935; Lazarsfeld, 1955; Blau & Duncan, 1967)，但近年來在心理與教育領域的應用更是頻繁於其他領域。其中最為人所知的，便是 Baron 與 Kenny (1986) 最常被引用的著作。在研究架構中，以 X 預測 Y 時，若將第三變項 Z 視為 X 對 Y 路徑上的中間傳遞者，則此架構被稱為中介模型，此時的第三變項 Z 則特稱為中介變項 (mediator, M)。在迴歸基礎上會得到三條路徑，分別是 $X \rightarrow Y$ 、 $X \rightarrow M$ 及 $M \rightarrow Y$ 的預測效果。而中介效果 (mediation effect) 則是指稱 $X \rightarrow M \rightarrow Y$ 的影響歷程，許多的檢定程序議題都聚焦於此效果的顯著性考驗上。

三、因果推論 (causal inference)

因果關係 (Cause and effect relationship) 是指一個事件 (原因, antecedence) 對第二個事件 (結果, outcome) 之間的作用關係。其中，後發生者被認為是前發生者的結果，亦即原因總是先於結果。在本研究所秉持的本體論觀點中，因果關係乃實質存在，但其全貌有待研究而逼近。因果推論，則是對因果關係的可能運作方式提出想像，本質上乃一種待否證的假設概念。對因果推論的否證，則是指探究此一假設是否成立獲得實證資料支持或因果強健程度的一連串哲學推理與實證 (統計推論) 過程。在因果推論上有二個主要探討的方向：其一乃因果傳遞的「機制」，在統計上即為中介 (mediation)，此為本研究所關心的焦點。其二乃因果傳遞的「時機」，在統計上即為調節 (moderation)，由於討論的方向差異較大，故不在本研究討論範圍內。

第二章 文獻探討

本章針對因果推論、中介模型及縱貫中介模型有關之文獻進行探討，以期建立本研究之論述框架。因此，就因果推論的相關概念、中介模型之定義與使用現況、縱貫設計的中介模型、自我迴歸取向的縱貫中介模型、強化因果推論的可能策略等五部分，做如下探討。

第一節 因果推論的邏輯與影響因素

壹、因果推論的邏輯

在社會科學領域中，所有學生在修習「實驗方法」課程時都曾聽聞「實驗方法是唯一能檢證因果關係的途徑」一說，此種說法並非絕對，而是相對於調查方法所言。倘若觀其言行，就能發現越是從事實驗法的人，越是強調因果關係的重要性，甚至認為發現因果關係是科學研究的唯一（或第一）優先目標。

在哲學上，「原因」經常被定義為「結果」的一個充分且必要的條件（sufficient and necessary condition）（Copi, 1986）。之所這樣論述是因為：1. 若原因僅為結果的充分條件，則可從原因推導結果，卻不能從結果來歸咎原因；2. 原因僅為結果的必要條件，則只能從結果來歸咎原因，卻不能由原因來推導結果；3. 唯有當原因是結果的充分且必要條件時，才能由原因歸咎結果，且能從結果推論原因。這樣的邏輯會導出一個結論：任何結果都應該只有一個充分且必要條件。然而 Mill（1973）卻對此提出異議，因為任何結果都只能有一個充要條件是很明顯的違反現實。故認為，原因是結果的一組必要條件，但結合起來後則成為充分條件（necessary but jointly sufficient）。例如，對於構成燃燒的條件有：可燃物、氧氣及到達燃點溫度，三者缺一不可，齊全時必會造成燃燒現象。

那麼人類是如何判斷何者為原因 X、何者為結果 Y？最直觀也最單純的方法即所謂的「歸納類化」（inductive generalization）。倘若能觀察到結果 Y 許多次，每次都發現 Y 伴隨著 X，就可推論出：下一次，若 X 出現，則相似的 Y 現象也會

出現，進一步類化到所有的 Y 現象，然後下結論：X 導致 Y。然而歸納法最大的缺陷在於，僅需一次的例外就可以推翻根據許多次觀察所做出的結論。另外，使用這種方法很可能讓觀察者傾向觀察到符合規則的事件而忽略不一致的證據。是故，歸納方法也不是適切的因果推論策略。

針對上述歸納法的缺點，Mill 提出五種方法來進行因果推論，其基本原理乃是現代實驗設計的想法來源。這五種方法分別是：一致法 (method of agreement)、差異法 (method of difference)、聯合法 (一致法與差異法)、殘餘法 (method of residuals)、與共變項法 (method of concomitant variation)。各種方法的細節請參考表 2-1。Mill 宣告其方法可以用來讓每個人都成為科學家，去「發現」或「證明」因果關係，但根據其內涵可發現，所有的歸納方法其實都是一種演繹法，必須要先滿足「 $ABC \rightarrow XYZ$ 」這一因果關係的前提假設 (assumption)，否則結論都很可能為錯誤。那麼「 $ABC \rightarrow XYZ$ 」的前提假設是怎麼生成的？只能從理論或知識而來，唯有理論堅實且知識廣博者才能提出合理的因果假設，此乃長期教育與努力之成果，要成為科學家沒有捷徑 (詹志禹，1993)。

究其本質觀之，Mill 測試因果關係的方法其實都是一種「相關」(或稱共變)：因與果之間的相關、不同條件下與結果的相關等等。另外也有一些哲學家提出因果推論的條件，例如 Hume (1739) 認為「原因必須發生在結果之前」；Kant (1781) 則提出「因果之間必須有滋生傳導 (generative transmission) 機制」；而 von Wright (1971) 則強調人「操弄原因能改變結果」，也正是實驗方法的原則。

表 2-1-1 Mill 所提五種歸納法的整理摘要表

名稱	內涵	舉例	缺點
一致法	$ABC \rightarrow XYZ$ $ADE \rightarrow XTU$ 因此, $A \rightarrow X$	二人食物中毒, 唯一都有吃下的是薯條, 可推論薯條是原因。	可能還有其他共同原因: 例如都沒洗手。
差異法	$ABC \rightarrow XYZ$ $BC \rightarrow YZ$ 因此, $A \rightarrow X$	一人食物中毒, 另一人否, 中毒者吃了雞翅, 可推論雞翅是原因。	中毒者可能沒洗手。
聯合法	$ABC \rightarrow XYZ$ $ABC \rightarrow XYZ$ $ADE \rightarrow XTU$ $BC \rightarrow YZ$ 因此, $A \rightarrow X$	實驗組-控制組, 差別只在接受操弄與否, 其他條件相同。	各種威脅實驗效度的因素。
殘餘法	$ABC \rightarrow XYZ$ 已知 $B \rightarrow Y$ 已知 $C \rightarrow Z$ 因此, $A \rightarrow X$	吃了三個藥丸, 嘔吐、腹瀉、發燒都痊癒, 已知前二個藥丸是治嘔吐與腹瀉, 則可推論第三個藥丸是治療發燒。	結論是可能而非必然。前二個藥丸也可能有交互作用。
共變法	$ABC \rightarrow XYZ$ 或 $ABC \rightarrow XYZ$ $A \text{ 增加} \rightarrow X \text{ 增加}$ $A \text{ 減少} \rightarrow X \text{ 增加}$ $A \text{ 減少} \rightarrow X \text{ 減少}$ $A \text{ 增加} \rightarrow X \text{ 減少}$ 因此, $A \rightarrow X$	想知道月球引力是否導致潮汐, 無法操弄月球的有無, 僅能觀察潮汐高低與月球引力的關係。據此推論月球引力是潮汐高低之原因。	有相關不一定有因果。

隨著科學哲學理論的發展，Einhorn 與 Hogarth (1986) 認為，人類在判斷 X 和 Y 之間因果關係上有六大影響因素：(一) X 的背景突出性 (difference in the background)；(二) X 必須先於 Y；(三) X 和 Y 之間的因果鏈強度 (causal chain strength)；(四) X 與 Y 的共變程度；(五) X 與 Y 在時間與空間上的接近性；(六) X 與 Y 的相似性。

第一個因素在概念上假設人的知覺是根據背景而來的，當背景改變時人所知覺的事件原因也會改變。技術上近似於 Mill 的聯合法，因果推論當中都假設了一組必要但不充分的條件，推論過程只是從這組條件中篩選出比較顯著的、特殊的條件。第二個因素則是時間順序，因為事件的順序是容易影響因果推論的主要理由。然而若時間順序概念又可分為邏輯時序與測量時序，邏輯時序是指邏輯上合理的先後順序，測量時序則是指實際測量時的先後順序，二者都可能被誤用。邏輯時序的誤用，例如雞啼總在日升之前，則雞啼為日升之原因。測量時序的誤用，例如在二十歲時測量人際支持，四十歲時測量心理健康，結果發現有顯著相關，故推論年輕時的人際支持會影響中年的心理健康；然而，人際關係跟心理健康或許早已存在，其先後順序與測量順序無關。第三個因素就是前述所談的滋生傳導機制，因果鏈的強度是指因果推論的合理性，也就是在闡述因果關係時，其邏輯架構與事實的可信程度。

第四個因素，可以統計學的概念加以闡述：

		Y 是否出現	
		是	否
X 是否出現	是	A	B
	否	C	D

在統計上，要判斷 X 與 Y 的關係時，A 和 D 代表一致的情況，B 和 C 則代表不一致的情況。倘若 A+D 越高且 B+C 越低，則 X 與 Y 的關係越高。然而一般人不一定能如此周全地考量，部分的人會僅依據 A 的高低來判定 X、Y 之間的關係。再者，倘若以 KAPPA 係數的觀點來看，就算是一致的部份也有可能是機率造成的，其概念本質上，還是屬於歸納方法，也依然受到相對的限制。

第五個因素，時間與空間上的接近性，則與第三個因素因果鏈的強度有些關係。倘若因果之間在時間或空間上相隔越遠，則其推論的合理性也越發降低，也不容易被發現。然而即使差距很遠，只要能被賦與一條合理的因果鏈則可建立起因果關係。反之，當 X 與 Y 在時間或空間上非常接近，則即使缺乏合理的因果推論機制，也還是很容易被認為有因果關係存在。

第六個因素，相似性，是指因果推論受到原因與結果的相似程度影響。相似性可以分為二個層次，其一為大小或強度的相似性，其二為物理性質的相似性。例如許多人都不相信第一次世界大戰的起因是一個小小的刺殺事件，事件本身只是個導火線，帝國在數百年間累積的數個經濟政治衝突才是真正的原因。細菌剛被發現時，人們也不相信這麼小的東西會導致尺寸數萬倍大的人類生病或死亡 (Einhorn & Hogarth, 1986)。此外，中醫向來有「以形補形、以臟補臟」說法，例如吃核桃補大腦，因核桃的外形與人腦類似或者吃豬肝補肝等等的食療傳統，這也是基於形狀相似的概念而來。

在這六大因素當中，Einhorn 等人 (1986) 認為前三者是判定因果關係的必要條件。倘若 X 跟背景一樣而非異常條件，或 X 後於 Y 發生，或 X 與 Y 之間沒有任何因果鏈，則人們不會做出 X 與 Y 之間的因果判斷。那在這三個因素中又以何者為最關鍵呢？或許是因果鏈的概念，因為從概念的廣度而言，因果鏈，也就是因果推論的整個機制就已經涵蓋背景突出性及時間先後順序，而後三者則是納入來檢驗整個因果鏈強度是否足夠的判斷依據而已。回顧前人研究後，White (1988) 也得出類似的結論：時空接近性、相似性、操弄、共變、因果鏈這幾個因素中，因果鏈可能是最基本的。

貳、影響因果推論的可能因素

在上述討論中，都是針對因果推論的基本邏輯和一般人在推論上的判斷項目，然而在現代科學中，統計方法學當中的統計推論 (statistical inference)，恐怕讓更多科學從業人員誤解了因果推論的內涵。

統計推論是指如何根據樣本數據去推斷母體特徵的方法，乃建立在對樣本數據進行描述的基礎上。在使用統計分析的過程中，最大的問題通常在於邏輯的欠缺及對於假設考驗的誤解。具體而言會有以下幾種誤解 (詹志禹，1997)：

(一) 相關分析能 (或不能) 證明因果關係？

在相關的資料中的確無法歸納出因果關係，同時也不能證明因果關係，但能否證 (refute) 或考驗因果關係存在的假設。其原因在於：有因果必有相關，有相關不一定有因果，無因果則一定無相關 (Mulaik, 1987)。從邏輯學來說，「若 P 則 Q」的等值句為「若非 P 則非 Q」，換言之，「若有因果則有相關」的等值句為「若無因果則無相關」，而非「若有相關則有因果」。倘若研究者根據理論推導出「X 會影響 Y」此一假設，分析資料後發現其迴歸係數達到顯著，能做的結論是「X 會影響 Y 這個因果關係的假設尚未被推翻，目前為止仍然是可接受的假設」。相反地，倘若分析資料後發現其迴歸係數未達到顯著，能做的結論則是「X 會影響 Y 這個因果關係的假設，根據現有的資料來看，是不被接受的」。由此可知，統計分析所能做到的是根據實際資料去否證因果關係的假設，而該假設仍是一個懸而未決的概念。

(二) 變項間具有因果關係才能進行迴歸分析或其他等值模型？

習慣上，在變異數分析 (ANOVA) 或迴歸分析 (含其延伸方法) 當中，研究者將自變項當作原因，把依變項當作結果。這是一種根據理論或邏輯推導而來的決定，而非統計學上的考量，甚至可以將毫無關係的變項投入分析方法當中。統計方法本身是中性的工具，它並不知道研究者腦中對於因果關係的想像。重點

在於，因果關係是一種尚待考驗的假設。此外，統計方法分析的是自變項與依變項之間的函數關係，並沒有判定因果方向的功能。

(三) 在統計程序中對調自變項與依變項就可檢驗何為因、何為果？

根據上述的討論，可導引出「因果方向的判定是由研究者決定」，而研究者乃依據理論基礎及時間順序等線索做為參考。分析結果僅呈現「自變項與依變項的共變關係，以及是否可推論至母體」。倘若因橫斷面設計而僅有邏輯時序而無測量時序時，這時很多研究者會將自變項與依變項調換過來。甚至在某些研究者考量之下，迴歸分析的「回溯預測」或是做完多變量變異數分析(MANOVA)後改做區別分析，都是合理的分析策略。這些做法在統計上完全合乎規定，並且從未有過「自變項一定是原因，依變項一定是結果」這樣的假設。這種假設來自於研究者的理論基礎、思考習慣或一般印象所致。

既然統計方法在於分析自變項與依變項的共變關係，則二者之間的函數關係是對稱的，也就是說對調變項在分析中的位置，應該產生相同的結果。倘若結果不同，研究者應該思考的是：(1) 二個變項的單位是否相同？(2) 二個變項的編碼方式是否相同？(3) 二個變項的遺漏值處理是否相同？而非「因果方向不一樣」，因為這是對稱的數學關係上不可能的選項。

(四) 基於理論或實驗設計之路徑分析或結構方程式模型可以協助因果推論？

路徑分析與結構方程式模型都是相關概念的延伸，本質上為共變數結構的分析，因此在第一點談論的想法依然可用。但因為以模型為基礎(model based)的方法包含了適合度(goodness of fit)概念，所以在此特別說明：倘若研究者提出的某個因果模式，其適合度數據都符合要求，也只能下結論說：「該因果模式沒有被否認，可以暫時被接受，直到另一個更合理、適合度表現更好的因果模式出現為止」。值得注意的是，「更合理」是指增強因果鍵強度，也就是指其背後的理論基礎更好；「適合度更好」是指統計上各種殘差值更小，表示資料符合理論模式的預期；這二者缺一不可，其角色功能相輔相成。

綜上所述，統計方法的角色並非「證明」因果關係，而是透過數學關係來模擬因果關係的可能強度，並透過「否認」的邏輯程序，讓資料來篩選「較佳的假設」。

小結：中介模型乃一種統計方法，在應用上無法自外於更久遠的知識論（epistemology）與方法論（methodology）架構。本節透過邏輯、影響因素及統計推論等角度，列出當吾人試圖透過中介模型此一統計方法進行因果推論時，可能遇到的錯誤或考慮不甚完備之處。因果推論是人類與生俱來、對於解釋環境的一種思考方式，乃受到諸多因素影響逐步建構（construct）而來，是一種可改變的心理歷程，故非有無的判斷，而是根據考量因素的多寡進行程度的判定。而因果推論的本質，是奠基於理論的假設，是一種工具性（instrumental）的概念，用以解釋事物的運作。考量因素越周全的因果推論將被視為因果鏈強度較佳的假設，也越值得後續的統計分析。釐清上述觀點後，方能對後續中介模型的設計有更深層的理解，也鋪平更進一步強化因果推論的道路。

第二節 中介模型的原理及其應用

壹、第三變項之角色功能

在社會科學研究當中，主要的目的經常在於透過 X 預測 Y 。然而在現實世界中，還可能存在著一個或多個可能的第三變項 Z ，在概念上會以四種方式影響 X 對 Y 的影響程度。在這當中，中介模型是其中一種看待第三變項的方式，是一種強調動態過程的說明機制，故為本研究的焦點所在。為在討論中介模式之前釐清脈絡，茲先說明第三變項影響方式的四種類型如下（邱皓政，2017）。各種類型的圖示及預測方程式請參考表 2-3-1 所示。

（一）無第三變項影響

當研究架構不（或無法）考慮第三變項時，可利用簡單迴歸進行分析，得到迴歸係數 c ，用以表示 X 對 Y 的影響程度。此時的 c ，在解釋上很可能有錯誤，

因為並沒有排除掉可能的干擾因素 (confounding)。

(二) 第三變項為控制變數 (control variable)

當研究者認為第三變項 Z 也會影響 Y ，但 Z 並非該研究所關心的變數時，最簡易的處理策略便是使用多元迴歸分析，將 Z 也看作與 X 同等的預測變項，置入方程式當中，一起做預測變項之用。此時的 Z 被稱為控制變數，透過 X 與 Z 之間計算相關 (共變) 程度 r ，來達到互相控制的目標。此時， X 對 Y 的影響改為 c' ，用以表達其數值是受到 Z 納入迴歸模型中而受到改變的量。同理，這時估計出的 Z 對 Y 的迴歸係數 b' ，也是一個受到 X 控制後所估計出的參數。納入 Z 之後，若原本的 c 比起控制 Z 之後的 c' 較大，則稱 c 為虛假關係 (spurious relationship)；反之，若 c 較小，則稱 Z 為壓抑變項 (suppressor)。

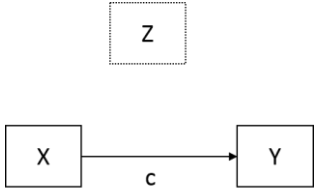
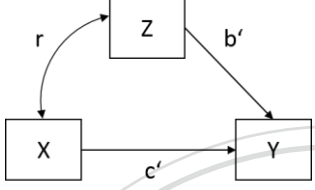
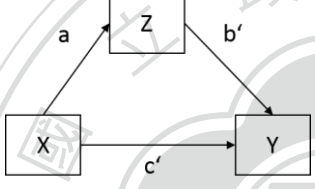
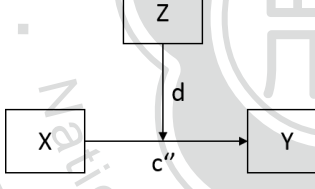
(三) 第三變項為中介變數 (mediator)

當研究者認為第三變項 Z 乃 X 對 Y 關係中的轉化機制傳遞者時，則稱第三變項 Z 為中介變數。 X 必先影響 Z ，再影響 Y ，是一個因果鏈傳遞的概念。在迴歸的估計上，將 Z 也視為 X 的依變項，此時 X 對 Z 為簡單迴歸，其迴歸係數 a 並無任何控制。接著，以 X 和 Z 為自變數來預測 Y ，此時，如同將 Z 做為控制變數時一樣，得到 Z 對 Y 的迴歸係數 b' ，及 X 對 Y 的迴歸係數 c' ，二個係數都是互相控制後得到的淨效果。

(四) 第三變項為調節變數 (moderator)

當研究者認為第三變項 Z 會影響到 X 對 Y 的預測程度時，則稱第三變項 Z 為調節變數。在概念上，可說 Z 是將 c'' 視作依變項，其迴歸係數則為 d 。在計算上，則是求取 X 與 Z 的交互作用 (interaction)，故在公式 (2-5) 當中，仍有 b'' 和 c'' 的估計。

表 2-2-1 第三變項 Z 影響方式示意圖

變項 Z 角色	模型圖示	預測方程式
無第三變項 影響		$\hat{Y} = \beta_{0c} + cX$ (2-1)
第三變項為 控制變數		$\hat{Y} = \beta_{0bc} + c'X + b'Z$ (2-2)
第三變項為 中介變數		$\hat{Z} = \beta_{0a} + aX$ (2-3) $\hat{Y} = \beta_{0bc} + c'X + b'Z$ (2-4)
第三變項為 調節變數		$\hat{Y} = \beta_{0bcd} + c''X + b''Z + dXZ$ (2-5)

資料來源：邱皓政（2017）。多層次模式與縱貫資料分析：Mplus8 解析應用（頁 342-343）。臺北市：五南。

貳、中介模型之原理與顯著性檢定議題

中介模型是一種將第三變項 Z 視為傳導機制的中間者的架構，在概念上符合前述因果鏈的看法，也就是因果推論的最核心關鍵：合理性。對於因果傳遞的合理性，所有研究者都會問：X 對 Y 的效果真的是透過二者間的中介變數 M 傳遞的嗎？這個問題可以分為哲學和統計二個方向探討：哲學關心的是因果機制的傳遞方式是否為真，而統計所關心的是有多少變異能被解釋（Pearl, 2012）。

從統計的觀點出發，中介模型的重點在於分析中介效果（mediation effect）

是否存在，也就是 X 是否透過 Z 進而影響 Y ，由於此時的 Z 特稱為中介變數，故後續都以 M 代稱之。一個典型的中介模型至少涉及三個變數：自變數 X 、中介變數 M 及依變項 Y 。為了估計中介效果存在與否，最直觀的作法是逐一估計每個迴歸係數。這也正是 Baron 與 Kenny (1986) 提出的經典做法，茲說明其步驟如下：

首先，檢測 X 對 Y 的簡單迴歸係數 c 是否達到顯著，若達到顯著則可進行下一步驟，此乃中介效果檢驗的大前提，也符合因果推論當中的「共變」原則。其次，檢驗 X 對 M 的簡單迴歸係數 a 是否達到顯著，若是則可進行下一步驟，也就是因果鏈的第一次延伸。第三步，檢驗 M 對 Y 的預測是否達到顯著，此時的估計乃是納入 X 與 M 的多元迴歸，故將得到 M 對 Y 的多元迴歸係數 b 及 X 對 Y 的多元迴歸係數 c' 。倘若 a 與 b 皆達到顯著則可宣稱中介效果存在。接著觀察 c 與 c' ，若納入中介變數 M 之後， c' 只是效果下降但仍然顯著則稱為部分中介效果 (partial mediation effect)；若 c' 不顯著 (完全消失) 則表示 X 的效果完全透過 M 傳遞到 Y ，稱為完全中介效果 (completed mediation effect)。

上述討論是針對中介效果是否存在，如果要計算出中介效果的大小，則可透過 X 對 M 、 M 對 Y 的二個迴歸係數之乘積 $a*b$ 得來，或是納入中介變數前後的 X 對 Y 的迴歸係數 c 到 c' 的縮減量得來 (MacKinnon, Warsi, & Dwyer, 1995)。表達如公式 (2-6)。

$$\text{中介效果量} = c - c' = a \times b \quad (2-6)$$

若從效果量拆解的角度觀之， X 對 Y 的簡單迴歸係數 c 原被稱為總效果 (total effect)，此時尚未納入中介變數 M 。中介途徑上的二個迴歸係數乘積 $a*b$ 為間接效果 (indirect effect)，納入中介變數 M 後， X 對 Y 的效果將從 c 變為 c' ，此時 c' 稱為直接效果 (direct effect)。上述效果具有加成性，可表達如公式 (2-7)

$$\text{總效果} = \text{直接效果} + \text{間接效果} = c' + a * b \quad (2-7)$$

此時，若間接效果或是 c 到 c' 的遞減量達到顯著水準，則可看作中介效果存在的證據。也正因為間接效果的統計顯著性是判斷中介效果存在與否的關鍵指標，故整個中介模型的發展歷史，始終圍繞著間接效果的顯著性考驗議題。

中介效果是由間接效果 a 與 b 的乘積估計而來，因此在顯著性考驗上，目標在於計算出 $a*b$ 的抽樣分配 (sampling distribution) 及其標準誤 (standard error)，以考驗 $a*b$ 是否為 0 或建立 95% 信賴區間 (confidence interval, CI)。最常被使用，也是多數 SEM 軟體 (LISREL、AMOS、EQS、Mplus) 內建的間接效果顯著性考驗方法是 Sobel (1982) 所建立起的公式，其本質乃 t 檢定，即透過 $a*b$ 的標準誤進行 $a*b$ 是否為 0 的 t 考驗，如公式 (2-8) 所示。

$$t = \frac{ab}{\sqrt{s_b^2 a^2 + s_a^2 b^2}} \quad (2-8)$$

在公式 (2-8) 當中，分母為聯合標準誤，乃透過 a 與 b 的標準誤求得。然而， $a*b$ 是二個迴歸係數的乘積，該項的機率分配並不服從常態分配，使得 Sobel 公式中的標準誤將是一種有偏的估計值 (biased estimator)，故 Sampson 與 Breuing (1971) 對標準誤提出修正。校正非常態問題後，Bobko 與 Reieck (1980) 則聚焦在樣本數小於 200 時可能產生的非正定問題上 (indefinite)，建議進行中介分析前，先把所有變數標準化，並利用三者的相關係數計算標準誤。

來到 1990 年代後，由於電腦運算的速度提昇，得以利用電腦來進行重複取樣 (resampling) 技術進而建立起參數的抽樣分配。根據中央極限定理 (Central Limit Theorem)，無論母體是否為常態，其多次重複取樣之樣本平均數所形成之抽樣分配必為常態，而抽樣分配之標準差則特稱為標準誤 (standard error)。拔靴法 (bootstrap method) 即依循此原理，讓電腦以抽樣的樣本資料為母體，進行上千次的置回取樣，所得到的 $a*b$ 抽樣分配，其標準差即為間接效果的標準誤 (Efron & Tibshirani, 1993)。此方法不需受

限於常態機率的要求，只要檢視 95%CI 是否包含 0，即可判斷間接效果是否顯著不為 0。由於學界開始重視 $a*b$ 不服從常態的事實，故此方法逐漸成為檢驗間接效果的主流技術，在近年的研究中受到大量的採用。

參、中介模型的使用現況及評論

由於中介模型將第三變數看作位居中間的效果傳遞者，在檢驗外顯行為或內在心理歷程的動態特徵上有很大的相似性，故受到心理、教育、管理等領域大量的應用。除了傳統的路徑分析 (path analysis) 取向外，結合潛在變數模式 (latent variable) 的潛在中介及考慮巢套資料 (nested data) 的多層次中介更是常見於當今主流學術期刊。更有甚者，結合中介與調節二者在同一模型當中的調節式中介 (moderated mediation, MoMe) 與中介式調節 (mediated moderation, MeMo)，其討論與應用更是在近年來蓬勃發展 (Hayes, 2013)。

誠然，間接效果的顯著性考驗是非常重要的議題。然而，中介模型的討論若僅在此之上，也就僅止於統計推論的層次。間接效果的討論比較近似於共變關係與因果鏈的強化，然而卻缺乏許多因果推論的要件，尤其是時間因素。

在缺乏時間因素之下，中介模型做為橫斷面調查方法的一種架構，在因果推論上有著極大的缺陷，也就是其內在效度 (internal validity) 不如實驗法的關鍵所在。在研究設計上，若要推論 X 是 Y 的因、Y 是 X 的果，一般而言，必須滿足三個條件：其一，X 與 Y 要有相關；其二，X 必須發生在 Y 之前；其三，X 與 Y 之間應無其他混淆變數。橫斷面研究之下的中介模式 (簡稱「橫斷中介」)，其影響效果的單一方向性，即是試圖在調查方法中做到因果推論的目標。然而，僅能滿足上述的條件一：相關而已。至於條件二的時間順序，橫斷中介僅能達到「邏輯時序」，也就是概念上的先後，此乃理論或邏輯推演而來。即使間接效果達到顯著水準，也只能下結論說「在單波測量的情況下，做這樣的因果關係假設，目前是可被接受的」。橫斷中介未能滿足更重要的「測量時序」，也就是實際發生的先後。而這樣的缺陷，即使是將變數之間的順序調換，也僅是邏輯時序的更改，

依然無法觸及到測量時序的概念。上述特性，乃是一般橫斷面調查方法在因果推論的最大致命傷。總言之，橫斷中介的設計僅是統計推論，而非因果推論。

第三節 縱貫設計的中介模型

在因果推論的過程當中，時間是最重要的因素。將時間因素納入中介模型時，要從二個面相來談，分別是邏輯時序與測量時序，二者必須同時滿足才能構成因果關係的充分條件。納入時間因素的研究方法即為縱貫面設計，納入時間因素的中介模型則稱為「縱貫中介」模型。本節將介紹常見的縱貫中介模型，包含序列設計取向、潛在成長模式取向，及自我迴歸取向，並說明其原理、應用狀況與設計上的限制，以做為後續強化因果推論的可能策略之討論基礎。

壹、縱貫設計中的時間議題

在因果的推論過程中， X 必先於 Y 、 Y 必後於 X ，此種關係稱為時序 (temporal sequence)，也是解釋時間效果的重要來源。時序可分為二種重要的概念：其一為邏輯時序，也就是根據理論、邏輯或過往證據推演出的變數間順序。其二為測量時序，也就是該變數所測量的時間點順序。在橫斷面研究 (cross-sectional research) 中，由於只有一次機會測量所有研究變數，故僅能做到遵守邏輯時序，無法遵守測量時序。如此一來，無論使用的統計方式有多麼精確，都無法說明 X 與 Y 之間的因果關係，頂多能描述二者之間的共變程度。而縱貫面研究 (longitudinal research) 中，因為有了多次測量變數的機會，能透過測量上的先後順序更完備地構築因果推論所需要的時序架構。然而，有多次測量並不保證測量時序的品質，還必須兼顧測量時間點、測量間距等安排上的細節。

測量間距，是指在不同測量波段之間間隔的時間長短，最佳的測量間距應該基於理論及前導研究 (pilot study) 所共同決定。要注意的是，測量間距的長短安排，將涉及到歷史事件、成熟等干擾因素的影響。在各波次的測量間距是否需相等的議題上，還是應回到理論或現象本身，針對特殊且有意義的時間點進行安排規劃。

雖然，縱貫設計中已納入時序概念，然而時間本身並非研究者所能操弄的變數，縱貫設計的討論核心在於如何安排適當的觀測角度，也就是測量時機（timing），此乃決定時效（effect of timing）的關鍵因素。關於如何考量最佳的時距（time duration）、延宕（lag）期數或測量時間點（time point）等議題，在心理學研究中仍然是個有待持續探索的領域（Cole & Maxwell, 2009; Reichardt, 2011）。

倘若研究者決定進行某（些）變數的重複測量，則隨著時間延續所得到的資料將形成時間序列資料。Kenny（1979）指出若時間序列資料能在跨時間下維持穩定而無系統性變化，則可稱為穩態（stability）。舉例來說可想成身高過了一定年齡後就不再增加，即呈現穩定狀態，也就是平均數/一級動差恆定的概念。換個角度來看，若某變數的跨時間相關程度越高，則表示穩態越高。反之，若跨時間相關程度低，則表示該變數在跨時間的測量上變異程度較大。當有二個以上的時間序列資料時，還有更進一步的弱定態（weak stationary）概念，也就是在穩態的基礎之上，時間序列之間的因果結構也維持不變，也就是同時維持一級與二級動差的恆定。（邱皓政，2017；Cole & Maxwell, 2003; Kline, 2011; Little, Preacher, Selig, & Card, 2007）。

僅在當 X 與 M 皆為穩態的情況下，橫斷面中介可以得到類似於縱貫面中介的結果，此為充分及必要條件。然而此項前提假設經常沒有被檢驗，因為檢驗必須在前導研究進行，這是絕大多數研究所欠缺的程序。當此假設被違反時，不論樣本規模有多大，橫斷面分析都將對 $X \rightarrow Y$ 的直接與間接效果產生高估或低估。而此類偏誤，在實務上，將導致橫斷面研究在 p 值或信賴區間 CI 的估計失去意義。因此，一個變數若在橫斷面研究中被發現是一個有效的中介變數，並不必然在縱貫面研究中。也會有同樣的表現。總結而言，橫斷面的中介設計在參數估計的正確性上，可說是全然的落敗於縱貫設計。

然而即使是納入時間因素的縱貫設計，若無重複測量的時間序列資料，在估計跨時間中介效果時，也可能遇到估計偏誤的問題。Mitchell 與 Maxwell(2013) 指出，跨時間中介效果的估計受到 X 與 M 二變數的穩態高低影響：若 X 的穩態大於 M，序列設計的偏誤較小；若 M 的穩態大於 X，則序列設計與橫斷設計的中介效果估計都有嚴重的偏誤。根據此結論，Mitchell 等人建議採用具有時間序列概念的自我迴歸取向的中介模型。

貳、序列設計

既然橫斷面取向有著許多的方法學的爭議，為何在明知縱貫面取向的好處後，橫斷面方法依然被大量採用呢？最有可能的答案即是成本考量，縱貫面方法的成本不僅僅是乘上測量波段的外顯成本，還要考慮樣本追蹤的維護、樣本流失的檢測及處理技術的升級等等隱性成本。倘若橫斷面與縱貫面設計之間的差異甚小，那研究者理所當然地，也不會有動機去選擇更高成本的做法。

然而縱貫面設計在強化因果推論一事上的好處，也是非常明顯的。為瞭解決橫斷中介在時間因素的不足，MacArthur 提出了補強版的序列設計中介模式 (sequential design mediation model)，也就是分別在前、中、後三個時間點測量 X1、M2 與 Y3，數字部分代表測量的時間點順序 (Kraemer, Essex, & Kupfer, 2008)。此時的第三時間點的依變項 Y3 是被第一時間點 (t=1) 的自變數 X1 及第二時間點的中介變數 M2 所共同預測，可表示如公式 (2-9)。

$$Y_{t+2} = cX_t + bM_{t+1} + \varepsilon Y \quad (2-9)$$

第二時間點的中介變數 M2 則是被第一時間點的 X1 所預測，可表示如公式 (2-10)。

$$M_{t+1} = aX_t + \varepsilon M \quad (2-10)$$

模式概念圖如下：

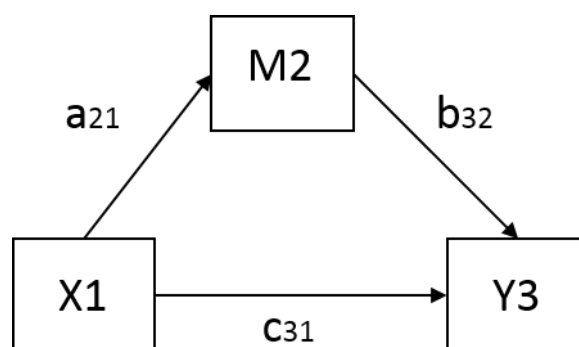


圖 2-3-1 三波追蹤資料之序列中介概念圖

(為求簡化，省略誤差項與符號)

然而，這項設計依然不能滿足因果推論的條件三，也就是未能解決混淆變數的問題。而在縱貫研究中，除了某變數的平均數如何隨著時間增減之外，同變數的不同波次測量的數值之間，也會有跨時間的相關。以時間序列分析 (time series analysis) 的術語而言，即是關於自我相關與自我迴歸的議題。這在因果推論上乃非常重要的切入點，畢竟，縱貫中介模型當中最大的混淆變數，很可能就是該變數本身的前一個狀態。

此外，序列中介模式在應用上，時間間隔的安排差異很大。例如 Carlson (2004) 討論青少年男女的身體意象研究時橫跨了一年的時間；Tein、Ayers 與 Wolchik (2006) 探究失去家庭成員的兒童與青少年橫跨了 11 個月的時間；Abramowitz、Nelson、Ryggwall 與 Khandker (2007) 針對強迫症機制的討論橫跨了 4 個月；Hampson、Goldberg、Vogt 與 Dubanoski (2006; 2007) 探討兒童期經驗與中年健康的關係研究甚至橫跨了 40 年。

因此，在討論 X1 對 M2、M2 對 Y3、X1 對 Y3 的效果時，由於時間範圍很大，導致將歷史事件、成熟效應、樣本流失等威脅內在效度的誤差來源，會一併介入模型中，導致估計上無法排除許多干擾。此外，在 Maxwell、Cole 與 Mitchell 等人的研究中，透過模擬資料的分析，發現除非序列設計的資料能維持穩態與均衡，否則無法透過序列設計來估計縱貫中介的效果，並建議應基於重複測量的基礎上

進行縱貫中介的分析(Maxwell, Cole, & Mitchell, 2011; Mitchell & Maxwell, 2013)。故此架構雖有時間向度介入，但在因果推論上卻因重重誤差而難以令人信服。

參、自我迴歸與交叉延宕

在橫斷面設計之下，抽樣過程若符合獨立且隨機的假設，則同一個變數的不同觀察值乃彼此獨立；例如某甲與某乙兩人的體重，原本就應該不會有任何關係。然而，在縱貫面設計之下，甲乙兩人的體重資料將被重複地測量，由於甲乙兩人每個時間點的體重數值都巢套在自己身上，因此具有共變關係。將甲乙兩人各自來看，不同時間點之間的體重數值會有相關，此種同一個觀察值內同變數跨時間的相關特稱為自我相關（auto-correlation）。

從結構上看來，若今天針對同一批受試者進行變數 Y 的四次測量，從第 t 期至第 t+3 期，中間將有三次時間間隔，此間隔稱為一期的時間落差（time lag）。每任二期測量都可能有關，故四次測量就可能有 $C_2^4 = 6$ 個自我相關係數，如圖 2-3-2 所表示。任何間隔 k 期的測量之間的自我相關係數可表示如公式 2-11：

$$\rho_{Y_t, Y_{t-k}}^k = \frac{cov(Y_t, Y_{t-k})}{\sqrt{\sigma_{Y_t}^2 \sigma_{Y_{t-k}}^2}} \quad (2-11)$$

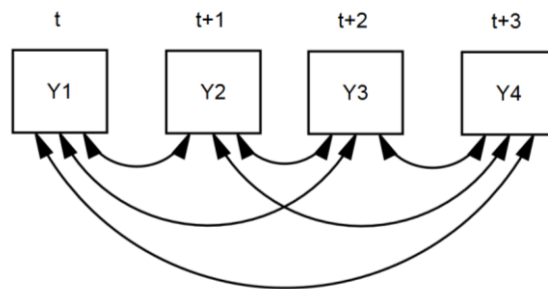


圖 2-3-2 四波重複測量的自我相關概念圖
(為求簡化，省略誤差項及其共變與符號)

公式 2-11 當中相關係數的冪次方 k ，表示測量之間的落後期數。落後一期時 $k=1$ ，此時的自我相關為 $\rho_{Y_t, Y_{t-1}}^1$ ；落後二期時 $k=2$ ，自我相關為 $\rho_{Y_t, Y_{t-2}}^2$ 。根據此定義， ρ^k 為落後 k 期的自我相關。落後一期者為 ρ^1 ，落後二期的自我相關為落後一期自我相關的連乘， $\rho^2 = \rho^1 \times \rho^1$ ，若落後三期則為 $\rho^3 = \rho^1 \times \rho^1 \times \rho^1$ 。由於相關係數介於 ± 1 之間，可看出最大的自我相關落在 $k=1$ 。落後期數越多則自我相關係數越小，最終可被忽略。一般情況下本研究皆採取落後一期，也就是一階的自我相關（first-order auto-correlation）及後續的自我迴歸。

在縱貫資料的分析上，影響變數當前狀態的最有效預測因子，當屬該變數的前一波狀態。舉例來說，預測明天體重最佳的變數並非今天吃了多少食物，而是今天的體重。是故，若要分析今天吃的食物如何影響明天的體重，就必須排除今天的體重此一變數的效果。以前一期資料去預測後一期的資料，這樣的觀念稱之為自我迴歸（auto-regression, AR）。可由公式 2-12 表示。

$$Y_{t+1} = \alpha + \theta_{t+1,t} Y_t + \varepsilon_{Yt+1} \quad (2-12)$$

在公式 2-12 中， α 為截距， θ 為自我迴歸係數（auto-regressive coefficient）， ε_{Yt} 為誤差項。其圖示如下。

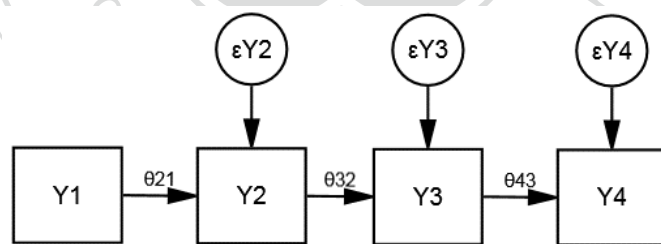


圖 2-3-3 四波重複測量的自我迴歸概念圖

上述的自我迴歸概念，展示了作為依變項的 Y 在一個時間序列的架構下，每一期的變化可由前一期來預測。另一方面，自變項 X 也可能是一個隨時間而變動的變數。以上述的體重預測為例，吃飯量的多寡也可以比照 Y 變數的 AR 架構，列出方程式如下：

$$X_{t+1} = \alpha + \theta_{t+1,t} X_t + \varepsilon_{Xt+1} \quad (2-13)$$

做為二個獨立存在的變數，一開始 X 與 Y 的自我迴歸及誤差都彼此獨立。然而一旦研究者透過理論推導後認為 X 與 Y 兩個時間序列之間可能有互相影響時，則可提出進一步的假設模型：將二個序列當作彼此的自變項與依變項，也就是 t 期的吃飯量可能影響 t+1 期的體重，反之亦然。當資料結構中存在著二個時間序列，可將兩個序列的前後各期當作彼此的 DV 與 IV，這樣的設定將呈現出交錯的路徑模式，稱為交叉延宕模式（cross-lag model, CLM）（Rogosa, 1980），如圖 2-4-4 所表示。圖中的交叉影響稱為交叉延宕效果（cross-lag effect），可由迴歸係數表示其效果。可用方程式 2-14 及 2-15 表示。

$$Y_{t+1} = \alpha_Y + \theta_{Y_{t+1,t}} Y_t + \theta_{Y_{t+1,t} X_t} X_t + \varepsilon_{Y_{t+1}} \quad (2-14)$$

$$X_{t+1} = \alpha_X + \theta_{X_{t+1,t}} X_t + \theta_{X_{t+1,t} Y_t} Y_t + \varepsilon_{X_{t+1}} \quad (2-15)$$

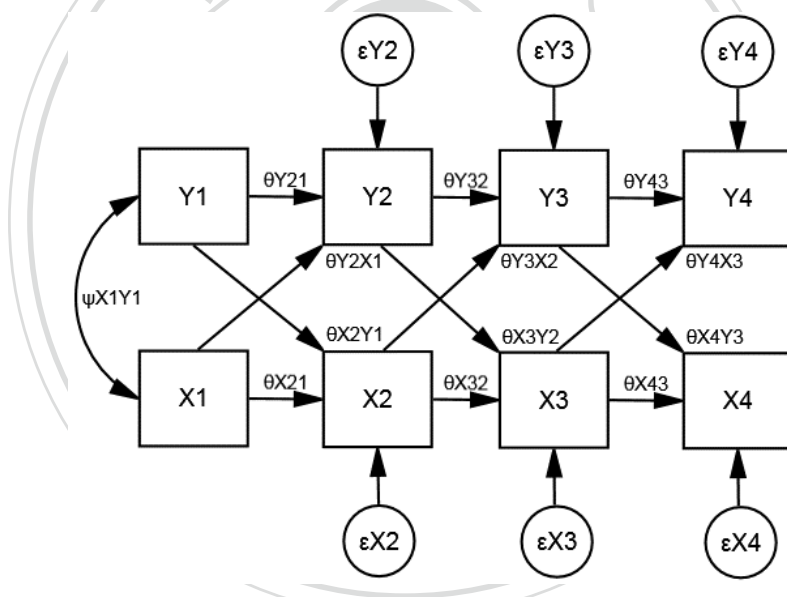


圖 2-3-4 四波重複測量的交叉延宕模式概念圖

（為求簡化，省略誤差項共變與符號）

在圖 2-3-4 可看到，若以 X 為吃飯量、Y 為體重，此模型的討論重點在於以 X_t 預測 Y_{t+1} ，也能以 Y_t 預測 X_{t+1} ，二者互為因果的一種假設。倘若交叉延宕效果的分析結果顯示迴歸係數全達到顯著正值，則表示在整個測量時間的範圍內，吃飯量與體重二者間是互相影響的機制，即吃飯會導致體重增加，體重增加後又吃更多的飯，呈現一個循環增加的態勢；若只有 $X \rightarrow Y$ 的係數達到顯著，則表示吃

飯會增加體重，但體重不會增加吃飯量；反之，亦然。因此，若能在交叉延宕模式中發現僅特定方向的係數有顯著，即可宣稱特定因果方向的假設可被接受（Marshall, Parker, Ciarrochi, & Heaven, 2014）。

在交叉延宕模式中，研究者真正關心的是交叉延宕效果，而自我迴歸則可做為一種控制的方法，使交叉延宕的效果能夠是跨時間影響的淨效果。在圖 2-3-4 可看到，若以 X 為吃飯量、 Y 為體重，此模型的討論重點在於以 X_t 預測 Y_{t+1} ，也就是當研究者要估計第二期的體重 Y_2 時，預測變數為 X_1 並以 Y_1 做為控制變數。在控制了前一期的體重後，就可以估計出前一期吃飯量的跨時間一階淨效果。至此，交叉延宕模式已經很大程度地符合因果推論中最重要三個條件：相關、時間順序與排除混淆。後續章節，將專節介紹將中介概念結合於此的設計。

肆、潛在成長模式

在前節當中曾提及過 X 、 M 、 Y 三個變數本身的穩態會影響跨時間中介效果的估計，但穩態的維持並非必要條件，原因在於穩態僅能反映平均數有跨時間不變性，且透過自我迴歸可以處理這部份的變異。然而，倘若各個變數的平均數有穩定變化的特性，不論是線性或二次變化，此時應採用動態觀點的潛在模式來估計模型參數，以避免嚴重的偏誤（余民寧，2013；邱皓政，2017；Curran, Howard, Bainter, Lane, & McGinley, 2014）。

潛在成長模式（latent growth model, LGM）又被稱為潛在改變量分析（latent change model）或成長曲線模式（growth curve model），本研究統一以潛在成長模式（LGM）稱之。其概念圖可表示如下：

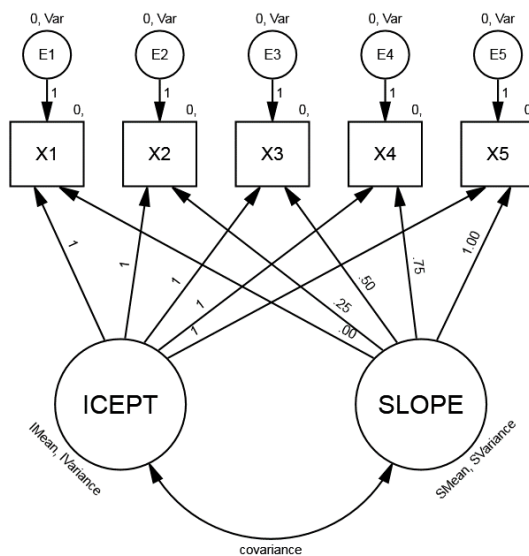


圖 2-3-5 五波重複測量的潛在成長模式概念圖

在圖 2-3-5 當中，X1~X5 為觀察變項，表示五次測量的值；變數 ICEPT 為潛在變項截距；變數 SLOPE 為潛在變項斜率；E1~E5 為測量誤差。

潛在成長模式的理論架構源自於結構方程式模型，可說是多向度測量模式的一種特例，透過限制因素負荷量來達成估計線性方程式的目的。例如，如果研究者要檢驗學童數學能力的變化，他必須透過估計程序確認隨著時間增加，數學能力將以何種函數變化。如果證據顯示一次方程式是較理想的遞變模式（也就是線性關係），那麼將會產生一個直線的一次方程式數學模型，其中截距(intercept)代表初始狀態的能力水準，斜率(slope)代表變動的比率，這兩個係數即是兩個用以顯示成長遞變的主要參數。雙因子的潛在成長模式則是估計平均數的起始狀態，也就是截距（又稱水準因子，level factor）之外，也檢驗了平均數的增減狀態，也就是斜率（又稱型態因子，shape factor）。相較之下，雙因子模型能夠提供更豐富的資訊，在 1980 年代被提出之後，便受到研究者的注意，並逐漸被發展成為縱貫研究資料的重要分析技術(McArdle, 1988; McArdle & Anderson, 1990; 余民寧, 2013)。在實際資料的分析上，倘若雙因子模型不適配，則可考慮單因子或三因子模型，也就是進行重複量數變異數分析或是增列二次曲線的可能性。

本研究所指稱之潛在成長模式皆以雙因子模型為主，探討的對象乃潛在變項截距（intercept，簡稱 I）與潛在變項斜率（slope，簡稱 S）二者。

結構方程式模型可分為結構模式（structure model）與測量模式（measurement model）二部分，而潛在成長模式則可視為測量模式的一種特例。因此，在研究概念的發想上，研究者的第一步通常是先針對重複測量的部份試圖去建立一條線性方程式，也就是估計出截距與斜率，此時，該模式又可稱為非條件化模式（unconditional model）；第二步驟則是納入其他共變數並搭配原本估計出的 I 與 S 變項，以進行結構模式的分析，此模式又可稱為有條件化模式（conditional model）（Meredith & Tisak, 1990; Willett & Sayer, 1994）。條件化模式所納入的共變項可分為二種，一種是隨時間而變的共變項（time-varying covariates），另一種則是不隨時間而變的共變項（time-invarying covariates）。

另一方面，倘若將條件化模式限定在中介模型上，則依照 I&S 變項的角色定位，可分為五類主要的模型，如表 2-3-1 所示。

（一）I&S 做為中介變數

此類模型將 I&S 變項看做是中介變數，探討 X 如何透過 I 或 S 進而影響 Y（Cheong, 2011; von Soest & Hagtvet, 2011）。在估計上，則是以 a_1*b_1 或 a_2*b_2 是否達到顯著，來判定 I 或 S 是否具有中介效果。例如：Simons、Chao、Conger 與 Elder (2001) 探討將親職品質的 I 與 S 做為中介變數 M，探討父母的教養策略不一致會如何影響少年期的犯罪行為。

（二）I&S 做為自變數

此類模型將 I&S 變數看作是自變數 X，探討初始狀態與變動速度何者對於中介變數 M 及依變項 Y 的解釋程度較大。在估計上，則是以 a_1*b 或 a_2*b 是否達到顯著，來判定 I 或 S 變項是否為中介模型的自變項。Reddy、Rhodes 與 Mulhall (2003) 即採用此模式，將教師對學生支持的 I 與 S 當作 X，進一步去預測中介變數自尊與依變項憂鬱間的關係。

(三) I&S 做為依變數

此類模型將 I&S 變成當作是依變項 Y，探討 X 與 M 如何影響特定變數的初始狀態與變動速度，是最常被使用的基本模型。在估計上，則是以 $a*b_1$ 或 $a*b_2$ 是否達到顯著，來判定該中介模式能否有效解釋 I 與 S 二個依變項。例如：Barnes、Reifman、Farrell 與 Dintcheff (2000) 以單次測量的父母支持做為 DV 及父母監控做為 M，探討前二者預測成年後酒精濫用的截距與斜率。

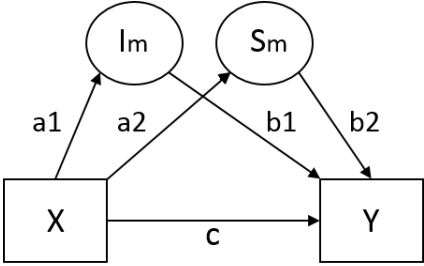
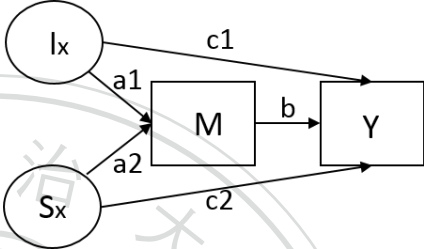
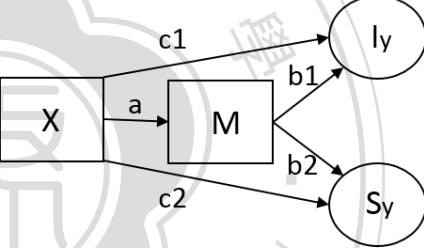
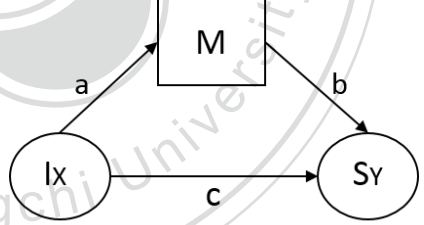
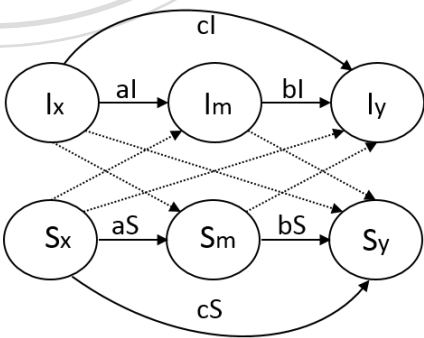
(四) I 為自變數，S 為依變數

此類模型是以 I 為 X、以 S 為 Y，探討截距如何轉變成斜率的過程。在估計上，則是以 $a*b$ 來判定 I 變項是否能透過中介變項，進而轉化為 S 變項。例如：Beauchaine、Webster-Stratton 與 Reid (2005) 在探討早發性行為問題時，將外向行為的 I 當作 X，親職參與做為 M，並以外向行為的 S 當作 Y，討論如何透過親職參與啟動外向行為的增加。

(五) I&S 做為自變數、中介變數與依變數

此類模型是以三條成長曲線為基礎，將 X、M、Y 各自視為一個獨立的變動系統，同時估計多條變動軌跡。這樣的平行歷程模型 (parallel process model)，不但能處理個體間的結構變異，也能處理個體內的成長軌跡。此時，估計出的中介模型，乃是不同時間點特定效果的整合效果。在估計上，則是以 $a_I * b_I$ 或 $a_S * b_S$ 是否達顯著，來判定變動軌跡中的中介效果是否存在。例如：Simons-Morton、Chen、Abroms 與 Haynie (2004) 發現，父母阻止吸菸的行為斜率，會透過同儕吸菸斜率的中介，進而影響到成年後吸菸量的斜率。但此中介現象，並不存在於截距變項上。

表 2-3-1 潛在成長模式之截距與斜率變數在中介模型中的角色

變項 I 與 S 角色	模型圖示
I & S 做為中介變數	
I & S 做為自變數	
I & S 做為依變數	
I 為自變數，S 為依變數	
I & S 做為自變數、中介變數與依變數	

資料來源：研究者自行整理。

註：為求簡化，省略誤差項及其共變與符號。

第四節 自我迴歸取向的中介模型

在前述的討論中，可得出自我迴歸取向的中介模型乃最大程度符合因果推論條件的結論。是故，本節將介紹自我迴歸取向縱貫中介模型，依重複測量的波次數目可以分為：僅二波測量的半縱貫設計與三波以上的全縱貫設計。後續章節，將說明其原理、應用狀況與設計上的限制。

壹、半縱貫設計

若研究者只有二波測量資料，則可以採用的中介模型分析策略，即被稱為半縱貫設計 (half-longitudinal design)，每個變項當下的狀態，都受到自身前一個時間點的影響。在排除 AR 的效果之後， $X_1 \rightarrow M_2$ 的效果與 $M_2 \rightarrow Y_2$ 的效果，即構成跨時間中介效果 (Cole & Maxwell, 2003)。然而，二波的 AR 中介，在邏輯架構上，卻有明顯的瑕疵。中介模型在時間概念上至少有三個相異的時間點，X、M、Y 三個變數是依序發生的，先有 X 再透過 M 影響 Y。然而二波的設計無法滿足這樣的推論需求， X_1 對 M_2 的效果方向是正確的，但重點在於 M_1 對 Y_2 的效果，讓此模型無法再跨出因果推論的下一步，也就是第三個時間點。換言之， $a_{21} * b_{21}$ 在測量點不足的情況下，難以被視為完整的縱貫中介效果。是故，嚴格探究下來，二波的 AR 中介，其檢驗的目標是 X、M、Y 三個變數間的效果，是否能在 AR 的前提下跨時間成立，而非完整地進行因果推論，故又被稱為部分縱貫中介 (partial longitudinal mediation)。其模式概念圖如下：

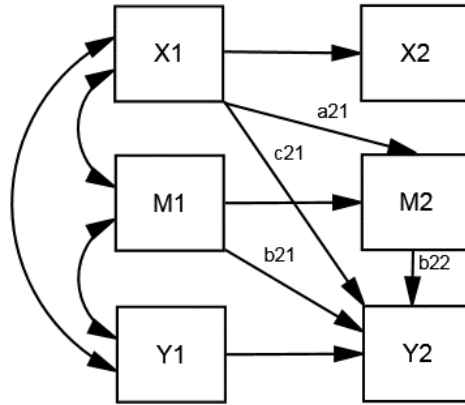


圖 2-4-1 二波追蹤資料之自我迴歸取向縱貫中介概念圖

(為求簡化，省略誤差項及其共變與符號)

由於重複測量的成本較高，在研究架構的考量下，研究者可以透過二波的重複測量就可以得到某種程度上跨時間效果的檢測，可做為更進一步的全縱貫設計之前導研究，可說是一種平衡性的考量。在分析中介效果上，有二種檢視的途徑，其一是以估計 a_{21} 與 b_{21} 是否達顯著，來判定縱貫中介效果是否存在 (Deacon, Kieffer, & Laroche, 2014; Spinrad, Eisenberg, Gaertner, Popp, Smith, Kupfer, & Hofer, 2007; Wood, Maltby, Gillett, Linley, & Joseph, 2008)；其二是以估計 a_{21} 與 b_{22} 是否達顯著來判定 (Feddes, Noack, & Rutland, 2009; Richter, Näswall, & Sverke, 2010)。

上述的第二種策略，當中的 b_{22} 其實是一個橫斷面的 $M \rightarrow Y$ 的係數，其實已經違反了縱貫設計的跨時間點效果的理念。倘若在半縱貫中介當中，納入 $X1 \rightarrow M1 \rightarrow Y1$ 以及 $X2 \rightarrow M2 \rightarrow Y2$ 的橫斷面影響做為控制變數，這符合 Mackinon (2008) 所提的雙向假設模型概念。然而，這種模型可能會有理論上的衝突，因為就邏輯上而言，它已經先否定橫斷中介的概念，另一方面，也可能造成縱貫中介效果的估計變異減少。因此，在實際分析上，自我迴歸取向的中介模型大多不納入橫斷面中介效果估計，而是採用殘差相關的方式取代。綜上所述，在僅有二波重複測量的情況下，建議採用第一種策略進行縱貫效果分析會是比较合理的選擇，也為下一節介紹的全縱貫設計提供進一步分析的基礎。

貳、全縱貫設計

基於修正 AR 半縱貫中介在第三時間點的推論困境，最合理的策略是 X、M、Y 三個變數皆進行三波測量，有著三條時間序列的此模型即稱為全縱貫設計 (full-longitudinal design)。以此概念為基礎的縱貫中介分析，將會是能同時滿足四個條件，且較具有因果推論性的設定架構 (Maxwell & Cole, 2007; Maxwell, Cole, & Mitchell, 2011; Mitchell, & Maxwell, 2013)。一般而言，Maxwell 等人的架構，將一個完整的縱貫中介架構為預測變數 X、中介變數 M、效標變數 Y 三者，都必須是測量三次的 3x3 九宮格設計 (邱皓政, 2017)。於此架構中，跨時間且兼具自我迴歸的中介效果，乃由 $a_{21} * b_{32}$ 所組成，也就是 $X_1 \rightarrow M_2 \rightarrow Y_3$ 的效果。至此，AR 取向的縱貫中介，擁有在邏輯時序與測量時序上都納入考量的設計。其模式概念圖如下：

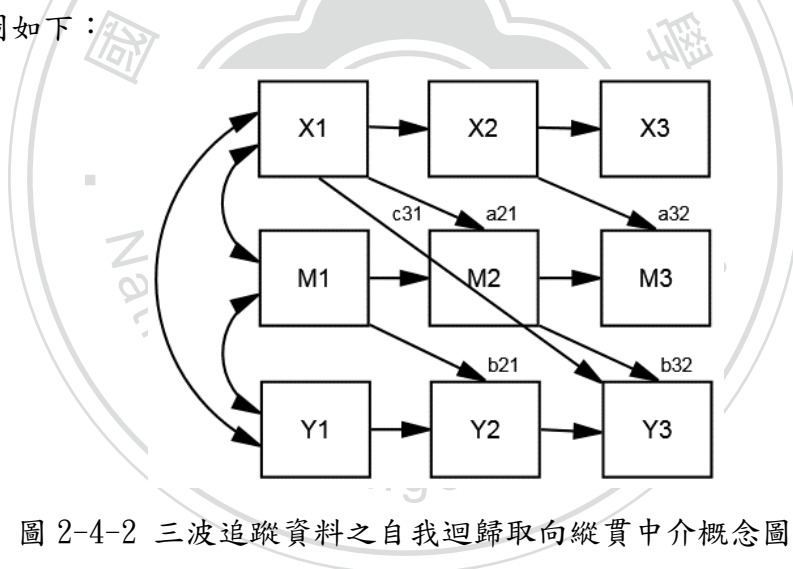


圖 2-4-2 三波追蹤資料之自我迴歸取向縱貫中介概念圖

(為求簡化，省略誤差項及其共變與符號)

從研究設計的觀點來看，對三個變數進行三波測量的時間與成本非常高，尤其是一份研究中可能還有其他變數的情況。然而，三波的全縱貫設計，因時間序列的存在，而能得到最完整的時序資料。

在縱貫設計之中，除了穩態、定態、時距、時點的議題之外，另一個也經常被討論的主題是延宕 (lag) 的期數問題。延宕是指在時間點 t 的自變數，對時間點 $t+n$ 的依變項產生效果時所間隔的測量點數目。換言之，時距與時點決定了整個研究的時間範圍長度，而測量次數的多寡及測量時間點的安排則決定了延宕的期數。在三波的全縱貫設計中，跨時間中介是指 $X_t \rightarrow M_{t+1} \rightarrow Y_{t+2}$ ，延宕期數為一期，也就是說自變數的影響力是發生在後一期的依變項上。然而在實務上，很可能因為測量時間點安排的緣故，導致 $X \rightarrow M$ 或 $M \rightarrow Y$ 的效果必須延宕二期或以上。此種情況下，三波的全縱貫設計就無法準確的估計跨時間中介效果。

縱貫資料展示了在原因與結果之間存在著延宕 (lag)，但效果很可能發生在各種延宕期數中。時間是持續流動且無法被操弄的，因此，研究者要根據理論、邏輯、前導研究及成本等因素，去決定測量的切入點 (時距&時點)。因此，可以推論：假設 $X_t \rightarrow M_{t+1}$ 的時間點時距是一年且測量二次，則 t 為現在、 $t+1$ 為一年後，則延宕期數為一期；若時距仍保持一年，但測量改成四次，此時的延宕期數就很有可能改變。因此，在時點的議題上，Reichardt (2011) 表示，若沒有在模型中納入足夠的時間點，即使是縱貫設計也仍會有偏誤，並據此認為三波重複測量的全縱貫設計根本，不足以估計縱貫中介 (仍有偏誤)。例如，任一時點的 M_t 可能不只受到 X_{t-1} 的影響，而是來自於所有先前時間點的 X_{t-n} 的總和。然而，研究者會根據自我相關隨著延宕期數增加而減低的概念，而選擇忽略；換言之，真正的因果關係，很可能超越研究者能夠檢驗的研究設計範疇。因此，Reichardt 並不相信估計中介效果上，有所謂正確或錯誤的延宕期數。將一段流動的時間切割成片段的測量點，本身就是一種化約論 (reductionism) 的概念，但也正因為測量的方式、次數會影響到中介效果的估計，在研究設計上，應以更多波次以求涵蓋的範圍擴大，來增進否證的效力。

第五節 強化中介模型因果推論的可能策略

從橫斷到縱貫的中介模式設計演變中，可以看出此乃交叉延宕（cross-lag）的延伸應用。然而，交叉延宕的檢驗當中，有個重要的觀念「若某個因果方向為真，則應會重複出現」並未被採用（Marshall et al, 2014）。在 Maxwell 等人的九宮格設計中，跨時間的中介效果僅需出現一次就宣稱其存在，這樣的想法對因果推論的正確性有所侷限（Reichardt, 2011），也無法驗證重複出現的假設。因此，本研究將擴大追蹤波段為五次，亦即 X、M、Y 都進行五次測量，將縱貫中介效果定義為 $a_{21} * b_{32}$ 、 $a_{32} * b_{43}$ 、 $a_{43} * b_{54}$ 。其模式的設定概念圖如下：

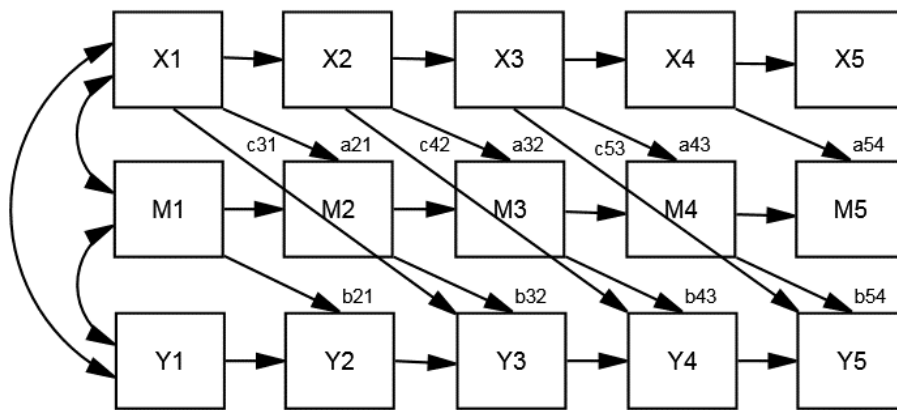


圖 2-5-1 五波追蹤資料之自我迴歸取向縱貫中介概念圖

（為求簡化，省略誤差項及其共變與符號）

此設計的優勢在於，對第二波開始的每一個變數，都會受到自身前一波狀態的預測，藉此設計排除掉自我預測的混淆效果。以第二波中介變數 M2 為例，排除了 M1 的影響後，即可得到第一波預測變數 $X1 \rightarrow M2$ 的淨效果。是故，三個縱貫中介效果 $a_{21} * b_{32}$ 、 $a_{32} * b_{43}$ 、 $a_{43} * b_{54}$ 都會是同時滿足三個條件的因果推論。

在影響方向的假設上，本研究僅有縱貫中介的假設，並不針對 Mackinnon (2008) 所提的雙重單向假設及交叉延宕假設進行驗證。理由在於：(1) 雙重單項假設包含了各波次內的橫斷中介 $X_t \rightarrow M_t \rightarrow Y_t$ 效果估計，這不合乎因果推論的時間先後概念，也徒增估計複雜性，本研究將以殘差相關取代；(2) 交叉延宕假設乃單向假設之延伸，除非單向假設無法被接受，才有需要進行交叉延宕假設的檢驗，以分析 $Y_t \rightarrow M_{t+1} \rightarrow X_{t+2}$ 這條中介機制的可能性。

在顯著性考驗部分，本研究將以前述拔靴法檢定三條中介效果的顯著性。倘若三條中介的顯著結果不盡相同，則表示此中介現象僅出現在特定的時間點。倘若三條縱貫中介效果皆為顯著，本研究將會依照定態的概念，進一步檢定 $a_{21} * b_{32}$ 、 $a_{32} * b_{43}$ 、 $a_{43} * b_{54}$ 等三條中介效果是否相等。技術上，將檢定 $a_{21} * b_{32}$ 、 $a_{32} * b_{43}$ 、 $a_{43} * b_{54}$ 被設為相等的模型是否適配。若適配，則表示此中介效果具有跨時間的穩定性；若否，則表示此中介效果在不同時間點的效果量不穩定。無論結果為何，對於任何實證研究所建構的理論模型，也就是其領域知識 (domain knowledge)，在內在效度上都能擁有極大的突破，也是對因果推論完備條件的強化。



第三章 研究方法

為達本研究之目的，乃探討相關理論與研究，並以資料分析演示，做為主要的研究方法。透過統計分析的過程與結果討論，建構自我迴歸取向的縱貫中介模型，並探討變項間的因果結構模式，以及縱貫中介模型在因果推論上所提供的資訊與限制。本研究取二份具有相異特性與主題的資料庫進行分析，以王金香(2010)所收集之國三學生數學模擬考資料為研究一，以中央研究院所釋出之「臺灣青少年成長歷程」資料庫為研究二，茲將研究資料與分析模型描述如下。

第一節 研究一資料說明

本研究透過實證資料之分析，進行縱貫中介模型的建構並探究因果推論上可能遭遇的問題與限制。研究一所使用之資料庫，為王金香(2010)博士論文所蒐集之五波重複測量之實證資料重新整理而得。該研究的原始目的在探討數學焦慮、數學學習動機與數學學業成就之縱貫模式，並以潛在成長模式做為分析架構，探討三個變項的變化趨勢及其跨時間關係。

在因果關係的推論上，學界認為數學焦慮與數學學習動機之間的關係有二種。部分研究指出，兩者之間具有正向相關，原因在於焦慮情緒具有激發(activate)的特質，可以引發人們的動機去處理負向事件；或是為了避免焦慮而被迫投入更多的努力去強化某些目標。在數學科目上，則是對數學的焦慮能夠引發學習數學的動機(李仁豪, 2005; Elliot & McGregor, 2001; Pekrun, Goetz, Titz, & Perry, 2002; Turner & Schallert, 2001)。然而，也有另一些研究顯示，數學焦慮與數學學習動機之間具有負向相關(魏麗敏, 1996; 蔡文標、許天威、蕭金土, 2003; 吳明隆、葛建志, 2006)。綜合上述文獻，本研究將透過不同因果順序的設定，套用橫斷中介、序列中介與縱貫中介的方式，逐一檢驗實證資料較支持哪一種因果推論，以及不同的中介模型在推論因果時的完備條件。

在研究樣本部分，透過立意抽樣，徵詢台中縣使用相同模擬考卷版本的八所學校同意，以其國中三年級學生進行問卷調查並記錄模擬考之數學科成績。每校

均取國三 3~6 個班級，共 32 個班級，1056 位學生，在模擬考前二週填寫問卷，為期五次。以此五次測量之縱貫資料做為正式樣本，在刪去有遺漏值的觀察值後，可用於統計分析的樣本為 987 人。

研究變項與測量工具部分，資料中共有三個研究變數，茲說明如下：

一、數學焦慮

數學焦慮變數，乃取自王金香（2010）自編「數學焦慮量表」中的變數資料。數學焦慮量表包含六個子因素，共 24 道題目，採李克特四點量尺，選項從「非常不同意」、「不同意」、「非同意」、「非常同意」分為四個等級，給予 1、2、3、4 分。得分越高者，即表示該個體所知覺到的數學焦慮程度越高。該量表之信效度表現良好，六個子因素的組合信度介於.75~.84，平均變異抽取量介於.43~.57，驗證性因素分析結果顯示資料與理論結構達到適配（ $\chi^2=1032.34^{***}$, $df=222$, $GFI=.93$, $CFI=.97$, $RMSEA=.058$ ）。根據王金香（2010）的理論架構，本研究以數學焦慮量表之平均數為自變項 X，並以五次測量結果，架構時間序列資料。

二、數學學習動機

數學學習動機變數，乃取自王金香（2010）自編「數學學習動機量表」中的變數資料。數學焦慮量表包含四個子因素，共 16 道題目，採李克特四點量尺，選項從「非常不同意」、「不同意」、「非同意」、「非常同意」分為四個等級，給予 1、2、3、4 分。得分越高者，即表示該個體的數學學習動機越高。該量表之信效度表現良好，四個子因素的組合信度介於.67~.86，平均變異抽取量介於.34~.60，驗證性因素分析結果顯示資料與理論結構達到適配（ $\chi^2=341.31^{***}$, $df=88$, $GFI=.96$, $CFI=.99$, $RMSEA=.052$ ）。根據王金香（2010）的理論架構，本研究以數學學習動機量表之平均數為中介變項 M，並以五次測量結果，架構時間序列資料。

三、數學學業成就

數學學業成就部分，為求控制所有樣本都是在同樣標準下評分，挑選使用相同版本模擬考卷的學校學生進行資料收集。數學學業成就乃取國三第二、三、四、五次模擬考與基本學力測驗為效標，依照樣本在此五次測驗的數學模擬考得分。分數越高，即表示其數學學業成就越高。每次測驗的數學試題約有 20 至 30 個題目，並依據答對題數給予固定的量尺分數。根據王金香（2010）的理論架構，本研究以數學學業成就為依變項 Y，並以五次測量結果，架構時間序列資料。

第二節 研究二資料說明

研究二的分析資料取自於「臺灣青少年成長歷程研究」，簡稱「臺灣青少年計畫」(Taiwan Youth Project, TYP)的長期追蹤研究資料庫，係由中央研究院社會學研究所家庭與生命歷程研究小組對臺北市、臺北縣及宜蘭縣的國中學生所進行的長期固定樣本追蹤研究設計，問卷內容包含家庭、學校和社區等三個社會制度為主軸。

研究二選取的研究樣本是以 TYP 資料庫所釋出的第一波、第四波及第八波國一樣本為主（伊慶春，2012a；伊慶春，2012b；伊慶春，2014），由於是固定樣本追蹤，故測量時間分別在國一、高一及大二。選取在第一波、第四波及第八波都有資料者，在刪去遺漏值後，總計共得到 1529 位有效樣本。背景變數方面，就讀國中所在地分別為臺北市 36.6%、新北市 37.5%及宜蘭縣市 26.0%。性別部分，男性 50.4%，女性 49.6%。

研究二參考賴俊生、謝亞恆（2017）的研究架構與分析結果，該研究以國三樣本的第三波資料為樣本，並以親子互動為預測變數 X、自尊為中介變數 M、憂鬱為 Y 進行橫斷中介模型的建構與檢驗。在理論背景方面，過往研究指出親子互動乃是由父母與子女間交互作用的歷程，彼此分享情感、意見、態度及喜好，知覺彼此傳遞的訊息與行為表現(吳欣璇、金瑞芝，2011)。親子互動愈佳，可增強

子女對自我的評價，進而提昇心理幸福感；反之，親子互動若差，可能不利於子女的人際關係適應與發展，更會造成情緒低落，長期來看將有害人格發展(黃韞臻、林淑惠，2012；顏綵思、魏麗敏，2005；Ferguson, 2013)。另一方面，青少年學生具有高度自尊時，能對自己產生較為積極正面的態度與評價，並與家庭、同儕等他人有良好的互動，相信自己具備解決問題的能力，因此較不容易產生情緒抑鬱的現象(Falci, 2008；林杏足，2003；陳毓文，2004；蘇曉憶、戴嘉南，2008；吳齊殷、黃鈺婷，2010；Dori & Overholster, 1999；Meadows, Brown, & Elder, 2006；Paxton, Neumark-Sztainer, Hannan, & Eisenberg, 2006)。

在研究變項的測量上，茲說明如下：

一、親子互動

親子互動的題目分為和母親互動及和父親互動，包含：會問他對重要事情的看法、會注意聽他的看法或想法、會關心他、讓他覺得窩心、以他為榜樣等五個題目。在測量尺度方面，第一波測量為七點尺度，分數愈高代表互動程度愈低。第四波及第八波資料則為四點量尺，分數愈高代表互動程度愈低。研究二以上述親子互動題目之平均數為預測變數 X ，並以三次測量結果架構時間序列資料。

二、自尊

自尊的題目包含：我沒有辦法解決自己的某些問題、我沒有辦法控制發生在我身上的事情、我有時候覺得自己很沒有用、有時候我認為自己一無是處等四個題目，三次測量皆為四點尺度，分數愈高代表程度愈低，由於都是反向問題，所以分數越高代表自尊越高。研究二以上述自尊題目之平均數為中介變數 M ，並以三次測量結果架構時間序列資料。

三、憂鬱

憂鬱的題目包含孤獨感和情緒抑鬱二大類。孤獨感包含：過去一個星期，你有沒有孤獨的情形？過去一個星期，你有沒有鬱卒的情形？過去一個星期，你有沒有擔心過度的情形？等三個題目。情緒抑鬱則包含：過去一個星期，你有沒有

感覺身體某些部位虛弱的情形？過去一個星期，你有沒有不想活了的情形？等二個題目。三次測量皆為五點尺度，分數越高表示憂鬱的情況越嚴重。研究二以上述憂鬱測量題目之平均數為效標變數 Y ，並以三次測量結果架構時間序列資料。

第三節 分析模型

本研究擬定以自我迴歸取向的縱貫中介模型做為分析五波次縱貫性資料的統計方法，並以單向中介假設 (single direction hypothesis) 做為模型設定之依據，以「auto-regression mediation with 5 waves, ARMe5」稱之，如圖 3-1。以下茲從理論概念與數理模式說明之。

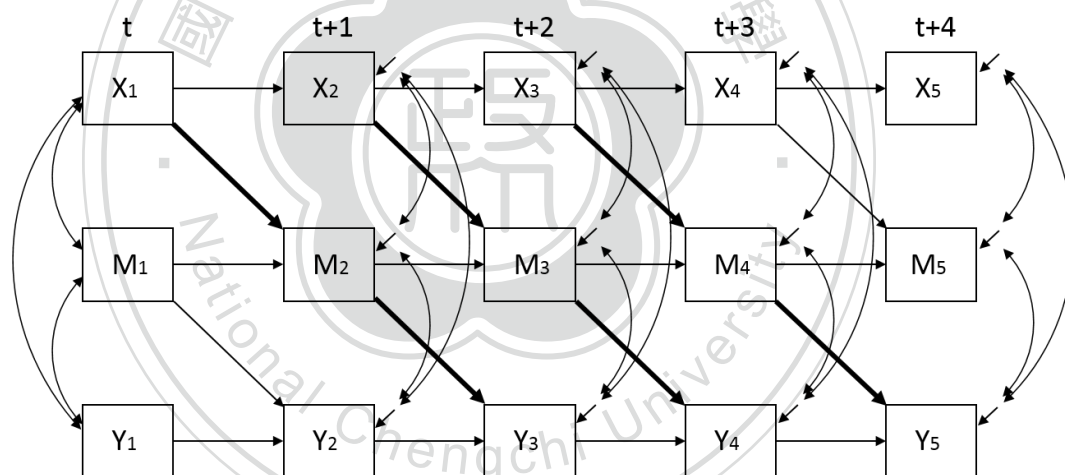


圖 3-3-1 單向 ARMe5 模型 (僅有縱貫中介)

(粗直線為構成間接效果的參數，省略 $X \rightarrow Y$ 的直接效果以利圖示)

壹、理論概念

根據文獻探討的結果，在估計中介模型的因果推論時有七個重要的考量：

(一) 納入測量時序的概念

此乃超越橫斷設計，踏入縱貫設計的第一步。除了透過理論與邏輯推導出變項在概念上的先後順序，也在實際的測量上有先後順序。換言之，要有相異時間點的 X_t 、 M_{t+1} 、 Y_{t+2} ，也就是滿足序列設計的需求。

(二) 必須有自我迴歸做為控制

在因果推論當中，為求取自變項的淨效果，必須要控制住混淆變項。在時間序列當中，依變項的最大混淆變項即為自身的前一期狀態。換言之，要估計 $X_t \rightarrow Y_{t+1}$ 就必須排除掉 $Y_t \rightarrow Y_{t+1}$ 的效果。因此，多個基於重複測量所建立的時間序列資料，乃自我迴歸做為控制的基礎條件。

(三) 增加測量波次

時間是一個連續的概念，且無法被操弄。研究者僅能決定觀察的次數，並透過各種方式去推估最佳的測量時點。在任何條件下，增加測量波次都有助於正確地估計縱貫中介的效果。

(四) 設定延宕期數為 1

延宕期數並非統計問題，而是測量與理論問題。在完美情況下，設定良好的測量時點安排，就可以確保延宕期數為 1。然而，完美並不可得，故需增加測量時點，以求更細緻的中介效果估計。因此，在沒有其他理由的情況下，僅需設定延宕期數為 1。

(五) 根據邏輯時序設定為單一單向中介假設

縱貫中介的基本架構是三條時間序列資料，故在效果的估計上有三：僅估計縱貫中介效果的單一單向假設、同時估計橫斷與縱貫中介的雙重單向假設，以及估計交叉延宕效果。根據邏輯及避免稀釋變異的理由，本研究捨去橫斷中介效果，僅估計縱貫中介效果。再者，基於研究一與研究二的相關理論推導，可知中介乃

特定方向，並不需要進行交叉延宕的分析。綜合上述理由，本研究採取單純估計縱貫效果的單一單向假設為模型設計之依據。

(六) 縱貫中介的效果可能會重複出現

基於研究理論的探討，研究者會想要知道縱貫中介的效果，在整個測量的時間範圍內，是僅出現一次或是會重複出現。縱貫中介效果的探索，無論其在不同時點間是否存在，對於理論與實務的因果關係而言，都將是很重要的進展。

(七) 若前述條件成立，進一步檢驗縱貫中介效果是否相等

倘若縱貫中介的效果僅在特定時間點當中出現一次，則表示該時間點乃重要的理論關鍵時間。若是縱貫中介的效果重複出現，則表示其為一種穩定出現的現象，此時可進一步檢驗中介效果是否跨時間相等。若是，即則表示此縱貫中介乃是一個弱定態的現象，其因果現象可發生在任何時間點之間。

貳、數理模式

本研究的 ARMe5 模型根據單一單向中介假設，將估計出多條橫跨二期的 $X_t \rightarrow M_{t+1} \rightarrow Y_{t+2}$ 縱貫中介效果，如圖 3-2-1 所表示，並可由方程式 3-1 至 3-3 等三條自我迴歸方程式定義之（邱皓政，2017）。

$$X_{t+1} = \alpha_{X_{t+1}} + \bar{\theta}_{X_{t+1},t} X_t + [\varepsilon_{X_{t+1}}] \quad (3-1)$$

$$M_{t+1} = \alpha_{M_{t+1}} + \theta_{M_{t+1},X_t} X_t + \bar{\theta}_{M_{t+1},t} M_t + [\varepsilon_{M_{t+1}}] \quad (3-2)$$

$$Y_{t+2} = \alpha_{Y_{t+2}} + \theta_{Y_{t+2},M_{t+1}} M_{t+1} + \bar{\theta}_{Y_{t+2},t+1} Y_{t+1} + [\varepsilon_{Y_{t+2}}] \quad (3-3)$$

在方程式當中， $\bar{\theta}_{X_{t+1},t}$ 、 $\bar{\theta}_{M_{t+1},t}$ 、 $\bar{\theta}_{Y_{t+2},t+1}$ 為自我迴歸係數（係數上方標示 \rightarrow 以利辨識）， θ_{M_{t+1},X_t} 和 $\theta_{Y_{t+2},M_{t+1}}$ 則為落後一期的縱貫直接效果。

特定時間點之間間接效果 (ind_{ab})，則可由兩個直接效果相乘得到，如方程式 3-4 表示。

$$ind_{ab} = \theta_{M_{t+1},X_t} \times \theta_{Y_{t+2},M_{t+1}} \quad (3-4)$$

再者，因自我迴歸以 X_t 、 M_t 、 Y_t 做為下一期 X_{t+1} 、 M_{t+1} 、 Y_{t+1} 的自變項，故第 t 期後的各期估計殘差間將有一定程度的相關，必須納入估計，第 $t+1$ 期的殘差變異與共變數矩陣可表示如下：

$$\psi_{t+1} = \begin{bmatrix} \psi_{XX(t+1)} & & \\ \psi_{XM(t+1)} & \psi_{MM(t+1)} & \\ \psi_{XY(t+1)} & \psi_{MY(t+1)} & \psi_{YY(t+1)} \end{bmatrix} \quad (3-5)$$

參、資料分析

本研究根據文獻探討結果，奠定研究架構與選取資料庫進行分析。描述統計與相關分析係採用 SPSS 22 版軟體，進階分析係以路徑模式檢驗各個潛在變數之間的直接效果與中介效果，並以拔靴法 (bootstrap method)、檢驗中介係數之顯著性，使用軟體為 AMOS 21 版。

第四章 結果與討論

本章旨在根據研究目的，針對實證調查資料所得結果加以分析討論，以瞭解各種中介模型在因果推論上的合理程度。本章主要內容包括：第一節呈現研究一：國三數學科模擬考資料分析結果。第二節呈現研究二：青少年成長歷程資料庫分析結果。

第一節 研究一：國三數學科模擬考資料分析結果

本節主要在呈現研究一所使用之國三數學科模擬考資料分析結果，分別就描述性統計、橫斷中介分析、序列中介分析、縱貫中介分析等四個部分進行探討，並針對上述分析之因果推論完備條件進行檢核。

壹、描述性統計

國三學生在數學焦慮、數學學習動機及數學成就上的描述統計方面，由於本研究的測量時間在國三共六次的模擬考範圍中，故共有數學焦慮 1~5（從第二次模擬考開始測量至第六次模擬考共計五次， $X_1 \sim X_5$ ）、數學學習動機 1~5（從第二次模擬考開始測量至第六次模擬考共計五次， $M_1 \sim M_5$ ）及數學成就 0~5（從第一次模擬考開始測量至第六次模擬考共計六次， $Y_0 \sim Y_5$ ）。第一次模擬考時只有測量數學考試成績，並無測量數學焦慮及數學學習動機，故不將 Y_0 納入分析架構當中。

由前述討論數學焦慮與數學學習動機的理论可知，焦慮與動機之間的因果順序是有不同說法的，故研究一的各種中介模型分析皆有二種版本：其一， $X \rightarrow M \rightarrow Y$ ；其二， $M \rightarrow X \rightarrow Y$ 。此外，過去研究指出，數學成績的低落與焦慮之間有相關存在，故研究一以第一次模擬考成績 Y_0 進行切割，以第一四分位差 Q_1 以下者為低成就考生為樣本進行後續分析。

首先呈現全體樣本描述統計資訊，包含平均數、標準差、相關係數等。在數學焦慮的表現上，X1~X5 的平均數介於 1.91 至 1.99 之間，標準差介於 0.45 至 0.48 之間，數據在跨時間上具有穩定性。在數學學習動機的表現上，M1~M5 的平均數介於 2.59 至 2.71 之間，標準差介於 0.54 至 0.56 之間，數據在跨時間上具有穩定性。在數學模擬考成績方面，Y0~Y5 的平均數介於 28.17 分至 33.99 分，標準差介於 5.29 至 8.44 之間，雖然平均數有上下波動，但標準差則隨著時間逐漸增加，顯示離散程度越來越大。在相關係數部分，同一個變項在不同時間點間的相關都達到顯著水準 ($p < .05$)，值得注意的是不顯著 ($p > .05$) 的相關係數大都集中在 X 與 M，也就是考試焦慮與動機的相關係數上，詳細數據請見表 4-1-1。

接著呈現低成就樣本的描述統計資訊，包含平均數、標準差、相關係數等。數學模擬考總分為 60 分，全體樣本之 Y0 平均數為 28 分，Q1 為 17 分，故研究一所定義的低成就者為第一次數學模擬考分數在 17 分以下者。在數學焦慮的表現上，X1~X5 的平均數介於 2.00 至 2.10 之間，標準差介於 0.44 至 0.46 之間，數據在跨時間上具有穩定性。在數學學習動機的表現上，M1~M5 的平均數介於 2.42 至 2.56 之間，標準差介於 0.55 至 0.61 之間，數據在跨時間上具有穩定性。在數學模擬考成績方面，Y0~Y5 的平均數介於 15.87 分至 21.21 分，標準差介於 5.29 至 8.44 之間，雖然平均數有上下波動，但標準差則隨著時間逐漸增加，顯示離散程度越來越大。在相關係數部分，同一個變項在不同時間點間的相關都達到顯著水準 ($p < .05$)，值得注意的是不顯著 ($p > .05$) 的相關係數大都集中在 X 與 Y，也就是考試焦慮與數學成績的相關係數上，詳細數據請見表 4-1-2。

表 4-1-1 全體樣本在研究變項上之描述統計 (N=987)

	平均數	標準差	X1	X2	X3	X4	X5	M1	M2	M3	M4	M5	Y0	Y1	Y2	Y3	Y4
X1	1.98	0.47	1														
X2	1.99	0.45	.73	1													
X3	1.91	0.48	.68	.73	1												
X4	1.97	0.46	.64	.70	.73	1											
X5	1.97	0.46	.57	.64	.66	.73	1										
M1	2.71	0.54	.11	.09	.15	.10	.10	1									
M2	2.66	0.55	.07	.11	.12	.10	.09	.72	1								
M3	2.62	0.56	<u>-.03</u>	<u>.02</u>	.10	<u>.03</u>	<u>.02</u>	.67	.70	1							
M4	2.63	0.55	<u>.00</u>	<u>.02</u>	.07	.08	.07	.63	.68	.73	1						
M5	2.59	0.54	<u>-.02</u>	<u>.02</u>	.07	<u>.05</u>	.11	.60	.64	.70	.69	1					
Y0	28.87	13.64	-.29	-.27	-.21	-.20	-.22	.24	.24	.30	.29	.28	1				
Y1	28.17	12.39	-.26	-.24	-.19	-.20	-.21	.23	.24	.30	.27	.27	.88	1			
Y2	29.23	13.05	-.27	-.25	-.19	-.20	-.21	.23	.26	.30	.29	.29	.89	.88	1		
Y3	33.99	14.62	-.27	-.25	-.20	-.21	-.20	.24	.28	.30	.29	.28	.89	.87	.89	1	
Y4	28.83	15.09	-.29	-.27	-.21	-.22	-.21	.23	.26	.31	.29	.28	.88	.86	.89	.90	1
Y5	31.78	15.28	-.28	-.25	-.20	-.20	-.20	.27	.29	.33	.31	.31	.89	.87	.88	.90	.90

註：底線標示之相關係數為未達顯著水準 ($p > .05$)，其餘相關係數皆達顯著水準 ($p < .05$)

表 4-1-2 低數學成就樣本在研究變項上之描述統計 (N=443)

	平均數	標準差	X1	X2	X3	X4	X5	M1	M2	M3	M4	M5	Y0	Y1	Y2	Y3	Y4
X1	2.10	0.46	1														
X2	2.10	0.44	.73	1													
X3	2.00	0.46	.63	.68	1												
X4	2.05	0.44	.62	.67	.67	1											
X5	2.05	0.44	.52	.58	.59	.67	1										
M1	2.56	0.58	.20	.22	.29	.18	.20	1									
M2	2.49	0.59	.20	.31	.30	.23	.22	.72	1								
M3	2.44	0.61	.11	.21	.33	.19	.17	.67	.72	1							
M4	2.44	0.55	.13	.20	.27	.23	.20	.65	.68	.73	1						
M5	2.42	0.58	.09	.18	.23	.16	.27	.62	.65	.68	.71	1					
Y0	16.20	5.29	-.18	-.12	<u>-.09</u>	-.13	-.12	.19	.14	.16	.13	.17	1				
Y1	17.86	6.71	<u>-.09</u>	<u>-.03</u>	<u>-.09</u>	-.13	<u>-.09</u>	.19	.18	.20	.13	.17	.45	1			
Y2	18.00	6.73	-.15	-.11	<u>-.06</u>	-.15	<u>-.08</u>	.18	.21	.23	.20	.22	.54	.55	1		
Y3	21.21	7.40	-.11	-.10	<u>-.06</u>	-.13	<u>-.04</u>	.21	.23	.20	.14	.18	.51	.53	.61	1	
Y4	15.87	8.51	-.18	-.16	-.14	-.18	-.11	.20	.19	.24	.18	.20	.52	.48	.62	.65	1
Y5	18.33	8.44	-.16	-.08	-.08	-.13	<u>-.05</u>	.27	.28	.29	.24	.27	.54	.53	.62	.64	.64

註：底線標示之相關係數為未達顯著水準 ($p > .05$)，其餘相關係數皆達顯著水準 ($p < .05$)

貳、橫斷中介分析

橫斷中介分析，係以同一波測量到的變項進行中介分析。由於在理論上存在二種解釋考試焦慮、學習動機的因果論述，故此部份的分析將分為二大類模型、分別進行 1~5 波次的檢驗，共計 10 個橫斷中介模型的分析。

橫斷模型 A：數學焦慮_t → 數學學習動機_t → 數學成就_t, t = 1~5

橫斷模型 B：數學學習動機_t → 數學焦慮_t → 數學成就_t, t = 1~5

橫斷中介的分析，將呈現 $X_t \rightarrow M_t$ 的標準化迴歸係數(γ_t)、 $M_t \rightarrow Y_t$ 的標準化迴歸係數(β_t)、中介效果($\gamma_t * \beta_t$)及其 95%信賴區間、BIC 值(χ^2)。中介效果的顯著性，將視其 95%信賴區間是否包含 0 而定，若包含 0 則判定為不顯著，若不包含 0 則判定為顯著。每一個波次的資料較為適配橫斷模型 A 或橫斷模型 B，則以 BIC 值較小者為適配最佳選擇。

在橫斷模型 A 的分析當中，各波的 γ_t 介於.20 至.33 之間，並且皆達到顯著水準($p < .001$)；各波的 β_t 介於.18 至.27 之間，並且皆達到顯著水準($p < .01$)；中介效果 $\gamma_t * \beta_t$ 介於.039 至.074 之間，由於其 95% CI 皆不包含 0 皆視為顯著。詳細資料請見表 4-1-3。

表 4-1-3 研究一各波次橫斷模型 A 中介分析

	γ_t	β_t	$\gamma_t * \beta_t$	95% CI	BIC
第一波	.20***	.19**	.039	.020~.075	38.88
第二波	.31***	.21***	.067	.040~.104	46.17
第三波	.33***	.20***	.066	.035~.104	38.36
第四波	.23***	.18***	.041	.020~.071	54.15
第五波	.27***	.27***	.074	.042~.114	39.07

註：ns $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$

在橫斷模型 B 的分析當中，各波的 γ_t 介於.20 至.33 之間，並且皆達到顯著水準 ($p < .001$)；各波的 β_t 介於-.05 至-.18 之間，僅第二波與第四波資料達到顯著水準 ($p < .05$)，其餘皆未達顯著水準 ($p > .05$)；中介效果 $\gamma_t * \beta_t$ 介於-.015 至-.033 之間，然而其 95% CI 在第一波、第三波及第五波包含 0 故視為不顯著。詳細資料請見表 4-1-4。

由於橫斷模型 A 與橫斷模型 B 二者之間使用的是同一批資料，故彼此可視為競爭模型。從橫斷模型 A 觀之， γ_t 的數據顯示數學焦慮對數學學習動機有正向預測能力且一致性地達到顯著。 β_t 的數據顯示數學學習動機對數學成績有正向預測能力且一致性地達到顯著。再者，中介效果的分析顯示，橫斷模型 A 的中介效果在各個波次都一致性的顯著，表示橫斷模型 A 背後所代表的理論在每一波次的測量資料中都有相似的解釋程度。

而從橫斷模型 B 觀之， γ_t 的數據顯示數學學習動機對數學焦慮有正向預測能力且一致性地達到顯著。 β_t 的數據則顯示出非常不一致的結果，在第一波、第三波及第五波當中數學學習動機對數學焦慮並無顯著的預測效果，而在第二波及第四波當中卻顯示數學焦慮對於數學成就有顯著的負向預測效果。再者，中介效果的分析顯示，橫斷模型 B 的中介效果僅在第二波、第四波成立，其餘波次的中介效果並未達到顯著，表示模型 B 對於解釋每一波次的測量資料並無相似的解釋程度。

在檢驗資料比較適配於哪一種競爭模型方面，採用 BIC 值為指標，BIC 值越小表示越具備簡效性。從每一波次當中，對比橫斷模型 A 與橫斷模型 B 的 BIC 值可發現，除去第四波的 BIC 值差異極小之外，其餘所有波次都是模型 A 的 BIC 值小於模型 B，顯示就每波次測量資料而言，選擇橫斷模型 A 會是較能夠通過否證的因果推論。各波次詳細 BIC 值請見表 4-1-3 及表 4-1-4。

表 4-1-4 研究一各波次橫斷模型 B 中介分析

	γ_t	β_t	$\gamma_t * \beta_t$	95% CI	BIC
第一波	.20***	-.09ns	-.019	-.046~.001	51.80
第二波	.31***	-.11*	-.033	-.067~-.005	61.99
第三波	.33***	-.06ns	-.019	-.047~.011	55.04
第四波	.23***	-.18***	-.041	-.068~-.021	56.82
第五波	.27***	-.05ns	-.015	-.045~.011	71.63

註：ns $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$

若從因果推論的完備條件論之，雖然橫斷模型 A 與橫斷模型 B 之間，透過上述標準化迴歸係數和中介效果可看出，橫斷模型 A 相對於橫斷模型 B 會是更佳解。相對來看，橫斷模型 A 有較小的 BIC 值，故其邏輯時序具有實證數據的支持，而橫斷模型 B 則在 BIC 的表現較差，其邏輯時序並不若橫斷模型 A 般獲得資料的支持。然而，二者都屬於橫斷中介模型，其研究設計僅有考量到「有共變關係」和「有邏輯時序」二點上，並未滿足「有測量時序」及「無混淆變數」的要求。換言之，無論在橫斷中介模型中獲得何種結果，即使完全符合研究者所預期的目標或得到良好的數據結果，也僅是滿足前二個條件，無論是橫斷模型 A 或橫斷模型 B 在橫斷的研究設計下，距離完備的因果推論還有非常大的改善空間。

表 4-1-5 研究一橫斷中介的因果推論完備條件檢核表

因果推論完備條件	檢核結果	說明
1. 有共變關係，變數之間具有顯著的相關。	具備	橫斷模型 A 與橫斷模型 B 的變數之間相關係數都達到顯著水準。
2. 有邏輯時序，透過理論或邏輯推演以建構理論模型。	部分具備	<p>1. 橫斷模型 A 與橫斷模型 B 各自有相對應的理論、邏輯架構及過往研究的證據引導。</p> <p>2. 在橫斷模型 A 當中，所有波次內的γ_t係數與β_t係數、中介效果等都具有一致且顯著的表現。但在橫斷模型 B 當中，各波次之間的γ_t係數與β_t係數、中介效果等在顯著性結果上有所差異。</p> <p>3. BIC 的比較顯示橫斷模型 A 的邏輯時序較為資料所支持。</p>
3. 有測量時序，在不同時間點測量研究變數。	不具備	橫斷模型 A 與橫斷模型 B 皆為橫斷中介模型，所有變數都在同一時間內測量。
4. 無混淆變數，有對自我迴歸現象進行控制。	不具備	橫斷模型 A 與橫斷模型 B 的所有變數皆無針對自我迴歸現象進行控制。

參、序列中介分析

序列中介分析，係以不同波次測量到的變項進行中介分析，中介模型當中的預測變數、中介變數及效標變數皆來自前、中、後三個不同波次的測量。如同前述，由於在理論上存在二種解釋考試焦慮、學習動機的因果論述，故此部份的分析也將分為二大類模型、分別設定啟始點為第 t 波、第 $t+1$ 波、第 $t+2$ 波進行檢驗，由於本資料只有五波資料，故啟始點為第一波至第三波，共計六個序列中介模型的分析。

序列模型 A：數學焦慮 _{t} → 數學學習動機 _{$t+1$} → 數學成就 _{$t+2$} ， $t = 1 \sim 3$

序列模型 B：數學學習動機 _{t} → 數學焦慮 _{$t+1$} → 數學成就 _{$t+2$} ， $t = 1 \sim 3$

序列中介的分析，將呈現 $X_t \rightarrow M_{t+1}$ 的標準化迴歸係數 ($\gamma_{t+1,t}$)、 $M_{t+1} \rightarrow Y_{t+2}$ 的標準化迴歸係數 ($\beta_{t+2,t+1}$)、中介效果 ($\gamma_{t+1,t} \times \beta_{t+2,t+1}$) 及其 95% 信賴區間、BIC 值 (χ^2)。中介效果的顯著性，將視其 95% 信賴區間是否包含 0 而定，若包含 0 則判定為不顯著，若不包含 0 則判定為顯著。每一個啟始點的資料較為適配模型 A 或模型 B，則以 BIC 值較小者為適配最佳選擇。

在序列模型 A 的分析當中，各啟始點的 $\gamma_{t+1,t}$ 介於 .20 至 .27 之間，並且皆達到顯著水準 ($p < .001$)；各啟始點的 $\beta_{t+2,t+1}$ 介於 .23 至 .24 之間，並且皆達到顯著水準 ($p < .01$)；中介效果 $\gamma_{t+1,t} \times \beta_{t+2,t+1}$ 介於 .046 至 .064 之間，由於其 95% CI 皆不包含 0 皆視為顯著。詳細資料請見表 4-1-6。

表 4-1-6 研究一各啟始點序列模型 A 中介分析

啟始點	$\gamma_{t+1,t}$	$\beta_{t+2,t+1}$	$\gamma_{t+1,t} \times \beta_{t+2,t+1}$	95% CI	BIC
第一波	.20***	.23**	.046	.021~.078	42.45
第二波	.21***	.24***	.051	.027~.085	53.49
第三波	.27***	.24***	.064	.036~.099	41.05

註：ns $p > .05$ ；* $p < .05$ ；** $p < .01$ ；*** $p < .001$

在模型 B 的分析當中，各啟始點的 $\gamma_{t+1,t}$ 介於.19 至.30 之間，並且皆達到顯著水準 ($p < .001$)；各波的 $\beta_{t+2,t+1}$ 介於-.10 至-.14 之間，並且皆達到顯著水準 ($p < .05$)；中介效果 $\gamma_{t+1,t} \times \beta_{t+2,t+1}$ 介於-.021 至-.041 之間，然而其 95% CI 在啟始點第一波、啟始點第三波包含 0 故視為不顯著，僅啟始點第二波的 95% CI 不包含 0 視為顯著。詳細資料請見表 4-1-7。

表 4-1-7 研究一各啟始點序列模型 B 中介分析

啟始點	$\gamma_{t+1,t}$	$\beta_{t+2,t+1}$	$\gamma_{t+1,t} \times \beta_{t+2,t+1}$	95% CI	BIC
第一波	.22***	-.10*	-.021	-.046~.002	57.09
第二波	.30***	-.14**	-.041	-.067~-0.018	58.96
第三波	.19***	-.13**	-.025	-.047~-0.001	77.76

註：ns $p > .05$ ；* $p < .05$ ；** $p < .01$ ；*** $p < .001$

由於序列模型 A 與序列模型 B 二者之間使用的是同一批資料，故彼此可視為競爭模型。從序列模型 A 觀之， $\gamma_{t+1,t}$ 的數據顯示數學焦慮對數學學習動機有正向預測能力且一致性地達到顯著。 $\beta_{t+2,t+1}$ 的數據顯示數學學習動機對數學成績有正向預測能力且一致性地達到顯著。再者，中介效果的分析顯示，序列模型 A 的中介效果在各個啟始點的中介模型中都一致性的顯著，表示序列模型 A 背後所代表的理論在每個啟始點的序列資料中都有相似的解釋程度。

而從序列模型 B 觀之， $\gamma_{t+1,t}$ 的數據顯示數學學習動機對數學焦慮有正向預測能力且一致性地達到顯著。 $\beta_{t+2,t+1}$ 的數據則顯示數學焦慮對於數學成就有顯著的負向預測效果。再者，中介效果的分析顯示，序列模型 B 的中介效果僅在啟始點第二波成立，其餘啟始點的中介效果並未達到顯著，表示序列模型 B 對於解釋每個啟始點的序列資料中並無相似的解釋程度。

BIC 值是由觀察值與期望值之間的差距計算而來，故 BIC 值越大表示理論與實際資料差距越大。在每個啟始點當中，對比序列模型 A 與序列模型 B 的 BIC 值可發現，所有啟始點都是序列模型 A 的 BIC 值小於序列模型 B，顯示就每個啟始

點的序列資料而言，選擇序列模型 A 會是較能夠通過否證的因果推論。各啟始點詳細 BIC 值請見表 4-1-6 及表 4-1-7。

若從因果推論的完備條件論之，雖然序列模型 A 與序列模型 B 之間，透過上述標準化迴歸係數和中介效果可看出序列模型 A 相對於序列模型 B 會是更佳解。相對來看，序列模型 A 有較小的 BIC 值，故其邏輯時序具有實證數據的支持，而序列模型 B 則在 BIC 值的表現較差，其邏輯時序並不若序列模型 A 般獲得資料的支持。然而，二者都屬於序列中介模型，其研究設計不但滿足橫斷設計「有共變關係」和「有邏輯時序」二點，還滿足「有測量時序」的要求，已經將因果推論中非常重要的時間因素納入，但依然無法回應「無混淆變數」的要求。

總體而言，相對於橫斷中介，序列中介在因果推論的完備條件上又往前踏出一步，在研究設計中落實了時間因素的考量。然而，如同文獻討論中所述，任何因果關係的推論都必須控制混淆變數，而最需要控制的變數則是該變數的前一期狀態，也就是自我迴歸，這是序列中介仍力有未逮之處。

表 4-1-8 研究一序列中介的因果推論完備條件檢核表

因果推論完備條件	檢核結果	說明
1. 有共變關係，變數之間具有顯著的相關。	具備	序列模型 A 與序列模型 B 在變數之間的關係數上都有顯著的情況。
2. 有邏輯時序，透過理論或邏輯推演以建構理論模型。	部分具備	<p>1. 序列模型 A 與序列模型 B 各自有相對應的理論、邏輯架構及過往研究的證據引導。</p> <p>2. 在序列模型 A 當中，所有啟始點的$\gamma_{t+1,t}$係數與$\beta_{t+2,t+1}$係數、中介效果等都具有一致且顯著的表現。在序列模型 B 當中，所有啟始點的$\gamma_{t+1,t}$係數與$\beta_{t+2,t+1}$係數在顯著性上有一致表性，但在中介效果的結果上有所差異。</p> <p>3. BIC 的比較顯示序列模型 A 的邏輯時序較為資料所支持。</p>
3. 有測量時序，在不同時間點測量研究變數。	具備	序列模型 A 與序列模型 B 皆為序列中介模型，所有變數都在三個不同的時間點進行測量。
4. 無混淆變數，有對自我迴歸現象進行控制。	不具備	序列模型 A 與序列模型 B 的所有變數皆無針對自我迴歸現象進行控制。

肆、縱貫中介分析

縱貫中介分析，乃結合自我迴歸控制與序列中介而成。中介模型當中的預測變數、中介變數及效標變數，都各自有完整的追蹤測量以形成時間序列資料。故在分析上，除了預測變數、中介變數及效標變數皆來自前、中、後三個不同波次的測量，每一個變數也都同時受到自身前一期狀態的預測，在排除自我迴歸的影響後，所得到即是跨時間的淨效果。

在縱貫中介的分析中，數學焦慮有從第二次模擬考到第六次模擬考共五波測量形成時間序列，數學學習動機也從第二次模擬考到第六次模擬考共五波測量形成時間序列，而數學成就則有從第二次模擬考到第六次模擬考共六波測量形成時間序列。如同前述，由於在理論上存在二種解釋考試焦慮、學習動機的因果論述，故此部份的分析也將分為二大類模型。縱貫模型 A 將以數學焦慮為預測變數 X，數學學習動機為中介變數 M，數學成就為效標變數 Y，如圖 4-1-1。

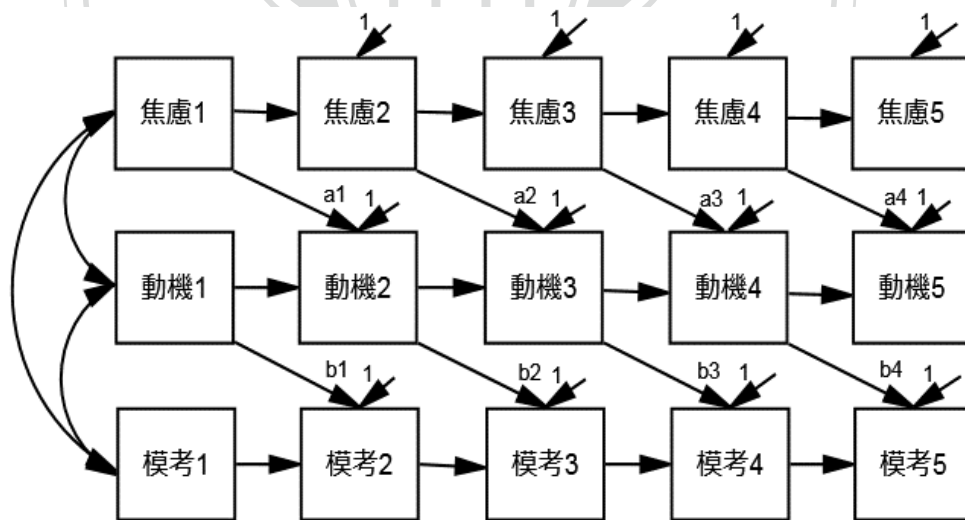


圖 4-1-1 研究二縱貫模型 A 中介架構概念圖

縱貫模型 B 則以數學學習動機為預測變數 X，數學焦慮為中介變數 M，數學成就為效標變數 Y。

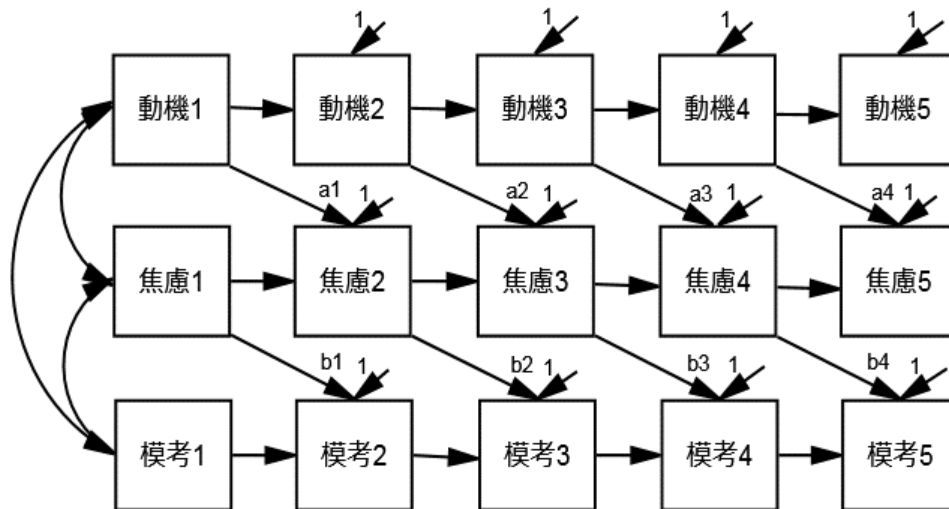


圖 4-1-2 研究一縱貫模型 B 中介架構概念圖

基於方程式 3-1 至方程式 3-3 所述，在縱貫中介分析當中，將檢視縱貫模型 A 與縱貫模型 B 的 $\bar{\theta}_{X_{t+1,t}}$ 、 $\bar{\theta}_{M_{t+1,t}}$ 、 $\bar{\theta}_{Y_{t+2,t+1}}$ 等自我迴歸係數，以及 θ_{M_{t+1},X_t} 和 $\theta_{Y_{t+2},M_{t+1}}$ 等落後一期的縱貫直接效果。接著檢視特定時間點之間的中介效果 (ind_{ab})，可由兩個直接效果相乘得到，如方程式 3-4 表示。

$$ind_{ab} = \theta_{M_{t+1},X_t} \times \theta_{Y_{t+2},M_{t+1}}$$

縱貫模型 A 的中介分析顯示，數學焦慮為預測變數 X 的四波自我迴歸係數 $\bar{\theta}_{X_{t+1,t}}$ 介於 .67 至 .73 之間，所有係數皆達到顯著水準 ($p < .001$)。數學學習動機為中介變數 M 的四波自我迴歸係數 $\bar{\theta}_{M_{t+1,t}}$ 介於 .70 至 .72 之間，所有係數皆達到顯著水準 ($p < .001$)。數學成就為效標變數 Y 的五波自我迴歸係數 $\bar{\theta}_{Y_{t+1,t}}$ 介於 .45 至 .61 之間，所有係數皆達到顯著水準 ($p < .001$)。整體而言，三條時間序列都呈現出穩定的自我相關。

在縱貫直接效果部分，前一期數學焦慮對後一期數學學習動機 $X_t \rightarrow M_{t+1}$ 的標準化迴歸係數 θ_{M_{t+1},X_t} 介於 -.01 至 .07 之間，僅有 $X_1 \rightarrow M_2$ 的係數 .07 達到顯著水準 ($p < .05$)，其餘的縱貫直接效果皆未達顯著水準 ($p > .05$)。此結果顯示，數學焦慮對數學學習動機的正向影響，在控制自我迴歸的效果之後，僅在第一波

的數學焦慮對第二波的數學學習動機成立。再者，前一期數學學習動機對後一期數學成就 $M_{t+1} \rightarrow Y_{t+2}$ 的標準化迴歸係數 $\theta_{Y_{t+2}, M_{t+1}}$ 介於.09至.15之間，所有係數皆達到顯著水準 ($p < .05$)。此結果顯示，數學學習動機對數學成就的正向影響，在控制自我迴歸的效果之後，在每個前一期數學焦慮對數學學習動機皆成立。

在縱貫中介效果部分，如同序列中介的概念，也是探討啟始點的數學焦慮透過後一期數學學習動機為中介變數進而影響後二期的數學成就 $X_t \rightarrow M_{t+1} \rightarrow Y_{t+2}$ ，唯此處的中介效果已經控制各個變數的自我迴歸影響，可視為序列中介效果的淨效果版本。由於縱貫中介必須包含三個相異的時間點，在五波資料的架構下，將有三次的縱貫中介效果檢驗，分別以第一波、第二波、第三波做為啟始點，縱貫中介效果以 ind_{ab} 表示。

首先，以第一波數學焦慮為啟始點的縱貫中介效果 ind_{ab} ，其未標準化係數為.132，95% CI 介於.007至.344之間，並未包含0，故判定為顯著。此結果顯示，在控制自我迴歸效果後，第一波數學焦慮將以第二波數學學習動機為中介進而影響第三波數學成就。再者，以第二波數學焦慮為啟始點的縱貫中介效果 ind_{ab} ，其未標準化係數為-.019，95% CI 介於-.257至.190之間，因包含0，故判定為不顯著。此結果顯示，在控制自我迴歸效果後，第二波數學焦慮並未以第三波數學學習動機為中介進而影響第四波數學成就。最後，以第三波數學焦慮為啟始點的縱貫中介效果 ind_{ab} ，其未標準化係數為.079，95% CI 介於-.078至.284之間，因包含0，故判定為不顯著。此結果顯示，在控制自我迴歸效果後，第三波數學焦慮並未以第四波數學學習動機為中介進而影響第五波數學成就。

綜合上述分析可知，縱貫模型 A 的中介效果僅發生在以第一波為啟始點的情況下，也就是第一波數學焦慮將以第二波數學學習動機為中介進而影響第三波數學成就。其餘測量波次，並未發現相同的縱貫中介現象。另一方面，在縱貫直接效果的分析上，也可發現數學學習動機對於數學成就有一致性的顯著正向影響。詳細數據請見表 4-1-9 並以圖 4-1-3 呈現。

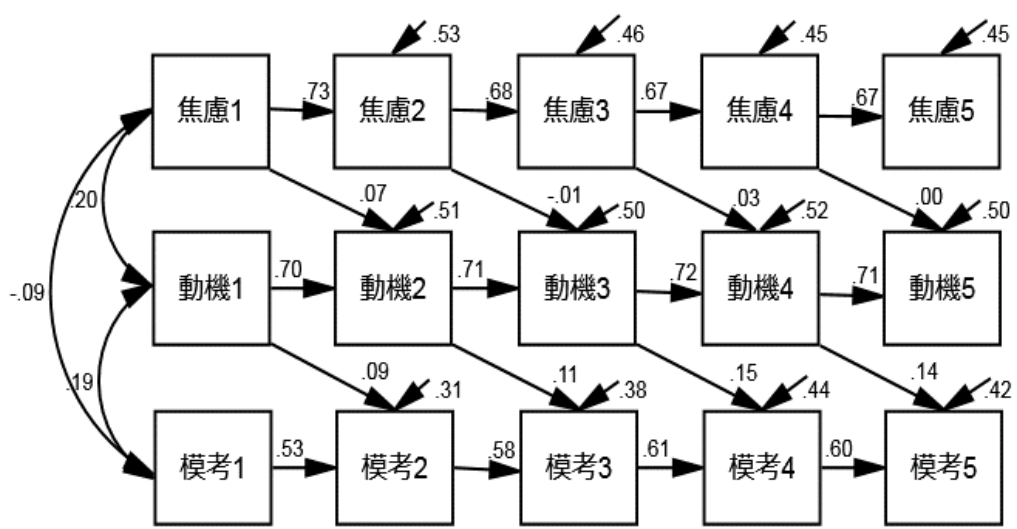


圖 4-1-3 研究一縱貫模型 A 中介分析結果圖

表 4-1-9 研究一縱貫模型 A 中介分析

啟始點	自我迴歸係數			縱貫直接效果		中介效果	95% CI	
	$\bar{\theta}_{X_{t+1,t}}$	$\bar{\theta}_{M_{t+1,t}}$	$\bar{\theta}_{Y_{t+1,t}}$	θ_{M_{t+1},X_t}	$\theta_{Y_{t+2},M_{t+1}}$		LL	UL
第一波	.73***	.70***	.53***	.07*	.09*	.132	.007	.344
第二波	.68***	.71***	.58***	-.01ns	.11**	-.019	-.257	.190
第三波	.67***	.72***	.61***	.03ns	.15***	.079	-.078	.284
第四波	.67***	.71***	.60***	.00ns	.14***			

註 1：ns $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$

註 2：中介效果及其 95% CI 為未標準化數值

縱貫模型 B 的中介分析顯示，數學學習動機為 X 的五波自我迴歸係數 $\bar{\theta}_{X_{t+1,t}}$ 介於.71 至.73 之間，所有係數皆達到顯著水準 ($p < .001$)。數學焦慮為 M 的五波自我迴歸係數 $\bar{\theta}_{M_{t+1,t}}$ 介於.65 至.71 之間，所有係數皆達到顯著水準($p < .001$)。數學成就為 Y 的六波自我迴歸係數 $\bar{\theta}_{Y_{t+1,t}}$ 介於.45 至.62 之間，所有係數皆達到顯著水準 ($p < .001$)。整體而言，三條時間序列都呈現出中高度的自我迴歸現象。

在縱貫直接效果部分，前一期數學學習動機對後一期數學焦慮 $X_t \rightarrow M_{t+1}$ 的標準化迴歸係數 θ_{M_{t+1},X_t} 介於-.03 至.10 之間，僅有 $X_1 \rightarrow M_2$ 的係數.08 和 $X_2 \rightarrow M_3$ 的係數.10 達到顯著水準 ($p < .001$)，其餘的縱貫直接效果皆未達顯著水準 ($p > .05$)。此結果顯示，數學學習動機對數學焦慮的正向影響，在控制自我迴歸的效果之後，僅在第一波的數學學習動機對第二波的數學焦慮與第二波的數學學習動機對第三波的數學焦慮成立。再者，前一期考試焦慮對後一期數學成就 $M_{t+1} \rightarrow Y_{t+2}$ 的標準化迴歸係數 $\theta_{Y_{t+2},M_{t+1}}$ 介於-.05 至-.13 之間，僅有以第一波和第三波為啟始點的係數皆達到顯著水準 ($p < .001$)。此結果顯示，數學考試焦慮對數學成就的負向影響，在控制自我迴歸的效果之後，僅在第一波及第三波為啟始點的情況下成立。

在縱貫中介效果部分，如同序列中介的概念，也是探討啟始點的數學學習動機透過後一期數學焦慮為中介變數進而影響後二期的數學成就 $X_t \rightarrow M_{t+1} \rightarrow Y_{t+2}$ ，唯此處的中介效果已經控制各個變數的自我迴歸影響，可視為序列中介效果的淨效果版本。由於縱貫中介必須包含三個相異的時間點，在五波資料的架構下，將有三次縱貫中介效果檢驗，分別以第一波、第二波、第三波做為啟始點，縱貫中介效果以 ind_{ab} 表示。

首先，以第一波數學學習動機為啟始點的縱貫中介效果 ind_{ab} ，其未標準化係數為-.064，95% CI 介於-.218 至.008 之間，因包含 0，故判定為不顯著。此結果顯示，在控制自我迴歸效果後，第一波數學學習動機並未以第二波數學焦慮

為中介進而影響第三波數學成就。再者，以第二波數學學習動機為啟始點的縱貫中介效果 ind_{ab} ，其未標準化係數為-.188，95% CI 介於-.430 至-.061 之間，並未包含 0，故判定為顯著。此結果顯示，在控制自我迴歸效果後，第二波數學學習動機將以第三波數學焦慮為中介進而影響第四波數學成就。最後，以第三波數學學習動機為啟始點的縱貫中介效果 ind_{ab} ，其未標準化係數為.025，95% CI 介於-.020 至.200 之間，因包含 0，故判定為不顯著。此結果顯示，在控制自我迴歸效果後，第三波數學學習動機並未以第四波數學焦慮為中介進而影響第五波數學成就。

綜合上述分析可知，縱貫模型 B 的中介效果僅發生在以第二波為啟始點的情況下，也就是第二波數學學習動機將以第三波數學焦慮為中介進而影響第四波數學成就。其餘測量波次，並未發現相同的縱貫中介現象。另一方面，在縱貫直接效果的分析上，也可發現變數之間並未呈現一致性的顯著結果。詳細數據請見表 4-1-10 並以圖 4-1-4 呈現。

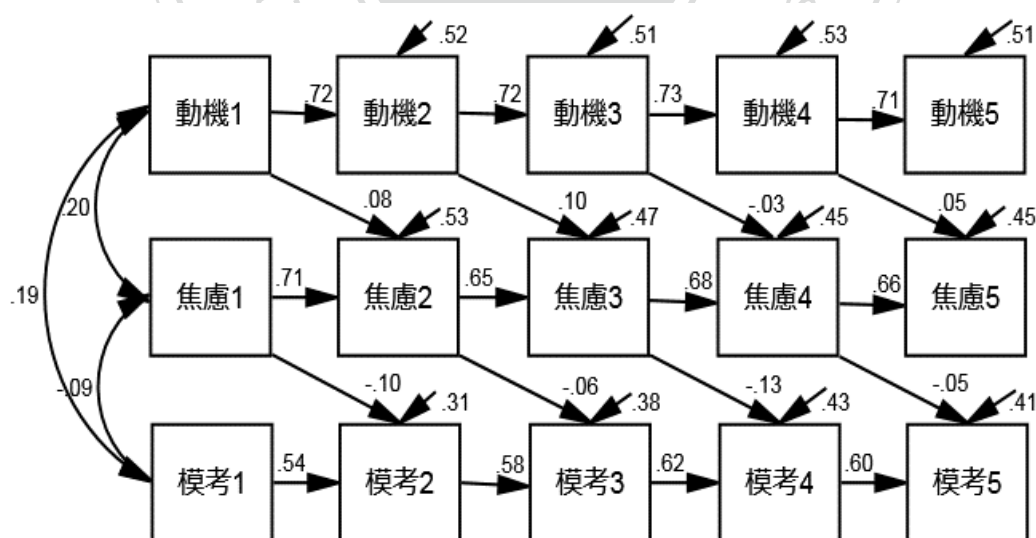


圖 4-1-4 研究一縱貫模型 B 中介分析結果圖

表 4-1-10 研究一縱貫模型 B 中介分析

啟始點	自我迴歸係數			縱貫直接效果		中介效果	95% CI	
	$\bar{\theta}_{X_{t+1,t}}$	$\bar{\theta}_{M_{t+1,t}}$	$\bar{\theta}_{Y_{t+1,t}}$	θ_{M_{t+1},X_t}	$\theta_{Y_{t+2},M_{t+1}}$	ind_{ab}	LL	UL
第一波	.72***	.71***	.54***	.08***	-.10***	-.064	-.218	.008
第二波	.72***	.65***	.58***	.10***	-.06ns	-.188	-.430	-.061
第三波	.73***	.68***	.62***	-.03ns	-.13***	.025	-.020	.200
第四波	.71***	.66***	.60***	.05ns	-.05ns			

註 1：ns $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$

註 2：中介效果及其 95% CI 為未標準化數值

由於縱貫模型 A 與縱貫模型 B 二者之間使用的是同一批資料，故彼此可視為競爭模型。不同於前述的橫斷中介及序列中介當中，總以模型 A（數學焦慮→數學學習動機→數學成就）的檢核結果較為完備的結論，在縱貫中介的架構下，縱貫模型 A 與縱貫模型 B 之間的完備條件判定，無法如之前一般篤定地偏向其中一方。

首先，從共變關係來看，縱貫模型 A 與縱貫模型 B 在數學焦慮、數學學習動機及數學成就的相關係數都呈現顯著，並且在各自的時間序列上都呈現顯著的自我迴歸現象，且標準化迴歸係數皆達到.50 甚至.70 以上。

縱貫模型 A 與縱貫模型 B 在邏輯時序上，同前述般都有對應的理論架構。然而縱貫模型 A 與縱貫模型 B 在縱貫直接效果及縱貫中介效果的表現上，卻都出現不一致的顯著性結果。縱貫模型 A 顯示考試焦慮僅在第一波時會影響到第二波的學習動機，而每一波的學習動機都會影響下一波的數學成就。縱貫模型 A 的縱貫中介效果則僅出現在以第一波考試焦慮為啟始點的情況。縱貫模型 B 也有類似的現象，但更不一致的是，學習動機對考試焦慮的縱貫直接效果僅在第一波與第二波成立。而考試焦慮對數學成就的縱貫直接效果則是僅以第一波及第三波考試。縱貫模型 A 的中介效果僅出現在第一波為啟始點上，而縱貫模型 B 則僅出現在以第二波為啟始點上。但相較於橫斷中介與序列中介，可發現在控制自我迴歸的效

果後，直接效果及中介效果都一致性的降低，甚至改為不顯著的情況。然而在較複雜的縱貫中介架構之下，縱貫模型 A 與縱貫模型 B 的 BIC 值分別為 966.10 及 960.15，二者的差異量非常微小。因此，很難透過 BIC 值的差異比較，來斷定何者是較能通過否證的因果模型。

然而，縱貫模型 A 與縱貫模型 B 都屬於縱貫設計，故都具有測量時序及控制自我迴歸。整體而言，比起前述的橫斷及序列設計都是較具備因果推論完備條件的設定。無論是否在縱貫模型上發現到不同於橫斷設計及序列設計的結果，都應看作較貼近真正答案的選項。

表 4-1-11 研究一縱貫中介的因果推論完備條件檢核表

因果推論完備條件	檢核結果	說明
1. 有共變關係，變數之間具有顯著的相關。	具備	縱貫模型 A 與縱貫模型 B 二者在變數間相關係數、時間序列的自我迴歸係數上都是一致性的表現。
2. 有邏輯時序，透過理論或邏輯推演以建構理論模型。	部分具備	1. 縱貫模型 A 與縱貫模型 B 各自有相對應的理論、邏輯架構及過往研究的證據引導。 2. 但二者在縱貫直接效果、縱貫中介效果的表現上都有不一致的顯著性結果。 3. BIC 的比較顯示二個模型之間的差異性極小。
3. 有測量時序，在不同時間點測量研究變數。	具備	縱貫模型 A 與縱貫模型 B 皆為縱貫中介模型，所有變數都在三個不同的時間點進行測量。
4. 無混淆變數，有對自我迴歸現象進行控制。	具備	縱貫模型 A 與縱貫模型 B 的所有變數皆有針對自我迴歸現象進行控制。

第二節 研究二：臺灣青少年成長歷程資料庫分析結果

本節主要在呈現研究二所使用之臺灣青少年計畫資料庫分析結果，分別就描述性統計、橫斷中介分析、序列中介分析、縱貫中介分析等四個部分進行探討，並針對上述分析之因果推論完備條件進行檢核。

壹、描述性統計

國一學生在親子互動、自尊及憂鬱上的描述統計方面，由於本研究的測量時間，係從國一開始，總計橫跨八年共三次的電話訪問範圍中，故共有親子互動 1~3 (分別是第一年、第四年及第八年共計三次，X1~X3)、自尊 1~3 (分別是第一年、第四年及第八年共計三次，M1~M3) 及憂鬱 1~3 (分別是第一年、第四年及第八年共計三次，Y1~Y3)。由過往研究可知，親子互動可透過自尊為中介進而對憂鬱產生影響，模型設定上以親子互動為預測變數 X、自尊為中介變數 M、憂鬱為效標變數 Y，因果推論的方向為 $X \rightarrow M \rightarrow Y$ 。

首先，呈現全體樣本描述統計資訊，包含平均數、標準差、相關係數等。在親子互動的表現上，X1 採用七點量尺，其平均數為 4.24、標準差為 1.41；X2~X3 皆採用四點量尺，X2 平均數為 2.88、標準差為 1.20，X3 平均數為 1.90、標準差為 0.59，數據在跨時間上有所變動。在自尊的表現上採用四點量尺，M1~M3 的平均數介於 2.47 至 2.50 之間，標準差介於 0.53 至 0.55 之間，數據在跨時間上具有穩定性。在憂鬱方面採用五點量尺，Y1~Y3 的平均數介於 1.33 分至 1.54 分，標準差介於 0.44 至 0.51 之間，整體來看憂鬱情況偏低。在相關係數部分，所有相關係數都達到顯著水準 ($p < .05$)，詳細數據請見表 4-2-1。

表 4-2-1 臺灣青少年計畫研究變項之描述統計 (N=1529)

	平均數	標準差	X1	X2	X3	M1	M2	M3	Y1	Y2	Y3
X1	4.24	1.41	1								
X2	2.88	1.20	.28	1							
X3	1.90	0.59	.25	.23	1						
M1	2.47	0.54	-.13	-.09	-.09	1					
M2	2.50	0.53	-.10	-.12	-.10	.26	1				
M3	2.47	0.55	-.10	-.11	-.19	.19	.34	1			
Y1	1.44	0.48	.15	.05	.15	-.25	-.18	-.15	1		
Y2	1.33	0.44	.08	.07	.13	-.10	-.27	-.20	.28	1	
Y3	1.54	0.51	.09	.09	-.26	-.11	-.15	-.35	.27	.29	1

註：所有相關係數皆達顯著 ($p < .05$)

貳、橫斷中介分析

橫斷中介分析，係以同一波測量到的變項進行中介分析。故此部份的分析將分別進行 1~3 波次的檢驗，共計 5 個橫斷中介模型的分析。

橫斷模型：親子互動_t → 自尊_t → 憂鬱_t, t = 1~3

橫斷中介的分析，將呈現 $X_t \rightarrow M_t$ 的標準化迴歸係數(γ_t)、 $M_t \rightarrow Y_t$ 的標準化迴歸係數(β_t)、中介效果($\gamma_t * \beta_t$)及其 95%信賴區間、BIC 值(χ^2)。中介效果的顯著性，將視其 95%信賴區間是否包含 0 而定，若包含 0 則判定為不顯著，若不包含 0 則判定為顯著。

在橫斷模型的分析當中，各波的 γ_t 介於-.12 至-.19 之間，並且皆達到顯著水準 ($p < .001$)；各波的 β_t 介於-.25 至-.35 之間，並且皆達到顯著水準 ($p < .001$)；中介效果 $\gamma_t * \beta_t$ 介於.032 至.067 之間，由於其 95% CI 皆不包含 0，皆可視為顯著。結果顯示，研究二所提出的橫斷中介模型在三波測量資料中都獲得支持，其直接效果與中介效果都達到顯著水準。在橫斷設計下，親子互動可透過自尊為中介進而對憂鬱產生影響，此一因果推論獲得支持。詳細資料請見表 4-2-1。

表 4-2-1 各波次橫斷模型中介分析

測量時間	γ_t	β_t	$\gamma_t * \beta_t$	95% CI
第一波 (第一年)	-.13***	-.25***	.032	.017~.048
第二波 (第四年)	-.12***	-.27***	.032	.017~.049
第三波 (第八年)	-.19***	-.35***	.067	.047~.091

註：ns $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$

若從因果推論的完備條件論之，從表 4-2-1 可看出所有相關係數皆達到顯著水準，故具備條件一之變關係。在各波的橫斷中介模型分析中，也都得到顯著的直接效果及中介效果，顯示研究二透過過往研究所建立的邏輯時序受到實證資料的支持。然而，橫斷中介模型在研究設計上僅有考量到「有共變關係」和「有邏輯時序」二點上，並未滿足「有測量時序」及「無混淆變數」的要求。換言之，無論在橫斷中介模型中獲得何種結果，即使完全符合研究者所預期的目標或得到良好的數據結果，也僅是滿足前二個條件，距離完備的因果推論還有非常大的改善空間。

表 4-2-3 研究二橫斷中介的因果推論完備條件檢核表

因果推論完備條件	檢核結果	說明
1. 有共變關係，變數之間具有顯著的相關。	具備	在橫斷模型當中，所有波次內的相關係數皆達到顯著水準。
2. 有邏輯時序，透過理論或邏輯推演以建構理論模型。	具備	1. 橫斷模型有相對應的理論、邏輯架構及過往研究的證據引導。 2. 各波次的 γ_t 係數與 β_t 係數、中介效果等都具有一致且顯著的表現。
3. 有測量時序，在不同時間點測量研究變數。	不具備	此模型為橫斷中介模型，所有變數都在同一時間內測量。
4. 無混淆變數，有對自我迴歸現象進行控制。	不具備	橫斷模型的所有變數皆無針對自我迴歸現象進行控制。

參、序列中介分析

序列中介分析，係以不同波次測量到的變項進行中介分析，中介模型當中的預測變數、中介變數及效標變數皆來自前、中、後三個不同波次的測量。如同前述，由於在理論上存在二種解釋考試焦慮、學習動機的因果論述，故此部份的分析也將分為二大類模型、分別設定啟始點為第 t 波、第 $t+1$ 波、第 $t+2$ 波進行檢驗，由於本資料只有五波資料，故啟始點為第一波至第三波，共計六個序列中介模型的分析。

序列模型：親子互動 _{t} → 自尊 _{$t+1$} → 憂鬱 _{$t+2$} , $t = 1$

序列中介的分析，將呈現 $X_t \rightarrow M_{t+1}$ 的標準化迴歸係數 ($\gamma_{t+1,t}$)、 $M_{t+1} \rightarrow Y_{t+2}$ 的標準化迴歸係數 ($\beta_{t+2,t+1}$)、中介效果 ($\gamma_{t+1,t} \times \beta_{t+2,t+1}$) 及其 95% 信賴區間、BIC 值 (χ^2)。中介效果的顯著性，將視其 95% 信賴區間是否包含 0 而定，若包含 0 則判定為不顯著，若不包含 0 則判定為顯著。

在序列模型的分析當中，親子互動 1 對自尊 2 的效果 $\gamma_{t+1,t}$ 為-.10 達到顯著水準 ($p < .001$)；自尊 2 對憂鬱 3 的效果 $\beta_{t+2,t+1}$ 為-.15 達到顯著水準 ($p < .01$)；親子互動 1 透過自尊 2 到憂鬱 3 的中介效果 $\gamma_{t+1,t} \times \beta_{t+2,t+1}$ 為.014，由於其 95% CI 皆不包含 0，皆可視為顯著。詳細資料請見表 4-2-4。

表 4-2-4 研究二序列模型中介分析

啟始點	$\gamma_{t+1,t}$	$\beta_{t+2,t+1}$	$\gamma_{t+1,t} \times \beta_{t+2,t+1}$	95% CI
第一波	-.10 ***	-.15***	.014	.007~.025

註：ns $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$

從序列模型觀之， $\gamma_{t+1,t}$ 的數據為顯示親子互動對自尊有負向預測能力且一致性地達到顯著，但親子互動乃反向計分，故概念上應為親子互動 1 能增加自尊 2。 $\beta_{t+2,t+1}$ 的數據則顯示自尊 2 對於憂鬱 3 就有顯著的負向預測效果，自尊越高者則憂鬱越低。再者，中介效果的分析顯示，序列模型的中介效果達到顯著，故概念上親子互動增加能影響自尊的增加，並進一步影響憂鬱的降低。

若從因果推論的完備條件論之，序列中介的三個變數之間都具有顯著的相關，滿足條件一需有共變關係的要求。再者，序列中介的邏輯時序有理論及過往研究為基礎，並在分析中獲實證得資料的支持。然而，序列設計不但滿足橫斷設計「有共變關係」和「有邏輯時序」二點，還滿足「有測量時序」的要求，已經將因果推論中非常重要的時間因素納入，但依然無法回應「無混淆變數」的要求。

總體而言，相對於橫斷中介，序列中介在直接效果及中介效果上都有降低的現象，其原因很可能來自於間隔的時間長達四年。然而，在研究設計中加入相異的測量時間點，此作法在因果推論的完備條件上又往前邁進。即使數據上顯示的效果量有所降低，卻是更為可信的結果。最後，如前所述，因果關係的推論都必須控制混淆變數，而最需要控制的變數則是該變數的前一期狀態，也就是自我迴歸，這是序列中介仍力有不足之處。

表 4-2-5 序列中介的因果推論完備條件檢核表

因果推論完備條件	檢核結果	說明
1. 有共變關係，變數之間具有顯著的相關。	具備	在序列模型當中，所有相關係數都達到顯著水準。
2. 有邏輯時序，透過理論或邏輯推演以建構理論模型。	具備	1. 序列模型有相對應的理論、邏輯架構及過往研究的證據引導。 2. 直接效果 $\gamma_{t+1,t}$ 係數與 $\beta_{t+2,t+1}$ 係數、中介效果等都具有一致且顯著的表現，數據結果支持邏輯時序的假設。
3. 有測量時序，在不同時間點測量研究變數。	具備	此模型為序列中介模型，三個變數分別在三個不同的時間點進行測量。
4. 無混淆變數，有對自我迴歸現象進行控制。	不具備	序列模型的三個變數皆無針對自我迴歸現象進行控制。

肆、縱貫中介分析

縱貫中介分析，乃結合自我迴歸控制與序列中介而成。中介模型當中的預測變數、中介變數及效標變數都各自有完整的追蹤測量以形成時間序列資料。故在分析上，除了預測變數、中介變數及效標變數皆來自前、中、後三個不同波次的測量，每一個變數也都同時受到自身前一期狀態的預測，在排除自我迴歸的影響後，所得到即是跨時間的淨效果。

在縱貫中介的分析中，親子互動有第一年、第四年和第八年共三波測量形成時間序列，自尊有第一年、第四年和第八年共三波測量形成時間序列，而憂鬱有第一年、第四年和第八年共三波測量形成時間序列。除了跨時間直接效果之外，第一波的親子互動、自尊及憂鬱也作為模型中的控制變數加上相關控制。縱貫模型 A 將以親子互動為預測變數 X，自尊為中介變數 M，憂鬱為效標變數 Y，架構設定如圖 4-2-1。

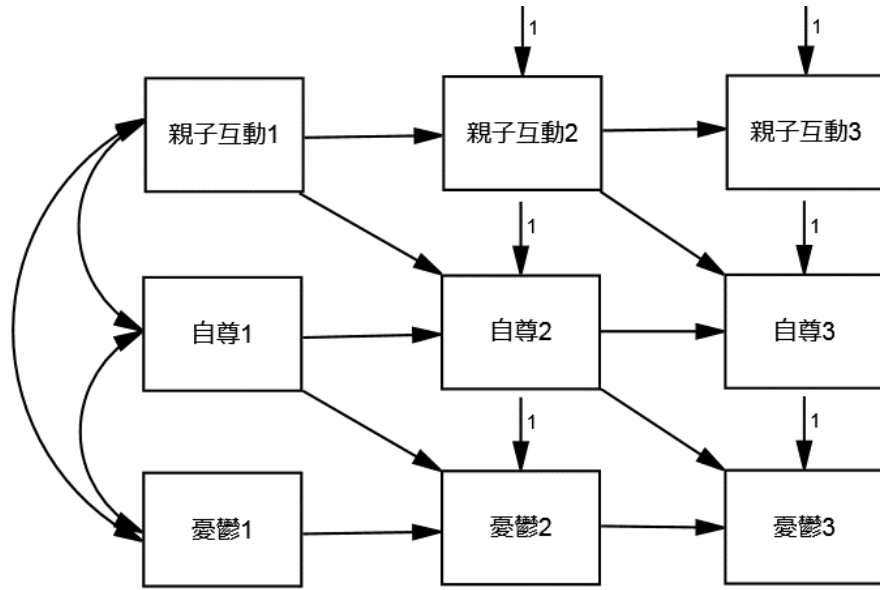


圖 4-2-1 研究二縱貫模型中介架構概念圖

基於方程式 3-1 至方程式 3-3 所述，在縱貫中介分析當中，將檢視縱貫模型的 $\bar{\theta}_{X_{t+1,t}}$ 、 $\bar{\theta}_{M_{t+1,t}}$ 、 $\bar{\theta}_{Y_{t+2,t+1}}$ 等自我迴歸係數，以及 θ_{M_{t+1},X_t} 和 $\theta_{Y_{t+2},M_{t+1}}$ 等落後一期的縱貫直接效果。接著檢視特定時間點之間的中介效果 (ind_{ab})，可由兩個直接效果相乘得到，如方程式 3-4 表示。

$$ind_{ab} = \theta_{M_{t+1},X_t} \times \theta_{Y_{t+2},M_{t+1}}$$

縱貫模型的分析顯示，親子互動為預測變數 X 的二波自我迴歸係數 $\bar{\theta}_{X_{t+1,t}}$ 為 .28 及 .23，所有係數皆達到顯著水準 ($p < .001$)。自尊為中介變數 M 的二波自我迴歸係數 $\bar{\theta}_{M_{t+1,t}}$ 為 .25 及 .34，所有係數皆達到顯著水準 ($p < .001$)。憂鬱為效標變數 Y 的二波自我迴歸係數 $\bar{\theta}_{Y_{t+1,t}}$ 為 .27 及 .27，所有係數皆達到顯著水準 ($p < .001$)。整體而言，三條時間序列都呈現出低度的自我迴歸現象。

在縱貫直接效果部分，前一期親子互動對後一期自尊 $X_t \rightarrow M_{t+1}$ 的標準化迴歸係數 θ_{M_{t+1},X_t} 皆為 -.07 和 -.07 皆達到顯著水準 ($p < .01$)。由於親子互動為反向計分，故此結果顯示，親子互動對自尊的正向影響，在控制自我迴歸的效果之後，在親子互動 1 \rightarrow 自尊 2 及親子互動 2 \rightarrow 自尊 3 皆成立。再者，前一期親子互動對

後一期憂鬱 $M_{t+1} \rightarrow Y_{t+2}$ 的標準化迴歸係數 $\theta_{Y_{t+2},M_{t+1}}$ 為-.03及-.07,僅有自尊2 \rightarrow 憂鬱3的 $M_2 \rightarrow Y_3$ 路徑達到顯著水準 ($p < .001$)。此結果顯示,自尊對憂鬱的負向影響,在控制自我迴歸的效果之後,僅有自尊2 \rightarrow 憂鬱3成立。

在縱貫中介效果部分,如同序列中介的概念,探討前一期親子互動透過後一期自尊為中介影響後二期憂鬱 $X_t \rightarrow M_{t+1} \rightarrow Y_{t+2}$,唯此處的中介效果已經控制各個變數的自我迴歸影響,可視為序列中介效果的淨效果版本。由於縱貫中介必須包含三個相異的時間點,在三波資料的架構下,僅有一次的縱貫中介效果檢驗,以第一波做為啟始點,縱貫中介效果以 ind_{ab} 表示。

以第一波親子互動為啟始點的縱貫中介效果 ind_{ab} ,其未標準化係數為.002,95% CI 介於-.001至.004之間,因包含0,故判定為不顯著。此結果顯示,在控制自我迴歸效果後,第一波親子互動並不以第二波自尊為中介進而影響第三波憂鬱。

綜合上述分析可知,縱貫模型並未發現顯著的中介效果,其結果不同於橫斷與序列中介。詳細數據請見表4-2-6,並以圖4-2-2呈現。

表 4-2-6 研究二縱貫模型中介分析

啟始點	自我迴歸係數			縱貫直接效果		中介效果	95% CI	
	$\bar{\theta}_{X_{t+1,t}}$	$\bar{\theta}_{M_{t+1,t}}$	$\bar{\theta}_{Y_{t+1,t}}$	θ_{M_{t+1},X_t}	$\theta_{Y_{t+2},M_{t+1}}$	ind_{ab}	LL	UL
第一波	.28***	.25***	.27***	-.07**	-.03ns	.002	-.001	.004
第二波	.23***	.34***	.27***	-.07**	-.07**			

註 1 : ns $p > .05$; * $p < .05$; ** $p < .01$; *** $p < .001$

註 2 : 中介效果及其 95% CI 為未標準化數值

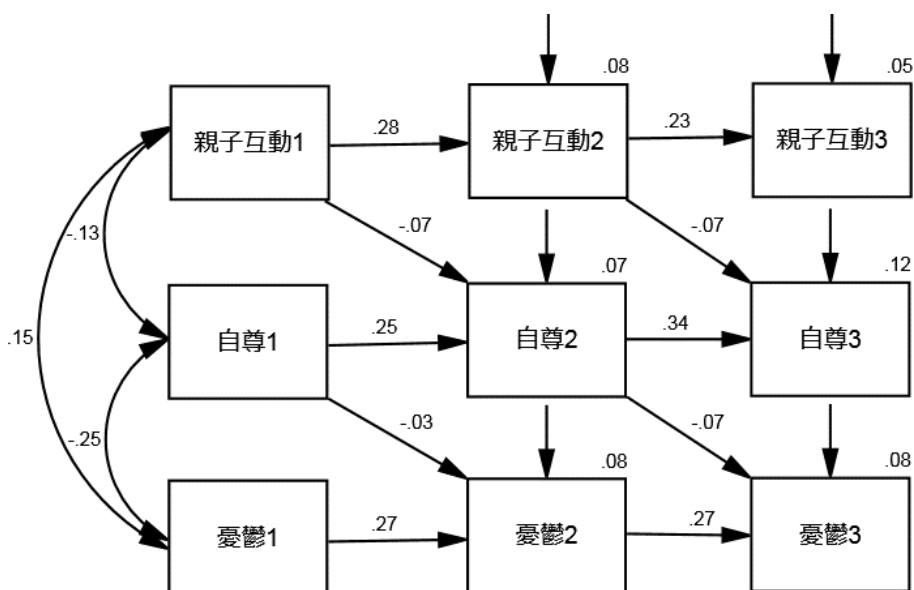


圖 4-2-2 研究二縱貫模型中介分析結果圖

在縱貫模型的因果推論完備條件方面，首先從共變關係來看，縱貫模型在所有研究變數之間都具有顯著相關，且在親子互動、自尊及憂鬱的時間序列上都呈現顯著的低度自我迴歸現象，標準化自我迴歸係數介於.23至.34之間。然而縱貫模型在縱貫直接效果及縱貫中介效果的表現上，卻都出現不一致的顯著性結果。縱貫模型顯示第一波和第二波的親子互動都會影響後一期的自尊，但只有第二波的正尊會影響後一期的憂鬱。再者，縱貫中介分析結果顯示為不顯著，表示親子互動並不以自尊為中介對憂鬱進行跨時間的影響。換言之，親子互動→自尊→憂鬱的因果推論，其邏輯時序雖有理論上的合理性，也通過橫斷中介和序列中介二種模型的否證，卻無法通過完整的縱貫中介模型之否證。

然而，縱貫中介設計在方法學上具有明顯的優勢，除了共變關係、邏輯時序外，也具有測量時序及控制自我迴歸。整體而言，比起前述的橫斷及序列設計都是較具備因果推論完備條件的設定。無論是否在縱貫模型上發現到不同於橫斷設計及序列設計的結果，都應視為在方法學上較為完備的答案。

表 4-2-7 研究二縱貫中介的因果推論完備條件檢核表

因果推論完備條件	檢核結果	說明
1. 有共變關係，變數之間具有顯著的相關。	部分具備	縱貫模型在所有變數之間的相關係數、自我迴歸係數都達到顯著水準。
2. 有邏輯時序，透過理論或邏輯推演以建構理論模型。	部分具備	1. 縱貫模型有相對應的理論、邏輯架構及過往研究的證據引導。 2. 直接效果 θ_{M_{t+1}, X_t} 和 $\theta_{Y_{t+2}, M_{t+1}}$ 並沒有一致性的顯著性表現，縱貫中介分析結果為不顯著，數據並不支持此邏輯時序。
3. 有測量時序，在不同時間點測量研究變數。	具備	此模型為縱貫中介模型，所有變數都在三個不同的時間點進行測量。
4. 無混淆變數，有對自我迴歸現象進行控制。	具備	縱貫模型所有變數皆有針對自我迴歸現象進行控制。

第三節 綜合討論

本研究旨在瞭解中介模型在因果推論上的架構，透過文獻討論整理檢驗各種中介模型因果推論完備條件的檢核清單，並用以檢視實際分析上的推論策略有哪些不足之處。在實際資料的分析上，取二種性質不同的資料進行檢驗。其一，以王金香（2010）針對國三學生在一年內的五波模擬考試所收集之資料為例，選取數學焦慮、數學學習動機及數學成就等三個變數，套用橫斷、序列及縱貫中介等三種模型進行分析並比較其結果，是為研究一。其二，以中央研究院釋出之「臺灣青少年成長歷程研究」在八年內的三波電話調查所收集之資料為例，選取親子互動、自尊與憂鬱等三個變數，套用橫斷、序列及縱貫中介等三種模型進行分析並比較其結果，是為研究二。分析完畢後，除了比較模型中數據的異同，並以檢

核清單說明各種模型在因果推論上的完備條件。綜合上述二個研究，本研究有三項特色，茲陳述如下：

第一，本研究所關切的議題，是在心理與教育領域當中使用頻率極高的中介模型之因果推論性，研究者以性質相異的二筆實證資料加以進行統計分析。過去研究多以橫斷設計為主，序列設計及縱貫設計在數量上極為稀少，雖主要是受限於時間及成本問題，但也有部分是對於方法學議題欠缺考慮。本研究不單以統計推論的角度出發，而是站在科學哲學的觀點上，以更基本的邏輯觀念，全面檢視整個因果傳遞機制在實證資料分析中如何確保正確性，避免中介模型的使用者根據不當的設計得到偏頗的結論而不自知。

第二，在選取數學模擬考試資料做為實證分析的範例中，其資料特性在於短暫且密集的測量，且理論上對於考試焦慮和學習動機何者為先，有不同的理論推導，故數學模擬考試的資料可以在完整的縱貫架構上，分析哪一套理論較為實際資料所支持。分析結果顯示，在橫斷中介及序列中介的架構下，二種模型的分析結果都一致地支持考試焦慮做為因，並促進學習動機進而提高考試分數的理論。然而，在五波縱貫中介的架構中卻有不同的結論。在單向假設之下，控制了因短時間密集測量而有中高程度的自我迴歸後，差異甚小的 BIC 值顯示兩種可能的影響途徑：其一，在第 0 波考試取得低分後，學生會在第一波時間點產生考試焦慮，進而提高第二波的學習動機，再正向影響到第三波的考試成績。其二，在第 0 波考試取得低分後，其第一波學習動機下降，但此效果不會馬上影響後一波的焦慮，而是透過自我迴歸延宕一波後再增強第三波的考試焦慮，進而負向影響到第四波的考試成績。綜上所述，研究一在方法學正確性較低的橫斷中介與序列中介當中，很明確地否證了學習動機為因的論點，而支持了考試焦慮為因的論點。然而，在方法學正確性較高的縱貫中介模型中，卻無法肯定地拒絕或支持其中一種因果推論。

第三，臺灣青少年計畫的資料庫，其特性在於橫跨時間較長，雖然該研究計畫截至 2018 年已經釋出 10 波測量資料，然而同時有測量到親子互動、自尊及憂鬱等三個變數的時間點僅有第一年、第四年和第八年。因此，研究二所使用的資料相對於研究一，橫跨了更大的時間間隔。相關的理論及過往研究，在同一個資料庫中選取國三樣本進行分析，結果顯示出在橫斷中介的架構下，親子互動會增加自尊並進一步降低憂鬱。在研究二當中，研究者以橫斷中介、序列中介及縱貫中介套用在完整版的縱貫資料庫中進行統計分析。結果顯示，在三個波次的橫斷中介分析中，在直接效果和中介效果上所能得到的最大影響程度，和賴俊生等人（2017）的研究結果相仿。然而在序列中介的架構下，無論是直接效果或中介效果，都有下降的現象。更甚者，在縱貫中介模型的分析中，親子互動無法透過自尊為中介對憂鬱產生影響。換言之，在方法學正確性較高的模型中，拒絕了過往研究所支持的假設。



第五章 結論與建議

本研究有三個研究目的，第一旨在釐清因果推論的哲學及統計基礎；第二旨在探討現存的中介模型在因果推論上的完備條件；第三旨在提出增強因果推論完備條件的可能策略。為達此目的，首先蒐集相關文獻，加以探討分析，以建立研究架構與因果推論完備條件檢核清單。接著選取二種性質不同的實證資料，進行分析與討論。故本章依據第四章研究結果與討論，就結果分析彙整為本章內容，共分為三節，第一節為本研究之結論，第二節為研究限制，第三節為本研究之建議。

第一節 結論

本節主要在呈現資料分析結果，分別就研究一及研究二的中介分析結果進行探討。

壹、數學焦慮、數學學習動機與數學成就之中介分析

一、橫斷中介模型

橫斷中介分析結果顯示，考量第一波到第五波的五個橫斷中介模型在直接效果顯著性、中介效果顯著及 BIC 值的結果，以數學焦慮為預測變數的橫斷模型 A，在數據表現上展現一致顯著的特性、較大的效果量和較小的 BIC 值。故實證數據較為支持以數學焦慮為預測變數、數學學習動機為中介變數，進而影響數學成就的邏輯時序，在因果推論完備條件上較優。

二、序列中介模型

序列中介分析結果顯示，考量以第一波到第三波為啟始點的三個序列中介模型在直接效果顯著性、中介效果顯著及 BIC 值的結果，以數學焦慮為預測變數的序列模型 A，在數據表現上展現一致顯著的特性、較大的效果量和較小的 BIC 值。故實證數據較為支持以數學焦慮為預測變數、數學學習動機為中介變數，進而影響數學成就的邏輯時序，在因果推論完備條件上較優。

三、縱貫中介模型

縱貫中介分析結果顯示，在控制自我迴歸現象後，考量以第一波到第三波為啟始點的三個縱貫中介模型在直接效果顯著性、中介效果顯著及 BIC 值的結果，縱貫中介模型 A 和 B 各自在不同的測量時間點展現出顯著的直接效果與中介效果，且 BIC 值非常接近。故在方法學上，雖然縱貫中介模型比起橫斷中介及序列中介有更好的因果推論完備條件，但實證數據卻無法在縱貫中介模型 A 與 B 之間做出絕對優劣的判斷。換言之，兩種因果推論的看法，在縱貫中介的架構上都尚未被明確地拒絕。

貳、親子互動、自尊與憂鬱之中介分析

一、橫斷中介模型

橫斷中介分析結果顯示，考量第一波到第三波的三個橫斷中介模型在直接效果及中介效果上都達到顯著，親子互動為以自尊為中介影響憂鬱的因果推論在三波測量資料中都獲得支持。

二、序列中介模型

序列中介分析結果顯示，考量從第一波為啟始點的序列中介模型在直接效果及中介效果上都達到顯著。故以親子互動為預測變數、自尊為中介變數，進而影響憂鬱的因果推論，在三波測量資料建構的序列中介模型當中獲得支持。

三、縱貫中介模型

縱貫中介分析結果顯示，在控制自我迴歸現象後，由於中介效果不顯著，實證資料並不支持親子互動為預測變數、自尊為中介變數，進而影響憂鬱的因果推論。換言之，在更有方法學優勢的情況下，縱貫中介模型得到不同於橫斷中介及序列中介的結論。

第二節 研究限制

壹、研究樣本的限制

第一，兩項子研究皆以國中學生為樣本，然而無論是焦慮→動機→成就，或是互動→自尊→憂鬱，兩種理論框架都應該收集更多元的樣本進行分析，以確定因果推論的假設在不同群體上的類推性。倘若樣本性質的差異能造成同一套因果推論有不同的結果，這也正說明研究因果傳遞時機的重要性，也就是調節的研究架構。

第二，在數學成就的分析中，僅選取第一次模擬考數學分析在 Q1 以下者。換言之，整個數學焦慮的因果推論僅是在低數學成就樣本進行否證，在全體或高數學成就樣本上，難以支持相同的因果推論，故在推論上需要特別注意。

第三，在青少年憂鬱的分析中，大多數樣本並沒有嚴重的憂鬱現象，很可能是效果偏低的可能原因之一。

第四，本研究提出以各個縱貫中介效果的相等做為因果推論完備條件的概念，並沒有機會在兩個子研究的資料中實際操作。推論原因可能在於不同領域 (domain) 當中，其資料的特殊性所致。兩個子研究的資料在標準化中介效果最大的橫斷模型中就已經低於.1，未來若要持續研究此概念的可行性，需要選取相關程度更高的研究變項。

貳、測量時間的限制

研究一在測量時間的特色上，是一年內密集地測量五次，故在自我迴歸係數上有較大效果的展現，然而從多元迴歸的角度來看，較高的自我迴歸係數限制了跨時間直接效果及中介效果的效果。而研究二則相反，由於資料庫並非每一波都有測量同樣的變數或題目，因此造就研究二有三年以上時距的資料結構。在長時間的間隔之下，很可能因為成熟效應、歷史效應等各種因素導致自我迴歸係數降低，也造成跨時間直接效果及中介效果的弱化。

測量時間點的設定，理應由該領域學者針對特定領域進行詳實的前導研究得來，然而在兩個子研究中僅是基於現實考量，如考試時間點及計畫執行時間來決定測量時機和間隔。研究變數並非從被測量的那一刻才存在，而是一直都存在。以研究一為例，數學焦慮及數學學習動機從小學一年級甚至更早之前就存在，兩者除了互相影響之外，也可能有各自形成的歷程與變化的軌跡。研究一的結果顯示，在一年六次考試的時間範圍內，數學焦慮為因與數學學習動機為因這兩種因果推論都難以被明確地拒絕。有一種可能性是，或許其中一個變數為因，但並不發生在本研究的測量時間內。在整段時間長河中，研究者以人為視角去擷取一段來觀察，這種方式本身就可能造成錯誤的詮釋。較嚴謹的做法是，透過前導研究針對該領域的現象進行解析，以決定測量的時間範圍、間隔及測量點數量。另一種可能，則是影響數學成就的機制，本質上可能是具有多個原因或多個中介變數，因此出現本研究效果量較低的情況。

參、分析方法的限制

本研究以路徑模型進行模式建構與分析，然而此統計方法並未考慮測量問題。雖然兩份資料庫在使用上都有先檢視原始文獻的信效度，但沒有以結構方程式模型分析，依然還是可能受到測量誤差變異的影響，導致直接效果與中介效果的估計有所偏誤。此外，在縱貫中介的架構下，並沒有針對控制變數的議題進行分析。控制變數在縱貫架構之下，如何測量與置入分析仍然是有待探究的議題之一。

第三節 研究建議

科學研究的目的即在於追求因果關係，過程中必須反覆檢驗許多因果推論的假設。越是逼近真正因果關係的因果推論，越能據以做出適切決策，使錯誤降低、遺憾減少。本研究的分析結果一再顯示，在橫斷設計中總是被支持的因果推論，不必然能在縱貫設計中得到一樣的結果；抑或是，在橫斷設計中經常勝過競爭對手的因果推論，在縱貫設計中也不必然繼續保有優勢。

用以檢驗因果推論的方法學論述不少，本研究專注於時間因素落實在中介模型當中的必要性，以及橫斷中介、序列中介及縱貫中介之間的因果推論完備條件討論。在縱貫追蹤架構及自我迴歸取向的典範之外，近年來開始有建議使用實驗性中介分析（experimental analysis of mediation）的呼籲出現，Bullock、Green 與 Ha（2010）指出非實驗資料經常有偏誤的情況產生，進而影響中介機制的估計。透過操弄中介變數，並觀察實驗組與控制組在直接效果與中介效果的差異，也是能更深刻理解因果傳遞機制的管道。這樣的研究設計策略，在概念上也涉及到調節式中介（moderated mediation）及多層次模型（multilevel modeling）的範疇（Bauer, Preacher, & Gil, 2006; Muller, Judd, & Yzerbyt, 2005）。

在實驗研究中，研究者所關注的焦點，在於實驗組受測者在效標分數上的得分是否顯著不同於控制組受測者。若差異分數達到顯著水準，則研究者即可宣稱實驗操弄對於效標分數之間具有因果關係的推論。然而，實驗組受測者僅接受過實驗組的操弄而沒有接受過控制組的操弄，反之，控制組受測者則是僅接受過控制組的操弄而沒有接受過實驗組的操弄。上述一般實驗操弄的作法稱為「事實」（factual），這樣的實驗設計並無法檢驗「反事實」（counterfactual）的可能性（邱皓政、孫國勛，2011）。反事實的部份是指，實驗組受測者若接受控制組操弄，而控制組受測者若接受實驗組操弄的結果（Morgan & Winship, 2014; Pearl, 2000; Rubin, 2004）。舉例來說，補習的效果通常是依據有參加補習者的成績顯著優於不參加補習者的成績，因此推論出補習造成較優成績之間的因果關係。然而，這樣的推論帶有選擇性偏誤，倘若原本參加補習者改為不參加補習，原本不補習者改為參加補習，若能得到相同結果，才真正能說明參加補習對於成績提高之間的因果關係（關秉寅、李敦義，2008）。

中介模型在方法學上仍有非常多待解的議題，例如進行反事實分析後的敏感度分析、在縱貫中介模型當中整合調節概念、多重中介變數的縱貫設計，或是巢套資料的情況下如何應用多層次模式到縱貫中介當中等。社會科學所探求的因果

關係，往往變數眾多、架構複雜，也因為如此，中介模型探索因果傳遞機制的角色更顯重要。然而，一個能完美檢驗因果推論的中介模型尚未出現，任何讀者或使用者都應該將中介模型的完善，想作是學術社群累積性的探索志業，想要獲得一個全面性的結論或完全正確的作法，是非常困難的。不論是實驗性中介、反事實分析等等精進因果推論的策略，無一不體現出抽象概念的深奧難解及實際作業的龐雜惱人。本研究受限於時間人力，僅能在真理的大海邊撈取一瓢飲，以因果推論中最重要時間因素切入探討，其餘議題的討論仍有待未來研究的進一步探索。

本研究雖引用相關文獻，用以說明橫斷設計及序列設計在因果推論上的缺失，但出發點並非要以現今的方法學發展否定過去學者們的努力。每一個時代都有其方法學的發展，學者們應用當時正確的方法進行嚴謹的研究，獲致重要結果即為其貢獻。學術的可貴在於每一代人的承先啟後。

現代科學的價值在於，承認人類感官的有限性而不再執著於證明某個觀點為具有無限性質的真理，而是讓各種觀點透過實證資料互相競爭，透過否證程序更貼切地解釋現象進而逐步地逼近真理。因果關係的確定是困難的，而因果推論的想像是豐富的，在追尋的道路上唯一有把握的，僅無簡單答案一事。

茫茫往代，既沉予聞；眇眇來世，倘塵彼觀也。

劉勰《文心雕龍》（序志）

一如過去學人們在尋真之道上的種種見解是如此絢惑我心；故也自覺此作或對後來者激起漣漪，若昨日煙塵。

參考書目

- 王金香(2010)。焦慮與動機影響數學學習之縱貫研究(未出版之博士論文)。國立政治大學, 臺北市。
- 伊慶春(2012a)。台灣青少年成長歷程研究: 國一樣本(J1)第四波, 青少年問卷(公共版)(C00179_3)【原始數據】取自中央研究院人文社會科學研究中心調查研究專題中心學術調查研究資料庫。
- 伊慶春(2012b)。台灣青少年成長歷程研究: 國一樣本(J1)第八波, 青少年問卷(公共版)(C00272_2)【原始數據】取自中央研究院人文社會科學研究中心調查研究專題中心學術調查研究資料庫。
- 伊慶春(2014)。台灣青少年成長歷程研究: 國一樣本(J1)第一波, 青少年問卷(公共版)(C00176_4)【原始數據】取自中央研究院人文社會科學研究中心調查研究專題中心學術調查研究資料庫。
- 余民寧(2013)。縱貫性資料分析—LGM的應用。臺北市: 心理。
- 吳明隆、葛建志(2006)。國民小學學生數學歸因信念, 數學態度, 數學焦慮與數學成就之相關研究。高雄師大學報, 21, 1-18。
- 吳欣璇、金瑞芝(2011)。華人家庭文化脈絡下新住民母親的親子關係: 以一越南出生與柬埔寨長大的新住民母親為例。應用心理研究, 51, 111-148。
- 吳齊殷、黃鈺婷(2010)。青少年初期身心健康變化及其動態影響變因之討論。中華心理衛生學刊, 23(4), 535-562。
- 李仁豪(2005)。以結構方程模式探討並驗證成就目標與學習行為的關係。教育與心理研究, 28(3), 551-580。
- 林杏足(2003)。青少年自尊量表編製報告。測驗學刊, 50(2), 223-256。
- 邱皓政(2017)。多層次模式與縱貫資料分析: Mplus8 解析應用。臺北市: 五南。
- 邱皓政、孫國勛(2011)。調查數據的事實與反事實研究: 以商管學生就學滿意度為例。教育與心理研究, 34(3), 1-28。

- 陳毓文(2004)。少年憂鬱情緒的危險與保護因子之相關性研究。**中華心理衛生學刊**，**17**(4)，67-95。
- 陳瑞麟(2017)。**科學哲學：假設的推理**。臺北市：五南。
- 黃韞臻、林淑惠(2012)。樂觀／悲觀對大學生學習倦怠的差異分析。**臺灣心理諮商季刊**，**4**(4)，14-36。
- 詹志禹(1993)。因果關係與因果推理。**國立政治大學學報**，**6**，1-15。
- 詹志禹(2002)。科學發現與知識成長：我們能從科學哲學當中學到什麼？**應用心理研究**，**15**，105-127。
- 蔡文標、許天威、蕭金土(2003)。影響國小數學低成就學生數學成就的相關因素之研究。**特殊教育學報**，**17**，1-37。
- 賴俊生、謝亞恆(2017)。快樂，男女有別嗎？青少年自尊對憂鬱情緒的性別差異。**高雄師大學報**，**43**，31-56。
- 顏綵思、魏麗敏(2005)。臺灣中部地區國中小學生自我概念、父母管教方式對攻擊行為影響之研究。**臺中教育大學學報：教育類**，**19**(2)，23-48。
- 關秉寅、李敦義(2008)。補習數學有用嗎？一個“反事實”的分析。**臺灣社會學刊**，**41**，97-148。
- 魏麗敏(1996)。國小學生學習動機、數學焦慮與數學成就之研究。**國民教育研究集刊**，**4**，133-155。
- 蘇曉憶、戴嘉南(2008)。青少年完美主義、自尊與其憂鬱傾向之相關研究。**諮商輔導學報：高師輔導所刊**，**18**，123-154。

- Abramowitz, J. S., Nelson, C. A., Rygwall, R., & Khandker, M. (2007). The cognitive mediation of obsessive-compulsive symptoms: A longitudinal study. *Journal of Anxiety Disorders, 21*(1), 91-104.
- Barnes, G. M., Reifman, A. S., Farrell, M. P., & Dintcheff, B. A. (2000). The effects of parenting on the development of adolescent alcohol misuse: A six-wave latent growth model. *Journal of Marriage and Family, 62*(1), 175-186.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.
- Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods, 11*, 142–163.
- Belsky, J., Pasco Fearon, R. M., & Bell, B. (2007). Parenting, attention and externalizing problems: Testing mediation longitudinally, repeatedly and reciprocally. *Journal of Child Psychology and Psychiatry, 48*(12), 1233-1242.
- Bentley, J. P., Brown, C. J., McGwin Jr, G., Sawyer, P., Allman, R. M., & Roth, D. L. (2013). Functional status, life-space mobility, and quality of life: A longitudinal mediation analysis. *Quality of Life Research, 22*(7), 1621-1632.
- Blau, P. M., & Duncan, O. D. (1967). *The American occupational structure*. New York, NY: Wiley.
- Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? *Journal of Personality and Social Psychology, 98*, 550–558.
- Burton, C. M., Marshal, M. P., Chisolm, D. J., Sucato, G. S., & Friedman, M. S. (2013). Sexual minority-related victimization as a mediator of mental health disparities in

- sexual minority youth: A longitudinal analysis. *Journal of Youth and Adolescence*, 42(3), 394-402.
- Caprara, G. V., Alessandri, G., & Eisenberg, N. (2012). Prosociality: The contribution of traits, values, and self-efficacy beliefs. *Journal of Personality and Social Psychology*, 102(6), 1289-1303.
- Carlson, J. D. (2004). Body image among adolescent girls and boys: A longitudinal study. *Developmental Psychology*, 40(5), 823-835.
- Carvalho, T., Alvarez, M. J., Barz, M., & Schwarzer, R. (2015). Preparatory behavior for condom use among heterosexual young men: A longitudinal mediation model. *Health Education & Behavior*, 42(1), 92-99.
- Chen, X., Stanton, B., Fang, X., Li, X., Lin, D., Zhang, J., & Yang, H. (2006). Perceived smoking norms, socioenvironmental factors, personal attitudes and adolescent smoking in China: A mediation analysis with longitudinal data. *Journal of Adolescent Health*, 38(4), 359-368.
- Cheong, J. (2011). Accuracy of estimates and statistical power for testing mediation in latent growth curve modeling. *Structural Equation Modeling*, 18(2), 195-211.
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, 112(4), 558-577.
- Crocetti, E., Klimstra, T., Keijsers, L., Hale, W. W., & Meeus, W. (2009). Anxiety trajectories and identity development in adolescence: A five-wave longitudinal study. *Journal of Youth and Adolescence*, 38(6), 839-849.
- Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change

- over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, 82(5), 879-894.
- Daniel, E., Madigan, S., & Jenkins, J. (2016). Paternal and maternal warmth and the development of prosociality among preschoolers. *Journal of Family Psychology*, 30(1), 114-124.
- Daniels, E., & Leaper, C. (2006). A longitudinal investigation of sport participation, peer acceptance, and self-esteem among adolescent girls and boys. *Sex Roles*, 55(11-12), 875-880.
- Deacon, S. H., Kieffer, M. J., & Laroche, A. (2014). The relation between morphological awareness and reading comprehension: Evidence from mediation and longitudinal models. *Scientific Studies of Reading*, 18(6), 432-451.
- Dinh, K. T., Roosa, M. W., Tein, J. Y., & Lopez, V. A. (2002). The relationship between acculturation and problem behavior proneness in a Hispanic youth sample: A longitudinal mediation model. *Journal of Abnormal Child Psychology*, 30(3), 295-309.
- Fisher, R. A. (1935). *The design of experiments* (1st ed.). Edinburgh, Scotland: Oliver and Boyd.
- Elliot, A. J., & McGregor, H. A. (2001). A 2×2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501-519.
- Falci, C. D. (2008). Gender trajectories of adolescent depressed mood: The dynamic role of stressors and resources. *Advances in Life Course Research*, 13, 137-160.
- Effron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

- Feddes, A. R., Noack, P., & Rutland, A. (2009). Direct and extended friendship effects on minority and majority children's interethnic attitudes: A longitudinal study. *Child Development, 80*(2), 377-390.
- Ferguson, G. M. (2013). The counterintuitive psychological benefits of intergenerational discrepancies in family prioritization for Jamaican adolescent-parent dyads. *Journal of Research on Adolescence, 23*(1), 35-44.
- Garst, H., & Fay, D. (2007). Making things happen: Reciprocal relationships between work characteristics and personal initiative in a four-wave longitudinal structural equation model. *Journal of Applied Psychology, 92*(4), 1084-1102.
- Hampson, S. E. (2008). Mechanisms by which childhood personality traits influence adult well-being. *Current Directions in Psychological Science, 17*(4), 264-268.
- Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2006). Forty years on: Teachers' assessments of children's personality traits predict self-reported health behaviors and outcomes at midlife. *Health Psychology, 25*(1), 57-64.
- Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2007). Mechanisms by which childhood personality traits influence adult health status: Educational attainment and healthy behaviors. *Health Psychology, 26*(1), 121-125.
- Harvey, E. A., Breaux, R. P., & Lugo-Candelas, C. I. (2016). Early development of comorbidity between symptoms of attention-deficit/hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *Journal of Abnormal Psychology, 125*(2), 154-167.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408-420.
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford.

- Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging, 18*(4), 639-657.
- Hume, D., Árdal, P. S., & MacNabb, D. G. C. (1739). *A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects*. London: John Noon.
- Idris, M. K., O'Driscoll, M. P., & Anderson, M. H. (2011). Longitudinal mediation effects of strain on the relationships between role stressors and employees' withdrawal responses. *Stress and Health, 27*(5), 403-412.
- Imai, K., Jo, B., & Stuart, E. A. (2011). Commentary: Using potential outcomes to understand causal mediation analysis. *Multivariate Behavioral Research, 46*(5), 861-873.
- Jang, H., Kim, E. J., & Reeve, J. (2012). Longitudinal test of self-determination theory's motivation mediation model in a naturally occurring classroom context. *Journal of Educational Psychology, 104*(4), 1175-1188.
- Kant I. (1781). *Kritik der reinen Vernunft*. Riga bei Johann Friedrich Hartknoch.
- Lazarsfeld, P. F. (1955). Interpretation of statistical relations as a research operation. In P. F. Lazarsfeld, & M. Rosenberg (Eds.), *The language of social research: A reader in the methodology of social research* (pp. 115–125). Glencoe, IL: Free Press.
- Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development, 31*(4), 357-365.
- Luyckx, K., Soenens, B., Goossens, L., & Vansteenkiste, M. (2007). Parenting, identity formation, and college adjustment: A mediation model with longitudinal data. *Identity, 7*(4), 309-330.

- Magidson, J. F., Blashill, A. J., Safren, S. A., & Wagner, G. J. (2015). Depressive symptoms, lifestyle structure, and ART adherence among HIV-infected individuals: A longitudinal mediation analysis. *AIDS and Behavior, 19*(1), 34-40.
- Mathieu, J. E., DeShon, R. P., & Bergh, D. D. (2008). Mediation inferences in organizational research: Then, now, and beyond. *Organizational Research Methods, 11*(2), 203-223.
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*(1), 83-104.
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*(1), 23-44.
- Maxwell, S. E., Cole, D. A., & Mitchell, M. A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research, 46*(5), 816-841.
- McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605.
- McLaughlin, K. A., Hatzenbuehler, M. L., & Hilt, L. M. (2009). Emotion dysregulation as a mechanism linking peer victimization to internalizing symptoms in adolescents. *Journal of Consulting and Clinical Psychology, 77*(5), 894-904.
- Meadows, S. O., Brown, J. S., & Elder Jr., G. H. (2006). Depressive symptoms, stress, and support: Gendered trajectories from adolescence to young adulthood. *Journal of Youth and Adolescence, 35*(1), 89-99.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*(1), 107-122.

- Mitchell, M. A., & Maxwell, S. E. (2013). A comparison of the cross-sectional and sequential designs when assessing longitudinal mediation. *Multivariate Behavioral Research, 48*(3), 301-339.
- Morgan S. L., & Winship, C. (2014). *Counterfactuals and causal inference methods and principles for social research* (2nd ed.). Cambridge, England: Cambridge University Press.
- Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*, 852–863.
- Negriff, S., Ji, J., & Trickett, P. K. (2011). Exposure to peer delinquency as a mediator between self-report pubertal timing and delinquency: A longitudinal study of mediation. *Development and Psychopathology, 23*(1), 293-304.
- Nielsen, K., Randall, R., Yarker, J., & Brenner, S. O. (2008). The effects of transformational leadership on followers' perceived work characteristics and psychological well-being: A longitudinal study. *Work & Stress, 22*(1), 16-32.
- Paxton, S. J., Neumark-Sztainer, D., Hannan, P. J. & Eisenberg, M. E. (2006). Body dissatisfaction prospectively predicts depressive mood and low self-esteem in adolescent girls and boys. *Journal of Clinical Child and Adolescent Psychology, 35*(4), 539-549.
- Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, MA: Cambridge University Press.
- Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys, 3*, 96-146.
- Pearl, J. (2012). The causal mediation formula—a guide to the assessment of pathways and mechanisms. *Prevention Science, 13*, 426–436.

- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91-105.
- Plotnikoff, R. C., Pickering, M. A., Flaman, L. M., & Spence, J. C. (2009). The role of self-efficacy on the relationship between the workplace environment and physical activity: A longitudinal mediation analysis. *Health Education & Behavior, 37*(2), 170-185.
- Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin, 88*(2), 245-258.
- Reddy, R., Rhodes, J. E., & Mulhall, P. (2003). The influence of teacher support on student adjustment in the middle school years: A latent growth curve study. *Development and Psychopathology, 15*(1), 119-138.
- Reichardt, C. S. (2011). Commentary: Are three waves of data sufficient for assessing mediation? *Multivariate Behavioral Research, 46*(5), 842-851.
- Reinboth, M., & Duda, J. L. (2006). Perceived motivational climate, need satisfaction and indices of well-being in team sports: A longitudinal perspective. *Psychology of Sport and Exercise, 7*(3), 269-286.
- Richter, A., Näswall, K., & Sverke, M. (2010). Job insecurity and its relation to work—family conflict: Mediation with a longitudinal data set. *Economic and Industrial Democracy, 31*(2), 265-280.
- Robitaille, A., Piccinin, A. M., Muniz-Terrera, G., Hoffman, L., Johansson, B., Deeg, D. J., & Hofer, S. M. (2013). Longitudinal mediation of processing speed on age-related change in memory and fluid intelligence. *Psychology and Aging, 28*(4), 887-901.
- Simons, R. L., Chao, W., Conger, R. D., & Elder, G. H. (2001). Quality of parenting as mediator of the effect of childhood defiance on adolescent friendship choices and

- delinquency: A growth curve analysis. *Journal of Marriage and Family*, 63(1), 63-79.
- Rubin, D. B. (2004). Direct and indirect causal effects via potential outcomes. *Scandinavian Journal of Statistics*, 31(2), 161-170.
- Shpitser, I. (2013). Counterfactual graphical models for longitudinal mediation analysis with unobserved confounding. *Cognitive Science*, 37(6), 1011-1035.
- Shrout, P. E. (2011). Commentary: Mediation analysis, causal process, and cross-sectional data. *Multivariate Behavioral Research*, 46(5), 852-860.
- Spinrad, T. L., Eisenberg, N., Cumberland, A., Fabes, R. A., Valiente, C., Shepard, S. A., & Guthrie, I. K. (2006). Relation of emotion-related regulation to children's social competence: A longitudinal study. *Emotion*, 6(3), 498-510.
- Spinrad, T. L., Eisenberg, N., Gaertner, B., Popp, T., Smith, C. L., Kupfer, A., & Hofer, C. (2007). Relations of maternal socialization and toddlers' effortful control to children's adjustment and social competence. *Developmental Psychology*, 43(5), 1170-1186.
- Stahl, D., Rimes, K. A., & Chalder, T. (2014). Mechanisms of change underlying the efficacy of cognitive behavior therapy for chronic fatigue syndrome in a specialist clinic: A mediation analysis. *Psychological Medicine*, 44(06), 1331-1344.
- Swart, H., Hewstone, M., Christ, O., & Voci, A. (2011). Affective mediators of intergroup contact: A three-wave longitudinal study in South Africa. *Journal of Personality and Social Psychology*, 101(6), 1221-1238.
- Taris, T. W., & Kompier, M. A. (2006). Games researchers play—Extreme-groups analysis and mediation analysis in longitudinal occupational health research. *Scandinavian Journal of Work, Environment & Health*, 32(6), 463-472.

- Tein, J. Y., Sandler, I. N., Ayers, T. S., & Wolchik, S. A. (2006). Mediation of the effects of the Family Bereavement Program on mental health problems of bereaved children and adolescents. *Prevention Science, 7*(2), 179-195.
- Turner, J. E., & Schallert, D. L. (2001). Expectancy–value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology, 93*(2), 320-329.
- von Soest, T., & Hagtvet, K. A. (2011). Mediation analysis in a latent growth curve modeling framework. *Structural Equation Modeling, 18*(2), 289-314.
- Wickrama, K. A., Lorenz, F. O., & Conger, R. D. (1997). Parental support and adolescent physical health status: A latent growth-curve analysis. *Journal of Health and Social Behavior, 38*(2), 149-163.
- Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10-18.
- Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116*(2), 363-381.
- Wills, T. A., Sargent, J. D., Stoolmiller, M., Gibbons, F. X., & Gerrard, M. (2008). Movie smoking exposure and smoking onset: A longitudinal study of mediation processes in a representative sample of US adolescents. *Psychology of Addictive Behaviors, 22*(2), 269-277.
- Wolchik, S. A., Tein, J. Y., Sandler, I. N., & Ayers, T. S. (2006). Stressors, quality of the child–caregiver relationship, and children's mental health problems after parental death: The mediating role of self-system beliefs. *Journal of Abnormal Child Psychology, 34*(2), 212-229.

- Wood, A. M., Maltby, J., Gillett, R., Linley, P. A., & Joseph, S. (2008). The role of gratitude in the development of social support, stress, and depression: Two longitudinal studies. *Journal of Research in Personality, 42*(4), 854-871.
- You, J., Lin, M. P., & Leung, F. (2015). A longitudinal moderated mediation model of nonsuicidal self-injury among adolescents. *Journal of Abnormal Child Psychology, 43*(2), 381-390.

