

國立政治大學「教育與心理研究」
2009年12月，32卷4期，頁1-22

刻板印象如何影響行爲？談刻板印象效果的內在歷程與影響

洪嘉欣* 陳皎眉**

摘要

本研究為單因子設計，並以理學院大學數學學科能力測驗成績在7級分以上的女大學生49名做為研究對象，操弄兩種刻板印象：「女性的數學較差」與「理學院的學生數學能力較好」，來探討被激發正向或負向刻板印象的個體之工作表現、工作選擇與自我能力評估，並驗證考試焦慮與工作記憶容量做為刻板印象效果的中介變項之可能性。

結果發現，當受試者被激發正向刻板印象時，其工作記憶容量會被提升，並選擇較困難的數學作業，但刻板印象威脅效果沒有顯現。而三組受試者，儘管在自陳式考試焦慮量表上沒有顯現出差異，但在脈搏測量上則顯現出組間差異。另外，本研究所提出的刻板印象效果之中介機制，並未在本實驗中得到支持。

關鍵詞：刻板印象威脅、刻板印象提升、工作記憶容量、考試焦慮、作業選擇

* 洪嘉欣：國立政治大學心理學系博士生

** 陳皎眉：國立政治大學心理學系教授

電子郵件：94752009@nccu.edu.tw

收件日期：2008.09.03；修改日期：2008.12.01；接受日期：2008.12.15

Journal of Education & Psychology
December, 2009, Vol. 32 No. 4, pp. 1-22

How Does Stereotype Affect Human Behavior? Stereotype Effect and Its Mechanism

Jia-Sin Hong^{*} Jeaw-Mei Chen^{**}

Abstract

Using single-factor design (gender-identity group, science-identity group, and control group), the present study intended to investigate both stereotype threat effects and stereotype boost effects on the working memory capacity, task performance, task selection, and self-evaluation. The authors also explored the possibility of anxiety and working memory capacity as the mediators of stereotype effects.

The results indicated the following: In the experiment, the participants in the science-identity group performed better in working memory capacity task and chose more difficult tasks to perform than the participants in control group. However, there were no such differences between the gender-identity group and control group. The results also revealed significant differences among three groups in biological anxiety measurement (pulsation), yet such differences were not found in the self-report anxiety measurement. Finally, the hypothesis that anxiety and working memory capacity were the mediators of stereotype effects was not supported by the current results.

Keywords: stereotype threat effect, stereotype boost effect, working memory capacity, test anxiety, task selection

* Jia-Sin Hong: Graduate student, Department of Psychology, National Chengchi University

** Jeaw-Mei Chen: Professor, Department of Psychology, National Chengchi University

E-mail: 94752009@nccu.edu.tw

Manuscript received: 2008.09.03; Revised: 2008.12.01; Accepted: 2008.12.15

壹、緒論

在大學校園中往往發現某些科系像是專屬於男性就讀，少有女性主修，如：數學、機械、土木等，反之，亦有某些科系像是專屬於女性，少有男性主修，如：護理、國文、英文等。而造成這種現象的原因除了男女在先天上的差異之外，主要還是因為男女性從小到大在社會化的過程中，因為接受到不同的對待而導致的結果，例如：在高中時男性被要求修習以木工、電器為主的工藝課，而女性則被要求修習以烹飪、裁縫為主的家政課，而造成這種不同對待的原因主要是來自於根深蒂固的性別刻板印象。

刻板印象指的是個體對於某一團體成員固定的、僵化的看法（陳皎眉、王叢桂、孫蒨如，2003；陳皎眉、孫曼暉，2006），這種看法有好有壞，並且在社會大眾中具有一定的普遍性。人們會依據這些刻板印象來決定要不要和某團體的某人交往或是跟此人的交往方式，屬於團體的個人往往也會受到此刻板印象的影響，而表現出相符合的行為，並更加深他人所抱持的刻板印象。

Steele與Aronson（1995）是最早探討「刻板印象威脅」的研究者，他們認為，當情境激發與個體有關的刻板印象，而此印象為負向時，個人會因為擔心自己的表現符合此負向的刻板印象，

而使得其表現下降。在他們的實驗中，讓黑人與白人受試者接受一困難的語文測驗。在「刻板印象威脅組」中，這份測驗被描述成「智力測驗」，而在「非刻板印象威脅組」中，這份測驗則被描述成與智力無關的「問題解決測驗」。結果發現，在刻板印象威脅組中，黑人的作業表現會因為受到種族刻板印象（黑人的智力較差）的影響而表現地較差，然而在非刻板印象威脅組中，黑人的表現則與白人沒有差異。而Steele與Aronson的另一個實驗則發現個體在受到刻板威脅的情況下，會花費較多的時間、用較不精確的方式來處理較少的題目，換言之，他們的努力並沒有得到相符的效能。這可能是由於他們必須在回答問題的同時，另外花費心力來調節因受威脅而起的挫折感。

Spencer、Steele及Quinn（1999）的實驗則顯示，當一個數學測驗在指導語中被描述為具有性別差異且女性會表現較差時，女性受試者的測驗表現會較男性受試者來得差，然而，當此測驗被描述為沒有性別差異時，則男女性在此測驗上的表現並無差異。Spencer等人認為這種效果是由於女性在接受測驗時，比起男性來說，須額外負擔對「女性的數學能力較男性來得差」的負向刻板印象，所以導致她們的數學表現較不理想。

然而，Shih、Pittinsky和Ambady

(1999) 的研究則發現有與刻板印象威脅相反的情形。他們認為個體同時會擁有對於多種不同團體的認同，而各個團體都伴隨不同的刻板印象，這些不同的刻板印象有正向的也有負向的，例如：對於亞裔的女大學生來說，她們同時具有兩種團體認同：女性與亞洲人，前者伴隨的是「女性數學較不好」的刻板印象，後者則是：「亞洲人的數學很好」。此時，當個體被激發不同的認同時，會隨著不同的認同而產生不同的行為。Shih等人以亞裔女大學生為受試者，並採用隱性激發 (subtle stereotype activation) 的方式，激發受試者性別或種族的刻板印象。此種隱性激發方式是請受試者填答有關不同族群認同的問卷，例如：在激發性別認同組，詢問受試者「所住的宿舍是單一性別或混合了不同性別？」；在種族認同組則是詢問受試者「妳的父母或家中長輩是否說除英語之外的語言？」；在無激發組則是詢問「妳滿意妳的手機業者嗎？」結果發現，被隱性激發亞裔認同的個體，其數學表現明顯比控制組（無激發任何認同）要來的好，然而，被激發女性認同的受試者，他們的數學表現則比控制組來的差。Cheryan 與 Bodenhausen (2000) 則將這種「因刻板印象而導致表現提升」的現象，命名為「刻板印象提升」(stereotype boost)。

除了與性別和種族相關的刻板印

象會產生刻板印象威脅效果，Croizet、Despres、Gauzins、Huguet、Leyens 和 Meot (2004) 的研究亦發現與科目 (subject) 相關的刻板印象亦可產生刻板印象威脅效果，他們針對主修不同科目（心理學或科學）的受試者進行研究，結果發現，當測驗被描述為是一個對數學及邏輯推理能力具診斷性的測驗時，主修科學受試者的表現會比主修心理學的受試者來得好，然而，當測驗被描述為非診斷性的測驗時，主修科學與主修心理學的受試者的表現則沒有差異。

因此，本研究的第一個目的是延續過往的國外研究，以臺灣的受試者來探討一個同時具有正負向刻板印象族群（理組女生），當被激發不同的所屬團體認同（性別／科目）時，他們的表現是否會有所不同。

另外，目前有關於刻板印象的研究大多著重於探討如何減除刻板印象威脅的影響，而對刻板印象為何會對個體產生影響之內在歷程的研究卻很少得到顯著結果。近來，Schmader 與 Johns (2003) 的實驗結果對此內在歷程提供一個可能的解釋。

Schmader 與 Johns (2003) 的實驗使用Turner與Engle (1989)「工作記憶容量測驗」(working memory capacity test) 替代傳統的數學測驗，並告知「刻板印象威脅組」的受試者此測驗能

反映受試者的數學能力，且男女性在數學能力的表現上有性別差異。而控制組受試者則被告知此測驗為測量工作記憶容量的測驗，而未提及男女在數學能力的表現上有無性別差異。

結果發現，在「刻板印象威脅組」中的男性受試者，其工作記憶容量測驗的表現與在控制組中的表現並無差異，然而，當女性受試者身處對其不利的「刻板印象威脅組」時，其表現會低於控制組，亦即此結果證實當個體在面臨刻板印象威脅時，其工作記憶容量會隨之降低。

接著，Schmader與Johns（2003）同時測量受到刻板印象威脅的女性受試者的工作記憶容量與數學能力，並採用Baron與Kenny（1986）的方法來檢驗工作記憶容量是否為刻板印象威脅與數學能力測驗表現間的中介變項。結果發現，刻板印象威脅的確對於工作記憶容量測驗與數學測驗分數有相當高的預測力，但當同時以工作記憶容量測驗分數與刻板印象威脅來預測數學測驗表現時，刻板印象威脅就不再是數學測驗表現的顯著預測變項，顯示工作記憶容量的確為刻板印象威脅與數學測驗表現的中介變項。也就是說，受到刻板印象威脅的個體因其工作記憶容量降低，而導致其數學能力表現較沒有被刻板印象威脅的個體來得不好。但是刻板印象提升效果是否會造成工作記憶容量上升，進

而導致作業表現上升則尚未有研究探討。因此，本研究的第二個目的是探討被激發正面刻板印象的當事人，其工作記憶容量是否會上升。

除了工作記憶容量可能影響作業表現外，什麼樣的機制影響工作記憶容量，也是值得進一步探討的主題。研究者認為考試焦慮或擔心被評價等情緒反應可能是刻板印象影響工作記憶的中介變項。一方面因為過去一些研究發現焦慮與擔心被評價是刻板印象與作業表現間可靠的調節變項（Marx & Stapel, 2006; Smith, 2004），另一方面，因為探討緊張程度與作業表現兩者之間關係的Yerkes-Dodson法則顯示，當面臨像考試這類的情境時，太緊張或太不緊張的人作業表現會不佳，因為太緊張會導致注意力窄化，太不緊張則導致注意力不集中，而會產生最佳表現的是中間程度緊張的人，這種狀態被稱為「最適度的緊張」（鄭昭明，2004）。

研究者認為當個體處於一個被評價的環境中，他的考試焦慮會提升，因此，需要花費一部分工作記憶容量來抑制考試焦慮。而若在此時被激發的是負向刻板印象，則其考試焦慮會更提升，需要花更多的工作記憶容量來抑制考試焦慮。反之，若激發的是正向刻板印象，則會降低考試焦慮，使其花較少的工作記憶容量來抑制考試焦慮。因此，研究者認為考試焦慮不只影響作業表

現，還會影響工作記憶容量。綜合以上所述，本研究的第三個目的是探討被激發正負面刻板印象的當事人，其考試焦慮是否會不同，以及考試焦慮是否為刻板印象激發與工作記憶容量的中介變項。

再者，Suen (2006) 的研究指出，不論正面或負面的刻板印象，除了會影響當事人的作業表現之外，還會影響他們對於自己表現結果的預期，這顯示刻板印象對當事人所造成的影响不只發生在作業表現上，但過往研究多半只專注在刻板印象對於作業表現的影響，而忽略了刻板印象的其他影響。因此，本研究一方面欲檢驗被激發刻板印象的當事人如果有機會選擇時，對於作業難度的選擇，另一方面則欲檢驗當事人對自己能力，以及對作業難度的評估。選擇「作業難度選擇」做為檢驗項目是因為研究者認為其會影響之後的學習意願。因為若當事人選擇一個比自己程度難度稍高的測驗時，他雖然會覺得此測驗困難，但相較於難度最高的測驗，當事人在此難度成功的機率仍是相當高的，倘若其挑戰成功，也會增強當事人對於與此作業對應的能力的自信（例如：數學作業對應數學能力），也會增加他對於此作業的興趣，也會更願意投注時間在此增進此能力上，而當下次遇到挑戰機會時，此成功的個人也會再挑戰稍高的難度，長遠下來，個人的能力

就會提升。相反地，當個人選擇一個很容易或極端困難的難度時，前者因為很容易達到目標，所以並不會增強其對於此能力的信心，後者雖然達成目標可以增強信心，但成功的可能性太低，對個人都是較為不利的。

而選擇「對作業難度評估」與「對自己能力的評估」做為檢驗項目的原因則是，研究者一方面想要了解刻板印象是否會對當事人在與能力相關的知覺上造成影響，另一方面則認為當事人若認為自己在群體中的表現不錯，就會更願意花時間增進自己的能力，也較願意面對與此能力相關的環境（例如：選與數學相關的科系）。相反地，若此人認為自己的能力不佳，則較有可能逃避與此能力相關的環境，如此一來，也沒有機會增進此項能力。因此，本研究的第四個目的是探討被激發正負面刻板印象的當事人，其「作業難度選擇」、「對作業難度評估」和「對自己能力的評估」是否會不同。

除此之外，雖然很多研究者都假設，刻板印象威脅會導致個人作業表現下降，而受威脅團體的作業表現下降則會反過來使得刻板印象被加深（Steele & Aronson, 1995），但這種循環卻尚未得到實證的支持。因此，本研究在實驗的最後也將測量受試者的「刻板印象相信程度」，來了解經過一連串模擬真實情況的實驗之後，刻板印象的相信程度

是否會被加深，此為本研究的第五個目的。

綜合以上所述，本研究欲提出一可能的內在機制歷程，以探討在正、負向刻板印象的激發下，是否會影響一個人的考試焦慮程度，而個人的考試焦慮程度是否會影響其工作記憶容量，進而導致其作業能力的表現受到影響。

另外，本研究還想探討，在正、負向刻板印象被激發的情況之下，受試者對於「作業難度選擇」、「作業難度評估」與「對自己能力的評估」是否會和沒有被激發刻板印象的受試者不同。而在經歷一連串實驗歷程後，受試者對於「刻板印象的相信程度」是否會增加。（研究架構圖如圖1）

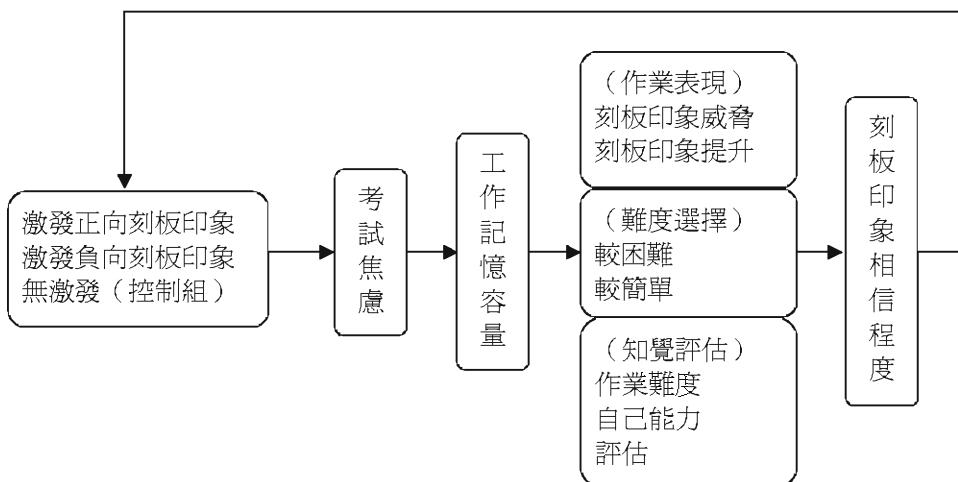


圖 1 研究架構圖

貳、研究方法

首先，為確定大學生的確存有「理學院的學生數學能力較其他學院學生來得好」的刻板印象，因此，請政治大學修習通識課「人際關係」的142名學生，對此刻板印象敘述句做相信程度的七點量尺評分，1代表非常不相信，7代表非常相信。結果發現，平均得分為5.36，標準差為1.27，證實此刻板印象

的確存在，遂進行以下研究。

一、研究對象

本研究的對象為國立政治大學理學院大學部且大學數學學科能力測驗成績在7級分以上的女學生共49人，每一個實驗組的受試者數為16~17人。其中，由於政治大學理學院除心理學系外的資科系與應用數學系的女學生較少，所以參與本次實驗的女學生大多所屬心

理學系（心理系：46名；應用數學系：3名）。而在年級方面，各年級參與實驗的人數差別都不大（一年級：14人；二年級：11人；三年級：13人；四年級：11人）。

二、實驗設計

本研究為單因子設計，獨變項為不同團體認同的激發，分為三組：性別認同組、理學院認同組、控制組，獨變項操弄參考Shih等人（1999）的方法，請各組受試者填答有關於其認同團體的10個問題，藉此激發她們對於目標團體的認同，以及相對應的刻板印象：「女性數學能力較差」或「理學院的學生數學能力較好」。例如：性別認同組的受試者需回答她們的性別、是否贊成男女分班或分校等問題；理學院認同組的受試者則需回答她們的學院別、共同必修科目等問題；而控制組的受試者則被詢問有關於其閱讀習慣之問題。

依變項共有八項指標：工作記憶容量測驗分數、數學測驗分數、考試焦慮量表分數、考試焦慮生理測量、數學測驗選擇難度、數學測驗難度評估、自我評估數學能力、刻板印象相信度。

三、實驗程序

當受試者進入實驗室後，研究者將其隨機分派至三組中的一組後進行個別施測。受試者被告知研究的目的是要

檢驗知覺處理速度與數學能力測驗間的關聯，所以受試者必須從事兩種工作（tasks），一是知覺處理工作，另一則是數學測驗。而考量受試者的生理狀態可能對研究結果有所影響，因此，在實驗的過程中，會進行四次的脈搏測量。但事實上，知覺處理工作為工作記憶容量測驗，而脈搏測量則為考試焦慮的生理測量。

當受試者閱讀完研究目的的指導語之後，實驗者會幫受試者進行第一次的脈搏測量。在脈搏測量結束後，研究者依其組別（性別認同組、理學院認同組、控制組）而給予不同的獨變項操弄問卷，並告知問卷的目的是在調查大學生活有關的一些問題，待受試者填答完操弄問卷後，則進行第二次的脈搏測量。

接著，請受試者仔細閱讀電腦螢幕上的知覺速度工作指導語，並依照指示開始作答。在此工作進行到一半之時，實驗者進行第三次的脈搏測量。

當受試者完成知覺處理工作之後，研究者詢問受試者：「如果有七種不同難度的數學測驗可以供您選擇填答，請問您願意選擇哪一份作答？」（難度為在100位臺灣的大學生中大約有20、30、40、50、60、70或80人答對）。待受試者回答之後，實驗者告知受試者，本次測驗為了要統一讓每位受試者都填答到相同難度的測驗，所以她

們後續所要接受的測驗難度和之前的難度選擇是沒有相關的，接著並給予「數學能力測驗指導語與練習題」，待其練習完後就進入正式數學能力測驗。

最後，當數學能力測驗結束後，進行第四次脈搏測量。接著，受試者被要求以七點量表（與「數學難度選擇」相同）來評估數學測驗的難度，並填答一份有關於她們的背景資料與實驗中主觀感受的問卷。此問卷內容包括：「考試焦慮量表」、「自我評估數學能力量表」、「做此測驗的盡力程度」、「學科能力測驗數學成績」、「是否接受過GRE相關測驗」、「刻板印象的相信程度量表」。實驗結束後，實驗者則會告知受試者本研究的真正目的，以及回答受試者與研究相關的問題。

四、實驗材料

(一) 工作記憶容量測驗 (working memory capacity test)

採用 Turner 與 Engle (1989) 之 Sentence-word span task，是一種受試者必須在同一時間從事兩種作業的雙重測驗 (dual task)。受試者將被呈現一個含15~20字的句子（例如：昨天晚上母親煮了麻油雞當作晚餐），受試者必須在5秒鐘內判斷電腦螢幕上呈現的句子是否有意義，並把答案記在答案紙的A部分（例：此為有意義句子，應寫下「O」），且記住此句子的最後一字

「餐」，接著從事下個判斷（例如：每個他們都吃會一起下午禮拜三茶；此為無意義句子，應寫下「X」，並記住「茶」字），接著再從事下個判斷（例如：昨天晚上大約12點出現了月全蝕；此為有意義句子，應寫下「O」並記住「蝕」字），直到出現實驗者所給予的線索鈴聲時，受試者必須把之前呈現各句的最後一字默寫在答案卷的B部分（例：「餐」、「茶」、「蝕」）。

本測驗的難度則是以需要背誦出來的字數決定，由三個到六個，每個難度各有五題，總共20個題組。

記分的方式採用「部分廣度計分」(partial-credit load scoring) (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005)，此計分法是不論受試者在句子判斷是否答對，只要能正確回憶出每個題組中幾個字就算得幾分，回憶字的順序則不拘，例如：答對三個字詞的題組則得三分，六個字詞的題組則得六分。而當受試者在五個字詞的題組只答對出四個字詞時，則得四分。本測驗最高可能總分為90分。

(二) 數學能力測驗

採用邱蜀娟 (2005) 經由GRE數學測驗，以及其自編題目中答對率15%~30%的題目所組成的「數學正式測驗」。此測驗共有16題，其中第1、2、4、6、7、10、12、13、16等9題為大小判斷題，選擇A表示認為左邊選項比右

邊選項大，選擇B表示認為右邊選項比左邊大，選擇C表示認為兩者一樣大，選擇D表示認為題目的訊息不足，無法判斷。而測驗的第3、5、8、9、11、14、15共7題為選擇題，受試者應從A、B、C、D、E五個選項中選取正確答案。受試者有15分鐘來完成所有題目。

為了避免受試者因不熟悉作答方式而答錯，研究者在進行正式測驗之前，給予題型的解說，以及6題練習題。練習題來自於邱蜀娟（2005）的「數學練習題」中的第1、2、3、13、15、18。

(三)考試焦慮量表

採用邱蜀娟（2005）以林莉萱（2003）參考鍾思嘉與龍長風（1984）翻譯 Spielberger、Gorsuch 和 Lushene（1970）的情境特質焦慮量表（State-trait anxiety inventory）所改編而成的「情境焦慮量表」，其內部一致性之 Cronbach α 值為 0.85。而由於本研究之受試者是在做完所有測驗後才填寫此量表，因此本研究者在每一個問題前均加入「在測驗的過程中」，以期受試者能回憶起在接受測驗時的感受，而此考試焦慮量表之內部一致性 Cronbach α 值為 0.89。

此量表共 8 題，其中第 1 題與第 4 題為反向題，受試者用 Likert 七點量尺的方式評估句子的符合程度，總得分愈高

代表考試焦慮愈高。

(四)考試焦慮生理測量

本研究在實驗中總共會進行四次脈搏測量，第一次脈搏測量在受試者讀完實驗指導語之後，第二次脈搏測量則在受試者填答完不同認同團體問卷後（獨變項操弄完），第三次脈搏測量則在工作記憶容量測驗進行至一半時，第四次脈搏測量則在受試者完成數學能力測驗後。

(五)自我評估數學能力量表

共 2 題，一題是請受試者用七點量尺的方式評估句子「整體而言，我覺得我的數學能力很好」的符合程度，得分愈高代表自我評估數學能力愈好。另一題是「在全臺灣的大學生中，我認為我的數學能力在 100 人中較____人為佳」。

(六)刻板印象相信程度量表

共 10 題，請受試者用七點量尺的方式評估她們認為下列刻板印象的敘述句為真實的程度，得分愈高代表愈相信此刻板印象為真。而研究者將只檢驗第 4 題「男性的數學能力比女性好」與第 7 題「理學院的學生數學能力較其他學生來得好」。

(七)個人資料

共兩題，一是請受試者回報其大學學科測驗的級分，另一題則是詢問受試者過去是否填答過 GRE 或類似測驗的數學試題。

(八)與研究相關問題

共兩題，請受試者用七點量尺的方式評估對於「知覺速度工作（工作記憶容量測驗）」與「數學測驗」的盡力程度，1代表完全沒有盡力，7代表完全盡力，得分愈高表示受試者愈盡力回答測驗問題。

參、研究結果

研究結果將分成三部分來呈現，第一部分為獨變項「不同認同團體激發」對各依變項的統計檢驗，第二部分則為控制變項的統計檢驗，第三部分則探討不同認同團體激發、考試焦慮生理測量、工作記憶容量測驗與數學能力測驗四者之間的關係。

一、依變項的檢驗

(一) 考試焦慮量表

研究者以每位受試者在考試焦慮量表上的總分做為依變項進行統計檢

驗。首先，所有受試者總分的平均數為 28.63，標準差為 8.15。其次，研究者以「不同認同團體激發」（性別認同組、理學院認同組、控制組）為獨變項，對「考試焦慮量表總分」做單因子 ANOVA 分析，結果發現三組平均數並沒有達到顯著差異 ($F_{(2,46)} = 2.011, p = .145$)。顯示三組在考試焦慮量表的得分上並沒有差異。

(二) 考試焦慮生理測量

研究者以「不同認同團體激發」為獨變項，以受試者「第二次」、「第三次」、「第四次」脈搏測量三者做為依變項，受試者「第一次脈搏測量」為共變項，進行 ANCOVA 分析，檢驗不同實驗階段三組受試者的考試焦慮是否有所不同。而各組受試者在控制第一次脈搏測量後的考試焦慮生理測量總分的平均數與標準差如表 1 所示：

表 1 各組調整後考試焦慮生理測量之平均數與標準差

| | 女性認同組 | 理學院認同組 | 控制組 |
|-----|----------------------|----------------------|----------------------|
| 第二次 | 76.32 (.83) n=16 | 73.84 (.81) n=17 | 76.59 (.832) n=16 |
| 第三次 | 75.91 (1.21) n=16 | 73.35 (1.17) n=17 | 76.16 (1.20) n=16 |
| 第四次 | 76.08 (1.55) n=16 | 71.95 (1.49) n=17 | 75.22 (1.54) n=16 |

註：括弧外為各組平均數，括弧內為各組標準差。

在進行 ANCOVA 分析之前，研究者先進行「組內回歸係數同質性」

（homogeneity of within-class regression coefficient）的假設檢定，結果發現，

獨變項與共變項之間並沒有交互作用效果存在 ($F_{(10,16)} = 1.494, p = .229$; $F_{(10,16)} = 1.238, p = .339$; $F_{(10,16)} = 1.365, p = .279$)，顯示各組內的共變項與依變項的線性關係具有一致性，符合假設檢定，可以進行ANCOVA分析。

而ANCOVA的分析發現，第一次測量脈搏對於第二、三、四次的脈搏測量皆具有顯著影響 ($F_{(1,45)} = 454.891, p = .000$; $F_{(1,45)} = 93.474, p = .000$; $F_{(1,45)} = 49.121, p = .000$)，而獨變項「不同認同團體激發」則在第二次脈搏測量上達到顯著水準 ($F_{(2,45)} = 3.469, p = .04$)，在第三、第四次脈搏測量上卻沒有達到顯著水準 ($F_{(2,45)} = 1.743, p = .187$; $F_{(2,45)} = 2.079, p = .137$)。顯示，儘管各組受試者在考試焦慮量表上的分數並沒有差異，受試者在獨變項操弄結束後的第2次脈搏測量上的確顯現出不同。

(三)工作記憶容量測驗

在所有樣本中，有一名受試者因為誤解指導語而未能正確完成此測驗，研究者在排除此人此項的測驗成績後，對剩下48人的資料進行分析。首先針對工作記憶容量測驗得分的分析發現，所有受試者的平均分數為71.50，標準差為10.92。接著，研究者以「不同認同團體激發」(性別認同組、理學院認同組、控制組)為獨變項，對「工作記憶容量測驗分數」做單因子ANOVA分析，結果發現獨變項「不同認同團體激

發」效果達到顯著 ($F_{(2,45)} = 4.005, p = .025$)。使用Tukey事後比較發現，理學院認同組的受試者在工作記憶容量測驗的表現的確要比控制組的受試者表現來得好 ($M_{理} = 75.76, M_{控} = 65.60, t(45) = 10.165, p = .021$)。而女性認同組的受試者表現和控制組的受試者表現則沒有差異 ($M_{女} = 72.50, M_{控} = 65.60, t(45) = 6.900, p = .160$)。顯示只有理學院認同組的受試者在工作記憶測驗上的表現比控制組的受試者來得好。

(四)數學難度選擇

所有受試者所選擇的數學難度之平均數為4.41，標準差為1.66。研究者以「不同認同團體激發」(性別認同組、理學院認同組、控制組)為獨變項，對「數學難度選擇」做單因子ANOVA分析，結果發現獨變項「不同認同團體激發」效果達到顯著 ($F_{(2,46)} = 3.658, p = .034$)。使用Tukey事後比較發現，在數學難度選擇上，理學院認同組的受試者比控制組的受試者更傾向選擇較難的數學測驗 ($M_{理} = 5.24, M_{控} = 3.88, t(46) = 1.360, p = .043$)。而女性認同組的受試者所選擇的數學難度則和控制組的受試者沒有差異 ($M_{女} = 4.06, M_{控} = 3.88, t(46) = .188, p = .939$)。

(五)數學能力測驗

研究者排除1名沒有大學數學學科

能力測驗成績者，以及8名曾經接受過GRE測驗的受試者後，對剩下40人的資料進行分析。結果發現，受試者在數學能力測驗上的平均答對題數為7.83題，標準差為2.55。而答對正確率（答對的題數除以嘗試回答的題數）的平均數為.58，標準差為.17。

研究者以「不同認同團體激發」為獨變項，以受試者在數學能力測驗的「答對題數」與「正確率」做為依變項，受試者「大學數學學科能力測驗成績」為共變項，進行ANCOVA分析。

在進行ANCOVA分析之前，研究者先進行「組內回歸係數同質性」的假設檢定，結果發現，獨變項與共變項之間並沒有交互作用效果存在 ($F_{(5,24)} = .723, p = .613$; $F_{(5,24)} = .712, p = .620$)，顯示各組內的共變項與依變項的線性關係具有一致性，符合假設檢定，可以進行ANCOVA分析。

而在進行ANCOVA分析之後發現，「大學數學學科能力測驗成績」在兩個依變項上皆有顯著的影響 ($F_{(1,36)} = 19.853, p = .000$; $F_{(1,36)} = 15.766, p = .000$)，控制「大學數學學科能力測驗成績」之後，獨變項「不同認同團體激發」不論在「答對題數」或「正確率」上皆沒有達到顯著水準 ($F_{(2,36)} = 1.478, p = .241$; $F_{(2,36)} = 1.089, p = .347$)，顯示不同認同團體的激發對於受試者在數學能力測驗上的表現沒有

影響。

(六)數學能力測驗難度評估

所有受試者對於數學能力測驗的平均難度評估為3.63，標準差為1.30。研究者以「不同認同團體激發」（性別認同組、理學院認同組、控制組）為獨變項，對「數學能力測驗難度評估」做單因子ANOVA分析，結果發現三組並沒有達到顯著差異 ($F_{(2,46)} = 1.419, p = .252$)。顯示三組受試者對於數學能力測驗的難度評估並沒有差異。

(七)自我評估數學能力

在所有樣本中，有一名受試者因為為僑生而沒有大學數學學科能力測驗成績，研究者在排除此人之後，對剩下48人的資料進行分析。結果發現，受試者在「整體而言，我覺得我的數學能力很好」的七點量尺評估的平均得分為4.06，標準差為1.19。而「在全臺灣的大學生中，我認為我的數學能力在100人中較____人為佳」的平均得分為61，標準差為16.93。

接著，研究者以「不同認同團體激發」為獨變項，以受試者在此兩題的回答做為依變項，受試者「大學數學學科能力測驗成績」為共變項，進行ANCOVA分析。

在進行ANCOVA分析之前，研究者則先進行「組內回歸係數同質性」的假設檢定，結果發現，獨變項與共變項之間並沒有交互作用效果存在 ($F_{(6,31)} =$

$F_{(6,31)} = 1.261, p = .303$ ），顯示各組內的共變項與依變項的線性關係具有一致性，符合假設檢定，可以進行ANCOVA分析。

而在進行ANCOVA分析之後發現，「大學數學學科能力測驗成績」對於兩個依變項皆有顯著的影響 ($F_{(1,44)} = 13.165, p = .001$; $F_{(1,44)} = 6.630, p = .013$)，而在控制「大學數學學科能力測驗成績」後，獨變項「不同認同團體激發」的效果則在這兩題上都沒有達到顯著水準 ($F_{(2,44)} = .151, p = .861$; $F_{(2,44)} = 1.300, p = .283$)。表示，不同認同團體的激發對於受試者自評之數學能力沒有影響。

(iv) 刻板印象相信程度量表

所有受試者對於刻板印象相信程度量表第4題「男性的數學能力比女性好」的七點量尺平均得分為4.69，標準差為1.43。第7題「理學院的學生數學能力較其他學生來得好」的平均得分為4.90，標準差為1.25。

研究者以「不同認同團體激發」（性別認同組、理學院認同組、控制組）為獨變項，對刻板印象相信程度量表的第4題與第7題分數為依變項，做單因子ANOVA分析，結果發現在這兩題上，三組受試者間的得分皆沒有達到顯著差異 ($F_{(2,46)} = 1.805, p = .176$; $F_{(2,46)} = 1.056, p = .356$)。顯示，三組受試者在完成實驗的各階段後對於兩種刻板印

象的相信程度都沒有差異。

二、控制變項的檢驗

盡力程度。研究者為了解三組受試者對於「工作記憶容量測驗」及「數學能力測驗」的盡力程度是否都相同，進行有關「盡力程度」二題項的統計檢驗。結果發現，受試者對於工作記憶容量測驗的盡力程度七點量尺的平均數為5.12，標準差為1.05。而在數學能力測驗盡力程度的平均數為5.24，標準差為.99。

研究者以「不同認同團體激發」（性別認同組、理學院認同組、控制組）為獨變項，受試者對於工作記憶容量測驗與數學能力測驗的盡力程度為依變項，做單因子ANOVA分析，結果發現，在這兩題上，三組皆沒有達到顯著差異 ($F_{(2,46)} = 1.037, p = .363$; $F_{(2,46)} = .046, p = .955$)。顯示，三組受試者對於工作記憶容量測驗或數學能力測驗的盡力程度都相同。

三、不同認同團體激發、考試焦慮、工作記憶容量與數學能力間關係

研究者以Baron與Kenny（1986）的方法來檢驗「不同認同團體激發」（獨變項操弄）、「考試焦慮生理測量」、「工作記憶容量測驗」與「數學能力測驗」四者之間的關係。首先檢驗

「考試焦慮生理測量」是否為「不同認同團體激發」與「工作記憶容量測驗」的中介變項。其次檢驗「工作記憶容量測驗」是否為「考試焦慮生理測量」和數學能力測驗的「答對題數」與「正確率」的中介變項。由於「考試焦慮生理測量」一共進行了四次，因此，研究者以最能代表獨變項「不同認同團體激發」之操弄效果的第二次脈搏測量；並控制代表個體基礎率的第一次脈搏測量後進行檢驗。

首先，研究者檢驗獨變項操弄對中介變項「考試焦慮生理測量」是否有顯著的預測效果，結果發現存有顯著效果 ($\beta = -.122, p = .022$; $\beta = .270, p$

$= .949$)。接著檢驗「考試焦慮生理測量」對「工作記憶容量測驗」是否有預測效果，結果並未發現有顯著效果 ($\beta = .137, p = .770$)。最後，檢驗「不同認同團體激發」是否對「工作記憶容量測驗」有預測效果，結果發現的確有顯著效果存在 ($\beta = .450, p = .008$; $\beta = .301, p = .068$)，而在排除「考試焦慮生理測量」之後，此顯著效果依舊存在 ($\beta = .439, p = .011$; $\beta = .292, p = .082$)，顯示，「考試焦慮生理測量」並非獨變項操弄與「工作記憶容量測驗」的中介變項，但受試者被激發的不同刻板印象的確會影響其工作記憶容量表現（見圖2）。

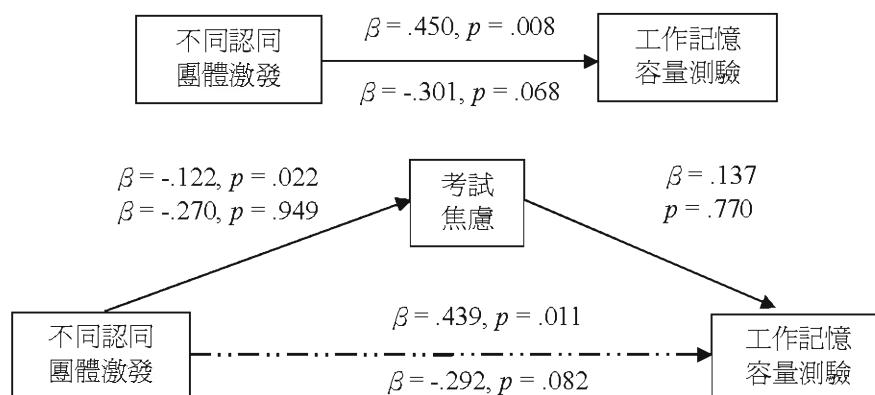


圖 2 不同認同團體激發與工作記憶容量測驗之中介變項模型

其次，研究者檢驗「考試焦慮生理測量」對「工作記憶容量測驗」是否有預測效果，結果並未發現有顯著效果 ($\beta = -.137, p = .770$)，接著檢驗「工作記憶容量測驗」是否能預測數學能力

測驗的「答對題數」與「正確率」，結果發現的確有顯著效果存在 ($\beta = .332, p = .021$; $\beta = .400, p = .005$)。最後，則檢驗「考試焦慮生理測量」是否能預測數學能力測驗的「答對題數」與「正

確率」，結果並未發現有顯著效果存在 ($\beta = -.325, p = .488$; $\beta = .311, p = .506$)，而在排除「工作記憶容量測驗」之後，亦未發現有顯著效果存在 ($\beta = -.278, p = .629$; $\beta = -.258, p$

= .555)，顯示，「工作記憶容量」並非「考試焦慮生理測量」與數學能力測驗的「答對題數」與「正確率」的中介變項，但其的確能預測受試者的數學能力測驗表現（見圖3）。

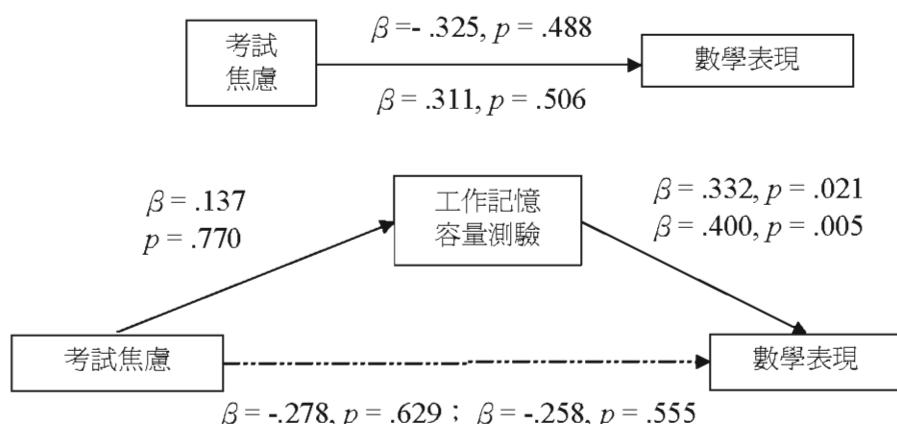


圖3 考試焦慮與數學表現之中介變項模型

肆、討論與建議

此段將分成兩部分討論，第一部分針對研究提出的問題進行討論，第二部分則探討本研究的限制，以及對未來研究的建議。

一、研究結果說明

本研究的目的是結合「刻板印象威脅效果」與「刻板印象提升效果」的概念，探討同時具有正、負刻板印象（理學院的學生數學能力較好；女性數學能力較差）的理學院女學生被激發不同的所屬團體認同（科目／性別）時，

對於受試者工作表現（數學能力測驗）、工作選擇（數學能力測驗難度選擇）與自我能力評估（數學能力評估）的影響，並驗證考試焦慮與工作記憶容量做為刻板印象效果的中介變項之可能性。本研究結果可分為四部分討論之。

第一，本研究同時採取正負向刻板印象操弄，探討先前研究尚未探討的議題：刻板印象提升效果是否也是因為工作記憶容量的上升而導致，並重複驗證刻板印象威脅效果會造成工作記憶容量下降之過去發現。結果發現，對於被激發正向刻板印象的受試者來說，他們的工作記憶容量與控制組的受試者相比

的確較好，且三組受試者對此測驗的自陳努力程度都相同，顯示刻板印象提升效果的確會造成工作記憶容量上升。而過去研究指出，高工作記憶容量的個體能較快、也較準確地想起與目標有關的知識、較能抑制不相關的訊息，也較會使用背景知識來解決問題（Barrett, Tugade, & Engle, 2004; Rosen & Engle, 1998）。因此，被激發正向刻板印象的受試者，對於問題解決的能力的確會提高。

然而，在另一方面，本研究卻無法重複驗證Schmader與Johns（2003）的發現，亦即，本研究發現對於被激發負向刻板印象的受試者來說，她們的工作記憶容量和控制組的受試者相比並沒有差異。研究者推估，這可能是由於本研究所找的受試者是政治大學理學院的女生，和Schmader與Johns所選的普通的大學女生及Shih等人（1999）所選的亞裔大學女生兩者相較，本研究的受試者本身可能就比較不相信「女生數學能力較差」刻板印象而選取理學院就讀（在本研究之刻板印象相信程度量表上，理學院的受試者對「男性的數學能力比女性好」的相信程度平均分數為4.69，顯示只有「稍微相信」）。因此，對於理學院的女生來說，他們本身就比較不會受到「女生數學能力較差」這個負面的刻板印象所影響，而導致本實驗的刻板印象威脅效果不論在「工作記憶

容量測驗」、「數學能力測驗」，以及「數學能力測驗難度選擇」上皆沒有顯現效果。

除此之外，本研究在「數學能力測驗」上沒有發現刻板印象效果，即三組受試者在「數學能力測驗」上的表現沒有差異，一方面可能是因為獨變項操弄的效果隨著時間的進行而減弱，而本實驗的實驗時間將近1小時，可能導致實驗後半的「數學能力測驗」、「數學能力測驗難度評估」及「對自己能力的評估」等依變項皆沒有顯現刻板印象效果。另一方面可能是因為，雖然邱蜀娟（2005）的「數學正式測驗」是選擇政治大學的學生平均答對率15%~30%的題目所組成（即，應算是偏難的題目），然而，本實驗的受試者（理學院女生）卻平均能答對近五成的題目（此測驗共16題中，受試者平均答對7.83題），亦即，對本研究的受試者來說這些題目並不太難。另外，檢視受試者們對「數學能力測驗難度」的七點量尺評估，發現他們的平均得分為3.63，亦顯示受試者主觀認為此測驗的難度為中間偏易。因此，亦有可能因為此測驗題目對於理學院的女生過於簡單而導致沒有顯現出刻板印象效果。

第二，研究者認為，考試焦慮應為刻板印象效果與工作記憶容量之中介變項，並採取兩種方式測量考試焦慮，一個為自陳式的「考試焦慮量表」，另

一個則是使用「脈搏」做為考試焦慮的生理測量。結果發現，儘管三組受試者在自陳式的考試焦慮量表上的回答並沒有顯著差異（平均得分為25.63、29、31.25），顯示受試者的自陳考試焦慮都偏低，然而，在四次的生理測量中卻發現，當研究者控制受試者的基礎率（第一次脈搏測量）後，獨變項「不同認同團體的激發」的確會對剛接受完操弄的受試者產生影響（第二次脈搏測量顯示出三組受試者間有差異）。然而，這種差異在第三次及第四次的脈搏測量並沒有發現。研究者認為這或許可以做為「受試者會付出心力來調節考試焦慮」的佐證，因為在實驗中，第二次與第三次脈搏測量的時間間隔不到10分鐘，但三組受試者在第二次脈搏測量中所展現出的差異在第三次脈搏測量卻消失了，而受試者在工作記憶容量上的表現卻顯現出刻板印象效果，證明此時獨變項「不同認同團體的激發」的效果尚未減弱，因此，研究者認為第二次與第三次脈搏測量間的差異應該是由於受試者付出心力來調節自己的考試焦慮，而非獨變項的操弄效果減弱所導致。然而，研究者採用Baron與Kenny（1986）的分析方法，結果並未發現「考試焦慮生理測量（第二次脈搏測量）」為獨變項「不同認同團體激發」與「工作記憶容量測驗」的中介變項。亦未發現「工作記憶容量測驗」為「考試焦慮生理測

量」與「數學能力測驗」的中介變項。顯示，儘管受試者會付出一些心力來調節考試焦慮，但是刻板印象直接對工作記憶容量造成影響的部分還是較大。

此外，本研究發現受試者在自陳式考試焦慮量表的表現上沒有差異，但生理測量卻顯現出組別差異。研究者提出四種可能的解釋，第一種是，因為刻板印象效果為一自動化的歷程，會不知不覺地影響刻板印象的當事人，因此儘管受試者沒有知覺到刻板印象的效果，她們亦可能會受到影響，換句話說，她們的考試焦慮表現在生理反應上，然而她們自己卻不見得知覺到，所以亦沒有顯現在自陳式的考試焦慮測量上。第二種，則可能是當受試者處於被刻板印象威脅或被評價的環境中，她們會啟動心理防衛機制來壓抑她們對於考試焦慮的知覺，因此在自陳式考試焦慮量表上無法顯現出組間差異。第三種，則是因為考試焦慮的自陳式問卷是在實驗的最後才給予受試者，而本研究因為實驗程序較長，此時獨變項操弄可能已經失去效果，因此受試者在此並未展現出刻板印象效果。第四種解釋則結合上段所述，受試者儘管會因為獨變項操弄會顯現出不同的考試焦慮，然而，受試者亦會自動調節考試焦慮以利於之後的工作，因此最後測量的自陳式考試焦慮量表並未顯現出效果。

第三，本研究除了探討刻板印象

效果對於作業表現的影響之外，並探討刻板印象效果對於「數學能力測驗難度選擇」、「數學能力測驗難度評估」、「對自己能力的評估」三方面的影響。在「數學能力測驗難度選擇」方面，本研究發現，對於被激發正向刻板印象的受試者來說，他們比起其他人的確會傾向選擇較困難的數學能力測驗。顯示，刻板印象當事人並不只是被動地被刻板印象影響作業表現，在刻板印象的效果下，這些當事人亦會主動地做出符合刻板印象的選擇。因此，在正向刻板印象影響下，理學院認同組受試者會更願意挑戰自己，倘若他們能一直受到這種刻板印象影響而持續挑戰自己，他們的能力也許會因此增強。然而，刻板印象威脅效果並沒有在「數學能力測驗難度選擇」上展現，研究者認為這可能是由於本研究受試者群本身的特性就較不會受到刻板印象威脅所導致。

而在「數學能力測驗難度評估」方面沒有顯現出刻板印象效果，研究者原先假設，對於那些受到刻板印象威脅效果影響而導致數學能力測驗表現較差的受試者來說，他們會認為測驗較困難，相反地，對於受到刻板印象提升效果而表現的較好的受試者則會認為此測驗較為簡單，然而，研究結果發現，對於三組受試者來說，他們在數學能力測驗上的表現並沒有差異，因此，可能導致三組受試者在「數學能力測驗難度評

估」上沒有顯現組間差異。而在「對自己能力的評估」方面亦沒有顯現出刻板印象效果，研究者認為這可能是因為，此依變項是在實驗的最後進行測量，而此時獨變項操弄效果已減弱，因此沒有顯現出組別效果。

最後，本研究原本預測，當受試者受到刻板印象效果影響而導致表現變好或變差，會使得他們對於「刻板印象的相信程度」加深。然而，研究結果並不支持假設。研究者推估，一方面是由於在實驗中他們無法知道其實際表現結果是否支持了刻板印象，亦即，女性認同組的受試者不知道他們的表現是否比其他男性差，理學院認同組的受試者則不知道自己的表現是否比其他學院的學生好，另一方面則可能是因為刻板印象的相信程度是一長久的歷程導致，因此無法被只有一次的實驗歷程所改變。

二、研究限制與對未來研究的建議

本研究的主要缺失在於儘管找到同時具備正向及負向刻板印象的受試者群（理學院女生），但是由於她們本身就是數學能力較好且較不易受到「女性數學比較差」的刻板印象影響的人，因此而導致刻板印象威脅效果並沒有顯現。所以，未來研究者應了解，即使受試者皆隸屬於相同刻板印象下的目標團體，但其中身處於不同的次團體，刻板

印象的效果亦會不同，換句話說，也就是即使受試者皆為女性，但「理學院的女性」比起「一般女性」（例如：政治大學的女學生）或者文學院的女學生來說，可能較不會被刻板印象「女性的數學能力比男性差」所影響。因此，未來研究者若想進一步探討相同刻板印象的激發，對於不同目標對象的影響差異，則可使用不同次團體（例如：文學院與理學院女學生）做為受試者，來比較刻板印象效果是否有所差異。

此外，由於本研究的受試者大多為政治大學心理系的女生，因此研究的外在效度較低，未來研究者應盡可能地擴大目標群體以提高研究的外在效度。

再者，由於本實驗的獨變項操弄和Shih等人（1999）相同，採取的是隱性激發，但相較於Shih等人在獨變項操弄完之後立即給予數學測驗，本研究則先讓受試者從事工作記憶容量測驗後再進行數學能力測驗，而實驗操弄的效果會隨著時間的進行而減弱，所以導致實驗後半的「數學能力測驗」、「數學能力測驗難度評估」，以及「對自己能力的評估」等依變項皆沒有顯現刻板印象效果。因此，研究者建議未來想要做這類實驗的其他研究者，應盡量縮短實驗時間，使獨變項操弄與之後的測驗之間最好在半小時之內，或者在中途重複再給予一次獨變項操弄。

第三，由於本實驗的受試者皆為

政治大學理學院的女學生，因此，本研究的外在效度較為不足。再者雖然參與實驗的人數有49人，皆達到各組應有的基本人數16人。然而，在進行中介分析時，由於有因為缺失數值或接受過GRE相關測驗而被排除的樣本，因此，實際分析的人數只有40人，導致統計考驗力不足，可能因此造成中介效果並未顯現。

另外，由於「數學能力測驗」之平均答對率對於不同目標群體皆會有所不同，而本研究忽略考量此因素，可能導致「數學能力測驗」過於簡單而未能在此測驗上顯現刻板印象效果。研究者建議未來研究者應謹慎測量並選取適合的難度，以控制此因素所造成的影響。

最後，本研究發現，在刻板印象效果中，受試者的生理焦慮與自陳式焦慮的結果相當不同，而不同階段的受試者其生理焦慮亦有所不同。因此，研究者除了建議其他研究者就生理與自陳式焦慮的歧異做進一步討論之外，亦建議其他研究者可以就不同時間階段的生理焦慮差異做進一步的探討。

參考文獻

- 林莉萱（2003）。探討術前音樂對體外震波碎石術病患的影響。臺北醫學大學醫學研究所碩士論文，未出版，臺北市。
- 邱蜀娟（2005）。性別刻板印象威脅與自我設限。國立政治大學心理學研究所碩

- 士論文，未出版，臺北市。
- 陳皎眉、王叢桂、孫蒨如（2003）。社會心理學（Rev. ed.）。臺北市：空大。
- 陳皎眉、孫曼暉（2006）。從刻板印象威脅談學業表現上的性別差異。教育研究月刊，174，19-30。
- 鄭昭明（2004）。認知心理學：理論與實踐。臺北市：桂冠。
- 鍾思嘉、龍長風（1984）。修訂情境與特質焦慮量表之研究。中國測驗學會測驗年刊，21（1），27-36。
- 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*, 1173-1182.
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553-573.
- Cheryan, S., & Bodenhausen, G. V. (2000). When positive stereotypes threaten intellectual performance. *Psychology Science, 11*(5), 399-402.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span task: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786.
- Croizet, J. C., Despres, G., Gauzins, M. E., Huguet, P., Leyens, J. P., & Meot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin, 30*(6), 721-731.
- Marx, D. M., & Stapel, D. A. (2006). Distinguishing stereotype threat from priming effects: On the role of the social self and threat-based concerns. *Journal of Personality and Social Psychology, 91*(2), 243-254.
- Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language, 39*, 418-436.
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*(3), 440-452.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*, 80-83.
- Smith, J. L. (2004). Understanding the process of stereotype threat: A review of mediational variables and new performance goal directions. *Educational Psychology Review, 16*(3), 177-206.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-28.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory (Self-evaluation questionnaire)*. Palo Alto, CA: Consulting Psychologists.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811.

Suen, M. (2006). *Stereotype-moderated mathematical performance in multiple category contexts*. Unpublished doctoral dissertation, University of Birmingham, Birmingham, UK.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28(2), 127-154.