

不同學習階段兩性數學成就差異之比較研究

宋曜廷

國立台灣師範大學
教育心理與輔導學系
sungtc@ntnu.edu.tw

陳慧娟

國立台灣師範大學
教育心理與輔導學系
t05004@ntnu.edu.tw

黃璫瑩

國立台灣師範大學
心理與教育測驗研究發展中心
lying@bctest.ntnu.edu.tw

摘 要

本研究的目的旨在探討性別與數學測驗難度之關係。研究採分層隨機抽樣方法，按都市化程度高、中、低比率抽取 94 年度兩次國中基測之重複考生各 10%，共計 12501 位考生為樣本。接著以縱貫研究法追蹤這群樣本三年後同時參加 97 年度大學學測與指考之考生，刪除不完整資料後，計自然組高中畢業生 3081 位；而社會組高中學生則有 3217 位。另外，本研究嘗試將國中基本學力測驗、大學學測及指考數學科測驗依通過率分為低、中、高三個難度的分測驗，分析兩性在不同學習階段、不同測驗難度下數學學習表現之差異情形。結果顯示：整體而言，以學習階段來看，無論國中或高中階段，男生在數學測驗平均數普遍高於女生；若依測驗難度進行分析，則發現在低難度測驗中，女生表現較佳，然而在高中階段由於測驗難度較高，男生在數學成就的表現又較女生優異。本研究結果大致符合過去的研究發現，最後根據文獻與研究發現提出建議以提升兩性在數學成就之表現。

關鍵字：數學學習、性別差異、測驗難度

一、研究動機與目的

由於數學成就對於學生後續學習、生涯選擇和專業成就有重要影響，所以在眾多學科中，數學測驗的性別差異始終是成就測驗關心且爭論的議題，(Fan, Chen, & Matsumoto, 1997)。特別是「性別平等」觀念近來備受重視，縮小男女生學習成就差距一直受到國際間教育界關注，許多國際性的學生學習成就調查研究，如 TIMSS (Trends in Mathematics and Science Study，簡稱 TIMSS) 和 PISA (Programme for International Student Assessment，簡稱 PISA) 都專章或專節報告男女生差異分析結果，反觀國內性別學習差異研究則較未受到相對應的重視。此外，測驗結果通常攸關考生的權益，因此測驗對不同背景應試群體的公平性不僅是測驗發展機構，也是一般會大眾關切的焦點。

有鑑於此，本研究企圖透過客觀嚴謹且經過標準化程序發展的測驗蒐集本土資料，由於國中基測、大學學測與指考，其應試之考生幾乎接近中學應屆畢業生的母群體數，這些測驗結果是了解台灣國、高中畢業生學習成就最豐富、也是最重要的資料庫。因此本研究期盼透過分析基測、學測與指考的資料深入了解國、高中畢業生數學科成就之性

別差異情形，提供教育實務與後續研究之參酌。

二、文獻探討

長久以來，性別間學習成就的差異是教育人員與學科專家關注的議題，儘管目前沒有特殊的理論可推測造成數學科性別差異的因素，然而相關的研究仍持續受到重視（Ryan & Fan, 1996）。過去後設分析研究的焦點多放在平均數的檢定（Gorard, Rees, & Salisbury, 2001），或以試題題型、試題內容來探討其差異。茲將有關研究說明如下。

Hyde、Fennema與Lamon（1990）採後設分析研究法整合了1963至1988年100篇性別與數學學習的相關研究並指出，整體而言，男生的數學表現優於女生。由發展階段來看在小學和中學階段，女生數學成就略優於男生，但沒有明顯的性別差異。在高中和大學階段，男生的表現優於女生。若就試題內容來看，女生在計算題表現較佳，男生在問題解決的表現較好，在數學概念理解上男女生則無明顯不同。Hyde 等人認為男生在高中和大學階段表現優於女生的原因，係因為高中和大學階段數學課程較強調問題解決所致。Willingham、Cole、Lewis 和Leung（1997）廣泛收集美國境內14種資料來源，包括標準化測驗、大型評量方案和數學資料庫，分析後也發現隨著年級增長，女生數學學習優勢表現見漸消退，而男生則逐漸增長的現象。根據美國國家教育成就評量（National Assessment of Educational Progress，簡稱NAEP）報告顯示，由1973至2004資料庫針對9歲、13歲、17歲三個年齡階段進行橫斷分析，結果均顯示男生數學成就普遍優於女生（引自Liu & Wilson, 2009）。Friedman（1989）整合77篇以中小學生為對象的研究進行分析，整體來說男女生的數學表現差距非常小，此外，Friedman 也指出近年來數學性別差距有縮小的趨勢。TIMSS最近兩次的調查（2003，2007）不管是國小四年級或國中八年級學生，其國際平均之數學成就對於性別並無顯著差異，TIMSS（2007）的資料雖顯示國中八年級女學生在整體數學成就表現明顯優於同年的男學生，但台灣、韓國、香港及日本依然維持與TIMSS（2003）一致，意即此亞洲四國之八年級男女學生在整體數學成就表現上並無明顯差異。而新加坡雖亦為亞洲國家，但其女學生在整體數學成就表現卻明顯優於男學生，原因值得進一步探討。不少學者指出數學科性別差異有所不同可能是各國社會、文化差異造成（Feingold, 1994；Hanna, 1989；Ethington, 1990）。國內簡茂發（1995）曾主持「國民教育階段學生基本學習成就評量研究」計畫，發現國小三年級女生在所有的分析項目的表現都顯著高於男生；國小五年級男生機率表現優於女生，女生數學解題表現優於男生，其餘項目無顯著性別差異出現；國中八年級男生在幾何、度量、代數和數學解題表現優於女生，其餘項目無顯著性別差異出現。總括而論，隨著年級增長，也出現女生數學優勢漸減，而男生漸長的現象。綜合上述，數學成就是否有性別差異存在？差異何時開始？原因為何？迄今仍難以獲得明確和一致的結論。

另外，Becker（1990）、Bielinski 與 Davison（1998, 2001）針對數學科試題難度與性別間的差異做探討，指出試題難度與性別間有交互作用存在。換言之，簡單的試題對女生來講更為容易，而複雜的試題對女生而言更加困難，並認為男女生在試題表現的差異與能力有關。Kyriakides 與 Antonious（2009）甚至認為試題難度是解釋兩性數學成就差

異的重要因素，並強調試題難度可取代試題內容、試題題型來解釋數學科性別間成就差異。

有鑑於過去研究通常採用橫斷法，即僅以不同年級學生樣本進行分析，如此可能受到其他無關變項（如課程教材、樣本特性、家庭結構等）的干擾，無法正確了解不同學習階段兩性數學成就差異的真正原因。本研究採用縱貫法蒐集同一位受試者國中基測數學成績及大學學測、指考數學成績，比較不同性別受試者在不同學習階段數學成就的差異，進而了解數學學習差異。由於過去研究者傾向於比較不同試題難度下性別表現之差異情形，然試題偏差亦會影響測驗結果，因此本研究將多個難度相近試題組合為三種測驗難度，在控制測驗難度下比較男女生的差異情形，嘗試回答性別間的差異是否因測驗難度而有所改變。

根據上述研究動機，本研究目的如下：

- （一）比較 94 年度參與國中基測男女生在基測數學科不同測驗難度（低、中、高）下之表現情形。
- （二）不同性別之高中自然組學生，在國中基測數學測驗、大學指考與學測數學測驗不同難度（低、中、高）下之表現是否有差異？
- （三）不同性別之高中社會組學生，在國中基測數學測驗、大學指考與學測數學測驗不同難度（低、中、高）下之表現是否有差異？

三、研究方法

（一）、研究對象

參與 94 年度兩次國中基測重複考生中按都市化程度高、中、低比率採分層隨機抽樣各抽取 10% 學生，共抽取 12501 位考生，並追蹤這些考生 97 年度學測與指考成績，刪除不完整資料後，同時參加學測與指考之自然組學生剩 3081 位，而同時參與學測及指考之社會組學生剩下 3217 位，表 1 為本研究抽樣學生之性別比例及人數分佈摘要表。

表 1 抽樣學生性別比例描述性統計摘要表

	94 年度國中基測	97 年度自然組	97 年度社會組
男性	6250 (50%)	2075 (67.3%)	1050 (32.6%)
女性	6251 (50%)	1006 (32.7%)	2167 (67.4%)
總數	12501	3081	3217

（二）、評量工具

本研究為追蹤不同性別學生在不同學習階段數學科表現之差異情形，並比較同一群學生數學科國中基本學力測驗及高中學科能力測驗數學科與數學科指定科目之學習成就表現，向「國民中學學生基本學力測驗工作推動委員會」及「大學入學考試中心」申請考生資料，資料操作方式說明如下：

（一）94 年度數學科兩次國中基本學力測驗（以下簡稱基測），題數皆為 33 題，題型為選擇題，合併兩次測驗後利用試題通過率參數將測驗分為易、中、難三部分，以比較男

女生在不同測驗難度的表現差異情形。

(二) 97 年度學科能力測驗數學科測驗(以下簡稱學測),題數為 20 題,依照試題通過率將測驗分為易、中、難三部分。

(三) 97 年度數學科指定科目考試(以下簡稱指考),分為甲、乙兩式。自然組考甲式;社會組考乙式。題型除了選擇題、填充題之外,還有實作題,由於實作題每題給分並不相同,且甲乙兩式皆只有 2 題,因此本研究並未將實作題納入分析,依照選擇題、填充題試題通過率將測驗分為易、中、難三部分。

由於抽樣自然組及社會組學生皆接受大學學測與指考測驗,若分別將學測及指考依難度分三個分測驗,每一式題目過少,因此本研究合併學測與指考測驗,再依通過率分為易、中、難三部分分析。

(三)、資料分析

本研究除了採用差異性檢定比較兩性在數學測驗的平均表現外,並利用迴歸模式進行分析,依變項為不同測驗難度之分測驗答對題數,自變項為性別、測驗總分,在國中基測測驗總分為兩次數學科基測成績加總;在大學學測及指考部分,由於學測計分採 1~15 級,為了讓兩次測驗比重較一致,以 100 除以 15 約等於 6.7,因此將大學學測級分乘以 6.7 之後與指考成績加總作為測驗總分。控制了測驗總分之後比較性別變項是否有顯著差異,若達顯著說明在此測驗難度下性別間達顯著差異,且此差異並非能力不同所造成。

四、結果與討論

(一)、國中基測性別差異分析結果

兩性國中基測數學科描述性統計及性別差異性檢定如下表 2,結果顯示男生平均數較高,變異也較大;且國中基測性別間之平均數達顯著差異。

表2 94基測兩次量尺總分之描述統計及性別差異性檢定

國中基測	性別	人數	平均數	標準差	t
第一次基測	男生	6250	41.555	13.350	14.241***
	女生	6251	38.251	12.581	
第二次基測	男生	6250	39.185	11.633	15.742***
	女生	6251	35.985	11.091	
兩次基測加總	男生	6250	80.740	24.039	15.556***
	女生	6251	74.236	22.693	

由於國中基測為中等難度偏易試題,若依照試題題數區分低、中、高三等分會使得困難試題的答對率過高,因此設定難易度分類方式如下:容易部分試題通過率為 65%以上,共 34 題;中等部分試題通過率為 45%至 65%之間,共 21 題;困難部分試題通過率在 45%以下,共 11 題,分別針對不同難度分測驗進行迴歸分析,結果如下表 3。

表3 性別在不同難度分測驗表現迴歸分析摘要表 (N=12501)

測驗難度	變項	非標準化係數		標準化係數
		B	Std	Beta
低難度	常數項	14.066	.106	
	94基測兩次量尺總分	.195	.001	.888***
	性別	.472	.044	.046***
中難度	常數項	.112	.066	
	94基測兩次量尺總分	.189	.001	.947***
	性別	.015	.027	.002
高難度	常數項	-2.241	.065	
	94基測兩次量尺總分	.102	.001	.847***
	性別	-.192	.027	-.034***

表3結果顯示控制學生基測總分，國中階段低難度及高難度測驗在性別間達顯著差異，其中低難度測驗女生表現較佳，而高難度測驗則是男生表現較佳。

(二)、學生在國中及高中階段性別表現分析結果

國中階段所有學生學習課程皆相同，高中卻只有高一課程是相同的，高二選組後自然組與社會組在數學科的學習內容就有所不同，Bridgeman 和 Wendler, (2005) 研究認為選課的不同會導致性別間的差異。為避免混淆變項影響推論，本研究特別針對高中選擇自然組及社會組學生分開檢驗數學科性別間差異。高中測驗採用大考中心學測數學科、指考數學科甲乙兩式，分別針對三份測驗試題通過率訂定測驗易、中、難三部分。由於學測與指考測驗難度及題數不同，分類方式有所差異，在學測部分通過率大於 50% 則為容易試題，通過率介於 30~50% 之間為中等難度試題，30% 以下為困難試題；在指考部分通過率大於 50% 為容易試題，介於 40~50% 則為中等難度試題，40% 以下則為困難試題。由於學測及指考題數較少，因此合併兩測驗進行分析，為了進行縱貫資料研究，將針對自然組及社會組學生基測及學測指考成績進行性別分析。

表4 自然組與社會組在國中基測及大學學測指考數學科測驗性別表現差異性檢定

組別	測驗	性別	人數	平均數	標準差	t
自然組	兩次基測加總	男生	2075	95.1812	15.13819	7.306***
		女生	1006	90.7227	16.23459	
	合併指考與學測	男生	2075	105.4770	41.49890	8.326***
		女生	1006	92.2907	40.64743	
社會組	兩次基測加總	男生	1050	81.7581	18.12815	2.432***
		女生	2167	80.1407	16.73248	
	合併指考與學測	男生	1050	75.5012	36.04171	1.276
		女生	2167	73.8044	33.89922	

表 5 自然組及社會組在不同測驗難度下之迴歸分析摘要表

測驗難度	變項	自然組		社會組	
		國中基測	學測指考	國中基測	學測指考
		標準化係數	標準化係數	標準化係數	標準化係數
低難度	總分	.768***	.801***	.832***	.836***
	性別	-.007	.007	.025**	-.031***
中難度	總分	.906***	.824***	.922***	.885***
	性別	.001	-.046***	-.005	-.032***
高難度	總分	.876***	.825***	.833***	.695***
	性別	.015	-.027***	.004	-.010
依變項: 低、中、高難度測驗答對題數				***p<.01	**p<.05

表4及表5分別分析性別在測驗的整體平均表現及檢定在不同難度測驗控制學生能力下，性別表現是否有顯著差異。表4結果顯示自然組學生在基測及指考學測均顯示男生平均成績顯著優於女生；社會組學生在基測結果顯示男生表現優於女生，在指考學測則未顯示性別差異。下表5將自然組與社會組學生兩次國中基測及大學學測與指考數學科難度分測驗進行迴歸分析，再控制學生能力下比較性別間是否達顯著差異，在國中基測學生數學能力以兩次基測量尺分數總分相加代表，而學測指考則以指考成績加上學測成績乘以6.7表示，使得學測及指考成績比重相當。

針對測驗難度迴歸分析結果整理至下表 6，國中全體抽樣在基測分析結果與文獻一致；自然組在國中基測未受不同測驗難度下而有所差異，但大學學測指考則顯示男生在中高難度測驗表現較佳；社會組基測則顯示低難度試題女生表現較佳，學測指考卻是男生在中低難度測驗表現較佳。自然組與社會組所呈現出來不同的組型可能與學生數學能力有關，而基測與學測指考分析結果差異可能由於基測題目較為容易，且縱貫研究樣本均為繼續升學學生，其國中數學程度比整體學生表現佳，由表 6 可知整體而言性別與測驗難度有交互作用存在。

表 6 不同難度下性別表現差異分析結果

階段性測驗	測驗 難度	國中全體抽樣 N=12501	自然組 N=3081	社會組 N=3217
94 年兩次國中基測	低難度	女生		女生佳
	中難度			
	高難度	男生佳		
97 年大學學測與指考	低難度	-----		男生佳
	中難度	-----	男生佳	男生佳
	高難度	-----	男生佳	

五、結論

McCornack 與 McLeod (1988)、Elliot 與 Strenta (1988) 研究發現若考慮選課的差異就能說明為何男生數學成績優於女生的(引自 Langenfeld, 1997)。因此本研究將選課差異考慮進來,並將自然組與社會組學生資料分開分析,台灣地區控制選組差異後,學生均接受相同的授課內容,分析結果顯示男女生表現仍因試題難度不同而有所差異,且高中階段男生表現較佳。不少研究者將男女數學科表現差異歸因為社會期許、性別刻板印象及學習策略所造成的(Benbow & Stanley, 1983)。

男生是不是天生數學就比女生好,這是一個爭論不休的話題。美國西北大學一項專門探討男女數學能力差異的研究計畫,他們以十五歲的學生為對象,採用世界經濟論壇性別平等指數,作為這項研究的評比標準。每隔三年進行一次全球調查,被調查的四十個國家當中,只有冰島是女生數學比男生強。另外十二個國家,男女數學能力幾乎是不分軒輊,這些國家相對來說,兩性平等也較為彰顯。換言之,性別愈平等的社會,這種差異愈小。

新進有關性別刻板印象的研究提出了「自我評估威脅」(self-evaluative threat)的觀點,認為當學習者處於有刻板印象的壓力情況下,他們通常會為了避免自己的表現與其所屬族群的負面刻板印象相符,而感受到自我評估的威脅,並因此導致表現不佳。至於這種性別刻板印象引發的威脅感是否只會在公眾場合發生呢?Inzlicht 與 Ben-Zeev (2003)的研究發現不然,其結果顯示不論是否會公佈考試成績,只要女性身處於可能激發其性別刻板印象的環境(如:周遭都是男性、體認到他人的期待、社會的觀感…),都會使她們在數學測驗上的表現低於其原有水準,Inzlicht 等人並進一步推測,這可能是因為性別刻板印象的威脅會促發受試者的自我表現動機以及自我評價的內在心理歷程,導致女性因承受較大壓力而抑制其成就表現。

總而言之,性別刻板印象使得女性容易低估自己的數學能力,有較高的測試焦慮,也較少朝數學專業領域進行更深入的鑽研及相關職業的發展已是司空見慣的事實。因此社會應建立更友善的環境幫助不同性別的學生發現自己的學習優勢,特別是給予女生更多的信心與支持,提供更多良師益友典範幫助她們在數學領域發揮所長。

參考文獻

Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science*, 222, 1029-1031.

Becker, B. J. (1990). Item characteristics and gender differences on the SAT-M for mathematically able youths. *American Educational Research Journal*, 27, 65-87.

Bielinski, J. & Davison, M. L. (1998). Gender differences by item difficulty interaction in multiple-choice mathematics item. *American Educational Research Journal*, 35(3), 455-476.

Bielinski, J. & Davison, M. L. (2001). A sex difference by item difficulty interaction in multiple-choice mathematics items administered to national probability samples. *Journal of Educational Measurement*, 38, 51-77.

Brideman, B., & Wendler, C. (2005). Characteristics of minority students who excel on the SAT and in classroom. Princeton, NJ: Educational testing Service.

Ethington, A. (1990). Gender differences in mathematics: An international perspective. *Journal for Research in Mathematics Education*, 21(1), 74-80.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116(3), 429-456.

Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research*, 59, 185-213.

Gorard, S., Rees, G., & Salisbury, J. (2001). Investigating the patterns of differential attainment of boys and girls at school. *British Educational Research Journal*, 27, 125 – 139.

Hanna, G. (1989). Mathematics achievement of girls and boys in grade eight: Result from twenty countries. *Educational Studied in Mathemetics*, 20, 225-232.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychological Bulletin*, 107(2), 139-155.

Kyriakides, L., & Antoniou, P., (2009). Gender differentces in mathematics achievement: an investingation of gender differentces by item difficulty interactions. *Educational Research and Evaluation*, 15(3). 223-242.

Langenfeld, T. E. (1997). Test fairness: Internal and external investigations of gender bias in mathematics testing. *Educational Measurement: Issues and Practice*, 16(1), 20-26.

Ryan, K. E., & Fan, M. (1996). Examining gender DIF on a multiple-choice test of mathematics: A confirmatory approach. *Educational Measurement : Issues and Practice*, 15(4), 15-20.

Willingham, W. W., Cole, N. S., Lewis, C., & Leung, S.W. (1997). Test performance. In W.W. Willingham & N. S. Cole (Eds.), *Gender and fair assessment* (pp.55-126).Hillsdale, NJ: Lawrence Erlbaum Associates.