

教育科學研究期刊 第六十七卷第一期

2022年，67（1），227-254

[https://doi.org/10.6209/JORIES.202203_67\(1\).0008](https://doi.org/10.6209/JORIES.202203_67(1).0008)



國中教育會考數學科的回沖效應初探

張銘秋

國立臺灣師範大學
心理與教育測驗研究發展中心

黃瓏瑩

國立臺灣師範大學
心理與教育測驗研究發展中心

陳佳蓉

國立臺灣師範大學
心理與教育測驗研究發展中心

陳柏熹

國立臺灣師範大學
教育心理與輔導學系
心理與教育測驗研究發展中心
暨學習科學跨國頂尖研究中心

曾芬蘭

國立臺灣師範大學
心理與教育測驗研究發展中心
暨學習科學跨國頂尖研究中心

摘要

本研究探討臺灣實施國中教育會考對數學科教學現場的回沖效應，透過教師問卷，比較國中數學教師在國中教育會考實施前後的教學差異，並調查其於衍生之相關議題的感受與態度。問卷調查結果顯示：一、近七成教師認同會花費較多的時間教授基測／教育會考常考的內容；二、教師在兩時期使用的教材仍以教科書與習作為主，約有六成教師經常使用測驗卷或考古題；教育會考時期非選擇題型式測驗卷或考古題的使用率較基測時期高（約高7.5%）；三、教學方法均以教師單向講述或師生雙向問答為主；四、教師在數學的閱讀理解能力重視程度比基測時期增加最多，其次為數學表達、溝通能力；五、學校評量各題型的比重未有明顯變化，但教育會考時期非選擇題比重有微幅提高；六、半數以上的教師不認同成績由百分等級制改為三等級，可降低學生學習壓力，且近七成教師認為易造成中等程度學生學習動機低落。由上述研究結果可知，國中教育會考數學科確實對國中數學教學現場同時產生了正向與負向的影響。

關鍵詞：回沖效應、高風險測驗、國中教育會考、數學非選擇題

壹、研究動機與目的

配合我國十二年國民基本教育（以下簡稱十二年國教）中國中畢業生得免試升學的理念，自2014年起取消國民中學學生基本學力測驗（以下簡稱基測）。為了維持教育品質，俾利國中教育順利銜接高中職或五專教育，由國立臺灣師範大學心理與教育測驗研究發展中心（以下簡稱師大心測中心）承辦國中教育會考（以下簡稱教育會考）。教育會考的測驗結果除提供教育部、各縣市教育局（處）、國高中及五專進行學力監控外，亦作為十二年國教實施初期免試超額比序的參考依據。基測採常模參照取向，考試結果僅能作為同一年度考生能力的比較之用，無法達到學力監控的目的，因此，教育會考改採標準參照取向，以A（精熟）、B（基礎）、C（待加強）三等級取代量尺分數，藉由減少各科成績等級適度減低考試壓力，並達到學力監控與提供具體學力訊息的目標（師大心測中心，2014）。

Hughes（2003）將回沖效應定義為測驗對教學與學習的影響，而與兩者最密不可分的幾個面向就是課程、教材、教學與評量。臺灣的國中小課程係以國家統一制定的課綱為課程發展依據，對大多數的學習領域或學科來說，教科書通常是教師教學及學生學習的主要材料與工具，教師的教學準備與授課模式往往深受所選用的教科書影響。在臺灣，各領域課綱規範了課程目標與學習重點內容，並透過教育部審定通過的教科書將其具體化。不同出版社的教科書編寫方式可能不同，但教育會考始終堅持考綱不考本的精神，因此學生使用任一版本教科書，皆足以作答會考試題。

以往基測數學科只有選擇題型，但無論是九年一貫課程綱要或十二年國教新課綱，數學領域除重視數學概念、演算、抽象和推理能力的培養，也開始重視數學溝通能力（教育部，2018）。其中，部分能力須藉由解題過程中各步驟合理性的探討，方能養成。若僅侷限於選擇題題型，就不容易有效評估數學溝通能力中與「表達」相關的重要學習能力指標。過去基測數學科受僅有選擇題的影響，教學現場過分強調選擇題的練習，甚至關注選擇題的解題技巧和方法，忽略演算或推論等相關能力的學習（師大心測中心，2018a）。教育會考數學科增加了非選擇題型，用以評估學生的數學溝通、表達等能力，以便更全面地瞭解學生的數學能力，積極引導教師的教學。教育會考已實施多年，但是否真的如其所言減輕了考試的壓力，以及教育會考數學科增加非選擇題型對國中數學教學與評量的影響（即回沖效應），值得進一步研究。

其次，教育會考數學科自2015年正式採計數學非選擇題成績以來，近4年（2015~2018年）數學科精熟等級的學生從2015年的15.20%增加至2018年的22.31%，增幅達7.11%；待加強等級的比例則是從2015年的33.22%，降至2018年的28.72%，降幅為4.50%。精熟等級的學生人數持續上升，且待加強等級的學生人數比例持續下降，代表學生整體數學能力有所提升，拔尖成效明顯優於扶弱，但全國仍有近三成的考生未達基礎能力等級的標準。由於教育會考成績等

級的資訊對於教育資源分配影響甚巨，各教育局（處）積極推動各項「減C計畫」以減少待加強等級學生人數比例，學校與教師對此採取的行動屬於回沖效應之一。

為此，本研究採問卷調查方式，藉以瞭解在基測改為教育會考、數學科加考非選擇題，以及在減C計畫之下，國中階段數學課程、教材、教學與評量的改變。此外，本研究挑選連續3年（2015～2017年）在教育會考數學科待加強人數百分比減少5%以上，且教育會考到考人數在150人以上學校的數學教師作為深度訪談的對象，目的是基於問卷調查的結果，更深入地蒐集數學教師的看法，據以提供教師與相關研究人員參考。目前國內探討高風險測驗對教學與學習回沖效應的研究屈指可數，本研究之研究結果可供未來教育政策擬定以及教學方法之規劃與推動的參考。

貳、文獻回顧

一、回沖效應

回沖效應（washback effect）亦譯為倒流效應或回流效應，最早出現於語言教學與測驗的文獻之中，某些研究者也會寫做“backwash”以描述考試或測驗帶來的影響。Hughes（2003）將回沖效應定義為測驗對教學與學習的影響，這種影響可能益於教學與學習，稱之為正向回沖效應（positive washback effect），但若妨礙了教學與學習，則為負向回沖效應（negative washback effect）。Alderson與Wall（1993）認為回沖效應指的是因為考試而迫使教師與學生做出他們本不需做的行為。Messick（1996）則是認為回沖效應指的是考試影響了教師與學習者，致使其做某些可以幫助或抑制學習行動的程度。換言之，測試的形式與內容直接影響教師對課程的規劃、教學和對學生的校內評估設計，即學生要考什麼，教師便教什麼，此一現象就稱為回沖效應（廖佩莉，2013）。

Pearson（1988）指出正向回沖效應是考試能帶動教學與課程的正面影響，即教師與學生對考試有良好與積極的看法，並達到學習與評量的預期目標。反之，負向回沖效應指的是考試帶來教與學的負面影響，如教學內容的窄化、學生大量練習考古題，或者學生與教師因考試而感到憂慮（廖佩莉，2013；Cheng & Curtis, 2004）。一般而言，愈高風險的測驗（high-stakes test）所產生的回沖效應愈大，反之，回沖效應愈小（Alderson & Wall, 1993; Shohamy et al., 1996）。

回沖效應相關研究所討論的情境都集中在課堂之上，但面向不同，大約可以歸納為課程、教材、教學法、評量、感受與態度。以下針對各面向分述之：

（一）課程

不同研究顯示與課程相關的回沖效應相互矛盾。多數的研究顯示考試對課程的內容有明顯的影響，主要是窄化了課程，教學只集中在考試的範圍內，甚至只著重在最常考的部分（廖

佩莉，2013；Alderson & Wall, 1993; Cheng, 1997），即只是為了測驗而教學。然而，Shohamy 等人（1996）卻表示考試對教學內容只有輕微的影響；Watanabe（1996）的研究也指出日本大學入學考試包含了閱讀與聽力，但並非所有教師都會教授這兩部分的內容。Rodríguez-Muñiz 等人（2016）針對西班牙大學入學考試數學科的回沖效應進行研究，他們分析了入學考試試題對課綱內容的覆蓋性，同時也對數學教師進行問卷調查。他們發現有大量的課綱內容未出現在入學考試中或有代表性不足的現象，同時，部分內容則是一直出現在入學考試中，代表入學考試窄化了課程，此為明顯的偏誤；而問卷結果顯示，教師雖然會盡量完成所有的課程內容，但當教學時間不足時，約有七成的教師會放棄比較少出現在考試範圍內的內容，然而也有近七成的教師表示只要考試設計得當，即使跳過某些內容不教，也不會對學生的學習產生太大的影響。

（二）教材

Cheng（1997）在香港的研究指出，「考試大綱影響香港中學的教學……，幾乎每所學校都必須為了學生而改變教科書內容」。Lam（1994）發現當考試修正時，教材的出版商就會引入一些創新的內容，也將教師變成教科書奴隸或考試奴隸，前者大量依賴教科書，後者則太過於重視考古題。Andrews 等人（2002）同樣提到教科書或參考書對香港的教育體系的重要性，教師大約花三分之二的時間在講解與考試相關的教科書或參考書。Cheng（1997）認為其中一個原因可能是因為香港的教科書不僅提供教學的資訊與活動，同時提出了建議的教學方法與時間分配。在考試教材中也提到時間因素，愈接近考試時間，考古題以及與考試相關的商業出版品的使用時間愈長（Alderson & Wall, 1993）。

（三）教學法

Alderson與Wall（1993）在斯里蘭卡的研究顯示，教師的教學方法不受考試內容的影響。Cheng（1997）指出公告了考試的修正內容後，雖然在教學內容上有改變，但這些改變並沒有改變教學方法。廖佩莉（2013）也指出因應考試的轉變，香港的課程內容與教材有明顯的變化，但課程規劃和教學策略方面並沒有很大的改變，仍是沿用慣用的教學法。Andrews 等人（2002）則是指出，考試的修正導致教師會特別教授某些考試技巧。Au（2011）也指出為因應高風險測驗，教學者已經不再教授學生不同的解法，而是教他們如何解讀題幹答對試題，並大量地練習考古題。侯傑泰與何穎欣（2008）的研究指出，學校運用大量時間去操練學生，學生的成績因此不斷上升，於是學校花更多時間去操練這個狹隘的考試內容。Rodríguez-Muñiz 等人（2016）的研究指出，由於每年入學考試的考題都是重複性的結構，讓教師與學生可以組織與規劃其教學及學習，此為正向的影響，但也因為是重複性的考題架構，師生能輕易地預測考試內容，因而將教學與學習的過程限制在練習解題，也就是為了考試而學習。

（四）評量

廖佩莉（2013）指出有近七成的教師在設計平常練習時，會參考以往全港性系統評估的試題或坊間的練習題；曾芬蘭等人（2019）的研究指出教育會考加考英語聽力後，學校定期評量中含英語聽力的比例較基測時期高約近10%。

（五）感受與態度

此部分分為學生與教師兩方面探討。面對測驗，學生感到焦慮、壓力、擔憂與不知所措（Mulvenon et al., 2005）。研究結果顯示自從有了《有教無類法案》（No Child Left Behind Act, NCLB）的高風險測驗後，學生的測驗焦慮增加了，測驗焦慮不只讓學生的測驗表現變得更糟（Ritt, 2016），而且也對學生的健康產生心理影響（Wren & Benson, 2004）。廖佩莉（2013）提到測驗對教師產生負向回沖效應，主要是增加了許多壓力，包括工作量增加、擔憂學生考試成績不佳，進而影響自身的評價、擔憂學生考試成績作為他用，進而衍生出對考試的負面觀感。

師大心測中心每年均針對當年度國中教育會考資料進行分析，提供教育部國民及學前教育署（以下簡稱國教署）全國性的分析結果以及各種學習扶助計畫之下學生、班級甚至學校表現的差異分析，亦提供國中、高中職、縣市政府教育局版本之分析結果。國教署與多數縣市政府教育局（處）會參考師大心測中心所提供的會考分析結果，作為全國性或區域性教育政策的擬定、教育資源分配、各項學習扶助計畫成效評估的參考（師大心測中心，2016，2017，2018b）。

二、基測改爲會考可能產生的可能效應

本節從過去高風險的基測、教育會考本身與相關考試及入學政策，說明升學考試與教育改革對教學現場所產生之直接或間接的影響。教育會考與基測之考試目的與用途並不相同，因此考試功能、計分方式、考試題型、辦理時間及結果呈現皆不相同，有關教育會考與國中基測之比較，可參見國中教育會考網站（師大心測中心，2014）。以下分別就測驗題型、標準參照計分，以及各縣市、國中階段學校之減C計畫等面向探討。

（一）測驗題型

基測數學科只考選擇題型，導致部分教學現場僅強調練習選擇題型，甚至只著重在教授解題技巧與訣竅，弱化演算或推論等相關能力的學習（師大心測中心，2018a）。許多數學教育專家都將教學的窄化歸因於考試題型單一化，教師除了課堂講授的內容外，平時的課室評量或是定期段考的題型也配合基測只採用選擇題。另一方面，教學與評量方法也會直接影響學生的學習，包括練習的題型、投入的時間、甚至學習興趣和動力。作為十二年國教教育改革後監測學術能力的重要機制，考量選擇題不能完整評估學生的表達能力，教育會考數學科增

加了非選擇題，以提供積極正向的測量導向教學（measurement-directed instruction），其目的是希望教師根據教育會考題型，調整教學和評量的方法。

教育會考因諸多考量（如試務及入學時程），目前僅有數學科包含非選擇題；不同於教育會考，針對學校平時課室評量，教師可利用學生在校期間學科的重要學習內容，搭配教學活動，透過多次、多元形式的評量，對學生的認知、技能、情意態度等各類能力，進行診斷性、形成性及總結性等各類低風險、甚至無風險的評量活動。如此一來，

結合多元評量和多次評量的成果，能夠得知各種面向的學生表現，並且能將評量的結果回饋教學以改進教學的品質，而非透過幾次總結性評量就決定了學生的學習成就。（宋曜廷，2012）

然而，即使教育會考已經實施多年，目前教師是否能在平時教學時選擇適當及多元的評量形式，還是值得進一步調查。

（二）標準參照計分

基測為常模參照測驗，除寫作測驗之外，每科均以1~80分的量尺分數（scale scores）標示考生測驗結果，各科量尺分數加總後轉換為總分PR值，作為申請入學、甄選入學或登記分發入學的依據。基測是學生之間相互的比較，再加上等級較多，容易產生競爭壓力。教育會考為標準參照測驗，以精熟、基礎、待加強等三等級取代量尺分數。標準參照測驗提供的回饋訊息不再僅僅是一個數字，而是呈現考生的能力和表現特點，使考生能夠瞭解自己的能力或學習成就表現，使每個國中畢業生、教師、學校、家長和教育主管機關能夠瞭解學生的學習品質，為下一階段的學習（高中、高職或五專）做準備。此外，為解決升學競爭較激烈區域免試超額時的抽籤問題，各科在維持三等級計分標準下，將精熟等級（A）前25%標示為A++，前26%~50%標示為A+；基礎等級（B）前25%標示為B++，前26%~50%標示為B+。等級加標示的作法，不僅保留教育會考標準參照測驗精神與功能，同時也能有效降低升學壓力。然而，欲落實十二年國教適性入學的精神，教育主管機關可視全國各區實施免試入學的成效，在適當時機減少標示數量（師大心測中心，2014）。

（三）各縣市、國中階段學校之減C計畫

自2014年起，教育會考取代了國中基測，成為十二年國教後監測學術能力的重要機制。根據《國民中小學成績評量準則》（教育部，2017）第14條第8款的規定，每年5月的考試結束後，學生、教師、學校、家長和教育主管機關都會利用教育會考的成績瞭解學生的學習品質。由於教育會考的結果對於教育資源分配有很大的影響，特別是針對「待加強」部分，各教育局（處）積極推動各種「減C計畫」和學習扶助計畫等方案，這些教育積極作為都是以教育會

考成績為依據。總的來說，全國國中各科教師必須瞭解教育會考命題與成績計算方式，以利輔導學生準備教育會考與回覆家長的諮詢。此外，由於教育部、教育局（處）和學校都非常重視教育會考的考試結果，教師必須回應學校當局的期望，採用不同的策略協助學生準備考試，這可能給教師帶來壓力。

以桃園市政府教育局為例，其要求轄內所有國中學校在開學期間進行會考試題分析及成績報告，擬定該市的減C增A目標，並作為學校及國教輔導團針對學力提升策略改進規劃參考，並直接將待加強級距為後25%學校列入輔導團重點輔導對象，並於108學年度起實施國中新生學力檢測，提供學生學力表現分析，增進到校輔導實益（詳情請見桃園市國民教育輔導團網頁：<http://ceag.tyc.edu.tw/ceag/>）。部分國、高中職也會透過會考分析結果來評估學生表現、調整課程安排、教學方法與診斷學生迷思概念。上述所提均為教育會考所產生的回沖效應。

Messick（1996）指出測驗實施一段時間後才會出現回沖效應。教育會考已實施多年，應適時檢視其對國中數學教學現場的影響。因此，本研究旨在調查與比較教育會考實施前後國中數學科在課程、教材、教學法、評量的異同，探討基測改為教育會考以及數學科加考非選擇題對國中數學教學現場的影響。

參、研究方法

本研究以問卷調查方式蒐集基測與教育會考時期，有關國中數學教學方式異同的相關資訊，並邀請九位第一線數學教師接受深度訪談，基於問卷調查的結果，更深入地蒐集數學教師的看法。此九位數學教師均任教於連續3年（2015～2017年）在教育會考數學科待加強人數百分比逐年減少5%以上，且每年到考教育會考數學科人數在150人以上的學校。九位數學教師中三位是領域召集人、三位教務主任、三位一般教師（一位是師大心測中心數學非選擇題的核心委員，兩位是評閱委員）。本研究引述訪談教師意見時，將以T1～T9為九位教師的代號（T9教師任教於偏遠學校）。

一、研究對象

問卷調查採用的是線上問卷的調查方式，分為一般數學教師問卷與數學科回沖效應問卷。首先發文至全國國中，由各校推派三位數學教師上網填答問卷。一般數學教師問卷填答完畢後，請教學年資5年以上者續填數學科回沖效應問卷，以確保填答數學科回沖效應問卷的教師同時具備基測與教育會考的教學經驗。本次發文947所國中，其中401所國中數學教師上網填寫問卷，以學校為單位而言，填寫率為42.34%。表1為教師問卷填答者基本資料統計表，一般數學教師問卷共計1,127人填寫，數學科回沖效應問卷則有893人填寫。男性數學教師填答比例較略高於女性教師、教學年資以11～15年占多數。

表1

教師問卷填答者基本資料統計

基本資料		一般數學教師問卷	數學科回沖效應問卷
性別	男	54.75	52.72
	女	45.25	47.28
教學年資	未滿5年	9.58	—
	5~10年	17.66	20.24
	11~15年	30.35	34.59
	16~20年	19.08	20.69
	21年以上	23.34	24.47

單位：%

由於教師問卷採線上自由填答，本研究透過任教學校區域與學校類型來驗證其代表性，教師問卷填答者任教學校之資料統計見表2。依據政府資料開放平台（<https://data.gov.tw/dataset/6088>）所取得之2017年國民中學名錄，北區學校占全體學校37.54%、中區學校占28.13%、南區學校占28.98%、東區學校占5.35%。問卷填答教師任教的學校區域，無論是一般數學教師問卷或數學科回沖效應問卷，教師任教學校的區域比例與全國學校差異不大，具有區域的代表性。在學校類型部分，教育部原就訂有偏遠與特偏國中、國小學校清單，本研究至教育部統計處偏遠地區中小地理資訊查詢系統（<http://stats.moe.gov.tw/remotegis/>）取得2016年學校名錄後，將偏遠與特偏兩類合併為一類，經統計非偏遠學校占全體國中74.55%，偏遠學校則占25.45%。對照本研究所蒐集之資料，偏遠學校教師的填答比例少於全國比例。

表2

教師問卷填答者任教學校資料統計

學校基本資料		一般數學教師問卷	數學回沖效應問卷	全國學校分配
學校區域	北區	38.27	38.27	37.54
	中區	26.50	27.38	28.13
	南區	29.65	30.19	28.98
	東區	5.57	4.15	5.35
學校類型	偏遠學校	19.48	17.91	25.45
	非偏遠學校	80.52	82.09	74.55

單位：%

二、研究工具

本研究共使用兩種教師問卷：數學科回沖效應問卷主要是根據文獻回顧中所提到的課程、教材、教學法、評量等面向，調查教師在基測時期與教育會考時期教學的差異；一般數學教師問卷則是針對教育會考所衍生的相關議題，如針對「減C計畫」採取的行動、感受以及教育會考對學生學習的影響；深度訪談則是基於兩種教師問卷調查的結果，採半結構式的訪談，期能更深入地蒐集數學教師的看法。以下分別說明兩種問卷的內容，以及深度訪談。

（一）數學回沖效應問卷調查

本次問卷調查所使用之研究工具為「國中基本學力測驗與教育會考實施時期之數學教學意見調查」，共計八大題35小題，主要調查的是「基測」和「教育會考」兩個時期，數學科教師在課程、教材、教學法、評量等四個面向上的差異；也設定開放性問題讓教師表達其他意見，以提供更豐富的資訊。數學回沖效應問卷的題目示例如表3所示。

表3

數學科回沖效應問卷題目示例

教學時對下列技能的重視程度	基測時期 (僅考選擇題)				教育會考時期 (考選擇題與非選擇題)			
	不重視	有點重視	重視	非常重視	不重視	有點重視	重視	非常重視
數學的閱讀理解能力	○	○	○	○	○	○	○	○
解題策略規劃	○	○	○	○	○	○	○	○
邏輯推理的合理性	○	○	○	○	○	○	○	○
解題過程的完整性	○	○	○	○	○	○	○	○
計算能力(如速度與正確度)	○	○	○	○	○	○	○	○
數學表達、溝通能力	○	○	○	○	○	○	○	○
解題技巧與訣竅	○	○	○	○	○	○	○	○

（二）一般數學教師問卷調查

本問卷蒐集教師對教育會考衍生議題所採取的行動與感受，以及各校「國中教育會考學力品質監控回饋報告資料」之運用。行動與感受部分主要包含兩大面向：1.對減C計畫的感受與態度；2.教師知覺教育會考對學生學習的影響。問卷均為單選題，採李克特式五點量表，從「非常不同意」到「非常同意」；國中教育會考學力品質監控回饋報告資料之應用主要是調查教師對心測中心所提供結果報告的運用方式。

(三) 半結構式深度訪談

本研究所採用的半結構式訪談法，針對前開兩種問卷的結果，諮詢九位數學教師的看法，主要問題包括：1.課程：文獻顯示考試對課程的內容有明顯的影響，主要是課程窄化，但本研究的問卷結果並非如此，請問您的看法為何？2.教材：請問您平時上課使用的教材有哪些？選擇使用這些教材的原因為何？3.教學法：基測與會考時期所使用的教學法有何差異？會考加考數非後，教學的改變為何？4.評量：會考加考數非後，評量的改變為何？5.衍生議題：會考採等級制，對教學與學習的影響為何？本研究每次訪談均錄音，並於訪談後轉錄為文字；引用受訪者訪談的內容時，均盡量呈現其原貌，相關性較低的內容予以省略。

肆、研究結果與討論

一、數學科教師在基測與教育會考時期教學方式的差異

本節呈現教育會考實施後，國中數學教師於基測與教育會考兩時期在課程、教材、教學法、評量等四個面向的異同。此外，研究者嘗試融入半結構式深度訪談的九位數學科教師的意見與作法。

(一) 課程

課程的部分，問卷調查結果顯示，超過八成的教師認同基測(81.63%)及教育會考(86.79%)的試題均能夠涵蓋考生在國中階段所學習的重要內容與能力；近七成教師認同會花較多的時間教基測／教育會考(68.42%、69.09%)常考的內容；六成左右的教師認為他們不會因為教學時數不足，所以不常出現在基測／教育會考(61.59%、61.81%)的內容就不教，但也有兩成的教師認同會有此現象。深度訪談時，九位受訪教師中，八位受訪教師均表示教學時不會跳過不考的部分，但一定會特別強調經常考的範圍；七位教師也表示基測或教育會考的試題都包含了重要的學習內容，所以也不太能跳過不用教的部分。例如其中一位教師表示：

由於數學的學科特性，不同單元間的學習有其連貫性，最基本、簡單的概念雖然不常考，但他們是學習進階概念的基礎，如果比較少考的基本概念不教，進階單元的學習就會出現問題。所以跳過某些考試不考的內容不教，對教師教學與學生的學習都沒有好處。但不可否認的，常考的內容一定會花比較多的時間教。(T2)

其他研究顯示考試對課程的內容有明顯的影響，主要是課程窄化，教學只集中在考試的範圍內，甚至只著重在最常考的部分(廖佩莉, 2013; Alderson & Wall, 1993; Cheng, 1997; Rodríguez-Muñiz et al., 2016)。然而，本研究問卷調查與深度訪談結果略有不同，由於數學的學科特性，讓教師不會窄化課程，也因為考試的試題都包含了數學重要的學習內容，即使花

比較多的時間在常考的範圍，也不至影響學生的學習，因此研究者認為教育會考並沒有導致數學科學習窄化的現象產生。

(二) 教材

教師使用的教材部分非常多元（詳見表4），但可看出仍以課本與習作為主，此與其他研究結果相同（Andrews et al., 2002; Cheng, 1997），兩者的使用率均高達九成以上；約有六成教師常會使用測驗卷或考古題；另有四成左右的教師會自編教材。值得注意的是，教育會考時期經常使用非選擇題型式測驗卷或考古題的比例較基測時期提升7.50%，其他教材使用頻率則皆相當接近。

表4

教學素材在課堂上的使用頻率

單位：%

	基測時期			教育會考時期		
	未曾使用	不常使用	經常使用	未曾使用	不常使用	經常使用
課本	0.67	7.28	92.05	0.67	7.05	92.28
習作中的選擇題	0.45	7.39	92.16	0.67	7.73	91.60
習作中的非選擇題	0.45	7.17	92.39	0.56	6.50	92.94
數位學習教材	5.26	65.29	29.45	5.38	63.49	31.13
自編教材	4.70	56.21	39.08	4.93	54.98	40.09
參考書	15.34	53.64	31.02	16.01	52.97	31.02
選擇題型式的 測驗卷或考古題	1.23	32.92	65.85	1.23	32.81	65.96
非選擇題型式 測驗卷或考古題	2.13	47.04	50.84	1.68	39.98	58.34

問卷結果顯示，兩時期的教學素材同樣以課本與習作等課內教材為主，而九位深度訪談教師也都會採用「精熟學習」策略，透過不斷地複習，讓學生熟練課程內容，再搭配課外教材做複習和延伸。其中一位受訪教師對教師高度使用教科書所提出的看法與Cheng（1997）提出的原因相似，即教科書不僅提供教學的資訊與活動，同時提出了建議的教學方法與時間分配，該位訪談教師指出：

國內教科書的編審都必須遵守一定的法規與程序，出版社都是邀請領域教授與現職教師組成團隊來編寫教科書，還要經過教育部的審查後才能出版。現在的教科書除

了提供各單元的教學目標、內容之外，也會提供建議的教學活動、方法與時間的配置。每一本教科書都是一群專業人士的心血結晶，教師個人的經驗很難超越，除了補充一些延伸內容，多數的課堂時間還是會用來講解教科書的內容。(T9)

(三) 教學法

在教學法的部分，兩時期各種教學模式的使用頻率大致相同，主要仍以教師單向講述或是師生雙向問答為主，分組能力教學的使用頻率較低（詳見表5）。此結果呼應了多數文獻研究結果，教學方法不受考試內容的影響，仍沿用過去的教學法（廖佩莉，2013；Alderson & Wall, 1993; Cheng, 1997）。八位深度訪談教師均表示基測與教育會考時期的課程內容並未有調整，最大差異在於教育會考時期加考非選擇題，然而在教授數學時，選擇題與非選擇題的解題方式都相同，差異在於解答的表示方法，所以沒有改變自身的教學方法。

表5

教學模式在課堂上的使用頻率

	單位：%					
	基測時期			教育會考時期		
	未曾使用	不常使用	經常使用	未曾使用	不常使用	經常使用
單向講述	0.34	11.53	88.13	0.34	12.87	86.78
問答法	0.22	17.58	82.20	0.22	16.01	83.76
同質性能力分組教學	13.89	71.00	15.12	13.77	70.44	15.79
異質性能力分組教學	6.05	66.97	26.99	5.38	65.18	29.45

根據問卷結果顯示，數學教師較少使用異質性或同質性的分組教學，但九位深度訪談的教師中有六位表示，平時上課時會依據學生的數學程度分組教學；其中三位教師採用異質性分組，將不同能力的學生分為同一組，並且在組內設置小教師制度，由於學生之間程度較為貼近，可以用對方容易理解的方式與語言相互教學，會比教師一一確認學生困難點後再解決的方式更具效益；三位教師提到學校申請政府經費補助採用兩班三組（或三班五組）的教學模式，讓不同程度的學生得到更適切的學習引導，達到適性教學與補救教學的雙贏目標。此外，師大心測中心受國教署委託，分析參與數學深根種子教師、數學科分組與補救教學、亮點教師、活化教學等多個不同專案教師之教學成效，發現所教授的學生會考數學成績顯著優於其他班級，相較於其他教師，專案教師多採用的是分組教學法（師大心測中心，2016·2017）。由此可知，分組教學確實對提升學生的數學學習成就有所助益，誠如其中一位受訪教師所說：

除分組教學之外，也會讓同學在分組內兩兩互教，類似學伴，讓會的教不會的。過程中會讓教導者先看對方不會的試題10分鐘，由教導者負責找這些問題的核心概念，也就是找出對方到底哪裡不會，只講解核心概念，其餘的試題讓被教導者自己嘗試解決。這樣方法可以提升教導者的程度、釐清概念，畢竟自己懂到要教到對方懂之間是有距離的，如果只是會解題，但概念講不清楚，講解時對方是不會懂的，如此便可強化教導者的能力，也可以解決被教導者的問題。另外，多數的題目還是要被教導者自己解決，所以被教導者也不會過分依賴對方，自己的能力確實也可以提升。(T3)

在各個數學技能的重視程度上（詳見表6），教師在數學的閱讀理解能力重視程度增加最多（非常重視比率增加8.4%）、其次為數學表達、溝通能力（非常重視比率增加6.84%）、解題過程的完整性（非常重視比率增加5.59%），其他數學技能重視程度則無太大差異。主要是因為基測時期題目均為選擇題型，教育會考則加入了非選擇題，且非選擇題命題趨向與生活情境結合，學生需先透過閱讀理解問題，才能進一步利用數學概念解題。三位受訪教師提及解選擇題時可以使用應試技巧，如代入法或刪去法，提升答對的機率，即使完全不理解題意，仍然可能猜對；九位受訪教師均表示，解非選擇題時若無法讀懂試題則無法作答，再加上近年來素養導向的命題趨勢，閱讀理解能力就變得愈來愈重要。其中一位受訪教師說道：

加考數非後就特別重視學生的數學表達與書寫能力。當中最重要的是要先確認學生能讀懂題目，因為現在的題目愈來愈生活化、敘述愈來愈長，所以能不能讀懂題目就很重要。有些學生數感很好，邏輯與計算能力也還可以，但就是題目讀不懂，這樣的學生在會考的題目上比較吃虧，所以會特別強調讀懂題目的重要性。主要的訓練方式是先讓學生圈出題目的關鍵字、條件與線索，請學生以自己的文字寫出題目到底要問什麼，再來解題。有些學生會圈出一堆莫名其妙與解題無關的字，或是沒有辦法以文字寫出題意，就會讓學生口述。這部分需要慢慢訓練，尤其是我們數學老師的文字功力也不是非常好，所以有點難度。(T5)

另一位受訪教師則是表示：

本校學生來自外配家庭與隔代教養的比例很高，近40%。這些學生通常閱讀會有問題，基本上都是作業看不懂，並不是概念不會或不會算，比如說不知道「和」是什麼意思，但改成「加起來」就知道。會先讓這樣的學生去上國文的補救教學。本校每週有一天的早自修要讀課外讀物、報紙，要寫讀後心得，這樣才能把學生的閱讀能力帶上來，其他學科才會一起進步。(T6)

表6

教學時對下列技能的重視程度

單位：%

	基測時期				教育會考時期			
	不重視	有點重視	重視	非常重視	不重視	有點重視	重視	非常重視
數學的閱讀理解能力	0.22	3.81	48.94	47.03	0.11	2.13	42.33	55.43
解題策略規劃	0.22	4.70	52.74	42.33	0.22	3.47	49.61	46.70
邏輯推理的合理性	0.22	2.69	47.93	49.16	0.11	2.24	44.90	52.74
解題過程的完整性	0.45	6.94	50.50	42.11	0.11	4.48	47.70	47.70
計算能力	0.56	12.32	54.87	32.25	0.78	11.31	54.31	33.59
數學表達、溝通能力	0.78	10.19	55.99	33.03	0.34	6.38	53.42	39.87
解題技巧與訣竅	1.57	14.33	55.32	28.78	1.34	13.66	54.09	30.91

(四) 評量

兩時期學校段考各題型比重的分配並未有顯著變化，仍以選擇題形式為主，但可發現教育會考時期非選擇題占20%以上的百分比微幅提高。基測雖未考非選擇題，但大部分學校段考仍然有此題型，其中有3.25%的學校在基測時期的段考未包含非選擇題，但教育會考實施後降為0.9%。訪談時，多數教師都表示即使基測時期未考非選擇題，學校段考也會包含非選擇題，其中一位受訪教師表示：

由於基測數學全部都是選擇題，雖然學校段考也會有計算題或填充題，但比重都很低，而聯合模擬考因為要符合基測的題型，就會全部都是選擇題。因為會考加考數非，所以校內段考更重視非選擇題，聯合模擬考也有非選，題型包括填充、計算與應用，雖然沒有硬性規定各題型比重，但非選不會超過40%。(T3)

另一位深度訪談教師則是表示：

本校在基測時期的段考就有計算題，只是沒有那麼重視，給分也是單純的對或錯。會考時期，數學老師要輪流出非選擇題，每兩週讓全校學生寫兩題數非，由出題老師提供評分標準，採部分給分方式，老師也會分析學生的作答反應，再針對錯的地方加強教學。平常教學的時候，會跟學生強調非選擇題的重要性，兩題占15分，一題就占7.5分，再加上現在數非比較簡單，會跟學生說有寫就有分，所以學生願意寫，就比較容易脫離C等級。(T6)

在非選擇題評分方式部分（詳見表7），兩時期評分方式百分比組型相似，半數以上學校屬於「有統一評分規範，取得評分共識後，各自評分」，近三成學校則是「有統一評分規範，教師各自解讀後評分」，僅少數學校屬於「無統一評分規範，各班教師自行評分」。細部比較兩時期數據，可以看到教育會考時期取得評分共識的教師增加了約6%，而未有共識自行評分的教師減少了約6%。

表7

學校段考中數學非選擇題的評分方式

	單位：%	
	基測時期	教育會考時期
無統一評分規範，各班教師自行評分	15.29	9.04
有統一評分規範，教師各自解讀後評分	28.94	29.38
有統一評分規範，取得評分共識後，各自評分	55.78	61.58

兩時期在批改數學非選擇題時，對不同作答內容的重視程度亦有所不同（詳見表8）。由於教育會考數學非選擇題重視的是「解題策略的適切性」及「表達」兩向度，若學生只寫答案但未呈現計算過程，因無法判斷學生數學知識的應用與解題思維過程的表達等能力，將評為0分。從問卷結果中可以發現教育會考時期對於「解題策略方向的正確性」以及「解題過程表達的合理與完整性」的重視程度較基測時期高，但兩時期對「最後答案的正確性」重視程度相似。

表8

批改數學非選擇題時各項作答內容的重視程度

	單位：%							
	基測時期				教育會考時期			
	不重視	有點重視	重視	非常重視	不重視	有點重視	重視	非常重視
解題策略方向的正確性	0.35	4.17	57.24	38.24	0.11	2.03	52.87	44.98
解題過程表達的合理性與完整性	0.35	4.29	56.55	38.82	0.23	1.80	51.18	46.79
最後答案的正確性	0.23	11.24	60.14	28.39	0.45	11.72	59.30	28.52

近年來，教育會考數學非選擇題型採用小題制，設計較為簡單的第1小題，希望能提升學生的作答意願。同時，許多教師也會透過教育會考網站所公布的數學非選相關資料（如評分

指引、各等級樣卷與說明)或參與正式考試閱卷作業,瞭解數學非選的閱卷方式與計分重點,再透過教師們在教學現場的引導,進而提升學生的學習意願與作答動機。其所產生的正向回沖效應便是數學非選的空白卷率呈現連續下降的趨勢,由2016年的31.08%,降至2017年25.65%,至2018年時更下降至20.19%(師大心測中心,2018b)。其中一位深度訪談教師談到他們學校面對數學非選的作法:

我們學校數學科每年都會派一位老師參加會考閱卷,並將接收到訊息,如閱卷方式、計分重點、怎麼寫就有分……等訊息帶回學校與其他數學老師分享,並請老師們在課堂上向學生宣導、說明、強調數學非選擇題的作答重點。也會叮嚀學生,如果不會寫就把題目抄兩遍,或是寫幾個學生認為跟解題有關的計算式。平常教學時也會跟學生強調數非有沾到邊就可能有分數,即使題目看不懂,還是要寫出幾個計算式,即使是把題目提到的數字湊一湊也可以,無論如何都不要空白。(T9)

綜上所述,相較基測時期,國中數學教師在教育會考時期的教學特色如下:1.由於教育會考的試題愈來愈生活化、素養化,教師因此更重視學生閱讀理解能力;2.教育會考加考非選擇題,學校段考評量題型、評分方式、非選評閱重點的調整,而對於評閱重視的內容也一併反映在教學內容中。雖然基測並沒有考非選擇題,但多數教師還是會教授,定期評量也會包含非選擇題,而教育會考加考非選擇題後,其教學時間與評量都有增加。這些教學內容的改變也與加考非選擇題的目的相呼應,正面引導教師的教學,強化學生數學溝通表達的能力,使學生對於數學的學習能夠更全面。

二、教育會考實施後教師對於所衍生議題的看法

教育會考除數學科題型上的變化之外,亦改變了測驗的計分方式,由基測的常模參照以PR值表示考生成績,改為標準參照測驗以等級表示考生表現。透過標準參照測驗,可以知道學生國中階段的數學能力,且希望藉此降低學生考試壓力。政府當局也積極推行「減C計畫」,希望學校及教師能針對數學能力低落且未達基礎標準的學生,給予更多教學資源,進而提升其能力,雖然用意良善,但實際運作上對於學生學習或教師心理感受為何,透過本研究教師問卷調查,進一步瞭解。

本問卷以五點量表方式詢問教師的意見,分為「對減C計畫的感受與採取的教學行動」以及「教師知覺教育會考對學生學習的影響」兩方面探討。為了方便說明,合併「非常不同意」與「不同意」,以及「非常同意」與「同意」成為三點量表。在衍生的減C計畫議題部分,由表9可知,55.91%的數學科教師認同他們針對減C計畫而採取的行動提升了學生的學習動機;67.53%的教師則是更加地精進其專業知能;40.20%的教師認為減C計畫對他們造成教學上的壓力,但只有31.23%的教師認為減C計畫讓學生有學習壓力。此外,七位訪談教師均表示由於縣

表9

對教育會考衍生議題的看法

	單位：%		
	不同意	沒有意見	同意
學校或我針對減C計畫而採取的行動，提升了學生的學習動機	14.55	29.55	55.91
為因應減C計畫，我更加地精進我的教師專業知能	9.85	22.63	67.53
減C計畫讓我有教學壓力	21.92	37.89	40.20
減C計畫讓我的學生有學習壓力	30.88	37.89	31.23
學校使用待加強學生的比例來評估我的教學績效	46.94	37.89	15.17
試題難度提高，學生鑽研難題的現象增加	39.66	27.33	33.01
成績分為三等級，可緩解學生間分分計較的競爭壓力	51.38	23.25	25.38
等級分得太粗，容易造成中等程度學生學習動機低落	10.29	20.50	69.21

市政府教育局（處）要求學校制定減C計畫，為能有效減C，教師們會使用心測中心所提供的教育會考分析結果，調整課程或教學方式，也會透過報告中的試題分析與選項分析找出需討論的試題與診斷學生的迷思概念。

另一方面，由於基測改為教育會考，試題難度由中間偏易改為難易適中，考生成績的表示方式也由量尺分數改為等級制，因而衍生數個議題，本研究就「鑽研難題現象」、「是否緩解競爭壓力」與「學生學習動機」等三個議題蒐集教師的意見。33.01%的教師認為試題難度提高，學生鑽研難題的現象增加；而原本教育會考成績分為三等級，目的是緩解學生分分計較的競爭壓力，但僅有25.38%的教師認同有此效果，超過一半的教師並不認同；69.21%的教師認為等級分得太粗，容易造成中等程度學生學習動機低落，較為可惜的是本研究並未蒐集學生對此三個議題的看法。其中一位訪談教師針對「是否緩解競爭壓力」與「學生學習動機」兩項議題表示了看法：

基測時期成績採PR制，學生會「分分計較」，教育會考採等級制，學生變成「題題計較」，尤其是等級邊緣的學生，只要差一題，成績就會天差地遠，所以並沒有因此降低學生的考試壓力。成績分為三等第是弊多於利，對A等級的學生而言還是很競爭。對B等級的學生來說確實會造成他們放棄學習，因為再怎樣認真也不會突破到A等級或掉到C等級，所以B等級的學生會向老師反應：「我不用學這麼多、這麼認真，反正考出來不會是A也不會是C」。(T2)

但另一位受訪教師則表示不同意見：

即使成績只分三等分，教學時我們會跟B等級學生說不要給自己設限，強調學習是過程，不能只重結果，所以B等級生也不會放棄學習。(T7)

教育會考的主要目的是作為國中畢業生學力檢定之機制，除可降低學生相互比較的競爭壓力、檢測學生基本能力之外，甚至也能與各先進國家（例如英國、美國、澳洲等）檢核學生學習狀況的理念和實務相接軌。為激勵學生學習，並協助教育部及各縣市教育局（處）解決免試超額時可能增額錄取過多的困境，在三等級的原則下，於精熟及基礎等級加上標示，共分為七個類別（師大心測中心，2014），也可以部分程度降低中等程度學生學習動機低落的問題。但不可否認的是，等級制確實讓部分中等程度的學生學習動機低落，但國中階段只是學習的中點，若學生在國中階段便自我設限，不利於學生後續階段的學習。

心測中心每年均針對當年度國中教育會考資料進行分析，提供各校「國中教育會考學力品質監控回饋報告」供各國中與高中職下載使用，報告內含該校學生在各科的成績等級人數百分比、試題通過率分析（分為該校、該縣市、全國三種）與選項分析（分為全校與待加強等級學生兩種）。2018年國中部分有95.10%下載了該資料。但本研究問卷結果顯示，25.20%的數學教師表示「不知道有這樣的報告」，23.96%的教師「知道這份報告但未曾看過內容」，28.04%的教師「看過本校的報告，但不曾使用該資料」，僅有22.80%的數學教師表示「曾經使用該校報告資料」。報告資料的使用部分，在有使用該資料的教師中，72.08%的數學教師表示「會利用此報告提供的資料，適當的調整教學內容與方法」，70.08%的教師會透過試題通過率分析，找出需要討論的試題。由於絕大多數的國中都下載了該報告，但卻僅有兩成多的數學教師表示曾用過該資料，此巨大落差需進一步注意。

伍、結論與建議

基測改為教育會考以及數學加考非選擇題確實為國中數學教學現場帶來影響，在教材、教學法及評量等各方面都產生了正向的改變，課程部分也未出現常見的課程窄化現象。研究者認為教育會考數學科加考非選擇題促進國中數學課程教學正常化，不再只著重選擇題解題能力的養成而已。以下從課程、教材、教學法及評量四個方面整理本研究的發現。

課程方面，問卷調查結果顯示，因為數學的學科特性，不同單元間的學習有其連貫性，所以教學時不會跳過不考的部分不教，但一定會特別強調經常考的內容，多數的教師也認為無論是基測或教育會考，試題都包含了數學重要的學習內容，因此沒有學習窄化的現象。教材方面，課堂教學仍以課本與習作為主，主要是因為教師認為教科書都是由專家學者群編製，除提供教學內容之外，也提出建議的教學活動、方法與時間的配置，因此使用率相當的高。此外，因應加考非選擇題，該形式的測驗卷或考古題使用比例比基測時期高出7.50%。在教學法部分則不受考試內容的影響，仍以教師單向講述或是師生雙向問答為主。須特別注意的是，

減C有成的深度訪談教師與專案教師多使用分組教學方式，深度訪談教師表示分組教學搭配小老師制度，可以同時提升不同程度學生的數學能力，且會比教師單一教授更具教學效率。除此之外，由於教育會考數學科的試題更加地生活化，促使教師更加注重學生數學的閱讀理解能力，而因為教育會考加考數學非選之故，數學表達與數學溝通以及解題過程完整性的重視程度也較基測時期增加。在評量方面，段考未包含非選擇題的比例從基測時期的3.25%下降至教育會考時期的0.9%，代表基測時期雖未有非選擇題，多數教師仍會教導與評量，只是愈接近正式考試，數學非選的教學與評量分量就會愈輕，而教育會考時期就會特別強調數學非選擇題。近年來，數學非選小題制的題型設計，以及國中教師在教學現場的引導，讓更多學生願意作答數學非選，其空白率因而下降也是正向回沖效應之一。在非選擇題的評分部分，教育會考時期的教師更加重視解題策略方向正確性與過程表達的合理性與完整性，學生可以獲得較完整的數學能力。

至於教育會考的負向影響，大致可以分為成績表示方式所造成的考試壓力，以及減C計畫與其因應策略造成的教學壓力兩大類。由於目前教育會考成績仍占免試超額比序三分之一的積分，對能力等級邊緣學生而言，以往基測時期分分計較的情形變成了題題計較，無形中考試壓力大增（Sung et al., 2016）。各縣市為了減少待加強等級的學生人數，督促各校提出減C計畫或策略，直接或間接地造成教師的教學壓力。數學科教師因為加考非選擇題及減C計畫，讓學生重複練習考古題，是否有過度操練的現象與相關衍生的問題，還需進一步研究。學生因為過度練習考古題而提高教育會考數學科成績，教師也認為此方法能快速診斷學生學習困難並改善考試成績，但學生的數學能力是否真的因此而提升，亦值得進一步探究。

那麼，究竟要如何強化教育會考帶來的正面回沖效應，淡化負面影響呢？以下是本研究提出的一些建議。

一、善用國中教育會考學力品質監控回饋報告資料

本研究從深度訪談中發現數學減C卓著的學校，其數學領域召集人與幾位核心教師會透過教育會考學力品質監控報告數據，找出該校學生表現比較差的試題，如平均通過率低於該縣市或全國的試題，搭配各題選項分析結果，進一步掌握學生的學習盲點，在領域會議時進行更多的專業對話，討論表現較差試題的成因，制定前類試題的教學方法或教學時應特別注意的事項，並提供給其他教師參考。除此之外，也將會考的題目融入平常上課，尤其是前述表現較差的試題，段考時亦加入會考題目（小幅修改或原題直接使用），藉此檢視所討論的教學改進方法是否有成效，若成效不彰則再討論改進。這些減C卓著學校的作法，充分體現國中教育會考「回饋學習成果，強化適性輔導」的重要目標，值得其他學校參考。雖然絕大多數國中學校（95.10%）都下載了該資料，但一般數學教師問卷顯示僅有22.80%的教師表示「曾經使用該校報告資料」，甚至有25.20%的教師表示「不知道有這樣的報告」，此實非我們所樂

見。此資料具參考價值，若能提升各校對此報告各項資料的運用，方能更有效率地提升教學品質。

二、提升數學試題的閱讀能力

素養是108課程綱要的主軸之一，根據張鎮華（2017）的解釋，數學素養包含：（一）數學學科知識的素養；（二）應用到學習、生活與職業生涯的素養；（三）正確使用工具的素養；（四）有效與他人溝通的素養。希望學生能在生活情境脈絡中遇到問題時，能找到問題與數學的關聯性，並運用數學知識、技巧，發揮思維判斷解決問題。因此，數學科試題的陳述將與生活情境更加結合，題目陳述也較長，學生須透過閱讀理解瞭解問題與數學的關聯性，由教師訪談得知部分學生會因閱讀理解能力不佳，縱使具有數學解題能力，卻無法找出題目文字中與解題相關的資訊，進而解決問題。因此，數學教師除了教導數學知識外，也需提升學生閱讀理解能力，其中有教師提供其教學經驗，引導學生閱讀數學試題，透過閱讀瞭解題目已知、未知、條件及關鍵字間的關係，並找出題目待解決的問題，進而運用數學知識解題，數學非選除了學生需要透過閱讀理解進行解題外，還需具備有效的溝通，表達其解題歷程的能力，因此除了數學學科知識外，教師也可強化學生閱讀理解與表達的能力，進而提升學生數學素養。

三、提升基礎等級學生的學習動機

無論是問卷調查或深度訪談均指出，教育會考三等級制度導致基礎等級學生學習動機低落，如何提升學習動機是現場教師皆需面臨的問題。國中教育會考數學科試題避免特殊繁瑣的解題方法，以核心、重要的知識與能力為評量重點，因此課堂的教學與評量可用一定比例基礎簡易的題型，讓學生建立信心。透過生活情境與數學內容的連結，也可與學生的興趣互相結合。此外，也可依據學生能力安排學習活動，使不同程度的學生都有體驗成功的機會，適時地給予讚美與鼓勵，強化師生間的互動，和諧融洽的上課氣氛有助於提升學生的學習動機。

陸、研究限制

本研究深度訪談之教師均來自數學減C成效卓著學校，但回沖效應本身應不分學校類型，故本研究深度訪談結果可能無法代表其他減C成效良好的學校或是一般學校數學教師的意見。此外，本研究因聚焦教師在基測改為會考、加考數學非選擇題以及減C計畫之下對教師教學的影響，並透過教師的意見來反映學生可能的感受，並未實際探究其對學生學習的影響。建議未來可透過如宋曜廷等人（2013）所編製的中學生考試壓力量表，評估會考是否確實如其宣稱之能適度地減低學生的考試壓力。

數學科部分試題在基測時期已就強調「生活化」，為因應新課綱強調素養取向的教學與評量，教育會考數學科試題命題設計也逐漸融入真實生活情境，讓學生透過解決問題展現「應用數學在生活情境的能力」，這類素養導向試題強調應用核心知識與技能解決真實情境脈絡中的問題，如此的轉變對國中數學教學現場的影響值得進一步探究。

誌謝

本研究由教育部補助國立臺灣師範大學高等教育深耕計畫「學習科學跨國頂尖研究中心」經費支持，特此致謝。

參考文獻

一、中文文獻

宋曜廷 (2012, 9月17日)。十二年國教要競爭力不要壓力。中國時報。http://only-perception.blogspot.com/2012/09/blog-post_18.html

【Sung, Y.-T. (2012, September 17). The 12-Year Compulsory Education appeals for competitiveness, not stress. *China Times*. http://only-perception.blogspot.com/2012/09/blog-post_18.html】

宋曜廷、趙子揚、王雅鈴、黃瓏瑩、陳佳蓉、曾芬蘭 (2013)。「中學生考試壓力量表」之編製及其信度與效度之分析。測驗學刊, 60(2), 291-318.

【Sung, Y.-T., Chao, T.-Y., Wang, Y.-L., Huang, L.-Y., Chen, C.-J., & Tseng, F.-L. (2013). The development of examination stress scale for junior high school students. *Psychological Testing*, 60(2), 291-318.】

侯傑泰、何穎欣 (2008)。學習回饋及系統監察：香港的經驗。教育研究與發展期刊, 4(4), 1-18。

【Hau, K.-T., & Ho, J. W. Y. (2008). Feedback and system monitoring of students learning: The Hong Kong experience. *Journal of Education Research and Development*, 4(4), 1-18.】

張鎮華 (2017)。數學學科知識也是數學素養 (數學素養系列之3)。高中數學學科中心電子報, 123。https://ghresource.mt.ntnu.edu.tw/uploads/1644995378956mONmCFXd.pdf

【Chang, G.-J. (2017). Fundamental mathematical knowledge is also part of mathematical literacy (Mathematical Literacy- Part 3). *The Newsletter of Mathematics Education Resource Center*, 123. https://ghresource.mt.ntnu.edu.tw/uploads/1644995378956mONmCFXd.pdf】

國立臺灣師範大學心理與教育測驗研究發展中心 (2014)。教育會考與國中基測比較表。https://cap.rcpet.edu.tw/background.html

【The Research Center for Psychological and Educational Testing, National Taiwan Normal University. (2014). *The comparison and contrast chart between Comprehensive Assessment Program for Junior High School Students and the Basic Competence Test for Junior High School Students*. https://cap.rcpet.edu.tw/background.html】

國立臺灣師範大學心理與教育測驗研究發展中心 (2016)。105年國中教育會考學力品質分析報告。作者。

【The Research Center for Psychological and Educational Testing, National Taiwan Normal University. (2016). *The analysis report of 2016 Comprehensive Assessment Program for Junior High School Students*. Author.】

國立臺灣師範大學心理與教育測驗研究發展中心 (2017)。106年國中教育會考學力品質分析報告。作者。

【The Research Center for Psychological and Educational Testing, National Taiwan Normal University. (2017). *The analysis report of 2017 Comprehensive Assessment Program for Junior High School Students*. Author.】

國立臺灣師範大學心理與教育測驗研究發展中心 (2018a)。數學科考試內容。https://cap.rcpet.edu.tw/test_4_4.html

【The Research Center for Psychological and Educational Testing, National Taiwan Normal University. (2018a). *Test specifications for the math test of the Comprehensive Assessment Program for Junior High School Students*. https://cap.rcpet.edu.tw/test_4_4.html】

國立臺灣師範大學心理與教育測驗研究發展中心 (2018b)。107年國中教育會考計分結果。

<https://ws.moe.edu.tw/Download.ashx?icon=.pdf&n=C612F6A0BC465D295A8E86BC9517043148B69F091504934381EA5F093DCCOD269DD10A5168B73EB1&u=C099358C81D4876C725695F2070B467E8B81ED614D7AF43EF5976C55CFC16E859844CB6EF80A45A2B1FC5A61A286592CCDCDBFDC3EF4C291B8073738BE2FC398F87CDOA27858A326CF48397566DA1298>

【The Research Center for Psychological and Educational Testing, National Taiwan Normal University. (2018b). *The results of 2018 Comprehensive Assessment Program for Junior High School Students*. <https://ws.moe.edu.tw/Download.ashx?icon=.pdf&n=C612F6A0BC465D295A8E86BC9517043148B69F091504934381EA5F093DCCOD269DD10A5168B73EB1&u=C099358C81D4876C725695F2070B467E8B81ED614D7AF43EF5976C55CFC16E859844CB6EF80A45A2B1FC5A61A286592CCDCDBFDC3EF4C291B8073738BE2FC398F87CDOA27858A326CF48397566DA1298>】

教育部 (2017)。國民中小學成績評量準則。 <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=H0070019>

【Ministry of Education. (2017). *Assessment criteria for high school and elementary school students*. <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=H0070019>】

教育部 (2018)。十二年國民基本教育課程綱要國民中小學暨普通型高級中等學校—數學領域。 <https://cirn.moe.edu.tw/Upload/file/27338/72246.pdf>

【Ministry of Education. (2018). *Twelve-year compulsory education curriculum outline of language area—Mathematic*. <https://cirn.moe.edu.tw/Upload/file/27338/72246.pdf>】

曾芬蘭、游羽萱、蔡逸凡、陳柏熹 (2019)。國中教育會考英語科聽力測驗實施的回沖效應。

教育科學研究期刊, 64 (2), 219-252。 [https://doi.org/10.6209/JORIES.201906_64\(2\).0008](https://doi.org/10.6209/JORIES.201906_64(2).0008)

【Tseng, F.-L., You, Y.-X., Tsai, I.-F., & Chen, P.-H. (2019). A pilot study of the washback effect of the incorporation of English listening test in the comprehensive assessment program for junior high school students. *Journal of Research in Education Sciences*, 64(2), 219-252. [https://doi.org/10.6209/JORIES.201906_64\(2\).0008](https://doi.org/10.6209/JORIES.201906_64(2).0008)】

廖佩莉 (2013)。倒流效應：香港全港性系統評估 (TSA) 對小學中國語文教師的影響。教育研究月刊, 228, 86-102。 <https://doi.org/10.3966/168063602013040228007>

【Liu, P.-L. (2013). The washback: The impact of the Hong Kong Territory-wide System Assessment (TSA) on primary Chinese language teachers. *Journal of Education Research*, 228, 86-102. <https://doi.org/10.3966/168063602013040228007>】

二、外文文獻

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129. <https://doi.org/10.1093/applin/14.2.115>

Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback: A case-study. *System*, 30(2), 207-223. [https://doi.org/10.1016/S0346-251X\(02\)00005-2](https://doi.org/10.1016/S0346-251X(02)00005-2)

Au, W. (2011). Teaching under the new Taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies*, 43(1), 25-45. <https://doi.org/10.1080/00220272.2010.521261>

- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38-54. <https://doi.org/10.1080/09500789708666717>
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 37-50). Lawrence Erlbaum Associates.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732980>
- Lam, H. P. (1994). Methodology washback an insider's view. In D. Nunan, R. Berry, & V. Berry (Eds), *Bringing about change in language education: Proceedings of the international language in education conference 1994* (pp. 83-103). University of Hong Kong.
- Mulvenon, S. W., Stegman, C. E., & Ritter, G. (2005). Test anxiety: A multifaceted study on the perceptions of teachers, principals, counselors, students, and parents. *International Journal of Testing*, 5(1), 37-61. https://doi.org/10.1207/s15327574ijt0501_4
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256. <https://doi.org/10.1177/026553229601300302>
- Pearson, I. (1988). Tests as levers for change (or "putting first things first"). In D. Chamberlain & R. Baumgartner (Eds.), *ESP in the classroom: Practice and evaluation* (pp. 98-107). Modern English Publications in Association with the British Council.
- Ritt, M. (2016). The impact of high-stakes testing on the learning environment. *Master of Social Work Clinical Research*, 658. http://sophia.stkate.edu/msw_papers/658
- Rodríguez-Muñiz, L. J., Díaz, P. Mier, V., & Alnoso, P. (2016). Washback effect of university entrance exams in applied mathematics to social sciences. *PLoS ONE*, 11(12), e0167544. <https://doi.org/10.1371/journal.pone.0167544>
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, L. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298-317. <https://doi.org/10.1177/026553229601300305>
- Sung, Y.-T., Chao, T.-Y., & Tseng, F.-L. (2016). Reexamining the relationship between test anxiety and learning achievement: An individual-differences perspective. *Contemporary Educational Psychology*, 46, 241-252.
- Watanabe, Y. (1996). Investigating washback in Japanese EFL classrooms: Problems of methodology. In G. Wigglesworth & C. Elder (Eds.), *The language testing circle: From inception to washback* (pp. 208-239). Applied Linguistics.
- Wren, D. G., & Benson, J. (2004). Measuring test anxiety in children: Scale development and internal construct validation. *Anxiety, Stress, and Coping*, 17(3), 227-240. <https://doi.org/10.1080/10615800412331292606>

Journal of Research in Education Sciences

2022, 67(1), 227-254

[https://doi.org/10.6209/JORIES.202203_67\(1\).0008](https://doi.org/10.6209/JORIES.202203_67(1).0008)

Washback Effect From Changes to Comprehensive Assessment Program Math Test for Junior High School Students

Ming-Chiu Chang

Research Center for Psychological and Educational Testing,
National Taiwan Normal University

Li-Ying Huang

Research Center for Psychological and Educational Testing,
National Taiwan Normal University

Chia-Jung Chen

Research Center for Psychological and Educational Testing,
National Taiwan Normal University

Po-Hsi Chen

Department of Educational Psychology and Counseling,
Research Center for Psychological and Educational Testing and
Institute for Research Excellence in Learning Sciences,
National Taiwan Normal University

Fen-Lan Tseng

Research Center for Psychological and Educational Testing and
Institute for Research Excellence in Learning Sciences,
National Taiwan Normal University

Abstract

Purpose

This study investigated the effects of incorporating constructed-response items into the math tests in Taiwan's Comprehensive Assessment Program for Junior High School Students (CAP). Using questionnaires, this study explored the differences in junior high school math teachers' overall approaches to teaching and compared the assessment approaches used in preparation for the Basic Competence Test for Junior High School Students (BCT) with those used for the CAP.

Literature Review

2.1 Washback Effect

Hughes (2003) defined the washback effect as the effect of a test on teaching and learning, which may be positive or negative. Studies on the washback effect have been classroom centered, focusing on the curriculum, materials, pedagogy, assessments, feelings, and attitudes.

2.2 Possible Effect of Adjusting the CAP

The purposes of the CAP and BCT are different; therefore, the functions of the exams, scoring methods, types of test items, test times, and presentation of the test results differ. For a comparison

of the CAP and BCT, please refer to the CAP website (<https://cap.rcpet.edu.tw/index.html>).

The goal of adding constructed-response items to the CAP math test was to induce positive measurement-directed instruction, with the expectation that teachers would adjust their teaching and assessment methods according to the type of items on the CAP. The Ministry of Education, county and municipal education bureaus, and junior high schools attach great value to the CAP results; thus, teachers must respond to the expectations of school authorities and use various strategies to help students prepare for the examination, which may put pressure on teachers.

Therefore, although the CAP has been administered for many years, whether math teachers can select appropriate and diverse forms of assessment in their regular teaching in response to the incorporation of constructed-response items is worth exploring.

Methods

3.1 Participants and Procedure

A survey was conducted with an online questionnaire, which was sent to all junior high schools, and three math teachers from each school were selected to complete the survey. After completing a general math teacher questionnaire, teachers with more than 5 years' teaching experience were asked to complete a math backwash questionnaire to ensure that those completing the questionnaire had teaching experience related to both the BCT and CAP.

3.2 Measures

This study used two types of teacher questionnaire and semistructured in-depth interviews as follows:

- (1) Questionnaire on the washback effect of the math test: This questionnaire had eight items and 35 subitems mainly investigating the curriculum, teaching materials, pedagogy, and assessment used by math teachers in preparation for the BCT and CAP.
- (2) General math teacher questionnaire: This questionnaire surveyed teachers' opinions regarding the CAP and their use of the Feedback "Report on Monitoring the Academic Quality of the CAP" in each school.
- (3) Semistructured in-depth interviews: Nine junior high school math teachers were asked in semistructured interviews for their comments on the results of the first two questionnaires.

Results

- (1) Approximately 70% of the teachers stated that they would spend more time on the instruction of test content covered by the BCT and CAP.
- (2) The teachers used textbooks as the main source of instruction for both exams, with additional practice of constructed-response items and sample questions from previous exams for the CAP

(approximately 7.5% higher).

- (3) Most teachers used one-way lectures or two-way teacher-student techniques to prepare students for both exams.
- (4) The teachers placed the greatest emphasis on reading comprehension skills for the CAP, followed by communication skills.
- (5) No significant difference was detected in the proportions of item formats on regular midterm and final exams, even with the slight increase in constructed-response items.
- (6) Over 50% of the teachers thought the change from a norm-referenced to a criterion-referenced scoring system would not reduce students' test anxiety, and nearly 70% of the teachers believed that the new scoring system would reduce learning motivation.

The findings of our study indicate a washback effect from the incorporation of constructed-response items into the CAP math test in junior high school math classes.

Conclusions and Suggestions

5.1 Conclusions

The characteristics of the math teachers' preparation for the CAP and BCT are compared as follows:

- (1) As the CAP items became more authentic and competency-based, teachers placed more emphasis on students' reading comprehension skills.
- (2) The addition of constructed-response items to the CAP led to adjustments in the types of test items and scoring methods used by teachers, with a greater emphasis on constructed-response items; this emphasis on constructed-response items was reflected in the teaching content.

Although the BCT contains no constructed-response items, most teachers nevertheless prepare students for this type of item; regular assessments now include constructed-response items. Such changes in the content of instruction achieve the intended purposes of adding constructed-response items to the test, that is, to guide teachers' instruction, strengthen students' ability to communicate and express themselves in math, and make students' learning of math more comprehensive.

5.2 Suggestions

- (1) Teachers should use the data from CAP test reports effectively to identify the test items on which students performed the most poorly and to further understand students' learning blind spots.
- (2) In addition to teaching mathematical knowledge, teachers should strengthen students' reading comprehension and communication skills to enhance their mathematical literacy.
- (3) To enhance the learning motivation of basic level students, a certain percentage of basic or easy items can be used in class to build confidence. Learning activities can be arranged in accordance

with student abilities so that students of different ability levels can experience success in math.

Keywords: washback effect, high-stakes test, the Comprehensive Assessment Program for Junior High School Students (CAP), mathematic constructed-response items