會考與段考評量導向下國中教師教學實踐的回沖效應、決策考量與實踐限制之研究

劉世雄*

摘要

研究目的

本研究探討國中教師在會考與段考並行的評量情境下，如何進行教學與評量實踐中的調整，並分析其中所呈現的回沖效應、決策考量與實踐限制。研究特別關注教師如何在兩種評量導向不一致的情況下，將不同評量要求轉化為可行的教學行動。

研究設計／方法／取徑

本研究採詮釋主義取向，以焦點團體訪談蒐集資料。研究者先透過前導研究修訂訪談題目，再邀請25位具會考科目任教與段考命題經驗之國中教師，進行五場焦點團體訪談。所得資料以主題分析法進行編碼、跨場次比較與整合，以歸納不同評量導向下的教學實踐經驗。

研究發現或結論

研究結果顯示，會考較引導教師朝向核心概念統整、跨單元連結與素養導向教學，段考則較聚焦課程進度、基礎能力與學生信心維持。教師的教學與評量調整，主要是在學習信心、跨文本閱讀能力、時間進度壓力與外部期待之間進行專業判斷。時間壓力、能力差異與制度規範構成兩種評量情境中的共同限制；家長期待則透過段考分數與校內排名形成較直接壓力，使段考命題與教學調整更易受即時成績評價制約。整體而言，教師教學實踐並非單純反映評量差異，而是在不同評量要求與情

境限制之間持續協調與轉化。

研究原創性／價值

本研究補充既有回沖效應研究較少處理的雙重評量情境，指出教師並非被動回應評量要求，而是在會考與段考並行脈絡中，依據學生學習、課程進度與外部條件進行專業判斷與取捨。研究亦顯示，段考回沖效應不僅來自評量內容本身，也受到家長對分數、排名及學校成績表現之社會評價機制影響，因而呈現不同於會考的情境特殊性。整體而言，回沖效應更適合被理解為嵌入情境脈絡的動態歷程。

教育政策建議與應用

研究結果顯示，多數限制與時間配置、制度規範及學生能力差異等結構性條件有關。因此，教育行政與學校推動素養導向教學時，不宜僅著眼於命題改革，也應調整課程節奏、評量制度與補救教學。此外，家長對段考分數與排名的高度關注，也顯示後續可增進其對評量取向與學生能力發展關係的理解。

關鍵詞：國中教育會考、校內段考、回沖效應、教師決策、教學實踐

DOI：10.6869/THJER.202512_42(2).0004
投稿日期：2026年4月24日，2026年6月11日修改完畢，2026年8月18日通過採用
*劉世雄，國立彰化師範大學師資培育中心教授，E-mail：shsiung@cc.ncue.edu.tw

壹、緒論

根據國立臺灣師範大學心理與教育測驗研究發展中心（2024）的報告，2024年國中會考約有一萬兩千名考生之五科成績均列為「待加強」等級，且英語、數學與自然科之待加強比例均超過20%。此一現象顯示，部分學生未達基本學力要求，並可能在知識理解、情境應用與遷移上面臨困難（Organisation for Economic Co-operation and Development [OECD], 2019, 2023）。在此情況下，會考可視為一項高風險（high-stakes）測驗，成為教師反思與調整教學的重要依據，以支持學生建立概念理解與知識應用能力（Boaler, 2016; Schwartz et al., 2011）。

此類評量結果對教學所產生的影響，可視為「回沖效應」（washback effect）（Cheng, 2005; Green, 2013）。會考命題強調「情境化分析」與「跨文本（文字、圖表、圖片、數據等）訊息擷取」能力，透過貼近生活的問題情境，引導學生理解學習內容與實際情境之關聯。教師須在教材、教學活動與評量設計中，強化學生之核心概念建構、生活情境連結與問題解決能力的培養。

既然教師可依學生的會考表現調整教學，理想而言，校內常態性評量（如段考）在取向與題目設計上亦應與會考命題形成連貫，作為學習歷程的銜接機制。已有研究指出，高風險標準化測驗會對教學內容、教學與評量方式產生引導作用，使教學逐漸朝向測驗取向發展（Au, 2007; Cheng, 2005; Green, 2013）。在臺灣108課綱實施後，校內評量已逐漸納入情境化、生活化與跨文本理解等素養導向題型；不過，教師在發展高層次評量任務時，仍可能受到多重條件限制。因受到學校脈絡、學科特性、命題習慣與時間壓力的影響，部分段考內容仍較偏向知識再現、程序性作答或基礎概念檢核，較難全面涵蓋分析、評估等高層次認知歷程，亦反映教師在實務中發展高層次評量任務所面臨的困難（Brookhart, 2010; Gulikers et al., 2018; Heitink et al., 2016）。上述顯示，大型測驗（如會考）與校內評量（如段考）在內容與認知要求上存在不一致，使教師在教學調整時須同時面對不同評量邏輯，可能影響教學設計的連貫性與學生學習經驗之整合。

過去關於回沖效應的研究，多聚焦大型考試對教學取向的影響，

較少探討教師在日常教學情境中如何回應不同評量系統的落差。近年來，本土研究雖已關注到國中教育會考與108課綱素養導向評量的實施困境，但對於教師在會考與段考雙重評量邏輯下之落差與回應的研究仍不多見。當兩者同時作用於教師時，其回沖效應可能相互強化、相互削弱，或在素養發展與分數期待之間形成拉扯。

相關研究指出，不同評量形式可能對教材選擇、教學活動與評量設計產生差異性的正、負向回沖效應。因此，理解會考與段考評量不一致對教學實踐的影響，須同時分析回沖效應、教師決策考量與實踐限制。

基於上述脈絡，本研究旨在探究教師因應會考與段考之教學實踐中的調整，並剖析其回沖效應、決策考量因素與實踐限制。具體研究問題如下：

一、在不同評量導向（會考與段考）下，教師在教材處理、教學活動與評量設計上的正、負向回沖效應有何異同？兩者呈現何種互動關係？

二、面對會考與段考兩種評量制度的不同要求時，教師如何進行專業判斷，並採取哪些策略性的教學應變與折衷？

三、教師在進行上述教學應變與折衷時，面臨哪些情境限制？這些限制在會考與段考評量導向下是否有所差異？

貳、文獻探討

一、評量結果回饋到教學活動的回沖效應

「回沖效應」源自應用語言學領域，用以說明測驗在設計與使用後，如何改變教師與學習者的認知、態度與行為（Alderson & Wall, 1993）。一般而言，回沖效應是指測驗對「教」與「學」所產生的影響，其作用可能促進或阻礙教學歷程（Messick, 1996）。此一影響主要發生於課堂層次，也可能擴及課程設計與教育制度，影響更廣泛的測驗（Bachman & Palmer, 1996; Wall, 1997）。因此，回沖效應可視為連結評量與教學實踐的重要機制。

在回沖效應的分析面向中，「方向性」為核心概念，常區分為「正

向」與「負向」回沖。當測驗能反映課程目標並提供有意義回饋時，可促進學習動機與理解，並引導教師發展較具深度的教學活動，此為正向回沖（Messick, 1996; Pearson, 1988）。相對地，當測驗偏重低層次知識或過度強調結果時，則可能產生負向回沖，例如縮減課程範圍、強化題型操練，使教學偏向應試訓練，並限制教學彈性（Madaus, 1988; Noble & Smith, 1994）。

然而，回沖效應並非單純的正負二分，而是一種情境依賴的複雜現象。即使施行相同測驗，在不同教師與學校情境中，亦可能產生不同類型與程度的影響（Alderson & Wall, 1993）。換言之，其作用不僅取決於測驗本身的設計，也與教師信念、教學經驗與制度脈絡相關，因此正負回沖可能同時存在。近年研究亦開始從教師實務角度進一步檢視回沖效應（Mizutani, 2009）。Green（2013）指出，回沖效應並非完全由測驗直接產生，而是涉及教師專業判斷與教學決策的中介作用。Rahman等人（2021）亦發現，部分教師對大型測驗持保留態度，並傾向依據自身教學經驗調整評量方式，顯示制度導向與教學導向之間存在一定的張力。相關本土研究與政策文件亦指出，素養導向評量雖有助於促進教師重新理解學科能力與命題取向，但在實務推動上仍面臨命題知能、時間不足、學生差異與學科特性等挑戰（王竹梅、丁一顧，2021；李偉俊，2023；張銘秋等人，2022；曾芬蘭等人，2019）。整體而言，既有研究仍多聚焦於高風險測驗，對不同評量情境中回沖效應的展現仍缺乏一致理解。

二、教師根據評量結果進行教學調整之專業判斷與考量因素

回沖效應對教學實踐的影響，關鍵在於教師如何理解並運用評量結果，以及進行教學調整時的考量。研究指出，評量並非直接決定教學實踐，而是透過教師的詮釋與專業判斷，轉化為具體的教學決策（Green, 2013）。從教師評量素養觀點來看，教師評量素養不僅涉及評量知識與技術的掌握，更是一種能在實務情境中結合教師信念、制度脈絡與教學行動的動態能力（Xu & Brown, 2016）。Looney等人（2018）亦指出，教師作為評量者，其專業判斷涉及對學習證據的詮釋，以及在課程要

求、學生需求與評量目的之間的取捨。據此，教師在會考與段考雙重評量情境下進行的教學調整，可理解為一種情境化的專業判斷歷程。

在教學調整方面，教師主要依據評量結果調整教材內容、教學活動與評量設計。首先，教師會依學生的學習表現調整教學重點，針對常見錯誤或學習困難進行補強；當評量與學習目標高度對齊時，有助於提升學生在大型測驗中的表現（Baird et al., 2017; Suurtamm et al., 2016）。其次，教師亦會調整教學策略以回應學生的理解落差（Heritage, 2013）。此外，教師也可能調整評量形式，使其更貼近學生學習狀況或與外部測驗對應。

然而，評量對教學的影響並非僅為單一維度的調整，而是受多重因素驅動。當教師或學生知覺評量具有高度重要性時，其回沖強度更為顯著（Toksöz & Kılıçkaya, 2017）；教師的評量素養、教學信念與情境條件，也會影響其解讀與運用評量資料的能力（Looney et al., 2018; Mizutani, 2009; Xu & Brown, 2016）。

教師對評量結果的解釋與歸因，亦會影響其教學調整方式。教師可能依據學生表現調整教學（Black & Wiliam, 2009; Heritage, 2013），在高風險測驗情境下也可能轉向以測驗內容為核心的教學（Madaus, 1988）。當同時面對外部測驗與校內評量時，教師須在制度要求與教學專業之間持續判斷，並權衡不同評量導向（Rahman et al., 2021）。

整體而言，教師依據評量結果進行教學調整，並非單純的結果回應，而是涉及對評量資訊的解讀、教學策略的選擇，以及多重情境因素的權衡。不同評量來源所引發的調整方式，亦反映教師在實務中持續進行判斷與取捨的歷程。

三、國中教師根據學生評量表現進行教學與評量調整的實踐限制

既有研究指出，教師實踐涉及制度規範、文化脈絡、個人信念與評量素養，並在多重目標與要求下進行取捨（Biesta et al., 2015; Mizutani, 2009; Priestley et al., 2015; Xu & Brown, 2016）。因此，教師依據學生評量表現進行教學與評量調整時亦可能受到多重條件限制。

教師在教學實踐中的調整，其限制主要來自三個層面。其一為評量設計層面，教師在進行校內評量時，常受到時間限制、命題習慣與升學壓力等因素影響，使高認知層次題型的發展與實施受到限制（Brookhart, 2010; Gulikers et al., 2018; Heitink et al., 2016）。其二為決策歷程層面，教師須同時考量學生表現、課程進度、評量要求與教學信念，使決策呈現多重交織（Looney et al., 2018; Mizutani, 2009; Xu & Brown, 2016）。近年研究亦指出，教師評量素養已從測驗知識與技能，轉向結合教師信念與專業實踐的動態能力；然而，仍可能受到考試導向文化、資源不足與教師訓練不足等因素限制（Pastore, 2023; Yen et al., 2024）。其三為情境條件層面，時間安排、制度規範、學生差異與同儕合作等因素，亦可能限制教學調整的幅度與方向（Biesta et al., 2015; Datnow & Park, 2019; Priestley et al., 2015）。不過，教師所採取的教材與教法安排仍可能影響學生學習表現。顏郁玲與陸偉明（2023）以臺灣教育長期追蹤資料庫分析國三學生與數學教師資料，發現學校數學教師使用多元教材教法對學生數學成就具有正向影響，並能調節學生努力程度與數學成就之間的關係。此一發現顯示，在結構性條件限制下，教師教材設計與教學方法仍是影響學生學習的重要因素。綜合而言，素養導向命題能否進一步轉化為教學實踐，不僅取決於命題技術，也受到教師專業判斷、教材教法安排與教學外在條件的影響。然而，既有研究多聚焦於高風險測驗對教學調整的引導作用（Cheng, 2005; Green, 2013），較少深入探討教師在日常教學情境中，如何回應不同評量系統之間的落差，亦較少關注其在調整歷程中的限制與挑戰。基於此，本研究以會考與段考作為兩種評量導向，比較教師在面對兩者時所產生教學調整及其限制的差異。相較於過去僅關注單一測驗回沖效應的研究，本研究進一步分析教師在雙重評量壓力下，如何在素養導向與成績導向之間進行教學與評量調整，並探討其在評量設計、決策歷程與情境限制上的異同。此一分析不僅可補充既有回沖效應研究於多重評量情境下的不足，亦有助於建構對教師教學調整歷程的整合理解。

四、國中教師在會考與段考的回沖效應之相關研究

當前已有研究探討學生評量表現如何回饋至教師教學設計。例如，張凱翔等人（2016）發現，將科學素養導向試題融入國中地球科學評量，有助於提升學生素養能力與學習表現；曾芬蘭等人（2019）亦指出，教師會因應會考命題方式調整教學策略，如增加聽力與口說教學，並提高評量中的聽力試題比例。以此而言，評量結果會引發教學策略調整，並可能促進素養導向與情境化測驗表現。

然而，回沖效應除具有正向作用外，也可能引發負向影響。張銘秋等人（2022）發現，教師為配合會考題型增加非選擇題比例，未必能減輕學生學習壓力，甚至可能降低中等程度學生的學習動機。Tirtaningrum與Agustinus（2015）指出，教師與學生常以通過測驗為主要目標，使教學偏向應試取向；Nurjati與Rahayu（2023）亦發現，大型英語評量雖可能提升學習動機與習慣，卻也限制教學設計彈性，使教學過度聚焦於評量標準。

上述研究顯示，當課程與教學內容與測驗命題理念一致，且教師依據評量結果進行教學調整並提升學習成效時，可形成正向回沖效應（Taylor, 2005）；反之，若教學僅聚焦於測驗題目而窄化課程內容，即使產生回沖，亦屬負向回沖效應（Brown, 2004）。

然而，既有研究雖已說明會考或特定評量題型可能對教學產生正、負向回沖效應（張凱翔等人，2016；曾芬蘭等人，2019；Cheng, 2005; Green, 2013），但較少同時納入不同評量系統進行分析。特別是在國中階段，教師如何在不同評量要求之間進行判斷與取捨，以及相關的專業判斷與情境條件限制，仍有待釐清。

參、研究方法與設計

根據研究目的與研究問題，本研究主要聚焦於教師如何詮釋學生在會考與段考中的表現，以及在教材處理、教學活動與評量設計上的回應方式，進一步探討其考量因素與所面臨的實踐限制。基於上述，本研究關注教師經驗的理解與詮釋，採取詮釋主義取向，並以焦點團體訪談蒐

集資料，以呈現其教學實踐經驗的觀點與意義。

一、研究參與對象

為提升訪談題目的適切性與資料信實度，本研究先邀請某直轄市某領域國中國教輔導團教師進行前導研究。研究者依據文獻探討發展訪談題目，並透過五位教師的討論與回饋，檢視題目是否貼近國中教育現場情境及教師對會考與段考的理解，據以修訂訪談題目。

完成訪談題目修訂後，由該輔導團教師協助推薦正式焦點團體訪談之參與者。本研究採質性取向，參與者之選擇並非追求統計代表性，而是採目的性取樣，取得與研究問題密切相關的經驗資料。因此，研究者請輔導團協助推薦參與者時，特別要求將下列條件納入考量：具備學生考試表現正、負向回沖經驗者、三年內曾任教國中三年級會考科目者、具校內段考命題經驗者，以及教學年資十年以上之教師，以增加參與者經驗的差異性與資料豐富性。最終共邀請25位國中教師參與焦點團體訪談，其中男性14人、女性11人，平均教學年資17年，且皆具國中三年級會考科目之任教經驗。

本研究參與教師皆符合上述取樣條件，並涵蓋不同會考科目。研究共進行五場焦點團體訪談，每場由五位教師組成，採多科跨校混合方式分組。由於本研究的分析焦點在於教師面對會考與段考雙重評量導向時所產生的教學調整、專業判斷與實踐限制，而非比較不同學科或學校類型之差異，因此未將學科別或學校性質設定為主要分析類別，以避免在樣本數有限及學校脈絡分類可能涉及主觀判斷的情況下產生過度推論。

每場焦點團體訪談時間約1.5小時。訪談前已完成錄音與研究倫理說明，並取得參與者同意及簽署同意書。訪談於2025年10月至11月間以實體方式進行，以促進互動與經驗交流，並提升資料蒐集的深度與完整性。

二、焦點團體訪談

根據文獻探討與研究問題，本研究將訪談題目區分為回沖效應、決策考量因素與實踐限制三個面向，並採取由經驗喚起逐步進入具體行動與反思分析之流程安排。首先，透過開放性問題，引導教師回顧其如何解釋學生在會考與段考中的表現；第二與第三題聚焦於不同評量導向下教材處理、教學活動與評量設計之調整；第四與第五題探討教師進行教學調整時的決策考量及其差異；第六與第七題則釐清教師在實踐過程中所面臨的情境限制，以及不同評量情境下的差異。訪談題目初稿經與輔導團教師討論與修正後，形成正式題目如下：

1. 當您看到學生在會考與段考中的表現時，通常會如何解釋其學習狀況？會從哪些面向來判斷？

2. 在會考導向下，您在教材處理、教學活動與評量設計上，會有哪些調整？原因為何？

3. 在段考導向下，您在教材處理、教學活動與評量設計上，會有哪些調整？與會考有何不同？

4. 在進行上述教學與評量調整時，您主要會考量哪些因素？例如學生表現、時間進度、考試壓力、學校要求或個人教學信念等。

5. 在會考與段考情境下，這些考量因素是否有所不同？請舉例說明。

6. 在進行教學與評量調整的過程中，您通常會遇到哪些限制？

7. 這些限制在會考與段考情境下，是否有所不同？

三、資料分析與詮釋

焦點團體訪談資料分析具有持續建構與修正的特性，研究者在資料蒐集過程中即同步進行初步分析，並透過反覆比較與詮釋逐步形成研究觀點。本研究依據訪談題目所涵蓋之核心面向，包括教學與評量設計的回沖效應、影響教學調整之決策考量因素，以及教學實踐所面臨的情境限制，分為兩個階段進行分析：

第一階段為訪談進行中的即時分析。研究者於主持焦點團體訪談

時，針對教師對教學活動、評量設計、決策考量，以及會考與段考正、負向回沖效應等議題的回應，進行初步整理與比較。研究者並於訪談過程中適時進行歸納與回饋，透過參與者的補充與修正，確認觀點之適切性，以提升資料的準確性。

第二階段為訪談後之整體分析。研究者將錄音資料完整轉錄後，依據研究問題所關注之三個核心面向，包括回沖效應、教學調整之決策考量因素，以及教學實踐之情境限制，進行資料編碼、比較與詮釋。本研究之分析歷程兼採開放編碼與主題整併。研究者先反覆閱讀逐字稿，標記教師對會考、段考、教材處理、教學活動、評量設計、學生表現、時間壓力與外部期待等相關陳述，形成初步開放編碼；其後再依研究問題進行編碼比較與整併，發展出主題與次主題。雖然本研究依研究問題設定以教材處理、教學活動與評量設計作為主要分析面向，但分析過程仍保留教師現場經驗中浮現之概念，以避免僅將資料套入既有框架。

在主題發展過程中，研究者特別關注不同評量導向（會考與段考）下所呈現的差異，例如教師在會考導向下較強調跨單元整合與閱讀理解訓練，在段考情境中則較聚焦於課程進度與基礎能力掌握，以及在評量設計上如何兼顧素養導向題型與學生信心維持。透過主題的持續比較與關聯性檢視，進一步建構教師在不同評量壓力下進行教學調整的整體意義架構。

在完成個別場次初步分析後，研究者進一步進行跨場次比較。研究者先整理五場焦點團體訪談各場次的主要主題與觀點，再檢視其在回沖效應型態、教學調整策略及背後考量因素上的一致性與差異性。例如，分析「閱讀理解與圖表判讀能力之強化」及「授課時數不足」是否為多數場次反覆出現的共通經驗；對於「校內輪流命題的限制」或「特定學生組成差異」等僅出現在部分場次的觀點，則回溯其發言脈絡與情境條件，以釐清其特殊性。

在此過程中，研究者亦持續修正與整合主題，透過不同場次資料的反覆檢視，確認主題的穩定性與解釋力，並降低單一場次資料可能造成的偏誤。對於參與者未形成一致看法之處，則保留其多元觀點，並結合具體教學情境加以說明，例如教師在「素養導向教學」與「考試成績壓力」之間的拉扯，以及在「提升能力」與「維持學生信心」之間的調整

策略，以呈現教師在會考與段考情境下的多樣決策歷程。

另外，在資料呈現上，本研究不將單一教師的發言直接視為全體觀點。若某一觀點於同一場次中獲其他教師的回應、補充或確認，研究者將其標記為該場次的共同觀點；若該觀點於不同場次反覆出現，則進一步歸納為跨場次主題。對於僅出現在單一教師或單一場次中的意見，則保留其情境脈絡，避免過度推論為全體教師共識。研究報告中僅呈現一位參與者的代表性語句，作為該觀點之引述，並註明其場次與教師代號，以利後續資料檢核。

為進一步呈現資料分析歷程，本研究整理一份編碼架構對照表（如表1所示），以說明逐字稿摘錄、初步編碼、次主題與主題之間的關聯。另為提升分析信實度，研究團隊共同檢視部分逐字稿、初步編碼與主題歸屬；若對編碼判斷出現分歧，則回到逐字稿脈絡檢視發言內容，並透過討論形成共識。考量研究結果中已呈現完整引文，表1僅列出部分次主題所對應之代表性逐字稿摘錄，完整語句與脈絡則進一步呈現於研究結果中。

表1

訪談資料主題分析與編碼架構對照範例

主題	次主題	初步編碼概念	逐字稿文本摘錄
回沖效應	教材處理：核心概念重組與課本範圍取捨	課綱、學習內容、核心概念、教材重組、命題依據	……會先了解課本單元在課綱中對應的學習內容與核心概念，再對照教科書，找出需要特別強調的部分，因為會考是依學習內容出題，不是依課本出題。（第一場次_D老師）
回沖效應	教學活動：跨文本能力訓練與進度壓力下的講述安排	會考趨勢、文本圖表判讀、訊息擷取、教學進度	現在會考趨勢出現很多圖表判讀與長篇閱讀，……會讓學生練習從文本中抓關鍵字。……教學進度就會完全趕不上，這是一個很大的兩難。（第三場次_A老師）
回沖效應	評量設計：情境化題型與基本分數配置	情境化題目、題目難度、表現分數、作答動機	……會考趨勢偏向生活化，……會比較傾向讓學生練習情境化題目。但如果題目太靈活，學生會考得很差，還是要給他們一些基本分數，否則他們會完全放棄作答。（第四場次_A老師）

四、研究信實度

詮釋性研究重視參與者經驗與意義建構，研究結果並非追求普遍性推論，而是在特定情境脈絡中形成對現象的理解。因此，本研究在參與者選擇、訪談題目設計與資料分析上，均以貼近教師實際經驗為原則，以提升研究信實度。參與者選擇已於前文說明，以下將著重說明訪談題目設計、資料蒐集歷程與資料分析程序。

在訪談題目設計方面，本研究依據文獻探討發展訪談架構，並透過前導性焦點團體訪談進行檢視與修訂，使題目貼近國中教育現場情境與教師實際經驗。正式訪談題目聚焦於教學與評量調整、決策考量因素，

以及會考與段考的回沖效應，引導教師說明具體作法與判斷依據，使資料能呈現實際教學情境，而不僅停留於概念性陳述。

在資料蒐集歷程中，本研究採焦點團體訪談，透過參與者之間的互動與討論，引出多元觀點。訪談過程中，研究者依據參與者的回應適時進行追問、釐清與歸納，並將初步整理之觀點回饋參與者確認。此一檢核主要針對研究者對發言內容的理解、共同觀點的歸納，以及不同意見之辨識，請參與者確認是否符合其原意，或進一步補充與修正。此項檢核範圍僅限於訪談當場形成之觀點摘要與語意理解，並非完整逐字稿或最終研究結果之審閱。

在資料分析方面，研究者完成個別場次的初步分析後，即進行五場訪談資料的交叉比對，以整合分析結果。訪談結束後，研究者回到逐字稿進行反覆閱讀與檢視主題，並透過研究團隊討論修正初步主題與詮釋。跨場次分析時，對於多數場次中反覆出現的觀點，確認其為共通主題；對於僅出現在部分場次的觀點，則回溯其發言脈絡與情境條件，以釐清其特殊性。同時，研究者亦檢視與主要主題不一致或未形成共識之觀點，作為例外案例加以分析，避免僅呈現支持主要詮釋的資料。此外，研究者在整體分析過程中持續修正與整合主題，並保留未形成一致意見之多元觀點，以呈現教師教學決策的複雜性。

綜合而言，本研究透過貼近實務的訪談設計、互動式焦點團體訪談的資料生成，以及跨場次比較、研究團隊討論、例外案例檢視與持續修正的分析程序，使研究結果得以呈現具情境脈絡的教師經驗，進而提升研究之信實度。

肆、研究結果與討論

一、雙重評量導向下的回沖效應轉化：教材、教學活動與評量設計的調整

根據焦點團體訪談資料的主題分析與跨場次比較結果可發現，教師在會考與段考兩種評量導向下皆會進行教學調整，但其方向與限制有所差異。會考導向較強調素養能力的整合發展，段考導向則較聚焦於教學

進度、基礎能力與學生當前學習表現的即時回應。以下將從教材處理、教學活動與評量設計三個面向加以說明。

（一）教材處理：核心概念重組與課本範圍取捨

在會考導向下，教師普遍強調核心概念的掌握，並依課綱重新解構教科書內容；相對地，在段考導向下，教材處理則較聚焦於課本範圍的熟練與基礎能力的鞏固。然而，兩者皆出現教材取捨的現象：在會考情境中，教師傾向針對常考概念加強提問並融入時事，使教材朝向廣度與整合發展；而在段考情境中，教師則受到固定範圍、共同進度與學生程度之影響，須先確保學生掌握基本內容，再逐步連結情境化或跨文本題型所須的教材理解。

> 因應會考，我通常在備課時，會先了解課本單元在課綱中對應的學習內容與核心概念，再對照教科書，找出需要特別強調的部分，因為會考是依學習內容出題，不是依課本出題。（第一場次_D老師）

> 我會針對會考較強調的考點加強提問，並融入時事，以提升學生的閱讀層次與教材廣度。因為現在會考很多題目都是素養導向，敘述很長，閱讀能力變得非常重要，我會帶學生讀新聞，增加他們對抽象概念的理解。（第五場次_A老師）

> 為了讓學生段考考得好，我們學校位處偏鄉，學生程度較弱，所以我主要還是以課本內容為主，讓學生有讀就有分，建立學習信心。（第四場次_A老師）

此結果顯示，教師在不同評量導向下對教材處理的判斷基準並不相同：會考導向較重視概念整合與應用情境的延伸，段考導向則較著重課程範圍的掌握與學生可達成的學習表現。進一步而言，教材取捨不只是「要教什麼」的內容選擇，而是教師面對會考情境閱讀與學生能力落差時，為維持學生可學習性所採取的策略性作為。

（二）教學活動：跨文本能力訓練與進度壓力下的講述安排

在教學活動方面，會考導向促使教師由講述轉向素養導向教學，例如培養閱讀理解、關鍵字擷取與圖表判讀等能力，並嘗試連結生活情境。然而，此一轉向亦伴隨「解題技巧化」現象，部分教學活動轉為應試策略訓練。相對地，段考導向下的教學活動較具即時回應特性，教師會依據學生表現調整教學步調與補強弱點，展現形成性評量的功能；但同時也受授課節數與進度壓力限制，理想的探究或討論活動往往被壓縮，而回歸以完成進度為主的講述教學。

> 現在會考趨勢出現很多圖表判讀與長篇閱讀，所以我在教學時會讓學生練習從文本中抓關鍵字。但這很花時間，如果每一課都這樣帶，教學進度就會完全趕不上，這是一個很大的兩難。（第三場次_A老師）

> 時間是一個很大的限制。雖然我們想嘗試分組討論或探究，但如果段考進度趕不完，就必須捨棄那些活動，回到最基本的講述，先把課文內容教完。……如果進度沒教完，學生在段考時就會遇到很大的挫折。（第一場次_C老師）

此結果顯示，教師在不同評量導向下對教學活動的安排具有不同取向：會考較引導教學朝向素養能力培養，但也可能使教學策略化；段考則較強調即時補強與進度完成，使教學更容易回到講述與內容覆蓋。換言之，教學活動的壓縮與轉向，反映教師在有限課時與評量壓力下，將素養導向要求轉化為可操作、可完成的教學形式之應變樣貌。

（三）評量設計：情境化題型與基本分數配置

在評量設計方面，會考導向促使教師減少記憶性題目，增加情境化與整合性題型，以提升評量層次；但其難度與閱讀負荷也可能對低成就學生造成壓力，因此教師須透過較簡化的題目維持其學習信心。相對地，段考則呈現「信心建立與鑑別需求並存」的特性，教師常以基本題

確保學生獲得基本分數，同時保留部分較高難度題目，以維持測驗的區辨力。

> 這幾年會考趨勢偏向生活化，所以我在教學和段考命題時，會比較傾向讓學生練習情境化題目。但如果題目太靈活，學生會考得很差，還是要給他們一些基本分數，否則他們會完全放棄作答。（第四場次_A老師）

> 我們也不希望外界覺得段考題目永遠都很簡單，所以這幾年命題時都會固定配置一部分趨近會考題型的題目，讓程度較好的學生也有挑戰的空間。（第一場次_D老師）

上述顯示，教師在評量設計上必須同時兼顧能力發展與學習支持：一方面依循會考趨勢提升題目情境化與整合性，另一方面保留一定比例的基礎題目，以回應學生程度差異並維持其作答動機。因此，基礎題與素養題的配置，並非僅是命題比例問題，而是教師在能力導向、分數表現與學生信心之間所形成的風險管理策略。

（四）討論

綜合教材處理、教學活動與評量設計三個面向之分析結果，可見會考與段考在回沖效應上呈現不同功能與樣態。為進一步呈現不同評量導向下的核心特徵，分析結果整合如表2：

表2

會考與段考評量導向下教師教學實踐回沖效應之比較

分析面向	會考導向	段考導向	共通特徵
教材處理	強調核心概念與時事連結，依考題趨勢重組教材內容	聚焦課本範圍與基礎能力，依進度與學生程度調整內容	皆出現教材取捨現象
教學活動	強化閱讀理解、圖表判讀與素養導向教學	依學生表現進行補強與調整，強調即時回應	受時間限制，理想教學活動易被壓縮
評量設計	提升題目情境化與整合性，強調能力導向	兼顧基本分數與區辨功能，維持學生信心	皆朝素養導向發展，但受現實條件限制
正向回沖	引導教學朝向素養能力培養與教材整合	促進形成性回饋與教學調整	皆有助於閱讀理解與圖表判讀能力的發展
負向回沖	教學易轉向解題技巧化，部分內容遭取捨	教學易受進度與制度限制而回歸講述	皆受時間不足與學生差異影響而形成教學限制
整體功能	提供素養導向教學的方向與長期規劃依據	回應短期學習須求並調節學生信心	共同形塑教師教學決策

表2顯示，教師在會考與段考情境中的教材處理、教學活動與評量設計，雖同樣受評量導向影響，但其正、負向回沖的表現並不相同。此一發現支持既有回沖效應理論對正、負向回沖的區分，也呼應測驗會引導教師調整教材、教學與評量設計的觀點（Cheng, 2005; Green, 2013）。不過，本研究進一步顯示，在會考與段考並行的情境中，回沖效應並非僅呈現正向或負向影響，而是反映於教材處理、教學活動與評量設計的不同轉化樣態。

在教材處理與教學活動方面，會考導向促使教師更加強調核心概念的掌握與教材內容的重組，並強化閱讀理解、圖表判讀與生活情境之連結。此一結果呼應Cheng（2005）和Green（2013）所指出高風險測驗對教學的導向作用，亦支持Pan（2009）所稱的系統性正向回沖。然而，教學實務中亦出現解題技巧化與內容取捨，依Brown（2004）觀點，這

屬於負向回沖，顯示素養導向在時間與考試壓力下仍可能轉化為應試策略。

相對地，段考在教學活動與評量設計上具有調節功能。雖然Brookhart（2010）指出校內評量常偏向知識再現，但本研究發現，教師會透過基礎題型維持學生基本分數與學習信心，並保留部分較高難度題目以維持區辨功能，顯示段考同時回應學生差異與須求，亦呼應Mizutani（2009）所強調教師信念在回沖實踐中的作用。

綜合而言，會考較提供素養導向教學的方向與長期規劃依據，但也伴隨解題技巧化與內容取捨；段考則較有助於即時補強、維持學生信心並回應差異，卻也容易產生在進度與制度壓力下回歸講述教學之現象。此一結果呼應Priestley等人（2015）以及Datnow與Park（2019）所指出教師決策與情境脈絡的密切關聯，亦顯示教師教學實踐並非僅由評量形式決定，而是受到評量導向、學生需求與情境條件交互影響。換句話說，會考與段考的回沖效應並非單純並列，而是呈現局部強化與局部削弱並存的互動關係。段考若納入情境化與跨文本能力要求，可延續會考的素養導向；若受進度、基本分數與學生信心維持制約，則可能削弱能力導向教學。從教學實踐來看，解題技巧化、內容取捨、講述教學與精熟操練，反映教師在有限時間、固定進度與學生能力落差中，將評量要求轉化為可操作教學形式的過程，並具體呈現在教材取捨、活動安排與題型配置之中。

二、會考與段考評量要求下教師教學調整之專業判斷與策略折衷

本節探討教師在會考與段考評量情境下進行教學調整時的主要決策考量及其差異。結果顯示，兩者皆受多重因素影響，但會考導向較偏重長期能力發展與績效考量，段考導向則較著重短期學習表現與學生信心的維持。

（一）在能力發展與學習信心之間調整教學難度

在教學調整過程中，教師普遍重視學生的學習表現與信心狀態。當學生在評量中表現不佳時，教師往往會調整教學難度或評量設計，以避免挫折感降低學習動機。在會考情境下，教師除掌握命題趨勢外，更關注如何協助低成就學生維持基本能力；在段考情境中，則較傾向透過調整題目難度與教學節奏，讓學生在短期內獲得成就感。

> 教學設計還是要從學生本身出發，看起點行為、先備知識到哪裡。程度不一樣，教學設計就會不一樣。我本身蠻常教到體育班學生，起點行為不同，教學設計就完全不同。（第五場次_E老師）

> 學生程度是我處理教材與教學的主要因素。若孩子跟不上教材，只是一直講課本其實沒有意義，教學必須讓孩子有回應，然後能夠藉此來增進他學習動力與信心，所以在段考上，我們會設計一些他們比較能理解的題目。（第一場次_C老師）

此結果顯示，無論面對會考或段考，教師皆將低成就學生的表現視為重要決策依據；差異在於，會考較著重在素養導向能力發展中兼顧低成就學生，段考則以維持其短期學習信心與成就感為出發點。

（二）在跨文本能力培養與單元內容掌握之間安排教學重點

教師在解讀評量結果時，也會關注學生是否具備閱讀理解、圖表判讀與關鍵訊息擷取等素養能力。當評量結果顯示學生在上述能力上有所不足時，教師會據此調整教學內容與活動設計，加強相關訓練。此一趨勢在會考導向下尤為明顯，教師較傾向從跨單元整合角度進行教學規劃；而在段考情境中，教師雖然同樣重視能力培養，但較常結合特定單元內容，呈現較細緻的教學調整。

> 以前我的教學可能就是從第一章、第二章、第三章這樣講下去，但現在因為會考趨勢，我會比較強調跨單元的整合。我會把相關概念

拉出來一起講，讓學生看到知識的關聯性，而不是只背課本裡的片段知識。（第二場次_C老師）

如果我只講大概脈絡，學生段考分數可能很差，所以我必須從單元內容講得很細。雖然這些細節對會考未必有幫助，但為了段考，還是要督促學生掌握。（第三場次_B老師）

教師在能力發展上採取兩種不同取向：面對會考時，較強調跨單元整合與能力遷移；面對段考時，則較重視單元內容的細節掌握，以協助學生回應教科書範圍內的評量要求。

（三）在長期能力培養與短期進度要求之間調整教學節奏

時間與進度是影響教師決策最普遍的因素。教師進行教學調整時，必須在深化能力培養與完成教學進度之間權衡。會考導向下的時間壓力具有長期性，教師較可能進行整體教學規劃，例如重新組織教材邏輯並安排素養題型練習；相對地，在段考情境中，時間壓力較偏向短期且即時，教師多採取重點補強與講述策略，以在有限課堂時間內完成評量範圍。

因為會考範圍很大，我會重新安排課本的教法，把相關單元挪動位置一起講，省下一些零碎時間，讓後面有更完整的時間做總複習。這種調整要從整學期甚至整學年的角度去看，才能在進度與深度之間找到空間。（第二場次_C老師）

段考最怕遇到連續假期或學校活動，進度就會趕不完。這時候我就必須縮短討論時間，趕快把該教的內容講完，或在考前最後一堂課快速補強重點。（第四場次_D老師）

此結果顯示，不同評量時程與功能定位，形塑了教師不同的時間運用邏輯：會考較引導教師進行跨單元與長期規劃，段考則使教師傾向短

期補強與進度導向的講述教學。

（四）在制度績效與家長分數期待之間調整教學與命題取向

教師的教學決策並非全然基於專業信念，亦深受行政環境與外部期待影響，包括教育政策、學校評量制度與家長期待等。在會考導向下，教師須回應政策績效指標，如提升減C比例或達成學力目標；在段考情境中，則須直接面對家長對分數變動與校內排名的關注。這些外部因素使教師在進行教學調整時，不僅要將學生學習成效列入考量，亦須在制度規範與環境壓力之間取得平衡。

> 會考是大家關注的指標，學校行政端也會分析會考成績。這讓我在教學上必須對應會考趨勢，如果學生跟不上，就會感到壓力，必須思考如何把會考所需的素養能力放進平時教學設計裡。（第四場次_A老師）

> 在我們這種學校，家長對分數變動非常敏感。如果段考出題方向變動太大，或平均分數掉太多，家長就會有意見。這限制了我們在段考中嘗試太創新的評量方式，命題時還是要顧慮家長的接受度。（第三場次_B老師）

整體而言，不同來源的外部壓力形塑了教師不同的決策邏輯：在會考情境中，教師較傾向將政策要求轉化為長期教學目標；在段考情境中，則較著重回應家長與校內單位對分數表現的即時期待，因而在命題與教學安排上採取較穩定以及可預期的做法。

（五）討論

綜合上述四項決策考量因素，本研究將教師在不同評量導向下的決策邏輯整理如表3：

表3

會考與段考評量導向下教師教學調整決策考量因素之比較

決策考量因素	會考導向	段考導向	共通特徵
學生學習信心	以提升整體能力為目標，並兼顧低成就學生	以維持學習信心與避免挫折為優先，據以調整難度	皆重視學生表現對教學調整的影響
跨文本閱讀能力發展	強調跨單元整合，聚焦閱讀理解與圖表判讀等能力	著重單元內容熟練與細節掌握，能力培養結合單元內容	皆重視跨文本閱讀相關能力
時間壓力與課程進度	偏向長期課程規劃，仍受整體進度限制	偏向短期補強，在有限時間內完成評量範圍	皆須在深度與進度間取捨
行政壓力與外部期待	受政策績效要求影響，須回應整體成績表現	受家長分數期待影響，須兼顧成績表現與教學評價	皆受制度與外部環境影響

表3顯示，教師因應會考與段考進行教學調整時，其專業判斷常受到學生表現、能力發展、時間進度與外部期待之交互影響。此一結果呼應Xu與Brown（2016）指出，教師評量運用深受學生表現、評量素養與情境條件影響，也呼應Mizutani（2009）關於教師信念在回沖實踐中具有調節作用的觀點。特別是學生學習信心與跨文本能力發展，已成為教師判斷評量要求能否轉化為學生可承受學習任務的重要依據。

進一步而言，教師在會考與段考之間形成的策略折衷，回應Green（2013）所強調的觀點：回沖效應並非由測驗直接決定，而是經由教師專業中介決策轉化為具體的教學實踐。會考導向使教師關注跨單元整合與能力遷移，段考導向則使教師回應單元內容、進度壓力與分數期待；此一差異亦與Brookhart（2010）及Gulikers等人（2018）所指出校內評量在發展高層次任務時所面臨的困難相呼應。

整體而言，本研究顯示，教師在會考與段考並行的評量情境下，並非只依單一評量結果直接調整教學，而是須在素養導向、成績壓力與學

生信心之間持續進行判斷與折衷。會考導向使教師較關注跨單元整合與能力遷移，段考導向則使教師更重視單元內容、進度壓力與分數期待，因而形成兼顧能力發展、學生信心與進度要求的策略性調整。

三、教師策略折衷中的情境限制：時間、學生差異與外部期待

本節聚焦於教師依據會考與段考結果進行教學與評量調整時所面臨的實踐限制。分析結果顯示，兩種評量情境下的限制來源大致一致，主要包括時間壓力、學生差異、行政要求與外部期待。這些因素影響教學調整的可行性，也使教師必須在理想教學與現實條件之間持續取捨。

（一）素養導向教學在課程進度壓力下的實踐限制

在教學實踐中，教師須在落實素養導向教學與完成課程進度之間取捨，且多半優先確保進度，使需高度師生互動的討論與探究活動受到壓縮。其原因在於，素養能力的培養需要較長時間累積，但課程時數有限。

> 會考強調的是長期能力，但在短短課堂時間內，又要顧及課程內容，又要帶學生做深度思考訓練，真的很困難。當我們想多帶一些素養導向的題目或思考時，會發現學生吸收需要時間，但進度卻一直逼著我們走，很難在有限時間裡把核心能力建構完整。（第一場次_E老師）

> 段考範圍固定、節數也固定，一旦學生吸收比較慢，我們就必須在教學方法上做犧牲。為了在段考前把考點講完，有時只能回到最快的講述法，把重點劃一劃。因為如果進度沒上完，對學生在段考的公平性上會出問題。（第三場次_C老師）

此結果顯示，在時間有限與進度壓力下，教師往往優先確保內容覆

蓋，而非持續深化學習歷程，使強調探究與歷程的素養導向教學較難充分展開，並較易轉向重點整理與講述式教學。

（二）學生能力差異的學習限制

學生能力的異質性，特別是閱讀理解能力的落差，使教師在推動素養導向教學時面臨深層挑戰。教師指出，會考題型所需的長文理解與跨文本資訊整合能力，對部分學力待加強的學生而言門檻過高。因此，在教學實踐中，教師須持續進行「降認知層次處理」，在維持學生可理解程度與符合評量要求之間取得平衡。無論在會考或段考情境中，教師皆須回應學生能力落差，確保不同程度的學生都能參與學習。

> 現在會考題目閱讀量很大，如果我直接拿很靈活的素養題給學生做，程度落後的孩子連題目都讀不完，甚至會直接放棄。我必須把教學層次拉低，先從最簡單的訊息抓取開始教。這種「降階」雖然會慢一點，但如果不做，素養題目對他們來說就只是天書。（第四場次_B老師）

> 段考我們還是會放比較多的基礎知識題，因為如果全部都考素養，後段孩子的分數會非常難看。所以平時上課會花很多時間在課本細節的解釋和基礎題目的練習，這種針對單元細節的精熟練習，對他們來說是比較抓得住的目標。（第二場次_A老師）

此結果顯示，學生能力差異對教學調整的影響並不因評量情境而有明顯不同。面對學習落差，教師的決策重點往往不在教學形式本身，而在於如何維持學習的可及性，並透過調整教學內容與學習門檻，使多數學生得以進入學習歷程。

（三）行政與政策的制度限制

教師的教學實踐並非全然自主，而是深受行政與政策要求制約。在會考導向下，教師須回應教育主管機關的績效檢核，例如各校「減C」

比例的學力監控；在段考情境中，雖未直接連結全國性績效指標，仍受校內制度規範與行政運作影響。在本研究中，不同評量情境下的行政與制度性限制並未呈現明顯差異。制度性安排，如統一的課後輔導進度或時間使用規範，普遍壓縮教師進行個別化補救教學或深化學習的專業空間，使教學調整受限於外在規範與績效要求。

> 現在行政端對會考成績的數據非常敏感，特別是考後的檢討會，我們會被要求分析為什麼某個級別的比例會變動。這種績效壓力會回推到教學實務上，讓你平時就不敢太放鬆，必須一直盯著進度與練習卷的表現。（第一場次_B老師）

> 有時候我們想針對低成就學生多留一點時間補救，但學校整體作息或課後輔導規定是統一的，很難有額外節數或空間做更彈性的調整。甚至也因為輪流命題，連段考命題的配分比例，行政端有時也會有指導意見。（第四場次_A老師）

上述資料分析結果顯示，行政與政策壓力不僅是教師決策時的考量因素，也進一步重塑教學調整的邏輯。在績效指標與制度規範的框架下，教師往往優先考量可被檢核與回應的教學結果，使原本較具彈性的教學選擇，轉向較為保守且規範導向的安排。進一步來看，段考進度、作息節數與規範等校內安排，會成為教師教學調整時所面對的情境限制，使教學較容易轉向完成既定進度、強化課本內容講述與採取較可預期的命題方式。

（四）家長期待形塑段考對教師教學調整的限制較會考顯著

在教學實踐中，家長對學生成績的關注主要透過段考分數與校內排名具體呈現，因而對教師的教學調整與命題決策形成較直接的限制。相較之下，會考雖具重要性，其結果多作為整體學力表現的參考，家長較少將其轉化為對教師的即時要求與壓力。因此，教師在日常教學與評量設計中，實際感受到的家長期待，主要集中於段考情境。

家長認為會考是整體結果，已經來不及改變，反而非常在意段考分數。如果我們嘗試較新的素養題，導致分數掉下來，家長就會有很大反彈。這讓老師在命題時不敢做太大變動，甚至在教學時會被迫多教解題技巧，確保學生分數不要太難看。這其實是一種無形的專業限制。（第三場次_E老師）

在學校裡，如果你為了自己班上的學生程度做了教學調整或簡化，但命題者不是你，學生考出來就會很慘。為了保護學生，你只能跟著大部隊的教學與命題方向走，這種集體性的競爭機制，縮小了個別老師依學生需求調整的空間。（第二場次_D老師）

此結果顯示，教師知覺到家長期待所形成的外部壓力，因而在教學設計與命題上更傾向採取穩健且可預期的作法，以避免分數過低可能引發的負面反應。換句話說，教學決策不僅受學生學習需求影響，也受到輪流命題與家長評價機制的限制，使段考情境下的教學調整更具風險控管取向。

（五）討論

本研究結果顯示，教師在教學與評量調整過程中所面臨的限制，主要被描述為與素養測驗特性、學生能力、行政壓力以及家長期待等結構性條件的交互作用有關。在會考與段考兩種評量情境中的限制大致相似，顯示教師面對的困難不完全來自評量形式本身，也來自教學現場的時間結構、學生學習條件與制度安排。

此一結果呼應Cheng（2005）和Green（2013）所指出高風險評量會引導教師關注能力導向教學，但也進一步顯示，若課程時間、學生基礎能力與校內制度未能提供相應支持，評量所指向的能力要求便難以在課堂中充分展開。此亦與Datnow與Park（2019）所指出制度性框架會影響教師實踐的觀點相符，顯示在教師經驗中，素養導向教學的實踐限制不因評量類型不同而有明顯差異。

不過，本研究也發現，家長期待展現出較明顯的情境差異。家長對

學生成績的關注多透過段考分數與校內排名具體呈現，因此，在教師的經驗中，較容易對教學與命題決策形成直接壓力；相較之下，會考雖具重要性，其壓力較少被教師描述為日常教學中的即時要求。此一差異補充了既有回沖效應研究較少觸及的社會評價面向，顯示段考的實踐限制不僅來自評量內容，也與分數、排名及家長評價在校內情境中的使用方式有關。

四、綜合討論

綜合研究結果可知，從教師經驗來看，教學實踐並非由單一評量直接驅動，而是在不同評量導向、教師專業判斷與情境限制交互作用下形成的動態歷程。本研究關注的不僅是不同評量如何產生回沖效應，更在於教師如何在兩種評量要求之間進行判斷、取捨與轉化。本研究結果進一步顯示，會考與段考的回沖效應並非單純相乘或相互抵消，而是呈現局部強化與局部削弱並存的樣態。段考若納入情境化與跨文本能力要求，可延續會考的素養導向；若受進度、基本分數與家長期待所制約，則可能削弱能力導向教學。

本研究的三個研究問題之間具有連動關係。會考與段考分別呈現不同的回沖效應，反映於教材處理、教學活動與評量設計；然而，這些調整並非評量特性的直接結果，而是教師綜合考量學生學習表現、課程進度、評量要求與外部壓力後所形成的教學決策。同時，這些決策又持續受到時間限制、學生能力差異、行政制度與家長期待等情境因素影響。換言之，回沖效應呈現的是教學方向，決策考量說明的是教師如何取捨，情境限制則界定了調整得以實現的條件。

本研究亦顯示，教師在理念上多能理解並傾向採用能力導向教學或素養導向命題，但在會考與段考兩種評量脈絡下，仍須面對時間不足、學生能力差異與制度規範等共同限制；其中，家長期待又主要透過段考分數與校內排名，形成較直接的教學與命題壓力。此一結果說明，教師的命題調整並非理念與實踐之間的單純落差，而是素養導向評量要求進入校內段考情境後所形成的專業取捨。進一步而言，教師在段考命題中未必排斥素養導向題目，而是須同時面對學生程度落差、輪流命題、集

體競爭與家長評價壓力的限制，因而審慎調整素養題型的比例與難度。因此，基礎題與素養題的配置，實為教師回應多重壓力的風險管理，而非單純的命題比例問題。

在研究貢獻上，本研究補充既有回沖效應研究較少處理的雙重評量情境。研究結果顯示，教師面對會考與段考並行時，並非被動地接受評量要求，而是透過對學生學習、課程進度與外部條件的綜合評估，將評量要求轉化為具體教學決策，並支持重新理解「評量影響教學」的線性觀點，指出回沖效應更適合被理解為嵌入情境脈絡的動態歷程。

綜合而言，會考與段考對教學的影響並非彼此分離，而是在同一教學現場中交錯作用，並透過教師決策判斷與情境回應，轉化為具脈絡性的教學實踐。此一發現不僅回應了本研究對評量如何影響教學實踐的關注，也為理解當前評量制度下教師教學行動的複雜性，提供了較具整合性的分析視角。

伍、研究結論與建議

一、研究結論

本研究探討教師在會考與段考並行的評量情境下，如何進行教學與評量調整。研究結果顯示，教師在兩種評量導向下皆會進行教學與評量調整，但其回沖效應、決策考量與實踐限制各有側重。整體而言，會考較能發揮能力導向與跨單元整合的引導作用，段考則較著重分數表現、進度掌握與學生學習信心的維持；教師在兩者之間，須依據學生表現、課程進度與外部要求持續進行專業判斷與調整。另一方面，時間壓力、學生能力差異與制度規範等結構性條件，在不同評量情境中大致相同，皆為構成教學實踐與調整的主要限制來源；相較之下，家長期待則透過段考分數與排名機制對教師產生較為直接的影響，形成相對顯著的外部壓力。

基於上述發現，本研究回應了評量不一致如何影響教學實踐的研究關注，指出教學調整並非單純反映評量差異，而是在共同結構條件下，經由教師專業判斷所進行的脈絡化轉化。此一結果顯示，理解當前教學

實踐，須超越單一評量導向的分析，轉而關注不同評量要求、教師決策與情境限制之間的交互作用，方能更貼近教學現場的實際運作。

二、研究建議

本研究發現，教師在會考與段考雙重情境下的教學調整，主要受時間、學生能力與制度條件等結構性因素制約，且多數限制在不同評量情境中呈現一致性；僅家長期待透過段考分數與排名機制展現較直接的關聯。未來研究可進一步探討家長是否理解會考題目與素養導向教學之間的關聯，也可檢視教師依據會考表現進行的教學調整，如何與學生理解及能力發展產生關聯。

在實務與政策建議方面，本研究顯示，教師教學調整所面臨的限制，多與時間配置、制度規範及學生能力差異等結構性條件有關。因此，教育行政機關與學校若欲推動素養導向教學，不宜僅著眼於命題改革，亦應同步檢視課程節奏、評量制度與補救教學條件。在教師教學實務方面，研究結果顯示，教師面對會考與段考並行時，關鍵在於能力發展、基礎鞏固、學生信心與進度壓力之間的取捨。因此，教師可檢視教材處理是否兼顧核心概念與課本範圍，教學活動是否將跨文本閱讀、圖表判讀與情境分析轉化為學生可進入的學習步驟，段考命題是否能在基礎題與情境化題目之間形成銜接，以支持學生逐步發展會考所需能力。

三、研究限制

在研究限制方面，本研究顯示，教師知覺會考與段考的回沖效應與其教學實踐及調整密切相關，且其中涉及多重考量因素。然而，教師對素養導向命題、評量結果解讀與教學調整的理解，仍可能因個人專業經驗而有所差異。因此，本研究結果應理解為特定參與教師在雙重評量情境下的經驗詮釋，不宜直接推論至所有國中教師。

另外，本研究雖以目的性取樣納入不同會考科目、具段考命題經驗，且曾經歷正向與負向回沖經驗之教師，以增加資料差異性，但訪談資料中亦出現教師提及「學校位處偏鄉，學生程度較弱」等情境經驗。

惟本研究並非以偏鄉學校、學校規模或行政負擔為主要分析類別，因此研究結果主要呈現參與教師在雙重評量情境下的教學調整經驗，尚不宜直接推論為不同學校脈絡下教師的普遍現況。

謝誌

本文為國科會專題研究計畫「國中教師在大型學習成就測驗回沖效應下教學與評量調整對學生表現提升之研究」之部分研究成果，計畫編號為 NSTC 114-2410-H-018-036-SS3，承蒙國科會經費補助，謹此致謝。

參考文獻

王竹梅、丁一顧（2021）。國中小素養導向評量實施之困境與反思。**臺灣教育評論月刊，10**(3)，1-7。

【Wang, Z.-M., & Ting, Y.-K. (2021). Challenges and reflections on the implementation of competency-based assessment in elementary and junior high schools. *Taiwan Education Review Monthly, 10*(3), 1-7.】

李偉俊（2023）。108課綱素養導向評量的實施成效與問題。**臺灣教育評論月刊，12**(3)，4-10。

【Li, W.-C. (2023). Implementation effects and problems of 108 curriculum competency-based assessment. *Taiwan Education Review Monthly, 12*(3), 4-10.】

國立臺灣師範大學心理與教育測驗研究發展中心（2024）。**113年國中教育會考各科能力等級加標示人數百分比統計表**。https://cap.rcpet.edu.tw/exam/113/113Table2.pdf

【Research Center for Psychological and Educational Testing, National Taiwan Normal University. (2024). *Percentage statistics of students reaching each proficiency level in each subject in the 2024 Comprehensive Assessment Program for Junior High School Students*. https://cap.rcpet.edu.tw/exam/113/113Table2.pdf】

張凱翔、莊舒婷、陳亞華（2016）。科學素養評量融入國中地球科學教學評量對國中教育會考自然科成績之影響。**科學教育月刊，390**，14-33。

【Chang, K.-H., Chuang, S.-T., & Chen, Y.-H. (2016). Incorporating scientific literacy assessment in teaching of earth science curriculum and its influence on Comprehensive Assessment Program for Junior High School. *Science Education Monthly, 390,* 14-33.】

張銘秋、黃璨瑩、陳佳蓉、陳柏熹、曾芬蘭（2022）。國中教育會考數學科的回沖效應初探。**教育科學研究期刊，67**(1)，227-254。

【Chang, M.-C., Huang, L.-Y., Chen, C.-J., Chen, P.-H., & Tseng, F.-L. (2022). Washback effect from changes to Comprehensive Assessment Program math test for Junior High School Students. *Journal of Research in Education Sciences, 67*(1), 227-254.】

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初探。**教育科學研究期刊，64**(2)，219-252。

【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.】

顏郁玲、陸偉明（2023）。學校教師因素對國三生數學成就之跨層次影響：考量參與數學補習的國中生個人因素。**清華教育學報，40**(1)，1-43。https://doi.org/10.6869/THJER.202306_40(1).0001

【Yen, Y.-L., & Luh, W.-M. (2023). Moderation effect of school teachers' teaching on mathematics achievement: Controlling for ninth graders' factors in a multi-level analysis of math cram schooling. *Tsing Hua Journal of Educational Research, 40*(1), 1-43. https://doi.org/10.6869/THJER.202306_40(1).0001】

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129.

Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeted washback—A case study. *System, 30*(2), 207-223.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*(5), 258-267.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford University Press.

Baird, J.-A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, Policy & Practice, 24*(3), 317-350.

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching, 21*(6), 624-640.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.

Boaler, J. (2016). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching*. Jossey-Bass.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD

Brown, D. H. (2004). *Language assessment principles and classroom practices*. Pearson Education.

Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.

Datnow, A., & Park, V. (2019). *Professional collaboration with purpose: Teacher learning towards equitable and excellent schools*. Routledge.

Green, A. (2013). Washback in language assessment. *International Journal of English Studies, 13*(2), 39-51.

Gulikers, J. T. M., Runhaar, P., & Mulder, M. (2018). An assessment innovation as flywheel for changing teaching and learning. *Journal of Vocational Education & Training, 70*(2), 212-231.

Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational*

Research Review, 17, 50-62.

Heritage, M. (2013). *Formative assessment in practice: A process of inquiry and action*. Harvard Education Press.

Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy & Practice, 25*(5), 442-467.

Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in curriculum* (pp. 83-121). University of Chicago Press.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*(3), 241-256.

Mizutani, S. (2009). *The mechanism of washback on teaching and learning* (Unpublished doctoral dissertation). University of Auckland, Auckland. Google Scholar.

Noble, A. J., & Smith, M. L. (1994). *Old and new beliefs about measurement-driven reform: The more things change, the more they stay the same* (CSE Report 373). University of California.

Nurjati, N., & Rahayu, E. M. (2023). Balancing scale of washback effect of English language assessment in junior high school level: A literature review. *Journal of Applied Linguistics and English Education, 1*(2), 1-9.

Organisation for Economic Co-operation and Development. (2019). *PISA 2018 results (Volume I): What students know and can do*. OECD Publishing.

Organisation for Economic Co-operation and Development. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education, PISA*. OECD Publishing.

Pan, Y. C. (2009). Review of washback and its pedagogical implications. *VNU Journal of Science, Foreign Languages, 25,* 257-263.

Pastore, S. P. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education, 8,* Article 1217167.

Pearson, I. (1988). Tests as levers for change. In D. Chamberlain & R. J. Baumgardner (Eds.), *ESP in the classroom: Practice and evaluation* (pp. 98-107). Modern English.

Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury Academic.

Rahman, K. A., Seraj, P. M. I., Hasan, M. K., Namaziandost, E., & Tilwani, S. A. (2021). Washback of assessment on English teaching-learning practice at secondary schools. *Language Testing in Asia, 11,* article 12.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*(4), 759-775.

Suurtamm, C., Thompson, D. R., Kim, R.-Y., Moreno, L. D., Sayac, N., Schukajlow, S., Silver, E., Ufer, S., & Vos, P. (2016). *Assessment in mathematics education: Large-scale assessment and classroom assessment*. Springer.

Taylor, L. (2005). Washback and impact. *English Language Teaching Journal, 59,* 154-155.

Tirtaningrum, M., & Agustinus N. (2015). Washback effect of the classroom test in junior high school. *Magister Scientiae, 37,* 41-52.

Toksöz, S., & Kılıçkaya, F. (2017). Review of journal articles on washback in language testing in Turkey (2010-2017). *Lublin Studies in Modern Languages and Literature, 41*(2), 184-204.

Wall, D. (1997). Impact and washback in language testing. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (pp. 291-302). Kluwer Academic.

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education, 58,* 149-162.

Yen, P. H., Binh, T. T. A., & Huong, T. T. T. (2024). Exploring the implementation and perception of competency-based assessment practices among Vietnamese EFL instructors. *Language Testing in Asia, 14,* Article 26.

Washback Effects, Decision-Making Considerations, and Practical Constraints in Junior High School Teachers' Instructional Practices Under the Assessment Orientations of the Comprehensive Assessment Program and School-Based Periodic Examinations

Shih-Hsiung Liu*

Abstract

Purpose

Previous studies on washback have predominantly examined the influence of large-scale examinations on instructional orientations, while giving comparatively limited attention to how teachers interpret and respond to discrepancies among assessment systems in everyday instructional contexts. Although recent local studies have addressed the Comprehensive Assessment Program (CAP) for Junior High School Students and the implementation challenges of competency-based assessment under the 2019 Curriculum Guidelines, less is known about how teachers perceive, negotiate, and respond to the divergent assessment logics of the CAP and school-based examinations. When these two assessment systems operate concurrently, their washback effects may reinforce or weaken one another, or generate tensions between competency development and score expectations. Prior research has further suggested that different forms of assessment may produce different positive and negative washback effects on material selection, instructional activities, and assessment design. Accordingly, understanding how inconsistencies between the CAP and school-based examinations shape instructional practices requires an integrated analysis of washback effects, teachers' decision-making considerations, and practical

constraints. This study therefore examines how junior high school teachers adjust their instructional and assessment practices within the concurrent assessment context of the CAP and school-based examinations, and analyzes the washback effects, decision-making considerations, and practical constraints involved in this process. In particular, the study investigates how teachers mediate divergent assessment demands and translate them into feasible instructional actions.

Design/Methodology/Approach

This study employed an interpretivist approach and generated data through focus group interviews. To enhance the appropriateness of the interview questions and the trustworthiness of the data, a pilot study was conducted with junior high school advisory teachers affiliated with a subject-area advisory group in a special municipality. The interview questions were developed on the basis of the literature review and subsequently refined based on discussions with and feedback from five teachers. Particular attention was given to whether the questions adequately reflected junior high school instructional contexts and teachers' understandings of the CAP and school-based examinations. The final interview protocol was structured to move from eliciting teachers' experiences to concrete instructional actions and reflective analysis. The opening question invited teachers to explain how they interpreted students' performance on the CAP and school-based examinations. Questions 2 and 3 examined adjustments in material treatment, instructional activities, and assessment design under different assessment orientations. Questions 4 and 5 explored teachers' decision-making considerations and how these differed across assessment contexts, while Questions 6 and 7 investigated contextual constraints encountered in practice and their variation across assessment contexts.

Twenty-five junior high school teachers with experience in teaching CAP subjects and designing school-based examinations were then recruited.

The participants included 14 male and 11 female teachers, with an average of 17 years of teaching experience ; all had experience teaching ninth-grade students in subjects assessed by the CAP. After informed consent had been obtained, five focus group interviews were conducted.

Data were analyzed using thematic analysis, involving coding, comparison, and integration. A single teacher's statement was not treated as directly representing the views of all participants. When a view was responded to, supplemented, or confirmed by other teachers within the same focus group, it was identified as a shared view within that group; when it recurred across different groups, it was further categorized as a cross-group theme. This analytic strategy was used to preserve the interactional nature of focus group data while avoiding overgeneralization from isolated statements. It also enabled the study to distinguish among views shared within a particular group, themes recurring across groups, and context-specific or exceptional cases. Through this process, teachers' experiences of instructional practice under different assessment orientations were identified.

Findings/Conclusions

The analysis yielded three major themes. First, with respect to material treatment, teachers engaged in the reorganization of core concepts and negotiated trade-offs in textbook coverage. Under the CAP orientation, teachers reorganized textbook content with reference to curriculum guidelines, core concepts, and frequently assessed ideas; under school-based examinations, they placed greater emphasis on textbook scope, fixed examination coverage, and students' attainable learning performance. Second, in relation to instructional activities, teachers incorporated cross-text competence training while also relying on lecture-based instruction under pressures related to content coverage and pacing. Competency-oriented teaching was often translated into more operationalized forms, such as keyword identification, graph and table interpretation, and guided

reading strategies, particularly when teachers had to complete the required instructional content within limited class time. Third, in assessment design, the findings revealed how teachers coordinated contextualized item types with the allocation of foundational scores. Teachers attempted to incorporate contextualized and integrative items, while retaining foundational items to sustain students' learning motivation, maintain their confidence, and avoid excessive experiences of failure. Overall, the findings indicate that the CAP tended to orient teachers toward core concept integration, cross-unit connections, and competency-based instruction, whereas school-based examinations were more closely associated with curriculum pacing, foundational learning, and the maintenance of students' learning confidence. Teachers' adjustments in instructional and assessment practices involved professional judgments concerning students' learning confidence, cross-text reading competence, content coverage and pacing pressures, and external expectations. The pressure to complete required instructional content, differences in student ability, and institutional regulations constituted common constraints across both assessment contexts. Parental expectations, however, generated more immediate external pressure through school-based examination scores and within-school rankings, making test design and instructional adjustments in school-based examination contexts more susceptible to achievement-based evaluation. This study addresses how assessment inconsistency shapes instructional practice, showing that instructional adjustment does not simply mirror differences between assessment systems but constitutes a contextualized transformation mediated by teachers' professional judgment under shared structural conditions. These findings suggest that understanding contemporary instructional practice requires moving beyond a single-assessment perspective and attending to the interactions among divergent assessment demands, teacher decision-making, and contextual constraints, thereby more accurately capturing how teaching operates in practice.

Originality/Value

This study extends existing washback research by foregrounding a dual-assessment context that has received limited scholarly attention. The findings demonstrate that teachers are not merely passive respondents to assessment demands; rather, within the concurrent institutional contexts of the CAP and school-based examinations, they exercise professional judgment and negotiate trade-offs in relation to student learning, curriculum pacing, and external conditions. The findings further challenge a linear conception of "assessment influencing teaching" and indicate that such a view is insufficient to explain the complexity of classroom practice. Washback should instead be conceptualized as a dynamic and contextually embedded process, with the washback effects of the two assessment systems forming an interactive configuration. These effects may sustain the orientation toward competency-based instruction when school-based examinations incorporate contextualized and cross-text demands, but may also attenuate this orientation when curriculum pacing, foundational score allocation, and parental expectations become dominant concerns. The study therefore reveals the intersection of the influences of the CAP and school-based examinations within the same instructional setting, as well as how these influences are mediated through teachers' professional judgment, decision-making, and responses to contextual conditions. Such intersecting assessment influences are subsequently transformed into situated instructional practices. By advancing this interpretation, the study offers a more integrated analytical perspective for understanding the complexity of teachers' instructional actions under the current assessment system.

Policy Implications and Applications

The study shows that the constraints teachers encountered when making instructional adjustments were largely rooted in structural conditions,

including pressure to complete required instructional content, institutional regulations, and differences in student ability. Accordingly, if educational authorities and schools seek to promote competency-based instruction, policy efforts should not be limited to item design reform, but should also address curriculum pacing, assessment systems, and the conditions under which remedial instruction is implemented. In terms of instructional practice, the findings indicate that when teachers face the concurrent demands of the CAP and school-based examinations, the central challenge lies in coordinating multiple instructional priorities, including competency development, the consolidation of foundational learning, students' learning confidence, and the need to complete required content within limited instructional time. Teachers may therefore examine whether their treatment of materials balances core concepts with textbook coverage, whether instructional activities translate cross-text reading, graph and table interpretation, and contextual analysis into accessible learning steps, and whether school-based examination design creates a progression from foundational items to contextualized tasks, thereby supporting students' gradual development of the competencies required by the CAP.

Keywords: Comprehensive Assessment Program (CAP) for Junior High School Students, school-based periodic examinations, washback, teacher decision-making, instructional practice



DOI：10.6869/THJER.202512_42(2).0004
Received: April 24, 2026; Modified: June 11, 2026; Accepted: August 18, 2026
* Shih-Hsiung Liu, Professor, Center for Teacher Education, National Changhua University of Education, E-mail: shsiung@cc.ncue.edu.tw