

國立臺東大學特殊教育學系，特殊教育中心
特殊教育學術研討會論文集，民 96，81-106 頁

語音合成與真人錄音對國小高年級低閱讀能力學生數學文字題解題表現差異之研究

黃巧雲

國立嘉義大學特教系

陳明聰

國立嘉義大學特教系

陳政見

國立嘉義大學特教系

摘 要

本研究旨在探討國小高年級不同閱讀與數學能力學生之數學文字題解題表現，是否因不同語音調整方式(語音合成與真人錄音)而有差異。研究者以 34 位國小高年級學生為對象，依其閱讀與數學能力分為三組，分別接受純文字、真人錄音、語音合成版線上數學成就測驗。主要研究結果如下：1. 國小高年級一般閱讀一般數學能力組與低閱讀低數學能力組學生在接受電腦化不同語音版本調整後，其對數學文字題解題表現並無影響，且調整前與調整後表現差異未達顯著；2. 低閱讀一般數學能力組學生在電腦化數學成就測驗文字題型之語音合成、真人錄音表現均較純文字佳，但表現差異未達顯著水準；3. 國小高年級不同能力學生之表現差異達顯著水準，但在電腦化數學成就測驗文字題型不同版本的測驗下，沒有交互作用。

關鍵字：測驗調整、語音調整、閱讀困難、語音合成、真人錄音

緒 論

一、研究動機

自一九八〇年代以來，逐漸盛行的融合教育（inclusive education）理念，主張讓身心障礙兒童在普通班接受教育，並提供必要的協助（何華國，2004）。融合教育是特殊教育的趨勢，在此趨勢下有越來越多的身心障礙學生在普通班中學習，而融合教育所涵蓋的是整個教育生態環境的調整和革新，如何有效的評量出在普通班身心障礙學生的學習成效對普通班老師是一大挑戰（鍾素香，2002）。根據民國 95 年 05 月 10 日修正的「國民小學及國民中學學生成績評量標準」第六條規定：「對於身心障礙學生之成績評量方式，應衡酌其學習優勢管道彈性調整之。」對特殊學生而言，一般標準化測驗分數的推論效度令人質疑，因為障礙影響他們在測驗上的表現，也無法真實反映測驗本身所要測的能力，因此需要透過調整，將影響表現的障礙移除，使個人顯現出真正的能力（張萬烽，2004）。因此如何利用考試調整方式協助身心障礙學生有效呈現其學習成果成為中要的議題。事實上，也已有許多的調整方式，不過目前測驗調整方式陳明聰、張靖卿（2004）將測驗調整方式歸納為六大調整類型：1.測驗呈現方式(presentation)、2.作答反應方式(response)、3.施測情境(setting)、4.測驗時限(timing)、5.時間安排(scheduling)及 6.其他(others)。

而根據 Sireci, Li 與 Scarpatti(2003)分析 150 份相關研究後發現，現行的調整方式以「測驗呈現方式」中的「語音調整」(oral accommodation)最為普遍。在國內方面，根據陳明聰、張靖卿（2004）調查則發現有六成四的受訪者認為：以錄音帶方式來呈現試題是可行的。然而，由於每一種調整方式都不同、施

測程序也不一致、所實施的對象不同且異質性也高，因此測驗調整的公平性仍充滿爭議（Sireci, et al., 2003）。而目前相當普遍的「語音調整」是否具有公平性？此乃研究者研究主要動機之一。

若以 Mayer(1992)從認知心理學的觀點，將數學解題歷程分為問題表徵(problem representation)與問題解決(problem solving)兩階段。其中，問題表徵是將問題的文字或圖案轉換為心理表徵，而此階段又包含兩個步驟：「問題轉譯」(problem translation)指解題者必須了解題意，並將問題從外在的文字表徵轉換為內在的心理表徵、和「問題整合」(problem integration)指解題者利用其基模知識(schematic knowledge)來統合題目中各條件的關係。因此，成功解題的第一步即是將題目文意正確轉譯為心理表徵。既使數學能力相同，如果閱讀能力不同，則其文字題解題歷程及結果也會有所差異。

由以上理論可知，在數學文字題解題歷程中數學解題及閱讀理解能力皆為成功解題的要素之一，但數學文字題解題最終目的在於測驗學生真正的數學能力，並不是閱讀理解能力，因此，宜試著藉由外在教學及評量方式的改變，來幫助低閱讀能力學生克服閱讀理解等無關測驗的干擾因素，以測得受試者真正數學能力。

隨著，電腦科技的發展，數位化語音讀題可能是一可行方式。數位化語音讀題依照目前科技發展有兩種語音錄製方式可選擇：一為真人錄音，另一為語音合成。使用語音合成與真人錄音兩者主要差異在於發聲的自然度與錄製時所投入的人力與資源，語音合成的發聲自然度雖比不上人工錄音來的自然，但現今科技發達，語音合成已發展至可依聽者較舒適聽覺的範圍改變其音調和音速。事實上，目前語音合成已大量且廣泛的使用在許多科技系統

上，例如：電話查詢系統、氣象查詢系統、鐵路訂票系統等，另外在輔助科技方面包括：導盲鼠、點字中文即時雙向轉譯系統、Victor Reader Daisy 有聲書撥放器等，可見語音合成已普遍的為大眾所接受。至於真人語音對評量調整效果研究部分，國外已有許多研究而國內只有一篇，林筱汶、陳明聰(2006)利用電腦播放之真人錄音檔搭配電腦化數學成就測驗，探討國小五年級低閱讀能力學生數學文字題解題表現。在國外則有 Kosciot 與 Ysseldyke 在 2000 年以錄音帶讀題的調整方式，針對郊區小學三至五年級一般、特殊學生所做的綜合成就測驗、及 Calhoon、Fuchs 與 Hamlett 在 2000 年以電腦呈現的人工音讀題的調整方式，針對九、十、十一、十二年級的學習障礙學生所做的電腦實作評量之研究(Sireci, Scarpatti & Li, 2005)，但對於語音合成的表現成效上則尚缺乏深入探討。雖然語音合成有發聲較不如真人錄音自然的缺點，但數學文字題題目簡短，若語音合成所獲得的效果能與真人錄音一樣，則是否可由較不費人力與資源且較具彈性並成本較低的語音合成來取代真人錄音，進行語音評量調整方式呢？此乃研究者研究主要動機之二。

二、研究目的與問題

根據前述的研究動機，本研究之目的在：探討不同類型學生(一般閱讀一般數學能力、低閱讀一般數學能力、低閱讀低數學能力)在純文字、語音合成及真人錄音三個版本數學文字題的表現差異。

根據前述的研究目的，本研究提出下列研究問題作為蒐集資料及實證的依據：

(一)不同類型學生(一般閱讀一般數學能力、低閱讀一般數學能力、低閱讀低數學能力)在純文字、語音合成及真人錄音三個版本數學文字題的表現是否有差異？

文獻探討

一、測驗調整的意義

測驗調整(test accommodation)，一般常用的名詞尚有 test adaptation、assessment adaptation、assessment accommodation。就測驗而言，測驗(test)、測量(measurement)、評定(assessment)、及評量(evaluation)四個名詞，常被許多教師及心理學者視為是同義詞，不加分別地交替使用(余民寧，2005)。不過在此仍需依其定義作區別，測驗是指一種特定行事的診斷程序，它是較大評量歷程的一部份，包括一組問題來產生一組分數或一些其他的數字結果；評量乃是使用測驗和其他測量學生成就和行為，以便做出教育性決定的歷程(張世慧、藍瑋琛，2004)。再就調整觀之，調整(accommodation)是指施測方式的調整，目的為在不影響評量工具效度的原則下，藉由施測方式的調整，確保評量對身心障礙學生所做評量結果之正確性；修正(modification)則不但施測過程改變，且測驗本身亦做更改，受試者即使經由評量調整，仍難以進行測驗，而需進一步對測驗的實施採取變通，而此一變通可能改變該項評量工具所預設之構念效度(胡永崇，2005)。

綜上所述，基於本研究的測驗內容為電腦化數學成就測驗，是在不改變測驗效度的原則下，利用語音調整的施測方式進行測驗，因此，本文將以「測驗調整」(test accommodation)來進行探討。

所謂測驗調整是在確保考試概念完整的調整下，移除與考試概念無關且會妨礙學生真正表現的不必要干擾(Sireci, Li & Scarpatti, 2005)。任何測驗之調整或改變，目的皆在於增進此一測驗對身心障礙者之測驗效度(validity)，但不可因為此一改變反而降低測驗之效度(胡永崇，2005)。根據陳明聰、張靖卿

(2004)綜合許多學者(Elliot, Thurlow, Ysseldyke, & Erickson,1997 ;Venn,2000)看法，測驗調整主要目的是要給予有特殊需求的學生一個平等的機會，在不受障礙影響下展現出他們的知識和技能，而且調整的結果也不可以影響原有測驗的效度，而一般學生使用測驗調整的方式也不會提高其成績。

由於部分考生與家長混淆的認知，認為考試調整目的是使身心障礙學生通過考試，所以當大考中心努力對調整的要求做公平的決策時，即遭遇到許多挫敗與敵意(謝秋梅，2004)。過去有許多教師常以「對其他學生不公平」作為反對為學障學生採取測驗調整措施之理由，但未採取測驗調整措施卻對學障學生「不公平」(胡永崇，2005)。因此，身心障礙學生考試調整在「公平性」與「執行層面」仍有許

多爭議的地方，有待工作人員進一步去釐清，而且政府相關單位也應該訂定或是修正法規，規劃更周全的考試調整措施，以嘉惠身心障礙學生，保障其考試權益(李佩蓉，2007)。

二、測驗調整的方式

根據陳明聰、張靖卿(2004)歸納出測驗調整類型，測驗調整方式包含六大調整類型：1. 測驗呈現方式(presentation)、2.作答反應方式(response)、3.施測情境(setting)、4. 測驗時限(timing)、5. 時間安排(scheduling)、6.其他(others)。表 1 為研究者根據上述六大調整類型，歸納(陳明聰、張靖卿，2004、張萬烽，2004、Fuchs, Fuchs, &Capizzi, 2005)將各調整類型之定義與改變方式整理如表 1：

表一 測驗調整的方式

類型	定義	改變方式
試題呈現方式	改變題目的呈現方式	用錄音帶、錄影帶呈現試題
		用電腦呈現試題
		用點字
		增加題目行距或減少每頁題數
		調整答案格的大小
		調整答案格的大小
		簡化或唸出指導語
		報讀題目
		在答案紙以提供箭頭或停止標誌提供線索
		將關鍵字詞畫線或加以標示
		提供額外例子
		放大題目
作答反應方式	改變學生原本作答方式	直接在題本上作答
		將反應的答案錄音後轉為文字
		使用打字、電腦文書處理或溝通板回答
		使用手語、口頭等替代性反應

表一 測驗調整的方式（續）

類型	定義	改變方式
施測情境	改變測驗所提供的環境	個別或小組施測 個別學習桌椅減少干擾 提供特殊燈光、特殊音響、調整式或特殊桌椅
施測時限	改變測驗的時間長度	允許彈性時間 允許延長時間 無時間限制讓學生完成所有題目 分段測驗
時間安排	改變測驗的時間	在一日中特定時間測驗 測驗分成數日完成 允許不同順序完成測驗
其他	部分無法歸類的調整方式	集中注意的提示 提供計算機、字典 用膠帶固定考試卷 提供模板以限制閱讀範圍

而實際應用上，研究者整理有關國內外升學考試身心障礙考生應考申請調整服務項目，包括九十五學年度指定科目考試、九十五學年度學科能力測驗、四技二專統一入學測驗、國民中學學生基本學力測驗、美國大學新生入學測驗（財團法人大學入學考試中心，2005a；財團法人大學入學考試中心，2005b；技專校院入學測驗中心，2006；教育部中部辦公室，2006；CollegeBoard.com，2006），將各入學考試調整方式做比較，可以發現最常見的調整項目為：

(1)測驗呈現方式：放大題目、點字、協助考生閱讀之輔具

(2)測驗時限：允許延長考試時間

(3)反應方式：放大答案格

而上述入學考試調整規定，雖然調整的類型與項目相同，不過每個考試的規定皆不一

致，如：呈現方式類型中的放大題目項目，在國民中學學生基本學力測驗與四技二專統一入學測驗的規定為放大 1.5 倍的試卷、九十五學年度學科能力測驗與指定科目考試皆放大為 A3 之試題本、美國大學新生入學測驗則為放大題目 (14pt；20pt)；時限類型中的允許延長考試時間項目，國民中學學生基本學力測驗規定為延長作答 20 分鐘、九十五學年度學科能力測驗為最多延長 20 分鐘應考時間、九十五學年度指定科目考試則最多延長 50 分鐘應考時間、四技二專統一入學測驗為每科目之考試時間依一般考試時間再延長 20 分鐘，但兩科目間休息時間減少 20 分鐘、美國大學新生入學測驗則未詳細規定只說明允許延長考試時間。

前述國外相關研究中最為常見的調整方式為語音報讀 (39%) 和延長時間 (24%)，而研究

中最常見的調整科目為數學、閱讀和自然，近60%的研究者將受試者鎖定在3年級至8年級的學生(Sireci, Scarpatti & Li, 2005)。

關於國內評量調整相關研究中，陳明聰與張靖卿(2004)對199位特殊教育工作者之意見調查，發現超過九成的受訪者認為「作答反應方式、試題呈現方式、施測情境安排、測驗時限調整」是可行的。而針對試題呈現方式，整體而言受訪者比較認為可行的調整方式為「放大題目、用點字、用錄音帶、增加題目行距或減少每頁題數」。

三、評量調整國內外之相關研究

目前國內關於語音評量調整的實徵研究僅有林筱汶、陳明聰(2006)針對國小五年級「低閱讀能力」學生，進行「電腦化語音讀題」調整方式的數學成就測驗，其他國內相關研究尚付諸闕如，且關於語音評量調整的研究中，近年來越來越多以錄音帶、錄影帶或是合成音呈現的調整方式。研究者針對國外語音評量調整相關研究作摘要表(附表1)，並依據語音調整方式作分類將研究結果敘述如下：

(一)紙本呈現加上口語報讀

1、對不同能力學生有不同影響

根據國外相關研究發現一般學生在調整後的表現不會因為不同方式的語音調整而有影響，但對於特殊及低閱讀能力學生皆有顯著的正面影響(Johnson, 2000; Weston, 2002; Tindal, Heath, Hollenbeck, Almond, & Harniss, 1998; Calhoon, Fuchs, & Hamlett, 2000; Huynh, Meyer, & Gallant, 2004)。

2、對所有學生皆有顯著正面幫助

Meloy、Dewille 與 Frisbie (2002)針對260名六、七、八年級的一般與閱讀障礙學生之自然、閱讀及數學方面，提供紙本呈現的口語報讀調整方式，結果發現不論在一般或閱讀障礙學生的表現上，皆提高其分數，且達顯著

水準。

3、對所有學生皆無顯著影響

Fletcher、Francis、Boudousquie、Copeland、Young、Kalinowski 與 Vaughn (2006)針對三年級一般和識字或解碼困難學生之閱讀方面進行語音調整，結果發現所有學生在語音調整前後其表現並無顯著差異。

(二)紙本呈現加上錄音帶或錄影帶讀題

1、對不同能力學生有不同影響

Kosciolet 與 Ysseldyke (2000)針對31名三至五年級一般與特殊學生，在閱讀成就測驗進行紙本呈現的錄音帶讀題調整方式之研究，結果發現一般學生不論在調整前或調整後的表現皆優於特殊學生，而特殊學生在調整後的分數提高，且達顯著水準，一般學生在調整前與調整後的分數並無顯著差異。

2、對不同年級學生有不同影響

Helwig 與 Tindal (2003)，在數學科利用紙本呈現的錄影帶讀題調整方式之研究，結果發現在語音調整前後的表現，對七、八年級學生無影響；四、五年級的特殊學生及四年級低閱讀高數學能力學生有顯著的正面影響；五年級低閱讀高數學能力學生無影響。

3、對所有學生皆無顯著影響

McKevitt 與 Elliott (2003) 針對79名八年級一般與特殊學生在閱讀方面實施紙本呈現的錄音帶讀題調整方式之研究，結果皆發現學生不論在調整前或調整後的表現並無顯著差異。

(三)電腦呈現加上人工音讀題

1、對不同能力學生有不同影響

林筱汶、陳明聰(2006)及 Calhoon、Fuchs 與 Hamlett (2000)，分別針對國小五年級一般及不同閱讀、數學能力學生在數學文字題與九、十、十一、十二年級的學習障礙學生在電腦實作評量進行電腦呈現的人工音讀題的調整方式之研究，結果發現對低閱讀一般數學能

力及學習障礙學生之表現有顯著正面影響，對閱讀與數學能力皆為一般或皆低者無影響，而對一般閱讀與低數學能力學生有負面影響。

(四)電腦呈現加上合成音讀題

1、對所有學生皆無顯著影響

根據 Brown 與 Augustine(2001)針對 206 名的一般與閱讀障礙學生之社會、自然，實施電腦呈現的合成音讀題調整方式，結果皆發現學生不論在調整前或調整後的表現無顯著差異。

2、對不同能力學生有不同影響

電腦呈現的合成音讀題的調整方式對於學習障礙學生之表現有顯著正面影響 (Calhoon, et al., 2000)。

測驗調整的公平性及不同調整方式對不同能力學生的幫助性，仍充滿許多爭議，由以上文獻綜合觀之，可發現不同語音的調整方式會有不同的研究結果，即使是相同的調整類型仍會導致不同的研究結果，目前尚無法對任何一種語音調整方式作定論。因此，關於語音調整的議題還有賴更多的教育工作人員做進一步的研究以釐清公平性及幫助性的問題。

研究方法

一、研究架構

為了達成研究目的，採用「二因子混合設計」(一因子獨立，一因子相依)，以探討「低閱讀能力、不同數學能力」、「不同語音調整類型」對學生數學文字題解題表現是否有顯著差異。茲將本研究架構說明如下：

(一)自變項：

本研究包含兩個自變項，分別為相依因子的測驗呈現方式和獨立因子的學生能力。測驗呈現方式依不同語音版本可分為「語音合成」、「真人錄音」、「純文字」三個水準；學生能力依不同閱讀與數學能力可分為「低閱讀一

般數學能力」、「低閱讀低數學能力」及「一般閱讀一般數學能力」三個水準。

(二)依變項：

數學文字題解題表現，是指研究者利用陳政見(2005)所編「國小高年級數學科成就測驗」試題，但由於本研究選取之語音合成軟體「文字 MP3 V2」無法辨識分數的正確念法，且本研究的主題為數學文字題解題表現，因此在刪除概念題及含有分數的題目後之試題內容，做為本研究電腦化線上測驗題目，受試者在電腦化數學成就測驗文字題型之語音合成、真人錄音及純文字平均得分，計分方法為答對一題得一分，總答對分數除以總題數之平均得分。

二、研究假設

根據緒論之研究目的，本研究提出下列之假設：

(一)探討不同能力學生 [一般閱讀一般數學能力、低閱讀一般數學能力、低閱讀低數學能力] 在純文字、語音合成及真人錄音三個版本上的表現差異。

三、研究對象

本研究之研究對象為「不同閱讀與數學能力之國小高年級學生」，依篩選流程可分為四個階段，依序為：抽樣、知會同意、依測驗得分分組、完成三次電腦測驗及問卷之施測對象，茲將說明於後：

(一)抽樣：

研究者先取得南部某縣市教育局同意之後，透過特教資源中心取得 2006 年通報為疑似學習障礙的國小學生所做閱讀及數學能力測驗之篩選結果，排除智能、感官、障礙因素後，取得高年級閱讀理解困難篩選測驗通過率在 .80 以下，疑似學習障礙學生之名單共 47 位，則此 47 位學生為「低閱讀能力者」。

(二)知會同意：

研究者請特教資源中心代為發給各校「研究參與邀請函」，並請各校轉發給上述 47 位學生，徵求學生及其監護人簽章同意。同意參與研究的學生共 22 人，包括低閱讀一般數學能力者 10 人、低閱讀低數學能力者 12 人，及由導師隨機抽樣一般閱讀一般數學能力者 22 人。

(三)依測驗得分分組：

為了解上述 22 位低閱讀能力學生的數學能力，研究者透過嘉義市特教資源中心取得 22 位學生鑑定時所接受柯華葳（1999）所編之「高年級閱讀理解困難篩選測驗」、「基礎數學概念評量」測驗得分。

並依據此測驗各分測驗之平均得分對照後，將低分者命名為「低數學能力組」，其他則為「一般數學能力組」。各項測驗得分，如表 2 所示：

表 2 閱讀與數學能力篩選測驗平均通過率

閱讀理解		基礎數學概念					
組別	平均通過率	九九乘法		空格運算		三則運算	
		分測驗	分測驗	分測驗	分測驗	分測驗	分測驗
		1	2	1	2	1	2
低閱讀與低數學（n=10）	0.37	0.69	0.98	0.25	0.51	0.11	0.31
低閱讀與一般數學（n=7）	0.32	0.70	0.94	0.62	0.98	0.41	0.46

(四)完成三次電腦測驗之施測對象：

上述 34 名學生中完成本研究三次電腦測驗，作為分析對象包括 (A) 一般閱讀 & 一般數學，17 人、(B) 低閱讀 & 低數學，10 人、(C) 低閱讀 & 一般數學，7 人。

四、研究工具

(一)電腦化數學成就測驗

指研究者利用陳政見(2005)所編「國小高年級數學科成就測驗」之試題，做為本研究電腦化線上測驗之題目內容。以下將就國小高年級數學科成就測驗及電腦化數學成就測驗之內容、計分、施測方式及信效度說明之。

1、國小高年級數學科成就測驗

(1)測驗內容

所有試題均為四選一的選擇題，總題數共

100 題，依據九年一貫的課程標準，以國民小學五、六年級數學領域為測驗編製學科。

(2)施測與計分方式

團體測驗方式施測。每答對一題得一分，答錯不倒扣，總分即為 100 分。作答時間為 50 分鐘。利用測驗總分，查對常模分數對照表，即可得知學生程度。

(3)信度與效度

a. 信度：

重測信度：相隔二週後由原校、原班級人數再測，求其前、後測之 Pearson 相關係數 ($p < .001$)，均達到 .001 顯著水準，顯示具有良好重測信度。

b. 效度：

將常模資料與三所南部國民小學 93 學年度下學期數學科成績以皮爾遜 (Pearson) 積

差相關求得相關係數結果：相關高達.807， $p < .001$ ；.762， $p < .001$ ；.751， $p < .001$ ，具效標關聯效度。

c. 難度與鑑別度：

全測驗試題難易度 P 介於.33～.71 之間，平均.49；常態化等距離難度指數 Δ 在 8.2～17.0 之間，平均 13.5，鑑別度 r (二系列相關) 在.21～.75 之間，平均.53；鑑別度 D 介於.20～.70 之間，平均.51。

由於研究之需要，研究者扣除概念題及含有分數的題目後，將試題分為三版本，每一版本有 20 題，因此另行比較三種版本之測考試題難易度。研究者隨機抽取南部一所國民小學高年級學生 20 名，以紙本呈現方式，進行測驗，限時 30 分鐘，時間一到就必須停止作答。作答結果以相依樣本單因子變異數分析，結果顯示三種版本間並無顯著差異 ($F = .27$ ， $p = .76$)。接著研究者進一步求三版本試題難易度相關，結果顯示三版本試題平均分數各為 13.1、12.15 與 12.65，從 Pearson' r 可知三版本之間的相關高達.864， $p < .01$ 、.786， $p < .01$ 與.758， $p < .01$ ，均達顯著水準，因此，三版本試題難易度趨於一致。

2、電腦化數學成就測驗

研究者把三種版本分配為純文字、真人錄音、語音合成三種版本。研究者利用 HTML、ASP 和 SQL 建立網路化的線上測驗，以利測驗的進行。以下說明測驗內容及執行方式。

(1) 測驗內容

a. 基本資料

包括學校、姓名、學號、性別、班別，以便依學生組別建檔。

b. 施測方式

請學生先在首頁輸入個人基本資料，並確定每位學生均了解操作與測驗方式後，即進行線上測驗。「純文字版本」僅以電腦螢幕呈現試題，不提供語音協助。而「真人錄音版本」、「語音合成版本」，於試題呈現時，由耳機同步播放所預錄之真人錄音檔及語音合成檔，一次朗讀一題，受試者皆須至少先聽完一次語音讀題後才能進行作答，若學生想再聽一次，可點選「再聽一次」來重複播放。作答時，試題一次出現一題，受試者可以滑鼠點選「上一題」或「下一題」的方式來修改或是作答下一題(如圖 1)，故學生可自行決定每題的作答時間、並隨時修改答案，唯每個版本作答時間皆為 30 分鐘，時間一到受試者就必須停止作答。

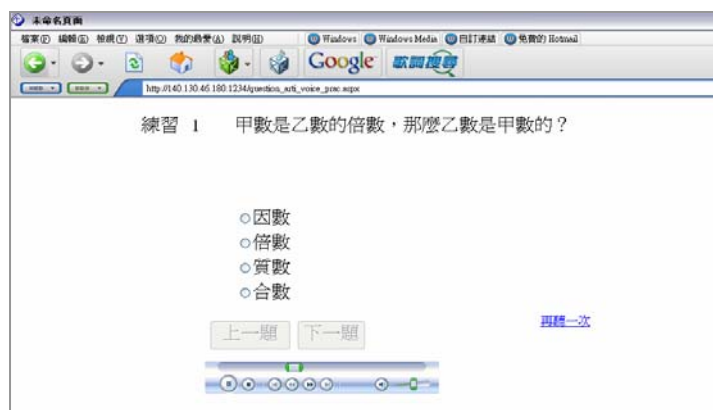


圖 1 電腦化數學成就測驗——語音版本畫面

c. 電腦化測驗之信度

為了解三種版本之難易度是否一致以及電腦化施測時間、流程是否適當，研究者隨機選取南部國民小學之高年級學生，共 23 名，施以電腦化純文字版本、真人錄音版本及語音合成版本數學成就測驗。

經由單因子變異數分析檢定接受純文字版本、真人錄音版本及語音合成版本之兩校學生，在三種版本試題之平均得分是否有差異。結果顯示未達顯著差異(純文字 $M=10.61$, $SD=4.48$; 真人錄音 $M=11.17$, $SD=4.28$; 語音合成 $M=11.17$, $SD=4.70$; $F=.40$; $p=.67$, $p>.01$)意即三種版本試題之間難易度趨於一致。

(2) 語音檔

a. 語音合成檔

研究者先從許多語音合成軟體的試聽之下，選出發音較自然也較為廣泛使用的兩種語音合成軟體：「Extra Speaker」及「文字 MP3 V2」。接著，商請 10 名大學生及一所國民小學 10 名高年級學生進行試聽，並為此兩種語音軟體之意見調查表(如表 3 所示)根據其發音及語調兩大向度，分別為它打上分數，由 1~10 分，1 分為最低分，10 分為最高分。結果顯示「文字 MP3 V2」除了在「斷詞的適切性」此項總分略低於「Extra Speaker」0.45 之外，其他項目得分都高出「Extra Speaker」。因此，研究者選用「文字 MP3 V2」為本研究使用之語音合成軟體。

b. 真人錄音檔

研究者商請一位音質較佳的大學生，先預錄試題 10 題，再請研究成員試聽並給予意見後，決定聲調、音速。再由該名大學生朗讀電腦化數學成就測驗真人錄音版本之試題，並錄製成聲音檔。此為本研究「電腦化數學成就測驗真人錄音版本」所使用之聲音檔。

表 3 「Extra Speaker」、「文字 MP3 V2」兩種語音軟體之意見調查表得分情形

向 度	「Extra Speaker」	「文字 MP3 V2」
◎發音		
清晰度	37.65	48.25
正確性(一般字)	39.70	46.40
正確性(破音字)	37.00	43.90
適合小學生的速度	40.90	42.35
斷詞的適切性	34.40	33.95
◎語調		
自然度	29.30	32.50

(3) 計分方式

測驗時所蒐集之資料將會以網路傳送方式，即時的建檔於資料庫中，再進行計分。各版本題數皆為 20 題，每答對一題即給一分，答錯則不給分，總分為 20 分。

五、實驗設計

本研究採「二因子混合實驗設計」(一因子相依，一因子獨立)，並透過隨機分派及對抗平衡設計(counter-balanced design)方法進行實驗(王文科，2002)。

(一) 實驗分組

研究者經篩選後所決定之正式施測對象共有 44 人，分為 A、B、C 三組。考量到每一位受試者皆接受三種實驗的處理，為防有「次序效應」(order effect)的產生，因此使所有實驗對象分別以不同順序來接受所有的實驗處理。

研究者將六所受試學校隨機分派為甲、乙、丙三組，每組各包括兩所學校。分派至甲組學生的施測順序為：電腦化數學成就測驗真人錄音版本→電腦化數學成就測驗語音合成版本→電腦化數學成就測驗純文字版本；乙組學生施測順序為：電腦化數學成就測驗語音合成版本→電腦化數學成就測驗純文字版本→

電腦化數學成就測驗真人錄音版本；丙組學生施測順序為：電腦化數學成就測驗純文字版本→電腦化數學成就測驗真人錄音版本→電腦化數學成就測驗語音合成版本。

(二)實驗地點與時間

為避免影響受試學生正常的上課時間，並顧及學校電腦教室的上課使用，研究者皆利用

早自修與午休時間於各校電腦教室，採用一人一機團體施測的方式。每位受試學生皆經過三次電腦化數學成就測驗，次施測的時間為 30 分鐘。

(三)實驗流程

茲將說明並繪製流程圖如圖 2。

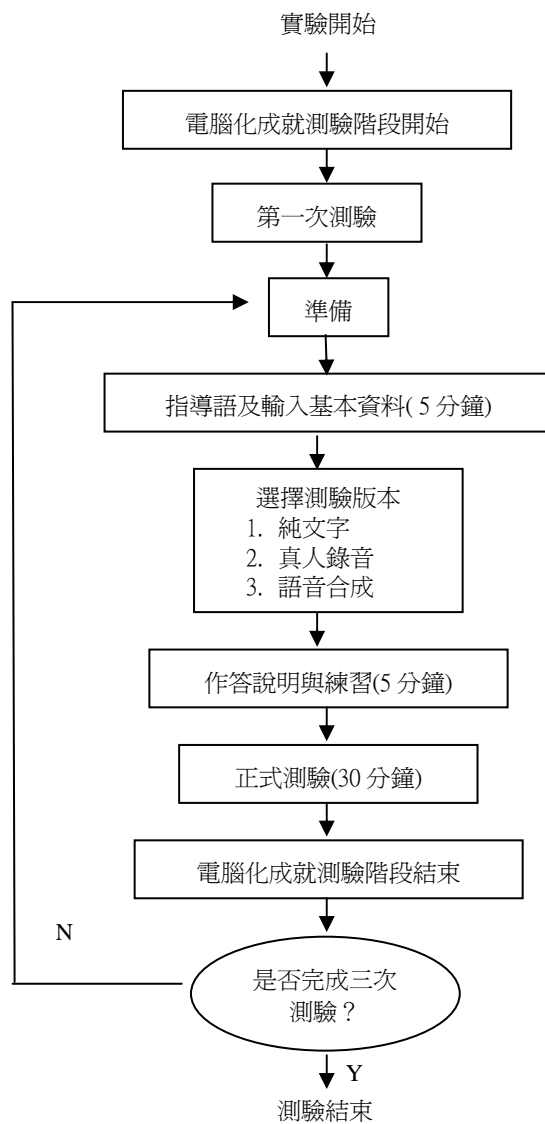


圖 2 實驗流程圖

結果與討論

一、探討不同能力學生在純文字、語音合成及真人錄音三種版本上的表現差異：

(一)不同能力學生在三種測驗得分表現。

三組學生在三版本得分經描述性統計分析整理如表 4。

表 4 各組學生之文字題解題表現摘要表 (n=34)

組 別	純文字版		真人錄音版		語音合成版	
	M	SD	M	SD	M	SD
A	11.12	3.10	11.88	3.66	11.82	3.66
B	5.60	1.84	5.50	1.90	5.50	2.07
C	5.86	1.77	6.43	4.08	6.00	2.94

A：一般閱讀一般數學能力 B：低閱讀低數學能力 C：低閱讀一般數學能力

根據表 4 得知：一般能力組學生的得分以真人錄音版最高、語音合成版次之、純文字版最低；低閱讀低數學能力學生的得分以純文字最高、真人錄音及語音合成得分則相同；低閱讀一般數學能力學生的得分以真人錄音版最高、語音合成版次之、純文字最低。

(二)比較不同能力學生在三種測驗得分表現比較分析。

為了解不同測驗呈現方式是否會影響學

生的解題表現，研究者採用「二因子混合設計」(一因子獨立、一因子相依)，以學生能力(獨立因子)、測驗呈現方式(相依因子)為自變項，語音合成版、真人錄音版及純文字版本得分為依變項進行分析。經 Box 法進行同質性檢定，結果符合變異數同質性假設($M = 20.86$ ， $p > .05$)。故可進行二因子變異數分析，並將其結果摘要於表 5。

表 5 學生能力、呈現方式二因子變異數分析摘要表

變異來源	平方和	自由度	均方和	F 值	顯著性
受試者間	1430.96	33			
學生能力	847.53	2	437.26	23.23	.00***
誤差	583.43	31	18.82		
受試者內	276.49	68			
呈現方式	2.57	2	1.28	.29	.75
呈現方式×學生能力	3.38	4	.85	.19	.94
誤差	270.54	62	4.36		
全體	1707.45	101			

*** $p < .001$

由表 5 可知國小高年級不同能力學生之表現差異達顯著水準，接著看主要效果，結果顯示不同測驗版本的表現差異未達顯著性。

即不同閱讀和數學能力之國小高年級學生，並未因不同測驗呈現方式，其數學文字題解題表現即有顯著差異。綜合上述研究結果得知，國小高年級學生在數學文字題之解題表現，會因為本身閱讀與數學能力不同而有顯著差異，但並不會因為測驗呈現方式不同而有顯著差異。

根據表 5 之平均數製成各組學生在文字題平均得分之折線圖。

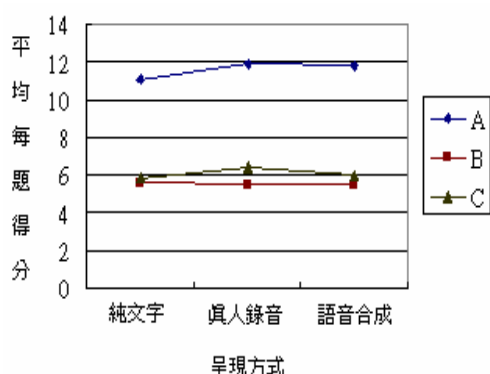


圖 3 各組學生在文字題平均-每題得分折線圖

註：A 表示「一般閱讀與一般數學能力組」

B 表示「低閱讀與低數學能力組」

C 表示「低閱讀與一般數學能力組」

由圖 3 可以看出，一般閱讀與一般數學能力組、低閱讀與低數學能力組、低閱讀與一般數學能力組，在不同測驗呈現方式下的文字題解題表現之間沒有交互作用。

研究結果發現，不同閱讀和數學能力之國小高年級學生，並未因不同測驗呈現方式，其數學文字題解題表現即有顯著差異，結果並未支持交互作用假設。不同閱讀和數學能力之國

小高年級學生對不同版本電腦化語音調整策略都接受，但不因能力、版本的不同而有顯著差異。

依據本研究結果，對照文獻探討所得，作一綜合性討論。

一、不同能力學生在電腦化數學成就測驗文字題型三種版本的表現差異

(一)一般閱讀一般數學能力組學生

無論是在純文字版本或是真人錄音、語音合成兩種不同呈現方式的語音調整，一般閱讀一般數學能力組學生表現皆優於其他兩組學生，且在三種版本的測驗表現上，對於一般閱讀一般數學能力組學生的表現均無顯著差異；表示語音調整策略具有公平性，一般學生不會因為有語音提示的評量調整提升其本身原本的答題表現，而影響了原測驗的效度，不同版本的語音也不會影響其解題表現。

(二)低閱讀低數學能力組學生

低閱讀低數學能力組學生在純文字版本或是真人錄音、語音合成兩種不同呈現方式的語音調整測驗上，表現都相當一致，並未因為有語音提示的評量調整而改變其原本的表現。表示雖然增加了語音提示的策略，企圖幫助此類學生解決在閱讀上的困難，但由於本身數學能力即低，因此語音提示還是未影響到該組學生的數學文字題解題表現。

(三)低閱讀一般數學能力組學生

低閱讀一般數學能力組學生在純文字版本或是真人錄音、語音合成兩種不同呈現方式的語音調整測驗表現上，在有語音的提示下該組學生的表現確實有提升但未達顯著水準 ($p=.94$)。

研究結果發現三種不同能力學生在電腦化數學成就測驗文字題型的表現會因學生能力不同而有顯著差異，但並不會因測驗呈現方

式不同，學生的表現差異即達顯著水準，且不同學生能力在不同測驗呈現方式下的交互作用未達顯著水準。其結果未能符合本研究假設：不同能力學生(一般閱讀一般數學能力、低閱讀一般數學能力、低閱讀低數學能力)在純文字、語音合成及真人錄音三個版本上的表現有顯著差異。

茲以過去研究結果進一步來討論如下：

一、與本研究結果一致的：

(一)學生經語音調整前、後的表現無顯著差異：

與 Brown 與 Augustine(2001)、McKevitt 與 Elliott(2003) 以及 Fletcher et al.(2006) 的發現一致，學生在調整前後的表現並無顯著差異。包括本研究在內，四個研究的研究對象皆為低閱讀能力學生，而調整方式有口語報讀、錄音帶及語音合成提示，不過除了本研究使用合成音與人工音做比較之外，其餘研究皆使用一種調整方式，而在調整科目及內容方面，則包括閱讀、社會、自然及數學文字題，其中多為以標準化測驗為測驗內容，僅 Fletcher et al.(2006)一篇為全州性成就測驗。

二、與本研究結果不同的：

(一)一般生與特殊生在語音調整下的表現皆有正面幫助：

Meloy et al. (2000)以口語報讀的調整方式，針對閱讀障礙學生，利用全州性成就測驗中的自然、數學與閱讀作為調整科目及內容，研究結果推斷出提供口語報讀的調整是缺乏公平性的。

綜上所述，可發現支持語音調整為有效的測驗調整方式的實徵研究多數是以大規模全州性的成就測驗作為研究測驗內容；相反的，

不支持語音調整為有效測驗調整論點的實徵研究中，其研究測驗內容多以自編成就測驗。根據胡永崇(2000)綜合許多相關研究顯示，「不專心」為鑑定出來之數學與閱讀障礙學生的最普遍特徵。

研究者在進行施測時最常被受試者所問到的問題即是「這個是考試嗎？」、「考不好會怎樣？會告訴老師嗎？」……等，關於會不會影響到他們在學校成績的問題，學生在此種心態之下作答，其作答動機即不高，容易在遇到題目較長或較難的題目時，就沒耐心思考或是選擇放棄，相對於美國大部分全州性成就測驗，關於能力檢定的實徵研究，受試者的作答動機即有很大的差異，而受試者的作答動機也是研究中相當重要的中介變項。因此，相關研究可能容易受到此中介變項左右而有不同的結果，未來研究者需特別注意學生動機此中介變項，對於研究結果所造成的影響。

(二)特殊學生在語音調整下的表現有正面幫助：

Johnson (2000)、Weston (2002)、Tindal et al. (1998)、Calhoon, Fuchs, & Hamlett (2000)、Huynh, Meyer, & Gallant (2004)、Kosciolet & Ysseldyke (2000)與 Kettler et al. (2005)，上述七份研究當中的研究對象除了閱讀障礙與學習障礙外，還廣泛的包括了有閱讀困難的所有特殊學生，而調整方式則有錄音帶、錄影帶、語音合成、口語報讀及以學生 IEP 所設計的調整方式，除了 Calhoon、Fuchs 與 Hamlett (2000)同時利用口語報讀、合成音讀題及錄影帶讀題相互比較外，其餘皆是以一種調整方式作為研究主題，而在調整科目及內容方面，包括有閱讀與數學科，其中多以大規模全州性的成就測驗，僅 Weston (2002)及 Calhoon、Fuchs、與 Hamlett (2000)的測驗內容為自編成就測驗。

本研究與國內唯一相關實徵研究林筱

汶、陳明聰(2006)之研究結果不同，該研究發現特殊學生在語音調整下的表現有正面幫助，且不同能力學生與不同測驗呈現方式達交互作用水準。

表 6 為本研究與林筱汶、陳明聰(2006)之比較、表 7 為本研究與林筱汶(2005)的低閱讀一般數學能力組學生數學能力之比較，綜合分析後可發現，本研究結果低閱讀一般數學能力組學生相較於無語音提示的純文字版本與在有語音提示的真人錄音版、語音合成版的表現上分數確有增加，但未達顯著差異；林筱汶(2005)針對五年級一般及不同閱讀、數學能力學生進行數位化語音讀題在數學文字題的調整，研究結果發現接受數位化語音調整之後，只對低閱讀且一般數學能力學生之文字題解題表現有顯著幫助，對閱讀與數學能力皆為一般或皆低者無影響，而對一般閱讀與低數學能力學生有負面影響。其研究結果在一般閱讀一般數學能力組、低閱讀低數學能力組學生相符，但對於低閱讀一般數學能力組學生則存在著有無顯著之差異。

在測驗內容方面林筱汶、陳明聰(2006)

所使用的是周台傑(2002)「國民小學數學成就測驗－四年級組卷」，對於研究對象針對五年級的學生，該測驗內容皆為學校上過的課程，而本研究所使用的陳政見(2005)「國小高年級數學科成就測驗」研究對象針對國小高年級學生，以五年級學生而言有約一半題目是六年級部分還未上過的課程，而六年級學生，由於本研究施測過程在上學期的學期中所進行，因此六年級學生還是有部分題目還未接觸過。再由表 7 來比較，可發現在均是接受柯華葳(1999)所編之「基礎數學概念評量」的測驗下，本研究與林筱汶、陳明聰(2006)在一般閱讀低數學能力組學生的數學能力相較之下偏低，因此雖有語音提示的幫助，克服了數學解題歷程中成功要件的第一階段學生閱讀題目能力，但在學生數學能力偏低且所測驗內容超出學生學習範圍的條件下，使得低數學一般閱讀能力組學生在有語音提示的真人錄音版、語音合成版的表現上分數確有增加，但未達顯著差異，因此即使給予語音提示仍無法對其數學文字題解題表現有顯著的幫助。或許學生雖符合分類標準，但學生性質仍有差異。

表 6 本研究與林筱汶(2005) 語音調整策略對不同能力學生數學文字題解表現成效之研究比較

	本 研 究	林 筱 汶、陳 明 聰(2006)
研究對象	國小高年級 一般閱讀一般數學能力組 低閱讀一般數學能力組 低閱讀低數學能力組	國小五年級 一般閱讀一般數學能力組 一般閱讀低數學能力組 低閱讀一般數學能力組 低閱讀低數學能力組
測驗內容	陳政見(2005) 「國小高年級數學科成就測驗」	周台傑(2002) 「國民小學數學成就測驗－四年級組卷」
研究結果	接受電腦化數學成就測驗文字題型 之語音合成、人工錄音後，低閱讀	接受數位化語音調整之後，只對低 閱讀且一般數學能力學生之文字題

表 6 本研究與林筱汶(2005) 語音調整策略對不同能力學生數學文字題解表現成效之研究比較(續)

本 研 究	林 筱 汶、陳 明 聰(2006)
研究結果(續) 一般數學能力組學生有幫助，但未達顯著差異；而一般閱讀一般數學能力組、低閱讀低數學能力組學生在接受數位化不同語音版本調整後，其對數學文字題解題表現並無影響。	解題表顯有明顯幫助，對閱讀與數學能力皆為一般或皆低者無影響，而對一般閱讀與低數學能力學生有負面影響。

表 7 本研究與林筱汶、陳明聰(2006)語音調整策略對不同能力學生數學文字題解題表現成效之研究低閱讀一般數學能力組學生數學能力之比較

	基 礎 數 學 概 念					
	九 九 乘 法		空 格 運 算		三 則 運 算	
	分測驗 1	分測驗 2	分測驗 1	分測驗 2	分測驗 1	分測驗 2
本研究	0.70	0.94	0.62	0.98	0.41	0.46
林筱汶、陳明聰(2006)	0.91	0.93	0.90	0.90	0.50	0.61

另外，值得注意的是在七名低閱讀一般數學能力組學生中，研究者觀察到其中一名數學能力較高的學生，其答對率在真人錄音版顯然較純文字與語音合成版高，而純文字與語音合成版得分則相當（三版本試題該生的得分各為，純文字：7分；真人錄音：15分；語音合成：8分。），表8為該生在基礎數學概念各分測驗得分與本研究低閱讀一般數學能力組

學生各分測驗的平均得分比較，該生在人工語音的提示下，對其數學文字題解題表現有顯著的幫助，但在語音合成的提示下則無顯著的效果，此現象似乎意味著並非所有學生皆適合語音合成提示的調整，因此若要選擇語音類型來作語音提示調整，還需審慎考慮學生的接受性。

表 8 該生在基礎數學概念各分測驗得分

	基 礎 數 學 概 念					
	九 九 乘 法		空 格 運 算		三 則 運 算	
	分測驗 1	分測驗 2	分測驗 1	分測驗 2	分測驗 1	分測驗 2
該生	1.00	1.00	0.875	1.00	0.40	0.40
本研究	0.70	0.94	0.62	0.98	0.41	0.46

而本研究在語音合成軟體的選用上，是依據軟體取得方式容易為優先考量。因此，僅選取「Extra Speaker」及「文字 MP3 V2」兩種語音軟體來做比較，此種選取方式稍嫌不夠謹慎，導致所選取的語音合成軟體所合成的聲音檔，可能並非為發聲最自然且音質最為清晰的語音。另外，本研究所使用的語音都是以女聲版本發音，若在真人錄音及語音合成版本中再加入男聲版本的聲音檔來做比較，本研究設計內容就能更為周詳。

此外，根據 Elliott、Thurlow、Ysseldyke 與 Erickson（1997）提出學生在首次接受考試調整，不應是在考試或是評量的情況下。考試調整應該融合到班級教學中，提供學生平常在教學中常使用的調整方式，讓學生在考試前有充分的時間學習和練習調整的方式（張萬烽，2004）。本研究在受試者皆是第一次使用電腦化數學成就測驗並加上數位化語音提示的調整方式下進行測驗，可能對於習慣了傳統紙筆測驗的受試者而言，還未能熟練此種作答方式，以致於影響作答結果。由三種版本之測驗試題難易度檢測中可發現，在第一次以紙本呈現方式的重測中，三版本試題受試者的平均得分各為，純文字版：13.1、人工錄音版：12.15 與語音合成版：12.65；而在第二次建立為網路化線上測驗加上數位化語音讀題的調整之下，進行重測，三版本試題受試者的平均得分各為，純文字版：10.61、人工錄音版：11.17 與語音合成版：11.17，受試者在第二次重測的得分表現上有明顯較低的情形，或許可由此來佐證前述所提出之，學生需在平時上課有充分的練習才能進行正式的考試調整，才能讓有特殊需求的學生在考試調整的幫助下，能不受自身障礙狀況的影響，表現出所具備的解題實力，以發揮考試調整真正的功用。

結論與建議

本研究旨在探討語音合成與人工錄音對國小高年級低閱讀能力學生數學文字題解題表現差異，研究內容為：不同能力學生在電腦化數學成就測驗文字題型三種版本的表現差異。研究者將於本章研究結果提出結論與建議。

根據研究結果，可歸納結論如下：

國小高年級不同能力學生在電腦化數學成就測驗文字題型不同版本的測驗下，未有交互作用。

本研究結果顯示，不論在語音調整前或是語音調整後的真人錄音版、語音合成版，一般能力組學生在數學文字題解題表現上並無顯著差異，且對於一般能力組學生不論是在真人錄音版或語音合成版，其表現不會受到調整策略的語音版本不同而有所差異。在語音調整後，不論是在真人錄音版或語音合成版調整策略的語音版本測驗下，低閱讀一般數學能力組學生的表現皆優於調整前，但其幫助卻未達顯著的差異。對於低閱讀低數學能力組學生，在語音調整前後的表現則都未有顯著的影響。而在交互作用方面，不同閱讀和數學能力之國小高年級學生，並未因不同測驗呈現方式，其數學文字題解題表現即有顯著差異。

根據前述研究結果與研究發現，就未來研究方向提出若干建議，以供參考。

一、研究上的建議

（一）比較不同年級學生之表現

本研究中的研究對象是指「國小高年級學生」，然而在過去研究中也有發現對不同年級學生會有不同影響，因此研究者建議可再依此，研究不同年級的研究對象，進一步分析是否會因研究對象的年齡而有不同效果。

(二) 比較不同程度學生之表現

本研究的研究對象中低閱讀一般數學能力組學生之數學能力普遍偏低，故調整效果可能因此受到影響，建議可將學生能力在分成數種能力群，比較調整效果是否受影響。

(三) 比較不同題目難易度是否影響學生之表現

本研究的試題內容難易度 P 介於 .33~.71 之間，部分題目相較於一般測驗難易度稍嫌困難，因此研究者建議可再利用不同難易度之測驗內容來比較研究對象表現是否受影響。

二、教學上的建議

(一) 利用電腦化線上測驗加上數位化語音提示的調整方式進行教學

本研究在受試者皆是第一次使用電腦化數學成就測驗加上數位化語音提示的調整方式下進行測驗，可能對於習慣了傳統紙筆測驗的受試者而言，還未能熟練此種作答方式，以致於影響作答結果，因此研究者建議可在進行施測前，先利用電腦化線上測驗加上數位化語音提示的調整方式進行教學，讓受試者先熟悉作答及調整方式，再來比較其調整效果是否受影響。

參考文獻

一、中文部分

王文科(2002)：教育研究法。台北：五南。

余民寧(2005)：教育測驗與評量-成就測驗與教學評量。台北：心理。

何華國(2004)：特殊兒童心理與教育。台北：五南。

李佩蓉(2007)。南部三縣市國中小普通班教師對身心障礙學生考試調整方式的看法與實施之調查研究。國立台南大學特殊教育學系碩士論文，未出版，台南。

周台傑(2002)。國民小學數學成就測驗。台北：教

育部特殊教育工作小組。

林筱汶、陳明聰(2006)。語音調整策略對不同能力學生數學文字題解題表現成效之研究，載於嘉義大學主編，2006年特殊教育國際學術研討會論文集（頁113-131）。嘉義：嘉義大學。

柯華葳(1999)。閱讀理解困難篩選測驗。行政院國家科學委員會特殊教育工作小組。

柯華葳(1999)。基礎數學概念評量。行政院國家科學委員會特殊教育工作小組。

張世慧、藍瑋琛(2005)。特殊教育學生鑑定與評量。台北市：心理。

張萬烽(2005)。特殊學生在考試上的調整。屏師特殊教育，8，18-27。

胡永崇(2005a)。學習障礙學生的評量調整措施。屏師特殊教育，1-9。

胡永崇(2005b)。評量調整在學習障礙學生之應用。載於洪儷瑜、王瓊珠、陳長益(主編)，突破學習障礙：評量與因應之探討(頁 69-94)。台北市：心理。

陳明聰、張靖卿(2004)：特殊教育工作者對身心障礙學生測驗調整意見之調查研究。特殊教育與復健學報，12，55-80。

陳政見(2005)。國小高年級數學科成就測驗。嘉義：嘉義市崇文國民小學委託專案。未出版。

鍾素香(2002)。彈性調整考試評量方式是否落實於學習障礙學生。行政院國家科學委員會專題研究成果報告（報告編號：NSC90-2413-H-110-003），未出版。

謝秋梅(2004)。考試調整對學習障礙學生考試表現之探討。台東特教，59-65。

技專校院入學測驗中心(2006)。四季二專統一入學測驗簡章。雲林縣：技專校院入學測驗中心。

財團法人大學入學考試中心(2005a)。財團法人大學入學考試中心九十五學年度指定科目考試簡章。臺北：財團法人大學入學考試中心。

財團法人大學入學考試中心(2005b)。財團法人大

學入學考試中心九十五學年度學科能力測驗簡章。臺北：財團法人大學入學考試中心。
教育部中部辦公室(2006)：彰化考區九十五年國民中學學生基本學力測驗簡章。教育部中部辦公室九十四年十二月廿七日教中(二)字第0940604526 號書函。

二、英文部分

- Brown, P. J., & Augustine, A. (2001, April). *Screen reading software as an assessment accommodation: Implications for instructions for instruction and student performance*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Elliott, J., Thurlow, M., Ysseldyke, J., & Erickson, R. (1997). *Providing assessment accommodations for students with disabilities in state and district assessments* (Policy Directions No. 7). Minnesota, National Center on Educational Outcomes. Retrieved May 23, 2007, from <http://education.umn.edu/NCEO/OnlinePubs/Policy7.html>
- Fletcher, J. M., Francis, D. J., Boudousquie, A., Copeland, K., Young, V., Kalinowski, S., & Vaughn, S. (2006). Effects of Accommodations on High-Stakes Testing for Students With Reading Disabilities. *Exceptional Children*, 72(2), 136-150.
- Fuch, L. S., Fuchs, D., & Capizzi, A. M., (2005). Identifying Appropriate Test Accommodations for Students With Learning Disabilities. *Focus on Exceptional Children*, 37(6), 1-9.
- Helwig, R., & Tindal, G. (2003). An Experimental Analysis of Accommodation Decisions on Large-Scale Mathematics Tests. *Exceptional Children*, 69(2), 211-225.
- Huynh, H., Meyer, J. P., & Gallant, D. J. (2004). Comparability of Student Performance Between Regular and Oral Administrations for a High-Stakes Mathematics Test. *Applied Measurement in Education*, 17(1), 39-57.
- Johnson, E. S. (2000). The Effects of Accommodations on Performance Assessments. *Remedial And Special Education*, 21(5), 261-267.
- Kettler, R. J., Niebling, B. C., Mroch, A. A., Feldman, E. S., Newell, M. L., Elliott, S. N., Kratochwill, T. R., & Boli, D. M. (2005). *Assessment for Effective Intervention*, 31(1), 37-48.
- Kosciulek, S., & Ysseldyke, J. E. (2000). *Effects of a reading accommodation on the validity of a reading test* (Technical Report 28). Minneapolis, MN:University of Minnesota, National Center on Educational Outcomes. Retrieved January , 27, 2007, from the World Wide Web:<http://education.umn.edu/NCEO/OnlinePubs/Technical128.htm>
- Mayer, R. E.(1992). *Thinking, problem solving, cognition*. New York:Freeman.
- McKevitt, B. C., & Elliott, S. N. (2003). Effects and Perceived Consequences of Using Read-Aloud and Teacher-Recommended Testing Accommodations on a Reading Achievement Test. *School Psychology Review*, 32(4),583-600.
- Meloy, L. L., Deville, C., & Frisbie, D. A. (2002). The Effect of a Read Aloud Accommodation on Test Scores of Students With and Without a Learning Disability in Reading. *Remedial And Special Education*, 23(4), 248-255.
- Sireci, S. G., Li, S., & S carpati, S. (2003). *The*

- effects of test accommodation on test performance: A review of the literature.* (Center for Educational Assessment Research Report no. 485). Amherst, MA: School of Education, University of Massachusetts.
- Sireci, S. G., Li, S., & Scarpati, S., (Winter 2005). Test Accommodation for Students With Disabilities: An Analysis of the Interaction Hypothesis. *Review of Educational Research*, 75(5), 457-490.
- Sireci, S. G., Scarpati, S. E., & Li, S., (2005). The Effects of Test Accommodation on Test Performance: A Review of the Literature1. *Center for Educational Assessment Research Report no. 485*, Amherst, MA: School of Education, University of Massachusetts Amherst.
- Tindal, G., Heath, B., Hollenbeck, K., Almond, P., & Harniss, M.(1998) Accommodating Students with Disabilities on Large-Scale Tests: An Experimental Study. *Exceptional Children*, 64(4), 439-450.
- Weston, T. J. (2002) *The Validity of Oral Accommodation in Testing*. NAEP Validity Studies.
- CollegeBoard.com. (n.d.). Retrieved August 30, 2006, from <http://www.collegeboard.com/ssd/student>

附 錄

附錄一：評量調整國內外之相關研究摘要表

作者	研究對象	調整方式	調整科目	研究方法	研究結果
林筱汶、 陳明聰 (2006)	五年級一般 及不同閱 讀、數學能力 學生共 126 名	電腦化語 音讀題	數學 (文字題8題)	實驗設計 (相依樣 本、對抗平 衡)、問卷	1.接受電腦化語音調整之後，只對 低閱讀且一般數學能力學生之 文字題解題表現有幫助且達顯 著水準，但對閱讀與數學能力皆 為一般或皆低者無影響，而對一 般閱讀與低數學能力學生則有 負面影響。 2.不同能力學生與不同測驗呈現 方式交互作用達顯著水準。 3.對電腦化的語音調整策略態度 為良好，但低閱讀能力的學生較 一般閱讀能力的學生肯定此策 略之「語音幫助程度」。
Fletcher et al. (2006)	三年級識字 或解碼困 難、一般學生 各 91 名	1.延長並將 考試時 間分成 兩段來 完成測 驗。 2.唸出語詞 給學生 聽。 3.念出題 幹，讓學 生獨立 閱讀文 章段落， 並盡可 能回答 學生理 解困難 的問題。	閱讀 (德州認 知與技能 評量)	實驗設計 (獨立樣本)	1.只有識字困難或解碼困難學生 在調整後的表現優於未調整的 表現，但並未達顯著水準。 2.一般學生在接受調整前後，表現 並沒有顯著差異。 3.調整對識字困難或解碼困難的 學生比對一般學生影響更大。

附錄一：評量調整國內外之相關研究摘要表（續）

作者	研究對象	調整方式	調整科目	研究方法	研究結果
Johnson (2000)	四年級一般、閱讀障礙學生共 115 名	口語報讀	數學（選擇題 40 題） 華盛頓州學生學習評(WASL)	實驗設計 （相依樣本）	在調整下一般生與閱讀障礙學生表現上皆有進步，且閱讀障礙學生的進步幅度較一般生高，然而交互作用假設未達顯著水準。但作者推斷此研究支持對閱讀障礙學生提供語音調整的提示。
Weston (2002)	四年級一般與學習障礙學生共 119 名	口語報讀	數學 (計算題 8 題、文字題 14 題)	實驗設計 （相依樣本、對抗平衡）	1.在調整下一般生與學障生成績上皆有進步，但學障生的進步幅度較一般生高，且閱讀能力高的學生，成績增加的較少。 2.不同能力學生與不同測驗呈現方式交互作用達顯著水準。 3.在調整下，學障學生在文字題增加的分數較計算題多。 4.報讀增加數學測驗對學障生的考試效度。
Tindal et al. (1998)	四年級一般與特殊學生共 164 名	1.直接於題本上作答 2.施測人員讀題兩次	1.閱讀 2.數學（選擇題 60 題） (大規模州測驗)	實驗設計 (獨立樣本)	1.直接於題本上作答的調整方式並無顯著影響。 2.調整對一般學生無顯著幫助，但對於特殊學生成績的增加則有達顯著水準，且不論是在調整或是未調整的成績上一般生得分皆高於特殊生。不同能力學生與不同語音測驗呈現方式交互作用達顯著水準。
Calhoon, Fuchs, &Hamlett (2000)	九、十、十一、十二年級的學習障礙學生共 81 名	1.教師讀題 2.電腦呈現+合成音讀題 3.電腦呈現+錄影帶讀題	數學實作 評量	實驗設計 （相依樣本）、問卷	1.經過調整後學生表現分數皆提高，且不論何種報讀形式的調整並無顯著差異。 2.有三分之二的學生表示較喜歡電腦報讀。
Huynh, Meyer, &Gallant (2004)	一組接受標準化測驗的一般學生 85638 名、一組接受標準化測驗的特殊學生 2642 名一組接受評量調整的特殊學生 934 名	口語報讀	數學南加州高中畢業試 (HSEE)	準實驗設計	1.接受調整的特殊學生組表現較未接受調整的特殊學生組好，且達顯著水準。 2.不同能力學生與不同語音測驗呈現方式交互作用達顯著水準。 3.一般學生組的表現皆優於其他兩組特殊學生。

附錄一：評量調整國內外之相關研究摘要表（續）

作者	研究對象	調整方式	調整科目	研究方法	研究結果
Kosciolet & Ysseldyke (2000)	郊區小學三至五年級一般與閱讀又困難的特殊學生共 31 名	錄音帶	閱讀 (加州成就測驗)	準實驗設計(相依樣本)、開放式問卷	1.一般學生不論在調整前或調整後的表現皆優於特殊學生，而特殊學生在調整後的分數提高，且達顯著水準，一般學生調整前與調整後的分數並無顯著差異。 2.不同能力學生與不同語音測驗呈現方式交互作用達顯著水準。 3.一般學生偏向喜愛未調整的考試方式。
Kettler et al. (2005)	119 名四年級一般與特殊學生，78 名八年級一般與特殊學生	依學生的 IEP 設計	數學、閱讀 (TerraNova 基本技能綜合考試)	實驗設計 (相依樣本)	1.在調整下四年級特殊學生測驗的分數進步較一般學生多，且調整交互作用假設達顯著水準。 2.閱讀測驗調整結果又較數學測驗調整結果更為顯著。
Helwig & Tindal (2003)	八年級特殊學生(學習障礙 70%、語言障礙 8%、嚴重情緒障礙 5%、智能障礙 5%；白種人 59%、黑種人 28%、西班牙裔 7%；男 63%、女 37%)及其導師、數學科教師、資源班教師	錄影帶讀題	數學 30 題	準實驗設計、教師問卷(事後回溯)	1.語音調整對七、八年級學生的表現無影響；四、五年級的特殊學生及四年級低閱讀高數學能力學生有顯著的正面影響；五年級低閱讀高數學能力學生無影響。 2.不同能力學生與不同測驗呈現方式交互作用假設未達顯著水準。 3.大體上，約 45%國小及 57%國中學生因教師錯誤判斷而申請了評量調整。
Meloy et al. (2002)	六、七、八年級的一般與閱讀障礙生共 260 名	口語報讀	自然、閱讀、數學 (愛荷華州基本技能測驗)	實驗設計 (相依樣本)	不論在一般或閱讀障礙學生的表現三科皆提高其分數，且達顯著差異此推斷提供口語報讀的調整給閱讀障礙學生是缺乏公平性的。
Brown & Augustine (2001)	一般與閱讀障礙學生共 206 名	電腦合成語音讀題	數會、自然	實驗設計 (相依樣本)	學生不論在調整前或調整後的表現並無顯著差異。
McKevitt & Elliott (2003)	79 名八年級一般與閱讀或語音有困難的特殊學生、48 名教師	錄音帶	閱讀 (TerraNova 基本技能閱讀考試)	實驗設計 (相依樣本)、訪談	1.學生無論在調整前或調整後的表現並無顯著差異。 2.延長時間為教師最常選用的調整方式。

附錄二：

有關國內外升學考試身心障礙考生應考申請調整服務項目，包括九十五學年度指定科目考試、九十五學年度學科能力測驗、四技二專統一入學測驗、國民中學學生基本學力測驗、美國大學新生入學測驗（財團法人大學入學考試中心，2005；技專校院入學測驗中心，2006；教育部中部辦公室，2006；CollegeBoard.com，2006），將各調整服務項目分述如下：

1. 國民中學學生基本學力測驗：

- (1) 特殊桌椅。
- (2) 行動不便者安排在一樓或有電梯之試場。
- (3) 現場報讀。
- (4) 可申請試卷包括放大 1.5 倍的試卷、點字試卷，及純文字點字試卷電子檔。
- (5) 延長作答 20 分鐘。
- (6) 重騰答案卷。
- (7) 准用盲用電腦及列表機(可自備但須附照片審核)。
- (8) 提早 5 分鐘進場。
- (9) 准用調頻助聽器、盲用算盤(可自備但須附照片審核)。
- (10) 准用擴視機、放大鏡、點字機(可自備但須附照片審核)。

2. 九十五學年度學科能力測驗：

- (1) 點字試題本或放大為 A3 之試題本。
- (2) 協助考生閱讀之輔具。
- (3) 最多延長 20 分鐘應考時間。
- (4) 原答案卡放大為 A4 之影本或空白答案紙作答選擇題或點字機作答。
- (5) 協助考生紀錄答案之輔具。
- (6) 提早 5 分鐘進場。
- (7) 其他功能性障礙所需之特別服務。

3. 九十五學年度指定科目考試

- (1) 點字試題本或放大為 A3 之試題本。
- (2) 協助考生閱讀之輔具。
- (3) 最多延長 50 分鐘應考時間。
- (4) 原答案卡放大為 A4 之影本或空白答案紙作答選擇題或點字機作答。
- (5) 協助考生紀錄答案之輔具。
- (6) 提早 5 分鐘進場。
- (7) 其他功能性障礙所需之特別服務。

4. 四技二專統一入學測驗

- (1) 將試場安排於一樓或適宜之試場為原則。
- (2) 視障生自行攜帶不做任何記號之檯燈或放大鏡。
- (3) 提供視障考生影印放大 1.5 倍之試題紙。
- (4) 每科目之考試時間依一般考試時間再延長 20 分鐘，但兩科目間休息時間減少 20 分鐘。
- (5) 以空白答案紙作答。

- (6)提早 5 分鐘進入試場準備。
- (7)其他項目。
- (8)宣佈事項寫在黑板上，並以紙版大字提醒。

5. 美國大學新生入學測驗

- (1)小組應試。
- (2)個別應試。
- (3)限制螢幕以減少分心。
- (4)提供特殊燈光。
- (5)提供特殊聲音處理。
- (6)提供調整式或特殊家具。
- (7)在監考人員的監考下選擇座位。
- (8)優先就坐。
- (9)放大題目(14pt；20pt)。
- (10)減少每頁題數。
- (11)閱讀機。
- (12)有顏色的作答紙。
- (13)在重要之作答方式或注意事項，可以螢光筆畫記的方式，提醒學生注意遵守。
- (14)以口語或手語方式呈現指導語。
- (15)擴視機。
- (16)擴大溝通器。
- (17)擴音機。
- (18)有顏色的覆蓋物。
- (19)點字。
- (20)點字圖解。
- (21)書面回答的點字設備。
- (22)題本外罩著塑膠版。
- (23)允許考試中間有小段休息。
- (24)允許延長考試時間。
- (25)專人抄錄。
- (26)專人口譯。
- (27)將答案錄音後轉成逐字稿。
- (28)使用電腦文書處理回答。
- (29)允許直接在題本上作答。
- (30)放大答案格。
- (31)分成數日考試。
- (32)在一日中的特定時間考試。

**The effectiveness of different speech output accommodations
(synthesized & digitized voice) on the performance of math test for
5th and 6th grade students with disability**

Qiao-Yun Huang
National Chiayi University

Ming-Chung Chen
National Chiayi University

Chen Cheng-Chien
National Chiayi University

ABSTRACT

This research aimed to explore the effectiveness of different speech output accommodations (synthesized & digitized voice) on the performance of math test for 5th and 6th grade students with disability. There were 34 students participated in this study. They were grouped into three groups based on their reading and math abilities. Each student took tests that displayed in three conditions, text only, digitized recoding and synthesized.

The finding of this study results are:(1) Students with general abilities have no effect on tests of three different types and students with poor reading and poor math abilities too. However, students with poor reading but general math abilities performed better on two different types of speech output accommodation, but there is no significant difference between text only, digitized voice and synthesized. (2) This paper didn't support the "interaction hypothesis".

Key words : test accommodation, speech output accommodation, reading disabilities, synthesized voice, digitized voice.