

教育科學研究期刊 第六十一卷第四期

2016 年，61 (4)，33-57

doi:10.6209/JORIES.2016.61(4).02



國小二至六年級朗讀流暢度篩檢準確度及切截點分析

蘇宜芬

國立臺灣師範大學
教育心理與輔導學系

李孟峰

國立臺灣大學
心理學系

張祐瑄

國立臺灣師範大學
教育心理與輔導學系

黃鈺茜

臺南市立土城高級中學

摘要

本研究主要目的是以閱讀的自動化工論為基礎，為國小二至六年級朗讀流暢度做篩檢準確度及最佳切截點分析。研究中，以閱讀理解對率作為效標變項，朗讀流暢度 z 分數作為預測變項，然後採用「接受者操作特徵」(Receiver-Operating-Characteristic, ROC) 分析，根據「ROC 曲線下面積」之大小，來檢測採用朗讀流暢度篩檢因認字效能不足導致閱讀理解困難之學童的準確度。此外，也根據 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率等數據，對朗讀流暢度做最佳切截點的分析。研究結果顯示，以朗讀流暢度來篩檢因認字效能不足導致閱讀理解困難的學童，具有中等的準確度。而且以二至六年級上下學期的資料而言，朗讀流暢度 z 分數最佳切截點都是-0.5。本研究也提供國小二至六年級上下學期的朗讀流暢度平均值資料，以供教育研究者及教師瞭解二至六年級學童認字效能平均成長情形之參考。

關鍵詞：切截點、朗讀流暢度、國小學童、閱讀理解、ROC 分析

通訊作者：蘇宜芬，E-mail: yifensu@ntnu.edu.tw

收稿日期：2015/08/03；修正日期：2015/12/07；接受日期：2016/02/03。

壹、緒論

流暢的認字技能是閱讀理解的基本條件，對於國小階段的學童而言，發展流暢的認字技能是重要的學習任務（Chall, 1996）。為什麼認字技能流暢是閱讀理解的基本條件，過去曾有三個理論加以闡釋。

最早的是 LaBerge 與 Samuels（1974）「閱讀的自動化訊息處理論」（a theory of automatic information processing in reading）。所謂「自動化」指的是：原本需要非常專注才能完成的技能，經過一段相當時間的練習之後，變成不需投注很多的注意力即可完成的歷程。達到「自動化」程度的技能有幾個特徵：即正確率高，速度快，不太費力、費神。認字技能若達自動化水準有助於閱讀理解的增進，主要是因為每個人的認知資源（如：注意力、工作記憶容量）都是有限的，當讀者的認字技能達到自動化程度時，他可以耗費較少的認知資源在認字歷程上，而有較多的認知資源用在理解上。相反地，如果讀者的認字技能未達自動化水準，他就需要耗費很多的認知資源在認字上，以致沒有多少剩餘的認知資源可以用來理解。因此，認字技能的自動化是良好閱讀理解的一個重要條件。

1980 年，Stanovich 以 LaBerge 與 Samuels（1974）的理論為基礎，加了些新的詮釋，提出「交互補償」（interactive compensatory explanation of reading fluency）理論。Stanovich 認為弱讀者（poor readers）與優讀者（good readers）主要的差異在於閱讀時對於文本（text）的處理方式不同。弱讀者在閱讀時，比較無法使用自動化的認字技能，而需使用較耗費認知資源的方式去辨識文字。為了補償文字辨識上的困難，弱讀者會運用文章脈絡輔助文字的辨識。然而，使用脈絡策略辨識文字是很耗費認知資源的，因此弱讀者就沒有足夠的認知資源用以理解文章。相反地，優讀者可使用自動化、較不耗費認知資源的方式去辨識文字，因此他們可以將較多的認知資源用於由上而下（top-down）的處理歷程以理解文章（Stanovich, 1980）。

此外，Perfetti（1985）也曾提出「語文效能理論」（verbal efficiency theory）來解釋認字效能對閱讀理解的影響。這個理論的主要內容是：讀者的注意力及工作記憶的資源是有限的，隨著閱讀能力的發展，讀者在認字的處理上會變得更有效能，因此可釋放注意力及工作記憶的資源，用以統整文章的命題與建構意義；相反地，閱讀能力弱的讀者由於認字效能低，需耗費較多的認知資源於認字處理上，以致無足夠的認知資源可用於文章的理解。

這三個理論，不論是 LaBerge 與 Samuels（1974）「閱讀的自動化訊息處理論」、Stanovich（1980）的「交互補償理論」、或是 Perfetti（1985）的「語文效能理論」，基本上均贊同人類的認知資源是有限的，且閱讀歷程包含認字及理解兩部分的處理。如果讀者在認字的處理上耗費太多的認知資源，就無足夠的認知資源可用於文章理解；但若讀者的認字處理不太需要耗費認知資源，則可將大部分的認知資源用於文章理解上。

三個理論的不同之處在於，Perfetti 的理論比 LaBerge 與 Samuels 的理論更具體地指出，認字效能低的讀者，需耗費較多的認知資源於認字處理上，所以比較沒有足夠的認知資源可用於文章意義的建構及文章命題的統整。Stanovich 的理論則是點出弱讀者與優讀者在閱讀處理歷程上的差異。弱讀者因為不具有自動化的認字技能，為了補償文字辨識上的困難，會運用文章的前後文脈絡輔助文字的辨識，但是這是一個比較耗費認知資源的文字辨識策略。由於弱讀者把大部分的認知資源用在文字辨識上，所以，比較沒有多餘的認知資源用於段落或文章整體意義的建構統整。相反地，優讀者具有自動化的認字技能，比較不需耗費認知資源去辨識文字，因此他們可以把較多的認知資源用於由上而下的處理歷程，也就是以文章的標題、次標題、或自行提取的上層概念，對文章內容進行統整性的理解。根據上述三個理論，認字技能是否達到自動化水準（亦即認字效能的高低、或認字技能是否流暢）是影響閱讀理解表現的重要因素，尤其影響文章意義的統整及深層的理解。

如果認字效能的高低是閱讀理解表現的重要影響因素，那麼認字效能如何測量呢？過去文獻中用來測量認字效能的方法大致可分為兩類，一類是以獨立於文章脈絡的字單 (word list) 來測量認字速度，另一類是選取文章請學生朗讀，測量其朗讀速度。

第一類的方法又有幾種不同的測量方式。第一種方式是請學生唸一份依難度排列的常用字單，然後測量其唸字速度，例如：Torgesen、Wagner 與 Rashotte (1999) 所編製「認字效能測驗」(Test of Word Reading Efficiency) 中的「常用字效能測驗」(Sight Word Efficiency subtest)，以及洪儷瑜、陳秀芬、王瓊珠與張郁雯 (2012) 所編的「常用字流暢性測驗」，即採此種方法。第二種方式是將一篇文章的字隨機排列，製成一份字單，請學生以最快的速度唸這份字單，然後測量其唸字速度，Jenkins、Fuchs、van den Broek、Espin 與 Deno (2003) 的研究即採用這種方法。第三種方式是將字一一呈現於電腦螢幕中央，請學生盡快唸出，然後測量其平均反應時間 (Schwanenflugel et al., 2006)。第四種方式是將真字與假字隨機一一呈現於電腦螢幕中央，請學生盡快按鍵判斷認得或不認得電腦螢幕中的字，然後測量其平均反應時間 (蘇宜芬、陳學志，2007)。這一類的方法可測量認字的處理速度，但無法評估到認字處理在認知資源上的耗費程度。此外，第三、四種方式涉及電腦程式的設計及平均反應時間的計算，對於第一線的教師而言，在測量工具的取得上比較不方便。

第二類的測量方法則是選取文章請學生朗讀，測量其朗讀流暢度。關於朗讀流暢 (fluent reading) 一般的定義為：能夠正確、快速且帶有聲音表情地朗讀文章 (National Institute of Child Health and Human Development [NICHD], 2000)。然而，如何客觀地測量「聲音表情」，有其困難度。Rice (1981) 的研究中，以三個向度測量朗讀的聲音表情，並考驗其信度。結果發現：重音 (stress) 的信度係數只有 .02，停頓 (pause) 的信度係數為 .13，音調 (pitch) 的信度係數為 .17。由於測量困難且不易具有可接受的信度，所以，過去在國內、外文獻中，關於認字效能的評量，最常被建議使用的指標為文章朗讀速率 (oral reading rate，也稱為朗讀流暢度，

oral reading fluency)，亦即測量學生朗讀文章時 1 分鐘所能唸對的字數 (Fuchs, Fuchs, Hosp, & Jenkins, 2001)。然而，不論國內或國外的研究報告，所提供的文章朗讀速率資料常常是年級常模 (亦即各年級的平均數、標準差)，以及文章朗讀速率和標準化閱讀成就測驗或標準化閱讀理解測驗的相關 (王梅軒、黃瑞珍, 2005; 葉靖雲, 1993, 1998; Fuchs, Fuchs, Eaton, & Hamlett, 2000; Fuchs, Fuchs, & Maxwell, 1988; Jenkins, Fuchs, Espin, van den Broek, & Deno, 2000)，或是朗讀流暢度對標準化閱讀理解測驗分數的預測力 (Jenkins et al., 2003; Schwanenflugel et al., 2006)，未見研究探討朗讀速度需達多少以上才能讓讀者對文章的閱讀達到較佳的理解水準。所以，上述兩項朗讀速度的測量資料在學理及應用實務上都有其不足之處。

如前所述，不論是 LaBerge 與 Samuels (1974)「閱讀的自動化訊息處理論」、Stanovich (1980)的「交互補償理論」、或是 Perfetti (1985)的「語文效能理論」，均認為認字技能達到自動化水準的讀者由於在認字處理上不太耗費認知資源，因此能有足夠的認知資源可用於理解文章；相反地，認字技能未達自動化水準的讀者由於在認字處理上需耗費大量認知資源，因此沒有足夠的認知資源可用於理解文章。所以在認字效能水準的評估上，除需考慮朗讀速度外，閱讀理解的表現也很重要，因為它反映了認字處理耗費認知資源的程度。Samuels (1988) 即曾建議應用雙重作業法 (dual-task method) 來評量認字自動化的程度，方法是測量學生的朗讀流暢度 (oral reading fluency)、閱讀理解及朗讀的抑揚頓挫表情。

「雙重作業法」是常運用於「自動化」研究中的一個方法，所根據的基本原理是：自動化歷程不會被一項需要耗費注意力的競爭作業所干擾或打斷 (Norman & Bobrow, 1975; Posner & Snyder, 1975; Shiffrin & Schneider, 1977)。雙重作業法典型的程序是：受試者要同時執行兩項不同的作業，而且這兩項作業的速度及正確率都要予以記錄。例如：其中一項作業是從 20 開始做加 3 的運算 (屬於認知作業)；另一項作業是根據三個不同頻率的聲音分別按三個不同的鍵 (屬於知動作業)。如果受試者在其中一項作業或這兩項作業的執行已達自動化水準，那麼在同時執行兩項作業時，彼此干擾的情形就會較少或完全沒有干擾。也就是在這兩項作業上的反應時間或錯誤率不會增加。相反地，如果受試者在其中任一項作業或兩項作業皆未達自動化水準，那麼干擾的情形就會發生，亦即反應時間或錯誤率就會增加。因此，雙重作業的干擾程度常被作為測量自動化程度的方法，尤其是用於注意力分配的測量。

Samuels (1988) 為什麼會建議應用雙重作業法來評量認字自動化的程度，這跟技能自動化的發展階段有關。LaBerge 與 Samuels 在 1974 年提出一個學習理論，認為技能自動化的發展是個三階段的歷程。其中，第一階段是從初學到精熟的歷程，這個階段的特徵是作業的執行仍需耗費大量的認知資源，但熟練度 (即正確率及速度的表現) 則持續地提高。第二階段則是從精熟發展到自動化的歷程。在此，精熟指的是作業的表現 (包括正確率及速度) 能持續維持在一定的高水準。這個階段的特徵是：由於正確率與速度已持續維持在一定的高水準，

所以進步比較小，但是，認知資源的耗費則顯著降低。第三階段則是自動化的階段，在這個階段只需耗費很少的認知資源即可達到穩定而高品質的作業表現。根據這個理論，要測量一項技能是否達到自動化的水準，只看正確率及速度的表現是不夠的，還要看認知資源（即注意力、工作記憶）的耗費程度，因此雙重作業法是比較合適的方法。由於在閱讀的過程中，涉及「認字」及「文章理解」兩個歷程，根據雙重作業法的邏輯，只要認字技能達到自動化的水準，那麼認字歷程的進行就不會干擾文章理解的執行，所以閱讀理解表現較佳。再者，由於讀者對文章內容有良好的理解，所以朗讀時也會有較貼切的斷詞及抑揚頓挫。然而，如前所述，「聲音表情」的測量有其困難，不容易具有可接受的信度，因此，本研究是以自同一篇文章測得的朗讀流暢度與閱讀理解程度進行分析，此乃不同於以往研究之處。

此外，2002年「美國教育進步評量」(National Assessment of Educational Progress, NAEP)曾對朗讀流暢進行研究，發現根據學生朗讀第1分鐘所測得的朗讀速度，往往比根據學生讀完整篇短文所測得的朗讀速度來得快(Daane, Campbell, Grigg, Goodman, & Oranje, 2005)，也就是前者容易高估學生的朗讀流暢度。因此，本研究所測量的朗讀流暢度是以學生唸完整篇短文的所需時間來計算。另外，依明尼蘇達大學(University of Minnesota)一系列有關朗讀流暢度的研究，顯示兩篇以上文章所測得的朗讀流暢度要比單篇文章所測得的速率來得穩定(Deno, 1985)，因此，本研究的朗讀流暢度均以學童朗讀兩篇文章的速率加以平均。

另外，就教育應用而言，如果認字技能流暢是閱讀理解的基本條件，那麼若有個參考的篩檢指標，提供教師瞭解哪些學生可能是因認字技能不夠流暢，導致閱讀理解困難的高危險群，那麼對於教師會很有幫助，有助於及早發現認字效能不足的學生，接受補救教學。關於理解困難的閱讀水準要如何界定，本研究採用美國閱讀教育學者所提出「挫折閱讀水準」(frustration reading level)的概念(Taylor, Harris, Pearson, & Garcia, 1995)。所謂「挫折閱讀水準」是指學生無法讀懂文章，且對於閱讀感到挫折。關於「挫折閱讀水準」的界定，在美國的標準化測驗及非正式閱讀評量(informal reading inventories)，大多採取閱讀理解對率低於50%為標準(Ekwall & Shanker, 1993; Hoover, Dunbar, & Frisbie, 2001; Johns, 1997)，只有很少數採取閱讀理解對率低於70%為挫折閱讀水準(Leslie & Caldwell, 2006)。本研究採取閱讀理解對率低於50%為挫折閱讀水準的界定標準，主要是基於兩點的考慮。第一個考量是，本研究在學童朗讀之後所採用的閱讀理解測驗，有60%的題目是屬於測量「字詞彙知識」或「字面理解」的題目，學生的答對率若在50%以下，意謂學生連文章裡的字詞彙知識或字面的理解都無法掌握完全。第二個考量是，答對率50%以下幾近於猜測，反映出學生對於文章的理解程度很弱，這樣的閱讀理解水準的確會令學生感到挫折，因此，以答對率低於50%作為「挫折閱讀水準」的界定標準，會比答對率低於70%更具有說服力。所以，本研究以閱讀理解對率為效標，採閱讀理解對率低於50%作為理解困難的界定水準，分析朗讀流暢度不同切截點(cutoff score)對於篩檢因認字效能不足導致閱讀理解困難之學生的效果，並依分析結果提出

最佳切截點的建議。

此外，在效度資料的提供上，過去，心理與教育領域的測量工具多以效標關聯效度提供效度證據。然而，效標關聯效度只提供某一群研究對象其預測變項與效標變項之間的關聯強度，但就篩檢工具或診斷測驗而言，讓使用者瞭解測量工具能夠篩檢出高風險學生（at risk student）的正確率有多高，也很重要。因此近年來，國外學者建議篩檢工具或診斷測驗的發展，應提供預測效度（predictive validity）的資料，也就是測量工具對於有學習困難風險的學生能正確檢出的程度有多高，而對於沒有學習困難的學生，不會錯誤檢出的程度又是多少（Rathvon, 2004）。因此，本研究除了仍提供朗讀流暢度與閱讀理解答對率的相關係數資料外，也提供預測效度的資料，包括以「接受者操作特徵」（Receiver-Operating-Characteristic，簡稱 ROC 分析）所得的「ROC 曲線下面積」（area under the ROC curve, AUC）之大小，來檢測以朗讀流暢度篩檢挫折閱讀水準學童的準確度。然後，再分析當「朗讀流暢度 z 分數」的切截點分別設在 0.0、-0.5、-1.0、-1.5 時，篩檢挫折閱讀水準學童的 χ^2 值、命中率（hit rate）、敏銳度（sensitivity, SEN）、錯誤拒絕率（false negative rate）及錯誤接受率（false positive rate），以進行朗讀流暢度最佳切截點的分析。

在上述預測效度的資料中，ROC 分析是用來檢測診斷工具或篩檢測驗準確度的方法，早期大多運用在醫學界，近年也逐漸為特教界所採用。透過 ROC 分析，可得到「接受者操作特徵曲線」（Receiver Operating Characteristic Curve，簡稱 ROC 曲線）。ROC 曲線是以「錯誤接受率」為橫座標，「正確接受率」為縱座標，由無數切截點數值所算出無數對「錯誤接受率」及「正確接受率」作圖所構成（王天苗、廖華芳，2007）。所謂「正確接受率」是指正確篩檢出問題個案之人數占所有問題個案數的比率，也就是「敏銳度」，以表 1 為例，就是「 $VP \div (VP + FN)$ 」。「錯誤接受率」則是被錯誤篩檢為問題個案的人數占所有非問題個案數的比率，以表 1 為例，錯誤接受率就是「 $1 - \text{特異度}$ 」，亦即「 $FP \div (FP + VN)$ 」。在 ROC 分析裡，「ROC 曲線下面積」的大小，可以反映篩檢工具的區辨力或準確度，AUC 面積愈大，表示篩檢工具的準確度愈高。一般而言，ROC 曲線下面積介於 .50~.70 時，表示診斷準確性較低；面積為 .70~.90 時，表示診斷準確性中等；面積為 .90 以上時，表示診斷準確性較高（林傑斌、林川雄、劉明德，2004）。

另外，以 ROC 曲線來判斷最佳切截點，通常切截分數設得愈高，敏銳度（即正確接受率）就愈高，但錯誤接受率也愈高；相反地，切截分數設得愈低，敏銳度就愈低，錯誤接受率也愈低。所以，在最佳切截點的選擇上，還需綜合考慮 χ^2 值、命中率及錯誤拒絕率。

χ^2 值考驗的是當朗讀流暢度的切截點設在某個分數時，這個切截點對於「高風險」及「非高風險」學生的分類，和以閱讀理解答對率對於閱讀「挫折」及「不挫折」學生的分類，兩者間是否一致。最佳的朗讀流暢度切截點對於「高風險」（at risk）及「非高風險」（not at risk）學生的分類，應該和以閱讀理解答對率對於閱讀「挫折」及「不挫折」學生的分類有很好的

表 1
預測效度指標

閱讀理解表現	根據朗讀流暢度所做的篩檢決定		預測效度指標
	高風險	非高風險	
閱讀理解答對率低於 50% (挫折閱讀水準)	正確接受的人數 (valid positive, VP)	錯誤拒絕的人數 (false negative, FN)	敏銳度 $VP \div (VP + FN)$
閱讀理解答對率等於 或高於50%	錯誤接受的人數 (false positive, FP)	正確拒絕人數 (valid negative, VN)	特異度 $VN \div (VN + FP)$

一致性。

「命中率」則是指正確接受和正確拒絕個案數占總篩檢人數的比率，以表1為例，命中率就是「 $(VP + VN) \div (VP + FN + FP + VN)$ 」。愈是準確的朗讀流暢度切截點，愈能正確檢出閱讀理解挫折的學生，且能正確排除閱讀理解沒有困難的學生，而命中率正是在檢視這個部分。此外，在切截點的選擇上，路君約（1989）曾提醒，在訂定切截點時，需特別注意「錯誤拒絕率」，也就是問題個案沒有被篩檢出來的人數占所有問題個案數的比率，以表1為例，就是「 $FN \div (VP + FN)$ 」。因為，如果施測目的是為了篩檢高風險學生，那麼篩檢工具及切截點就不僅要有高「敏銳度」，而且被「錯誤拒絕」的人數比例需盡量減到最低。因此，本研究也會評估不同切截點的「錯誤拒絕率」。

歸納以上說明，在預測效度的分析中，本研究會先以「ROC 曲線下面積」來檢測以朗讀流暢度篩檢挫折閱讀水準學童的準確度是否在中等以上（即 .70 以上）。在最佳切截點的選擇上，會以 χ^2 值考驗當朗讀流暢度的切截點設在某個分數時，這個切截點對於「高風險」及「非高風險」學生的分類，和以閱讀理解答對率對於閱讀「挫折」及「不挫折」學生的分類，兩者之間是否一致。此外，也會評估哪個切截點具有較高的命中率與敏銳度（即正確接受率）、較低的錯誤接受率與錯誤拒絕率，以作為最佳切截點的建議。

此外，不論是一般的閱讀教學或是補救教學，教師也需要一些研究數據，提供他們瞭解國小二至六年級學童認字效能平均成長情形之參考。但是，過去國內的朗讀速率研究所提供的多是以橫斷研究法所得的單一時間點資料（王梅軒、黃瑞珍，2005；葉靖雲，1993，1998）。因此，本研究也對二至六年級學童，在同一學年的兩個學期分別進行施測，以蒐集在學年內朗讀流暢度成長資料，提供教育研究者及教師瞭解學童認字效能成長之參考。

綜上所述，本研究主要目的是：一、以閱讀的自動化理論為基礎，為國小二至六年級分析朗讀流暢度不同切截點對篩檢因認字效能不足導致閱讀理解困難（也就是挫折閱讀水準）之學童的效果，並依分析結果提出最佳切截點的建議。二、提供國小二至六年級的學年內朗讀流暢度成長資料，以提供教育研究者及教師瞭解學童認字效能平均成長情形之參考。

貳、方法

一、研究對象

本研究二、三年級的研究對象來自新北市的七所國小，四、五、六年級的研究對象來自新北市的六所國小。各年級經家長兩個學期都同意子女參與，且每個學期均完成團體測驗及個別測驗的有效受試人數分別為：二年級 144 人（男生 70 人、女生 74 人）、三年級 146 人（男生 77 人、女生 69 人）、四年級 148 人（男生 76 人、女生 72 人）、五年級 172 人（男生 93 人、女生 79 人）、六年級 166 人（男生 77 人、女生 89 人）。各年級研究對象第一學期的平均年齡分別是：二年級 7 歲 9 個月、三年級 8 歲 7 個月、四年級 9 歲 8 個月、五年級 10 歲 7 個月、六年級 11 歲 8 個月。

二、研究工具

為瞭解本研究所測得的朗讀流暢度與標準化閱讀理解測驗的相關，並與過去這方面的研究做比較，因此，參與本研究的學童除了兩個學期都需接受朗讀流暢度的個別測驗外，也需接受標準化閱讀理解測驗的團體施測。茲將本研究所採用的研究工具說明如下。

（一）國民小學二至六年級閱讀理解篩選測驗

本研究所採用的標準化閱讀理解測驗是柯華葳與詹益綾（2006）編製的「國民小學二至六年級閱讀理解篩選測驗」。這份測驗的編製目的為測量國小二至六年級學童的閱讀理解能力，以作為篩選閱讀障礙者的初步評估工具。測驗內容包含「部分處理」（多義字題、命題組合）及「本文處理」（句子理解、短文理解）兩部分。二至六年級的題數介於 19~32 題之間。每個年級都有 A、B 兩個複本。本研究的各年級在第一學期均以複本 A 施測，第二學期均以複本 B 施測。此測驗各年級 A、B 兩複本的 Cronbach's α 信度係數介於 .70 至 .86 之間。效度資料方面，在常模樣本中，各年級高分組學童在不同類型題目上的得分均比低分組高。另外，此測驗與洪儷瑜、張郁雯、王瓊珠與陳秀芬（2006）所編製「識字量評估測驗」的得分相關，介於 .53 至 .70 之間。

（二）朗讀流暢度及閱讀理解個別測驗

在朗讀流暢度個別測驗部分，二、三年級的文章及閱讀理解題目選自蘇宜芬、陳柏熹與洪儷瑜（2010）所編製的「國民小學二、三年級閱讀理解成長測驗」，至於四、五、六年級的文章及閱讀理解題目則選自蘇宜芬、洪儷瑜、陳心怡與陳柏熹（2015）的「閱讀理解成長測驗（四至六年級版）」。

蘇宜芬等（2010）所編製的「國民小學二、三年級閱讀理解成長測驗」及蘇宜芬等（2015）

出版的《閱讀理解成長測驗（四至六年級版）》，目的都是為國小二至六年級各發展六套閱讀理解複本測驗。其中，「國民小學二、三年級閱讀理解成長測驗」每套複本測驗均有 16 題「部分處理」題及兩個「短文理解」題組。不論「部分處理」或「短文理解」，都含有測量字詞彙知識、字面理解、主旨大意與推論理解的題目。由於國小二、三年級的閱讀發展階段還在「學習閱讀」(learning to read)，說明文對於此階段的學童而言較難，因此「短文理解」的短文均為記敘文。二年級的文章長度為 60-100 字，三年級的文章長度為 80-120 字。每篇文章之後均附有五個題目，題目架構為：第 1 題是測量「字詞彙知識」；第 2-3 題為測量「字面理解」；第 4 題測量「主旨大意的理解」；第 5 題測量「推論理解」。每道題目均有四個選項，其中只有一個選項是正確答案。二、三年級每份複本的 Cronbach's α 信度係數介於 .62 至 .83 之間。本研究根據預試結果，從二、三年級的「短文理解」題組中，分別為每個年級挑選難度接近的四個題組，其中兩個題組用於第一學期，另兩個題組用於第二學期。本研究二、三年級每個學期所使用的題組與「國民小學二至六年級閱讀理解篩選測驗」(柯華葳、詹益綾，2006)的效標關聯效度介於 .55 至 .66 之間，均達 .001 顯著水準。

至於「閱讀理解成長測驗（四至六年級版）」，每套複本測驗均有一篇短記敘文（200-300 字）、一篇長記敘文（500-700 字）、一篇短說明文（200-300 字）、一篇長說明文（450-700 字），共計四篇閱讀理解測驗。此測驗的命題架構為：每篇文章之後都有 10 個題目，1-4 題測量「字詞彙知識」、5-6 題測量「字面理解」、7-8 題測量「主旨大意的理解」、9-10 題測量「推論理解」。每道題目均有四個選項，其中只有一個選項是正確答案。四、五、六年級每份複本的 Cronbach's α 信度係數介於 .74 至 .91 之間。本研究根據第一次預試後的結果，從四至六年級的題組中，分別為每個年級挑選難度接近的短記敘文及短說明文題組各兩組，每個學期均採用一個短記敘文題組及一個短說明文題組施測。本研究四、五、六年級每個學期所使用的題組與「國民小學二至六年級閱讀理解篩選測驗」(柯華葳、詹益綾，2006)的效標關聯效度介於 .54 至 .74 之間，均達 .001 顯著水準。

三、研究程序

每位學童在第一學期及第二學期均需接受一次團體測驗及個別測驗。團體測驗部分是施測「國民小學二至六年級閱讀理解篩選測驗」(柯華葳、詹益綾，2006)。個別測驗部分，二、三年級為兩篇短的記敘文，四、五、六年級為一篇短的記敘文及一篇短的說明文。在兩篇正式文章施測之前會先有一篇短文的練習題組，以確認受試者是否瞭解測驗程序。

個別測驗施測時，主試者會告訴學童：「在這個測驗中，我會請你唸兩篇文章。當你在唸文章的時候，請你一邊唸一邊認真地去瞭解文章的意思。每篇文章唸完之後，會有 X 個（二、三年級為五個，四、五、六年級為 10 個）閱讀測驗的題目要請你回答，所以你必須認真地讀。」指導語說明之後，如果受試者沒有疑問，就先做練習的題組。主試者確定受試者理解測驗方

式之後，再做正式測驗。

當受試者在朗讀文章時，主試者要記錄受試者唸錯之處以及朗讀所需時間，以計算受試者 1 分鐘唸對的字數（words correct per minute），並同時錄音，以備資料檢核之用。受試者唸完文章之後，主試者就收回文章，把該篇文章的閱讀理解題目拿給受試者作答。受試者答完題目之後，再依同樣程序進行下一篇。

受試者在作答時不能參閱文章，此作法乃參考 Samuels（2002）的建議以及 Johns（1997）的方式。此方式背後的邏輯是：認字效能高的學童由於其認字處理不太占用認知資源，因此能將較多的認知資源投注於理解，所以在閱讀過程中對於文章的意義、內容，能夠形成較佳的心理表徵，故即使在不參閱文章的情形下，他們也能依內在的心理表徵來作答，且能有較佳的答題表現。只是，為了降低學童記憶能力的個別差異因素影響答題表現，所以本研究選取短文作為測驗材料，以減輕學童記憶上的負荷。

每位學童受測所得資料包括：（一）「國民小學二至六年級閱讀理解篩選測驗」的答對率；（二）個別測驗中兩篇短文的閱讀理解答對率；（三）從個別測驗的兩篇短文所測得的平均朗讀流暢度（即朗讀文章時 1 分鐘所能唸對的字數）。其中，二至六年級兩個學期朗讀流暢度計算上的評分者一致性介於 98% 至 99% 之間。

四、資料分析

本研究除了分析二至六年級上、下兩學期，在「國民小學二至六年級閱讀理解篩選測驗」、兩篇短文的閱讀理解個別測驗之答對率，以及朗讀流暢度的平均數、標準差之外，也提供二至六年級各學期朗讀流暢度的測量標準誤及上、下兩學期的差異標準誤等資料。此外，本研究也分析二至六年級各學期朗讀流暢度與閱讀理解篩選測驗答對率，以及個別測驗閱讀理解答對率的相關，以便與過去研究做比較。

關於以朗讀流暢度篩檢挫折閱讀水準學童最佳切截點的分析，由於篩檢測驗通常是以平均數或平均數以下幾個標準差作為切截分數，因此，本研究在此部分，以「朗讀流暢度的 z 分數」為預測變項，「個別測驗閱讀理解答對率是否低於 50%」為效標變項，進行分析。首先，參考王天苗與廖華芳（2007）的作法，以「接受者操作特徵」分析所得的「ROC 曲線下面積」之大小，來檢測以朗讀流暢度篩檢挫折閱讀水準學童的準確度。然後，再分析當「朗讀流暢度 z 分數」的切截點分別設在 0.0、-0.5、-1.0、-1.5 時，篩檢「個別測驗閱讀理解答對率低於 50%」學童的 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率。最後則根據 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率，綜合評估哪一個朗讀流暢度的 z 分數是最佳切截點。

參、結果與討論

一、二至六年級兩學期各項測驗結果的平均數與標準差

本研究二至六年級學童於上、下兩個學期，在「國民小學二至六年級閱讀理解篩選測驗」答對率、兩篇短文的個別測驗閱讀理解答對率，以及朗讀流暢度的平均數、標準差，如表 2 所示。其中，除了五年級之外，二、三、四、六年級兩個學期朗讀流暢度平均數的差異都達 .001 顯著水準。二、三、四年級下學期的平均數均高於上學期，但六下的平均數反而低於上學期，五年級的兩個學期則無顯著差異。關於為什麼六下的朗讀流暢度平均值反而低於上學期，由於本研究兩個學期所使用的文章難度接近，因此推測可能與六下接近畢業，學童可能心情較浮動，影響專注與投入程度有關。

表 2

二至六年級各學期閱讀理解篩選測驗答對率、個別測驗閱讀理解答對率及朗讀流暢度的平均數與標準差

年級	學期	閱讀理解篩選測驗答對率		個別測驗閱讀理解答對率		朗讀流暢度	
		平均數	標準差	平均數	標準差	平均數	標準差
二	上	0.72	0.20	0.59	0.23	121.69	35.17
	下	0.78	0.18	0.67	0.22	127.94	34.38
三	上	0.63	0.19	0.64	0.22	135.09	36.36
	下	0.74	0.18	0.66	0.22	154.11	36.60
四	上	0.64	0.19	0.62	0.17	144.08	30.69
	下	0.64	0.18	0.67	0.19	161.74	31.13
五	上	0.65	0.17	0.61	0.14	166.07	31.91
	下	0.61	0.18	0.64	0.16	168.25	29.42
六	上	0.67	0.15	0.61	0.18	174.23	36.50
	下	0.70	0.15	0.61	0.15	168.52	33.42

再者，就五、六年級四個學期的朗讀流暢度平均值來看，朗讀流暢度的成長到高年級有趨緩的情形，這和陳惠琴與楊憲明（2008）的研究結果類似。在陳惠琴與楊憲明的研究中，二至六年級學童所讀的文章都一樣，是四篇低年級水準的文章（敘事體與說明體各兩篇），其結果顯示：朗讀流暢度在四年級之前成長較快，四年級之後則成長速率趨緩。

另外，本研究三下的朗讀流暢度平均值反而高於四上的平均值，一方面可能與研究樣本

有關，根據表 2，三年級樣本上、下學期閱讀理解篩選測驗答對率的平均值（0.69），高於四年級樣本平均值（0.64）；另一方面也可能與閱讀材料有關，如果個別測驗閱讀理解答對率可反映文章難度，那麼三下個別測驗所閱讀的文章對三年級學童而言的難度，可能比四上個別測驗閱讀的文章對四年級學童而言的難度略為容易。

與葉靖雲（1998）的分析結果相較，該研究是在中部地區選取二、四、六年級學童，在第一學期以國語課本中的兩篇記敘文，記錄第 1 分鐘唸對字數所得的朗讀流暢度，本研究二、四、六年級第一學期所測得的朗讀流暢度略低於葉靖雲所測得的速率，此情形和王文婷與蘇宜芬（2005）的發現一致。形成此差異的原因可能有二：（一）如前所述，「美國教育進步評量」在 2002 年的研究發現，根據學童朗讀第 1 分鐘所測得的朗讀速度，往往比根據學童讀完整篇短文所測得的朗讀速度來得快（Daane et al., 2005）；（二）葉靖雲所採用的文章為國語課本裡的文章，且都是記敘文，但本研究所採用的文章均非國語課本的文章，除了二、三年級都用記敘文外，四、五、六年級每學期均採一篇記敘文及一篇說明文，且指導語中要求學童一邊唸、一邊認真地去瞭解文章的意思，並讓學童知道每篇文章唸完之後，會有幾個閱讀測驗的題目要回答，因此學童必須一邊讀一邊盡力理解，這些差異有可能使得本研究所測得速率略低。

二、二至六年級各學期朗讀流暢度與閱讀理解篩選測驗答對率及個別測驗閱讀理解答對率的相關分析

根據表 3，除了五年級上學期朗讀流暢度與個別測驗閱讀理解答對率的相關為 .31 ($p < .001$)，其餘各年級兩個學期的朗讀流暢度與閱讀理解篩選測驗答對率，以及個別測驗閱讀理解答對率的相關都介於 .41~.65 之間，且均達 .001 顯著水準。過去研究中，葉靖雲（1998）的二、四、六年級及陳惠琴與楊憲明（2008）的二至六年級，朗讀流暢度與標準化國語文成就測驗、或標準化閱讀理解測驗得分的相關也是在 .43~.66 之間。因此，本研究在朗讀流暢度與標準化閱讀測驗的相關分析結果與過去研究相似。至於朗讀流暢度與唸同篇文章之後所測得閱讀理解分數的相關，過去國內尚無此方面的研究，但在國外，Valencia 等（2010）的研究中，以二、四、六年級學童為對象，以問答題測閱讀理解程度，結果兩者相關在 .12~.29 之間。在這部分，本研究測得朗讀流暢度與唸同篇文章之後閱讀理解表現的相關，則高於 Valencia 等人的研究。

三、以朗讀流暢度篩檢挫折閱讀水準學童最佳切截點的分析

本研究以「朗讀流暢度的 z 分數」為預測變項，以「個別測驗閱讀理解答對率是否低於 50%」為效標變項，對各年級兩個學期的資料進行 ROC 分析，以「ROC 曲線下面積」的大小，來檢測以朗讀流暢度篩檢因認字效能不足導致閱讀理解困難（亦即達挫折閱讀水準）之學童的準確度。此外，也分析當「朗讀流暢度 z 分數」的切截點分別設在 0.0、-0.5、-1.0、-1.5 時，

表 3

二至六年級各學期朗讀流暢度與閱讀理解篩選測驗答對率及個別測驗閱讀理解答對率的相關

年級	學期	朗讀流暢度與閱讀理解篩選測驗 答對率的相關	朗讀流暢度與個別測驗閱讀理解 答對率的相關
二	上	.53	.53
	下	.42	.46
三	上	.65	.56
	下	.56	.54
四	上	.49	.47
	下	.50	.42
五	上	.49	.31
	下	.54	.49
六	上	.51	.41
	下	.52	.53

註：表內相關係數均達 .001 顯著水準。

篩檢「個別測驗閱讀理解答對率低於 50%」學童的 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率，然後根據這些數據綜合評估最佳切截點。

首先，二上的 ROC 曲線下面積為 .74 ($p < .001$)，下學期為 .79 ($p < .001$)。在一般情形下，ROC 曲線下面積介於 .50~.70 時，表示診斷準確性較低；面積為 .70~.90 時，表示診斷準確性中等；面積為 .90 以上時，表示診斷準確性較高（林傑斌等，2004）。因此，就二年級的資料而言，以朗讀流暢度來篩檢挫折閱讀水準學童，上、下兩學期都具有中等的篩檢準確性。

接著，在切截點的選擇上，路君約（1989）曾提醒，在訂定切截點時，需特別注意「錯誤拒絕率」，也就是問題個案沒有被篩檢出來的人數占所有問題個案數的比率。如果施測目的是為了篩檢高風險學童，那麼篩檢工具及切截點要有高「敏銳度」，且被「錯誤拒絕」的人數比例需盡量減到最低。依此原則，根據表 4 的結果，二年級上、下學期都是當朗讀流暢度 z 分數為 0.0 時，有最高的敏銳度（.78 與 .81）、最低的錯誤拒絕率（.22 與 .19），但是命中率並不理想（.59 與 .56）。因此，若從命中率 .70 以上的幾個切截點做比較，朗讀流暢度 z 分數為 -0.5 時，敏銳度最高（上、下學期分別為 .59 與 .71），錯誤拒絕率最低（.41 與 .29），而上、下學期的 χ^2 值分別為 17.34 及 15.32，均達 .001 顯著水準，錯誤接受率則分別為 .22 與 .28，在可接受的水準。因此，二年級上、下學期朗讀流暢度 z 分數最佳切截點為 -0.5，也就是朗讀流暢度分別是每分鐘唸對 104 字與 111 字（根據表 4 四捨五入至整數）。

表 4

二年級上、下學期切截點設在不同朗讀流暢度 z 分數的 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率

學期	朗讀流暢度 z分數	朗讀流暢度 原始數值	χ^2 值	命中率	敏銳度	錯誤拒絕率	錯誤接受率
上	0.0	121.69	10.50**	.59	.78	.22	.48
	-0.5	104.11	17.34***	.73	.59	.41	.22
	-1.0	86.52	20.43***	.78	.41	.59	.08
	-1.5	68.94	10.79**	.77	.16	.84	.02
下	0.0	127.94	7.83**	.56	.81	.19	.48
	-0.5	110.75	15.32***	.72	.71	.29	.28
	-1.0	93.56	17.25***	.84	.43	.57	.09
	-1.5	76.37	20.91***	.88	.29	.71	.02

** $p < .01$. *** $p < .001$.

就三年級的資料而言，上學期的 ROC 曲線下面積為 .71 ($p < .001$)，下學期為 .79 ($p < .001$)。因此，以朗讀流暢度來篩檢三年級挫折閱讀水準學童，上、下兩學期都具有中等的篩檢準確性。

在最佳切截點的分析上，根據表 5，三年級上、下學期都是當朗讀流暢度 z 分數為 0.0 時，有最高的敏銳度 (.77 與 .84)、最低的錯誤拒絕率 (.23 與 .16)，但是命中率並不理想 (.60 與 .61)。因此，若從命中率 .70 以上的幾個切截點做比較，上學期當朗讀流暢度 z 分數為-1.0 及-1.5 時，命中率为 .70 以上(分別為 .73 與 .81)，但敏銳度都低於 .50(分別是 .29 與 .13)，錯誤拒絕率則皆高於 .50(分別為 .71 與 .87)，不適合作為切截點。但當朗讀流暢度 z 分數為-0.5 時，命中率为 .69，接近 .70，敏銳度 .58、錯誤拒絕率 .42， χ^2 值為 9.92，達 .01 顯著水準，錯誤接受率為 .28。因此，以 z 分數-0.5 作為切截點，優於 z 分數 0.0、-1.0 及-1.5。

至於下學期的資料，當 z 分數為-0.5、-1.0、-1.5 時，命中率都在 .70 以上，但其中 z 分數-0.5 有較高的敏銳度 (.76)、較低的錯誤拒絕率 (.24)， χ^2 值為 20.56，達 .001 顯著水準，錯誤接受率則為 .28，在可接受的水準，所以三年級下學期的最佳切截點也是 z 分數-0.5。也就是，三年級上、下學期朗讀流暢度 z 分數最佳切截點均為-0.5，亦即朗讀流暢度分別是每分鐘唸對 117 字與 136 字(根據表 5 四捨五入至整數)。

就四年級的資料而言，上學期的 ROC 曲線下面積為 .73 ($p < .001$)，下學期為 .78 ($p < .001$)。因此，以朗讀流暢度來篩檢四年級挫折閱讀水準學童，上、下兩學期都具有中等的篩檢準確性。

表 5

三年級上、下學期切截點設在不同朗讀流暢度 z 分數的 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率

學期	朗讀流暢度 z分數	朗讀流暢度 原始數值	χ^2 值	命中率	敏銳度	錯誤拒絕率	錯誤接受率
上	0.0	135.09	10.69**	.60	.77	.23	.44
	-0.5	116.91	9.92**	.69	.58	.42	.28
	-1.0	98.73	2.90	.73	.29	.71	.16
	-1.5	80.55	10.69**	.81	.13	.87	.01
下	0.0	154.11	13.40***	.61	.84	.16	.44
	-0.5	135.81	20.56***	.73	.76	.24	.28
	-1.0	117.51	27.77***	.84	.52	.48	.09
	-1.5	99.21	3.13	.83	.08	.92	.02

** $p < .01$. *** $p < .001$.

在最佳切截點的分析上，根據表 6，四年級上、下學期都是當朗讀流暢度 z 分數為-0.5 時，命中率在 .70 以上（上、下學期分別是 .70 與 .72），且有較高的敏銳度（兩學期都是 .58）、較低的錯誤拒絕率（兩學期都是 .42）， χ^2 值分別為 11.03 ($p < .01$) 及 12.52 ($p < .001$)，錯誤接受率則分別為 .26 與 .25，在可接受的水準。因此，四年級上、下學期朗讀流暢度 z 分數的最佳切截點為-0.5，也就是朗讀流暢度分別是每分鐘唸對 129 字與 146 字（根據表 6 四捨五入至整數）。

就五年級的資料而言，上學期的 ROC 曲線下面積為 .76 ($p < .001$)、下學期為 .78 ($p < .001$)。因此，以朗讀流暢度來篩檢五年級挫折閱讀水準學童，上、下兩學期都具有中等的篩檢準確性。

在最佳切截點的分析上，根據表 7，五年級上、下學期都是當朗讀流暢度 z 分數為-0.5 時，命中率在 .70 以上（上、下學期分別是 .73 與 .76），且有較高的敏銳度（.69 與 .60）、較低的錯誤拒絕率（.31 與 .40）， χ^2 值分別為 21.58 ($p < .001$) 及 26.10 ($p < .001$)，錯誤接受率則分別為 .26 與 .19，在可接受的水準。因此，五年級上、下學期朗讀流暢度 z 分數最佳的切截點為-0.5。另外，根據表 2，五年級上、下學期的朗讀流暢度成長並不明顯，所以兩個學期可採用同樣的篩檢標準，也就是每分鐘唸對的字數低於 150~154 字（根據表 7 四捨五入至整數）的學童，需加以留意。

就六年級的資料而言，上學期的 ROC 曲線下面積為 .79 ($p < .001$)，下學期為 .89 ($p < .001$)。因此，以朗讀流暢度來篩檢六年級挫折閱讀水準學童，上、下兩學期都具有中等的篩檢準確性。

表 6

四年級上、下學期切截點設在不同朗讀流暢度 z 分數的 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率

學期	朗讀流暢度 z 分數	朗讀流暢度原始數值	χ^2 值	命中率	敏銳度	錯誤拒絕率	錯誤接受率
上	0.0	144.08	15.93***	.62	.84	.16	.44
	-0.5	128.74	11.03**	.70	.58	.42	.26
	-1.0	113.39	1.50	.72	.26	.74	.16
	-1.5	98.05	3.20	.78	.13	.87	.04
下	0.0	161.74	15.93***	.62	.84	.16	.44
	-0.5	146.18	12.52***	.72	.58	.42	.25
	-1.0	130.61	14.60***	.79	.39	.61	.10
	-1.5	115.05	3.20	.78	.13	.87	.04

** $p < .01$. *** $p < .001$.

表 7

五年級上、下學期切截點設在不同朗讀流暢度 z 分數的 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率

學期	朗讀流暢度 z 分數	朗讀流暢度原始數值	χ^2 值	命中率	敏銳度	錯誤拒絕率	錯誤接受率
上	0.0	166.07	13.71***	.60	.81	.19	.45
	-0.5	150.12	21.58***	.73	.69	.31	.26
	-1.0	134.16	17.10***	.80	.41	.59	.11
	-1.5	118.21	5.59*	.81	.16	.84	.04
下	0.0	168.25	18.64***	.65	.77	.23	.39
	-0.5	153.54	26.10***	.76	.60	.40	.19
	-1.0	138.83	13.59***	.76	.33	.67	.09
	-1.5	124.12	12.55***	.77	.21	.79	.04

* $p < .05$. *** $p < .001$.

在最佳切截點的分析上，根據表 8，六年級上、下學期在朗讀流暢度 z 分數為-0.5 與-1.0 時，命中率都在 .70 以上（.73~.89 之間），且有較佳的敏銳度（敏銳度均大於 .60）、較低的錯誤拒絕率（錯誤拒絕率均小於 .40）， χ^2 值都達 .01 以上顯著水準，錯誤接受率在 .10 至 .26 之間，是可接受的水準。只是，如果從「讓錯誤拒絕的人數比例盡量減到最低」的角度考量，那麼 z 分數-0.5 是相對較佳的切截點，且與五年級的資料做綜合考量，也是比較合理的切截點。

表 8

六年級上、下學期切截點設在不同朗讀流暢度 z 分數的 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率

學期	朗讀流暢度 z 分數	朗讀流暢度原始數值	χ^2 值	命中率	敏銳度	錯誤拒絕率	錯誤接受率
上	0.0	174.23	8.34**	.57	.86	.14	.45
	-0.5	155.98	9.31**	.73	.64	.36	.26
	-1.0	137.73	22.05***	.84	.64	.36	.14
	-1.5	119.48	9.20**	.90	.21	.79	.03
下	0.0	168.52	15.54***	.61	.94	.06	.42
	-0.5	151.81	26.27***	.79	.81	.19	.21
	-1.0	135.10	44.85***	.89	.75	.25	.10
	-1.5	118.39	44.52***	.93	.44	.56	.02

** $p < .01$. *** $p < .001$.

另外，由表2可知，朗讀流暢度平均值到五、六年級的成長趨緩，六下朗讀流暢度平均值與上學期相較，沒有提升反而略降，由於兩個學期所採用的文章難度接近，根據表2六年級兩個學期的個別測驗閱讀理解答對率亦無顯著差異，因此，六下朗讀流暢度的略降，可能與六年級學童即將畢業，心情較為浮動，專注力稍受影響有關。因此，六年級兩個學期可採用同樣的篩檢標準，也就是每分鐘唸對的字數低於 152~156 字（根據表 8 四捨五入至整數）的學童，需加以留意。

根據表 4~8 的分析數據，二至六年級上、下學期，當朗讀流暢度 z 分數為-0.5 時，雖然敏銳度在 .58~.81 之間，但錯誤拒絕率也有 .19~.42，錯誤接受率則在 .19~.28 之間。原因是朗讀流暢度為閱讀理解的基本條件，但非充分條件，因為閱讀理解的影響因素，除了朗讀流暢度之外，還有推論能力、與文章主題有關的背景知識、閱讀策略運用能力等。所以，就教育應用而言，在測量朗讀流暢度之後，還是必須問學童一些與文章有關的閱讀理解題目，以排除被錯誤接受（即被錯誤篩檢為問題個案）的學童，也可篩檢出錯誤拒絕（即閱讀理解落後但未被篩檢出來）的學童。

此外，即使本研究在指導語中，已告訴學童朗讀之後需回答一些閱讀理解的題目，但仍可能有學童為求唸得快，而忽略了理解；相反地，也可能有學童為求仔細理解，而放慢了朗讀速度，這也是影響錯誤拒絕率及錯誤接受率的因素。對於前一類的學童，應不需擔心其朗讀流暢度不足的問題，但是，在其閱讀理解程度的部分，需輔以平時課堂觀察，或再以多篇分次的閱讀理解測驗加以確認。至於後一類的學童，「放慢閱讀速度」可能是他們所採取讓自己讀得仔細的策略；只是，也需留意當文章長度增加之後，他們是否仍能保有一樣的閱讀理解表現。若以一位中文閱讀流暢，而英文閱讀沒有中文閱讀那麼流暢的大學生為例來說明，

這位大學生在閱讀一般中文報紙文章時，可能不論文章長短，都可以讀得流暢，且理解得很好；但在讀英文報紙文章時，需放慢速度才能讀得懂，且當文章長時，不僅需放慢速度，還需輔以其他策略，才能幫助自己理解。這是因為當認字效能相對地沒那麼高時，在閱讀時，認字的處理就會占據相對多的認知資源，所以能用於理解的認知資源就不那麼寬裕。當文章短時，也許還可以負荷；但當文章長時，吃力的感覺就會變得明顯。

另外，「正確檢出」及「錯誤拒絕」的學童所需要的補救教學也有所不同。以朗讀流暢度「正確檢出」的學童，其閱讀理解表現落後的原因，可能主要來自認字速率低、認字效能不佳，因此所需的補救教學就需提高其認字效能，「重複閱讀法」(repeated reading)是可參考的方法之一(吳宜貞，2007；Samuels, 1997)。至於「錯誤拒絕」的學童其閱讀理解表現落後的原因，主要不是認字效能不佳，可能是其他原因，如前述的推論能力弱、背景知識不足、或不曾運用閱讀策略等，其所需的補救教學就要提升其推論能力、充實背景知識及學習有效的閱讀策略(柯華葳，2010；蘇宜芬，2004)。

此外，凡是測量難免會有誤差，因此，本研究也提供二至六年級各學期朗讀流暢度的測量標準誤，以及上、下兩學期差異標準誤的分析數據(見表 9)，讓教師們在做教育決定或安置時(例如：篩選哪些學童認字效能落後需要接受補救教學)，有更精確的資料作為判斷依據。

表 9

二至六年級各學期朗讀流暢度的複本信度、測量標準誤及上、下兩學期的差異標準誤

年級	學期	複本信度	測量標準誤	上下學期差異標準誤
二	上	.85**	13.62	18.73
	下	.86**	12.86	
三	上	.87**	13.11	17.10
	下	.91**	10.98	
四	上	.90**	9.71	12.74
	下	.93**	8.24	
五	上	.76**	15.63	18.19
	下	.90**	9.30	
六	上	.94**	8.94	13.84
	下	.90**	10.57	

** $p < .01$.

肆、結論與建議

本研究主要是以閱讀的自動化理論為基礎，分析國小二至六年級以朗讀流暢度篩檢因認字效能不足導致閱讀理解困難之學童的效果，並對最佳切截點提出建議，另外，也提供國小二至六年級上、下學期朗讀流暢度成長資料，以供教育研究者及教師瞭解二至六年級學童認字效能的平均成長情形。研究結果顯示：二至六年級學童朗讀流暢度有隨著年級成長的趨勢，只是到了五、六年級的四個學期間，成長速度趨於緩和。至於以朗讀流暢度篩檢挫折閱讀水準學童的效果，根據「ROC 曲線下面積」的分析數據，二至六年級上、下學期的資料都有中等的篩檢準確性。至於切截點的分析，根據 χ^2 值、命中率、敏銳度、錯誤拒絕率及錯誤接受率做綜合考量，二至六年級上、下學期均以朗讀流暢度 z 分數-0.5 較佳。若以朗讀速率表示，各年級的切截點是：二年級上、下學期分別是每分鐘唸對 104 字與 111 字，三年級上下學期分別為每分鐘唸對 117 字與 136 字，四年級上、下學期分別為每分鐘唸對 129 字與 146 字，五年級則為每分鐘唸對 150~154 字，六年級是每分鐘唸對 152~156 字。當學童的朗讀流暢度在上述切截點之下，且其閱讀理解對率在 50% 以下時，那麼就表示這位學童的認字效能不足，導致他的閱讀理解困難，因此需要提升認字效能方面的補救教學。

另外，在教育應用時，為了提高找出認字效能不足學童的篩檢準確度，研究者建議測量朗讀流暢度時，至少要以兩篇以上（含兩篇）的年級水準文章測量，且學童讀完文章後，必須回答一些與文章有關的閱讀理解問題，以評估其閱讀理解是否落後。再者，本研究所提供的測量標準誤，也可讓教師們在做教育決定時，能將測量誤差因素納入考慮，這是過去國內、外相關研究常忽略之處。

此外，本研究為了讓教育研究者及教師瞭解二至六年級學童朗讀流暢度成長情形，所以，提供了二至六年級上、下學期的平均數與標準差。然而，在教育應用上，研究者並不建議以此平均數作為各年級朗讀流暢度的「通過標準」(benchmark)。從表 4~8 的資料來看，當朗讀流暢度的 z 分數為 0.0 時，各年級的錯誤接受率在 .39~.48 之間，也就是閱讀理解在挫折水準以上的學童中，有四成的學童其朗讀流暢度是低於平均數的。所以，表 2 的資料比較適合用來瞭解二至六年級學童在不同學期之間認字效能的平均成長情形，至於各年級朗讀流暢度的「通過標準」或「最低門檻」則有待未來研究進一步的探討。

最後，對於本研究中兩個尚無法釐清之處，研究者對於未來此方面的研究，有以下建議：

一、本研究所採用的朗讀文章為課外的材料，且是以學童唸完全文所需時間來計算每分鐘唸對幾個字，而所測得的朗讀流暢度則比葉靖雲（1998）所測得的速率略低。葉靖雲的研究是採用兩篇國語課本的記敘文，且只記錄第 1 分鐘唸對的字數。因此，未來研究可進一步探討，如果採用國語課本中的文章為朗讀材料，且指導語中讓學童知道，唸完文章後需回答幾

個與文章有關的閱讀理解問題，那麼，以學童唸完全文所需時間，和只記錄第1分鐘唸對字數比較起來，所計算的朗讀速率是否會有差異，是否會如美國 NAEP 在 2002 年的研究結果一樣，發現後者傾向於高估學童的朗讀流暢度 (Daane et al., 2005)。

二、本研究在二至六年級的取樣上，採橫斷研究的方式，因此，當較低年級的平均表現優於較高年級時，較難釐清這是取樣上的問題，還是鄰近年級之間在所研究的能力上成長趨緩，甚至不再成長。因此，未來這方面的研究，可考慮至少在年段之間採縱貫研究的方法，也就是從三年級追縱至四年級，五年級追縱至六年級。另外一個可能的方式是，如果採用橫斷法的方式取樣，那麼各年級學童除了讀年級水準的文章外，也都讀兩篇低年級水準的文章，然後以這兩篇共同文章所測得的朗讀流暢度數據進行各年級的等化。

綜合言之，本研究提供臺灣第一份二至六年級上、下學期朗讀流暢度成長的資料，也是第一個研究分析以朗讀流暢度篩檢認字效能不足導致閱讀困難之學童的篩檢準確度及最佳切截點。希望本研究所提供的分析資料，對篩檢認字效能落後的國小學童有所幫助，也期待所提出的未來研究建議有助於此方面研究的進展。

誌謝

本研究承蒙科技部之經費補助（計畫編號：NSC98-2410-H-003-028、NSC99-2410-H-003-039）。研究過程中，感謝魏頡、項瑞芳、張菡穎、黃郁惠、林依曄、雒怡寧等協助施測，以及王磊、廖子慧、張璣勻、王姝云、林雅芳等協助資料登錄與校對。此外，感謝劉子鍵教授在資料分析上提供寶貴建議。最後，也感謝兩位審查委員提供精闢的修改意見。

參考文獻

一、中文文獻

- 王文婷、蘇宜芬（2005，1月）。國小二年級學童認字自動化指標之研究。論文發表於台灣心理學會第四十四屆年會，桃園市。
- 【Wang, W.-T., & Su, Y.-F. (2005, January). *Indicators of automaticity in word recognition for second graders*. Paper presented at the 44th annual meeting of the Taiwanese Psychological Association, Taoyuan, Taiwan.】
- 王天苗、廖華芳（2007）。嬰幼兒綜合發展測驗之判定準確度及切截點分析。《特殊教育研究學刊》，32（2），1-15。doi:10.6172/BSE200706.3202001
- 【Wang, T.-M., & Liao, H.-F. (2007). Assessment accuracy and cut-off points of comprehensive developmental inventory for infants and toddlers (CDIIT). *Bulletin of Special Education*, 32(2), 1-15. doi:10.6172/BSE200706.3202001】
- 王梅軒、黃瑞珍（2005）。國小課程本位閱讀測量方法之信度與效度研究。《特殊教育研究學刊》，29，73-94。doi:10.6172/BSE200509.2901004
- 【Wang, M.-H., & Huang, R.-J. (2005). The reliability and validity of curriculum-based reading measures on elementary school students. *Bulletin of Special Education*, 29, 73-94. doi:10.6172/BSE200509.2901004】
- 吳宜貞（2007）。重複閱讀在教學的實用性。《師友月刊》，478，69-71。
- 【Wu, Y.-C. (2007). The application of repeated reading. *The Educator Monthly*, 478, 69-71.】
- 林傑斌、林川雄、劉明德（2004）。SPSS 12 統計建模與應用實務。臺北市：博碩文化。
- 【Lin, J.-B., Lin, C.-X., & Liou, M.-D. (2004). *Statistic module and application of SPSS 12*. Taipei, Taiwan: DrMaster Press.】
- 洪儷瑜、陳秀芬、王瓊珠、張郁雯（2012）。常用字流暢性測驗編製研究。《測驗學刊》，59（2），247-276。
- 【Hung, L.-Y., Chen, H.-F., Wang, C.-C., & Chang, Y.-W. (2012). Construct study of the assessment of sight-word reading and fluency. *Psychological Testing*, 59(2), 247-276.】
- 洪儷瑜、張郁雯、王瓊珠、陳秀芬（2006）。識字量評估測驗。臺北市：教育部特殊教育小組。
- 【Hung, L.-Y., Chang, Y.-W., Wang, C.-C., & Chen, H.-F. (2006). *Assessment test of Chinese character lists for graders*. Taipei, Taiwan: Department of Student Affairs and Special Education, Ministry of Education.】
- 柯華蕓（2010）。閱讀理解教學。載於王瓊珠、陳淑麗（主編），*突破閱讀困難：理論與實務*（pp. 167-185）。臺北市：心理。
- 【Ko, H.-W. (2010). Instructions of reading comprehension. In C.-C. Wang & S.-L. Chen (Eds.), *Overcoming reading difficulties: Theories and practical applications* (pp. 167-185). Taipei, Taiwan: Psychological.】
- 柯華蕓、詹益綾（2006）。國民小學二至六年級閱讀理解篩選測驗。臺北市：教育部特殊教育小組。
- 【Ko, H.-W., & Chan, Y.-L. (2006). *Reading comprehension screening test for second to sixth graders*. Taipei, Taiwan: Department of Student Affairs and Special Education, Ministry of Education.】
- 陳惠琴、楊憲明（2008）。國小學童文本朗讀速度分析之研究。《特殊教育與復健學報》，18，1-30。

【Chen, H.-C., & Yang, H.-M. (2008). A study on oral reading rate in elementary school students. *Bulletin of Special Education and Rehabilitation*, 18, 1-30.】

葉靖雲 (1993)。課程本位閱讀測驗的效度研究。《特殊教育學報》，8，273-323。

【Yeh, C.-Y. (1993). The validity of curriculum-based reading measures. *Journal of Special Education*, 8, 273-323.】

葉靖雲 (1998)。課程本位閱讀測驗的效度研究。《特殊教育與復健學報》，6，239-260。

【Yeh, C.-Y. (1998). The validity of curriculum-based reading comprehension tests. *Bulletin of Special Education and Rehabilitation*, 6, 239-260.】

路君約 (1989)。《心理測驗》(上冊)。臺北市：中國行為科學社。

【Lu, C.-Y. (1989). *Psychological testing* (Vol. 1). Taipei, Taiwan: Chinese Behavioral Science Corporation.】

蘇宜芬 (2004)。閱讀理解的影響因素及其在教育上的意義。《教師天地》，129，21-28。

【Su, Y.-F. (2004). Sources of individual differences and its educational implications in reading comprehension. *Teacher's Chamber*, 129, 21-28.】

蘇宜芬、洪儷瑜、陳心怡、陳柏熹 (2015)。《閱讀理解成長測驗》(四至六年級版)。臺北市：中國行為科學社。

【Su, Y.-F., Hung, L.-Y., Chen, H.-Y., & Chen, P.-H. (2015). *Progress monitoring test of reading comprehension for grade 4-6 students*. Taipei, Taiwan: Chinese Behavioral Science Corporation.】

蘇宜芬、陳柏熹、洪儷瑜 (2010)。《閱讀理解成長指標之建立》。臺北市：國立臺灣師範大學教育評鑑與發展研究中心。

【Su, Y.-F., Chen, P.-H., & Hung, L.-Y. (2010). *Development of indicators for progress monitoring on reading comprehension*. Taipei, Taiwan: Center for Educational Research and Evaluation, National Taiwan Normal University.】

蘇宜芬、陳學志 (2007)。認字自動化指標之建立與信效度研究。《教育心理學報》，38 (4)，501-514。doi:10.6251/BEP.20070125

【Su, Y.-F., & Chen, H.-C. (2007). Indicators of character decoding automaticity: Reliability and validity. *Bulletin of Educational Psychology*, 38(4), 501-514. doi:10.6251/BEP.20070125】

二、外文文獻

Chall, J. (1996). *Stages of reading development* (2nd ed.). Orlando, FL: Harcourt Brace.

Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Oranje, A. (2005). *Fourth-grade students reading aloud: NAEP 2002 special study of oral reading (NCES 2006-469)*. Washington, DC: National Center for Educational Statistics, U. S. Department of Education, Institute of Education Science.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52(3), 219-232.

Ekwall, E. E., & Shanker, J. L. (1993). *Reading inventory* (3rd ed.). Boston, MA: Allyn & Bacon.

Fuchs, L. S., Fuchs, D., Eaton, S., & Hamlett, C. L. (2000). *Relation between reading fluency and reading comprehension as a function of silent versus oral reading mode*. Unpublished manuscript.

- Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Study of Reading, 5*(3), 239-256. doi:10.1207/S1532799XSSR0503_3
- Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*(2), 20-28. doi:10.1177/074193258800900206
- Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2001). *Iowa tests of basic skills*. Itasca, IL: Riverside.
- Jenkins, J. R., Fuchs, L. S., Espin, C., van den Broek, P., & Deno, S. L. (2000, February). *Effects of task format and performance dimension on word reading measures: Criterion validity, sensitivity to impairment, and context facilitation*. Paper presented at the Pacific Coast Research Conference, San Diego, CA.
- Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719-729. doi:10.1037/0022-0663.95.4.719
- Johns, J. L. (1997). *BASIC reading inventory: Pre-primer through grade twelve and early literacy assessments* (7th ed.). Dubuque, IA: Kendall/Hunt.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293-323. doi:10.1016/0010-0285(74)90015-2
- Leslie, L., & Caldwell, J. (2006). *Qualitative reading inventory-4: A valid research-based measure of reading improvement*. Boston, MA: Allyn & Bacon.
- National Institute of Child Health and Human Development. (2000). *Report of the national reading panel: Teaching children to read* (NIH publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Journal of Cognitive Psychology, 7*(1), 44-64. doi:10.1016/0010-0285(75)90004-3
- Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rathvon, N., (2004). *Early reading assessment: A practitioner's handbook*. New York, NY: Guilford Press.
- Rice, N. K. (1981). *The relationship between oral reading fluency and other reading behaviors among first-grade children* (Unpublished doctoral dissertation). Michigan State University,

- Lansing, MI.
- Samuels, S. J. (1988). Decoding and automaticity: Helping poor readers become automatic at word recognition. *The Reading Teacher*, 41(8), 757-760.
- Samuels, S. J. (1997). The method of repeated readings. *The Reading Teacher*, 50(5), 376-382.
- Samuels, S. J. (2002). Reading fluency: Its development and assessment. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 166-183). Newark, DE: International Reading Association.
- Schwanenflugel, P. J., Meisinger, E. B., Wisenbaker, J. M., Kuhn, M. R., Strauss, G. P., & Morris, R. D. (2006). Becoming a fluent and automatic reader in the early elementary school years. *Reading Research Quarterly*, 41(4), 496-522. doi:10.1598/RRQ.41.4.4
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84(2), 127-190. doi:10.1037/0033-295X.84.2.127
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16(1), 32-71. doi:10.2307/747348
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: PRO-ED.
- Taylor, B., Harris, L. A., Pearson, P. D., & Garcia, G. (1995). *Reading difficulties: Instruction and assessment* (2nd ed.). New York, NY: McGraw-Hill.
- Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly*, 45(3), 270-291. doi:10.1598/RRQ.45.3.1

Journal of Research in Education Sciences

2016, 61(4), 33-57

doi:10.6209/JORIES.2016.61(4).02

Assessment Accuracy and Cut-Off Points of Oral Reading Fluency for Grade 2-6 Students

Yi-Fen Su

Department of Educational Psychology and Counseling,
National Taiwan Normal University

You-Hsuan Chang

Department of Educational Psychology and Counseling,
National Taiwan Normal University

Meng-Feng Li

Department of Psychology,
National Taiwan University

Yu-Chien Huang

Tainan Municipal Tu Cheng High School

Abstract

This study evaluated the assessment accuracy of oral reading fluency. Additionally, this study focused on determining the cut-off point for identifying poor reading comprehension caused by low decoding efficiency in students of grades 2-6. Receiver operating characteristic (ROC) analysis was adopted to evaluate the assessment accuracy of oral reading fluency. The area under the curve (AUC) of the ROC indicated that oral reading fluency exhibited medium accuracy in screening. Based on analyses of χ^2 , the hit rate, sensitivity, false negative rate, and false positive rate, a z score of oral reading fluency below -0.5 was identified as the optimal cut-off point for screening students with poor reading comprehension caused by low decoding efficiency. In addition, the mean scores of oral reading fluency for students in grades 2-6 are provided to inform teachers and researchers of the level of oral reading fluency in elementary-school students.

Keywords: cut-off points, elementary-school students, oral reading fluency, reading comprehension, receiver operating characteristic analysis

