

國立臺灣師範大學教育心理與輔導學系  
教育心理學報，民 96，38 卷，4 期，501—514 頁

# 認字自動化指標之建立與信效度研究

蘇宜芬

陳學志

國立台灣師範大學  
教育心理與輔導學系

本研究的主要目的為探討以字彙辨認作業中的兩項測量指標做為國小學童中文認字自動化評量指標之可行性。其中一個指標為平均反應時間，另一個指標為以刺激字複雜度為預測變項，以反應時間為效標變項之迴歸線的斜率。研究中以重測信度、複本信度、及效標關連效度對這兩項指標進行考驗。研究結果顯示字彙辨認作業中的平均反應時間具有良好的信度與效度，適合做為認字自動化的評量指標。但字彙辨認作業中的斜率則不具穩定度，也缺乏良好的效度，故不適合做為認字自動化的評量指標。此研究結果可提供認字自動化評量系統之建構的參考。

**關鍵詞：**自動化、國小學生、評量指標、認字

閱讀是一項重要的基本能力，這項基本能力不僅影響學生在國語文科目上的表現，也影響學生在社會、自然、數學等學科的學習。而其中「認字」又是閱讀歷程中的一個重要成分（Roth & Beck, 1987；Stanovich, 1982），因此有關認字歷程的基礎研究與應用研究一直受到認知心理學家與教育學者的重視。過去的研究顯示：認字能力的高低與閱讀習得的快慢有密切關係（Bertelson, 1986；Biemiller, 1977-1978；Curtis, 1980；Juel, Griffith, & Gough, 1986；Stanovich, 1982, 1986）；而且認字能力的發展也有助於閱讀理解能力的增進（Calfee & Piontkowski, 1981；Herman, 1985；Stanovich, 1985）。

認字能力的發展並非達到精熟（亦即正確率高）即可，認字速度的快慢也會影響閱讀理解的程度。Paris 與 Carpenter（2003）曾報導兩篇相關研究，這兩篇相關研究均指出：認字正確率與文章理解只有在小學三年級之前有顯著相關，三年級之後兩者的相關就不顯著了。美國國家教育發展評量（National Assessment of Educational Progress，NAEP）計畫中的一項朗讀研究也發現：就國小四年級而言，朗讀正確率與朗讀流暢度兩者間無顯著相關；但是，朗讀流暢度與閱讀理解間卻有顯著相關（Pinnell et al., 1995）。另外，Rasinski 與 Padak（1998）的一項研究也發現：被轉介需要接受閱讀補救教學的學生大多主要的困難是反映在閱讀流暢度上，而較少是在認字正確率或是文章理解上。

其他一些研究亦指出：認字技能自動化對於閱讀理解的增進是很重要的（Adams, 1990；Laberge & Samuels, 1974）。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，變成不需要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。認字技能自動化之所以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，而可以有較多的認知資源用在理解上。相反的，如果讀者的認字技能未達自動化程度，則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（Perfetti, 1985；Samuels, 1985, 2006）。因此，認字技能的自動化是良好閱讀理解的一個重要條件。

在閱讀能力習得的有關研究中，學者們也發現有個稱為「馬修效應」（Matthew effect）的現象，也就是認字能力弱的學生，由於不太讀得懂文章，因此平常閱讀的文章也就比較少，以致於其閱讀能力落後同儕越來越多；相反地，認字能力強的學生，因為大多數閱讀理解表現也比較好，因此平常閱讀的書籍或文章比較多，所以其閱讀能力也越來越好（Goldman & Pellegrino, 1987；Jones, Torgesen, & Sexton, 1987；Torgesen, 1986）。因此認字能力的診斷與補救教學在基礎教育中是非常重要的。

然而，就認字的評量工具而言，目前已發展出來的測驗大多是著重在正確率與字彙量的了解，對於認字自動化的評估工具，除了 Su（1997）曾嘗試找出學童認字自動化的評估指標之外，鮮少有其他研究進行此項評量工具的發展。因此本研究的目的即以 Su（1997）的研究為基礎，嘗試找出國小學童適用的中文認字自動化評量指標，並據以做為建構認字自動化評量系統之參考。以下則就國內外有關認字技能自動化及認字自動化評量指標的研究做進一步的討論。

### 一、認字技能自動化的有關研究

在認字技能自動化的研究方面，Samuels, Laberge, 與 Bremer（1978）曾進行一個研究，以國小二、四、六年級閱讀能力中等的學生，及一組大學生為受試者，研究認字單位如何隨著認字能力的發展而改變。在這個研究中，每位受試者都要接受一個字彙分類作業（word categorization task）。這個作業是以電腦控制進行。電腦螢幕一次會呈現一個字，這些刺激字有些是三個字母長，有些是四個字母長，有些是五個字母長，有些是六個字母長，而刺激字的字頻均控制在一定的範圍之內。受試者需按鍵回答所看到的字是動物類的字，或是非動物類的字。

研究結果發現：受試者的反應時間除了隨著年級遞降之外，年級與字母長度之間也具顯著的交互作用。二年級學生的反應時間會隨著字母長度增加，六年級學生與大學生的反應時間則不會隨著字母長度增加，四年級學生的表現則介於兩者之間。Su（1997）重覆 Samuels 等人（1978）的研究程序，也得到類似的結果。由此結果看來，二年級學生在認字時，似乎傾向於一個字母一個字母地處理；六年級學生與大學生則傾向於以整個字為處理單位；四年級學生則介於兩者之間，有些短的字或容易的字，就以整個字為處理單位，至於長一些或難一些的字，就以字母或字母串為處理單位。

McCormick 與 Samuels（1980）以閱讀能力中等與閱讀能力較低的二年級學生進行一個類似的研究。結果發現：閱讀能力較佳的讀者相較於閱讀能力較低的讀者在認字時，傾向於以較大的單位去處理單字。其他如 Santa（1976-1977）及 Mason（1976）的研究也有相似的結果。

此外，一些使用眼球追蹤法所進行的研究也指出：閱讀能力的高低對認字單位的大小會有所影響。Taylor、Frackenpohl 及 Pettee（1960）發現：小學一年級學生平均看一個字會有兩個目光駐留；但十一年級的學生看一整個字只需要一個目光駐留。Buswell（1937）的研究也指出：低閱讀能力者在每一行的目光駐留數是高閱讀能力者的四到七倍。

由這些研究結果可以得知：就英文讀者而言，認字單位是會隨著讀者認字能力的增加而呈發展上的變化。剛開始時，閱讀的初學者傾向於以像字母這樣的小單位來處理認字的問題；然而，隨著認字能力的增進，讀者會越來越傾向於以整個字做為認字的處理單位。

至於在中文認字能力的發展上是否也有類似的現象呢？Su 與 Samuels（2007）的研究曾探討此問題。這個研究的實驗一中，以小學二、四、六年級閱讀能力中等的學生，及一組大學生為受試者，

每位受試者都要接受一個字彙辨認作業 (lexical recognition task) (Hue & Tzeng, 2000)。這個作業是以電腦控制進行。電腦螢幕一次會呈現一個字，這些字中有 15 個是低筆劃的 (筆劃數介於 4-7 劃之間，平均筆劃數為 5.53)，有 15 個是中筆劃的 (筆劃數介於 11-15 劃之間，平均筆劃數為 13.4)，有 15 個是高筆劃的 (筆劃數介於 20-23 劃之間，平均筆劃數為 21.33)。這 45 個字均為小學二年級水準的高頻字。受試者需按鍵回答「認得」或「不認得」螢幕上的字。為避免受試者不認真做判斷，只一味地按「認得」鍵，所以在高、中、低三個筆劃組的刺激字中均各又安排 15 個假字 (亦即看起來很像中文字，但實際上並不存在於中文裡的字)，共計 90 個刺激字，以隨機的順序出現。

研究結果發現：小學二年級學生在字彙辨認作業中，其認字反應時間隨著刺激字筆劃數的增加而遞增；但小學四年級、六年級的學生及大學生的認字反應時間則不隨著刺激字筆劃數的增加而改變。由此研究結果看起來：小學二年級學生認字的處理單位似乎與筆劃數或刺激字的複雜度有關，因此其認字反應時間隨著刺激字筆劃數的增加而遞增；但小學四年級、六年級的學生與大學生則傾向於以整個字做為認字的處理單位，所以其認字反應時間則不隨著刺激字筆劃數的增加而改變。此外，各年級的認字平均反應時間則隨著年級的增高而遞減。換言之，隨著認字技能越熟練 (亦即自動化程度越高)，認字的平均反應時間會越低，而且反應時間也比較不會隨著刺激字筆劃數的增加而遞增，亦即在橫軸為刺激字複雜度的反應時間圖中，斜率會越低。

此外，楊憲明 (民 87) 的研究中曾以國小五年級閱讀能力中等以上的學生以及閱讀障礙學生為研究對象，結果發現閱讀能力中等以上的國小五年級學生，在字彙判斷作業 (lexical decision task) 的反應時間不受字元複雜度 (即筆劃多寡) 所影響；但閱讀障礙學生的反應時間則會受字元複雜度的影響。根據這個研究結果，閱讀能力中等以上的國小五年級學生似乎傾向於以整個字元做為認字的處理單位，而閱讀障礙學生的認字單位則與筆劃有關。

因此，在中文認字能力的發展上似乎也具有與英文認字技能發展類似的現象，亦即：文字的初學者傾向於以較小的單位來進行認字的處理；然而，隨著認字能力的提高或自動化程度的增加，讀者會越來越傾向於以整個字元做為認字的處理單位。另一方面，隨著認字技能的提升，認字的反應時間也越低。

## 二、認字自動化的評量指標

在認字自動化的評量指標研究方面，Su (1997) 曾以 Samuels 等人 (1978) 的研究為基礎，嘗試找出一般教師可在教室中使用的認字自動化評量指標。由於在 Samuels 等人 (1978) 的研究中，學生的反應時間除了隨著年級的增高而遞減之外，年級與單字長度之間也具有顯著的交互作用，亦即：二年級學生的反應時間會隨著單字長度增加，六年級學生與大學生的反應時間則不會隨著單字長度增加，四年級學生的表現則介於兩者之間。換言之，隨著認字技能越熟練 (即自動化程度越高)，在字彙分類作業的結果中，以橫軸為刺激字複雜度的反應時間圖之斜率也會越低。因此這個字彙分類作業中的平均反應時間與斜率應可做為認字自動化的參考指標。Su (1997) 的研究中即以這兩項指標做為效標變項，分別看幾個預測變項 (如：標準化認字測驗的得分，文章朗讀速率，文章朗讀的聲音表情分數，文章朗讀後的閱讀理解分數，及朗讀後的文章內容回憶率等) 與這兩項效標變項的關係。結果發現：就二年級而言，朗讀的聲音表情分數與字彙分類作業中的斜率有顯著相關，而朗讀速率則與字彙分類作業的平均反應時間有顯著相關；就四年級而言，則朗讀速率與字彙分類作業的斜率及平均反應時間均有顯著相關。而朗讀速率也是目前美國特教界推廣運用的「課程本位評量」(curriculum-based measurement, 簡稱 CBM) 中，信度與效度最佳的一個認字能力或閱讀能力評量指標。只是朗讀速率的測量必須由老師一對一施測，就師資人力而言並不經濟 (Fuchs & Fuchs, 1992)。因此若能發展出可以在電腦上進行的認字自動化評量指標，則對老師與家長而言將是很有幫助的工具。

由於如前所述，在中文的研究中，Su 與 Samuels (2007) 也發現認字自動化程度越高的學生，在字彙辨認作業中的認字平均反應時間會越低，而且以刺激字複雜度為預測變項，以反應時間為效標變項之迴歸線的斜率也會越低。因此本研究嘗試以 Su (1997) 及 Su 與 Samuels (2007) 的研究為基礎，探討以字彙辨認作業中的平均反應時間與斜率做為認字自動化指標的可行性，檢驗這兩個指標的信度與效度，並據以做為建構認字自動化評量系統之參考。

另外，認字自動化並非「全有或全無」的能力變化，而是在認字能力發展中對於學過或接觸過的文字，其認字速度相對性的改變。例如：就一個國小五年級閱讀能力中等的學生而言，他對於國小二年級程度文字的辨識可能已達自動化的水準，但是對於五年級程度的文字可能尚不熟悉或是尚未達到自動化水準，因此這位學生對於二年級程度的文章應該可以輕鬆自如地閱讀並理解，但對於五年級程度的文章在閱讀理解上可能就會比較費力。為了初探性地探討字彙辨認作業中的平均反應時間與斜率是否能夠反映國小學童認字自動化的程度，本研究以國小二年級水準的高頻字做為刺激字，並以初學這些字的二年級學童以及熟悉這些字的四年級與六年級學童分別代表不同認字熟練程度的受試者，藉以檢驗字彙辨認作業中的平均反應時間與斜率是否適合做為認字自動化的指標。

## 方 法

### 一、研究對象

本研究自台北縣市共 16 所國小選取二年級學童 63 位，四年級學童 64 位，六年級學童 63 位做為研究對象。這些學生均為導師所推薦認字能力在班上屬中等程度的學生。由於認字反應時間的分析必須在受試者認真辨認刺激字，而且正確率在一定程度以上始具有意義。因此在本研究中，正確率低於 .70 的受試者資料不納入分析。各項分析的有效樣本人數將於結果部分說明。

### 二、刺激字的選取

本研究共選取兩套複本的刺激字。複本 A 乃以 Su 與 Samuels (2007) 實驗一中所使用的 45 個真字及 45 個假字做為刺激字。其中的 45 個真字是從國立編譯館 (民 88) 的「國民小學常用字彙研究字頻總表」中選取高頻次低複雜度的字 15 個，中複雜度的字 15 個，及高複雜度的字 15 個 (見附錄一)。在此所謂的高頻次乃指 100 萬字元的計次中，有 200 次以上使用出現次數的詞彙 (Cheng, 1982)，而「國民小學常用字彙研究字頻總表」所收錄語料的字彙總頻次為 1,419,219。至於低複雜度字元則指筆劃數介於 4-7 劃之間的字，中複雜度字元為筆劃數介於 11-15 劃之間的字，高複雜度字元則是筆劃數介於 20-23 劃之間的字。複本 A 三個不同筆劃水準刺激字之筆劃與頻次的平均數、標準差如附錄一所示。本研究選擇高頻次字元做為刺激字的原因是考慮小學二年級學童對中頻次或低頻次的字元不熟悉，可能導致高錯誤率，而影響反應時間的分析。此外，為避免受試者不認真做判斷，只一味地按「認得」鍵，因此在刺激字中也加入低、中、高複雜度的假字各 15 個，共計 90 個刺激字，以隨機的順序出現，但是假字的反應時間並不予以分析。

複本 B 則根據 Su 與 Samuels (2007) 實驗一刺激字的選取原則另外挑選 45 個真字及設計 45 個假字做為刺激字 (見附錄二)。複本 B 三個不同筆劃水準刺激字之筆劃與頻次的平均數、標準差亦如附錄二所示。

### 三、研究程序

本研究所採用的字彙辨認作業為一個需在電腦上進行的測驗。在正式測驗中，每個嘗試之前，

受試者都會先在電腦螢幕中央看到「+」，這時受試者若準備好要做這一題，就按空白鍵。按了空白鍵後，電腦螢幕中央就會出現一個字，如果受試者認得這個字（例如真字），就要按「○」（即鍵盤上的「/」鍵）；如果不認得這個字（例如假字），就要按「X」（即鍵盤上的「Z」鍵）。當受試者按了「○」或「X」的鍵後，螢幕中央的字就會消失掉，而「+」又會出現在螢幕中央。如果受試者遲遲未作答，則過了5秒鐘，螢幕中央的字就會消失掉，螢幕又會回到「+」的畫面，準備進行下一個嘗試。在這個測驗中，受試者在每一題的反應對錯、反應時間、及整體平均反應時間、斜率都會自動紀錄在電腦檔案中。

爲了探討字彙辨認作業中的平均反應時間與斜率做爲認字自動化指標的可行性，本研究對這兩個指標進行信度與效度的考驗。在信度研究部分，本研究以間隔一週的重測信度（郭生玉，民74），及複本A與複本B的複本信度考驗這兩個指標的穩定性。重測信度部分之所以採取一週的間隔時間，主要的考慮是：本研究所探討的認字自動化指標屬於一種能力指標，它會隨著學童在國語課或國語作業上的練習量多寡而呈現程度上的變化，因此兩次施測的間隔時間如果太長，可能會因間隔期間學童認字練習量或中文字接觸量的介入，而使得第二次施測時，學童的認字熟練度呈現水準上的變化，導致信度係數除反映指標穩定度之外，也受到學童在兩次施測間隔期間認字練習量的影響。故本研究之重測信度兩次施測時間採取一週的間隔。

此外，本研究也以效標關聯效度進行效度考驗。本研究採用的兩個效標爲：黃秀霜（民90）所編製的「中文年級認字量表」，及柯華葳（民88）所編製的「閱讀理解困難篩選測驗」。選擇這兩個測驗的得分做爲效標的原因爲國內目前尚無評估中文認字自動化程度的工具，而通常認字自動化程度較高的學生其識字量也較高，閱讀理解的表現也較好，因此本研究採用「中文年級認字量表」及「閱讀理解困難篩選測驗」做爲效標。此外，「閱讀理解困難篩選測驗」的題本有二、三年級版（18題）與四、五、六年級版（20題），兩個版本有4道重疊一樣的題目，由於本研究的受試者包含二、四、六年級的學童，因此本研究乃將兩版本的題目彙整爲34題的題本，做爲效度考驗的工具。

每位參與研究的國小學生均需接受兩次測驗。第一次測驗時，小朋友要做字彙辨認作業的複本A及複本B，以及「中文年級認字量表」與「閱讀理解困難篩選測驗」。間隔一週之後的第二次測驗，則是再做一次字彙辨認作業的複本A及複本B。

## 結 果

本研究的結果將分成三部分做分析，分別是：一、檢驗 Su 與 Samuels（2007）實驗一的結果是否可重複驗證；二、考驗認字自動化指標的信度；三、考驗認字自動化指標的效度。

### 一、檢驗 Su 與 Samuels（2007）實驗一結果的可重複驗證性（replicability）

由於本研究以 Su 與 Samuels（2007）的研究爲基礎，根據不同年級受試者在該研究實驗一中的反應組型，探討以字彙辨認作業中的平均反應時間與斜率做爲認字自動化指標的可行性，因此本研究第一部分的分析需先檢驗該研究實驗一的結果是否可重複驗證。Su 與 Samuels 的實驗一發現受試者在字彙辨認作業中的反應時間呈現年級與刺激字複雜度的交互作用，亦即小學二年級學生的認字反應時間隨著刺激字筆劃數的增加而遞增；但小學四年級、六年級的學生及大學生的認字反應時間則不隨著刺激字筆劃數的增加而改變。此外，各年級的認字平均反應時間則隨著年級的增高而遞減。本研究的複本 A 乃採用 Su 與 Samuels 實驗一的刺激字，爲檢驗該實驗的結果是否可重複驗證，本研究以受試者在複本 A 的反應時間資料進行 3（年級）X3（刺激字複雜度）二因子混合設計變異數分析。其中「年級」變項有三個水準，分別是小學二、四、六年級；「刺激字複雜度」變項有三個水準，分別

是低、中、高複雜度。由於認字反應時間的分析必須在受試者認真辨認刺激字，而且正確率在一定程度以上始具有意義，因此正確率在 .70 以下的受試者資料在本研究中不納入分析。有效樣本為二年級學童 44 位，四年級學童 63 位，六年級學童 63 位，共計 170 位受試者。

分析結果為：「年級」的主要效果 ( $F(2, 167) = 99.45, MSE = 126,292.21, p < .001$ )、「刺激字複雜度」的主要效果 ( $F(2, 334) = 14.45, MSE = 15,160.23, p < .001$ )、「年級」與「刺激字複雜度」的交互作用效果 ( $F(4, 334) = 3.71, MSE = 15,160.23, p < .01$ ) 均達顯著水準。至於各年級在「刺激字複雜度」的單純主要效果分析則顯示：二年級 ( $F(2, 334) = 6.06, MSE = 15,160.23, p < .008$ ) 與四年級 ( $F(2, 334) = 13.31, MSE = 15,160.23, p < .001$ ) 的單純主要效果達顯著水準，但六年級則未達顯著水準 ( $F(2, 334) = 2.89, MSE = 15,160.23, p = .057$ )。各年級在各刺激字複雜度水準的平均反應時間與標準差如表 1 所示。此分析結果與 Su 與 Samuels (2007) 實驗一的結果一致，亦即初學識字或認字技能較不熟練的學生其認字反應時間隨著刺激字筆劃數的增加而遞增；但較高年級或認字技能較熟練的學生其認字反應時間則不隨著刺激字筆劃數的增加而改變。另外，各年級的認字平均反應時間則隨著年級的增高而遞減。各年級的反應時間組型如圖 1 所示。

## 二、認字自動化指標的信度考驗

根據前述的分析，字彙辨認作業中的平均反應時間與斜率似可做為認字自動化的指標，因此本研究第二部分的分析以間隔一週的重測信度，及複本 A 與複本 B 的複本信度考驗這兩個指標的穩定性。由於認字反應時間的分析必須在受試者認真辨認刺激字，而且正確率在一定程度以上始具有意義，因此正確率低於 .70 的受試者資料在本研究中不納入分析。依此原則，複本 A 的重測信度分析，有效樣本為 168 人，複本 B 的重測信度分析，有效樣本為 155 人。另外，在複本信度部分，以受試者在第一次受測時的複本 A 與複本 B 資料做分析，有效樣本為 164 人。

表 1 各年級在各刺激字複雜度水準的平均反應時間與標準差

		低複雜度	中複雜度	高複雜度
二年級 ( <i>N</i> = 44)	平均數	1377.90	1410.07	1468.05
	標準差	273.99	278.39	294.75
四年級 ( <i>N</i> = 63)	平均數	1018.24	984.70	1095.07
	標準差	194.06	208.85	270.96
六年級 ( <i>N</i> = 63)	平均數	873.18	823.93	864.97
	標準差	184.76	172.18	193.12

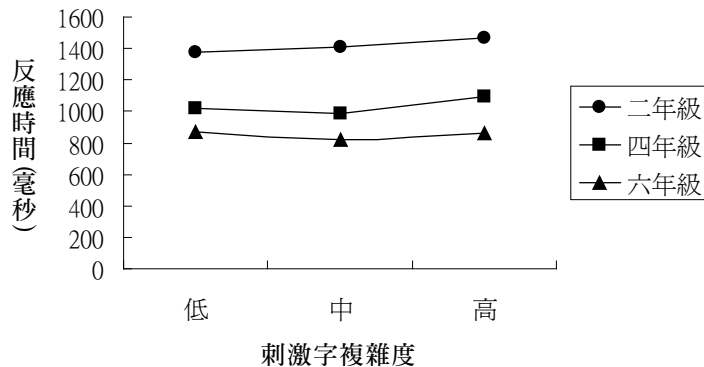


圖 1 各年級在各刺激字複雜度水準的反應時間組型

進行資料分析時，首先就有效樣本計算每位受試者在字彙辨認作業複本 A 與複本 B 的平均反應時間，並計算在複本 A 與複本 B 中以刺激字複雜度為預測變項，以反應時間為效標變項之迴歸線的斜率。由於有些受試者的斜率為負值，而從 Su (1997)、Su 與 Samuels (2007)、以及本研究第一部分的分析來看，負值的斜率似也反映受試者的反應時間不隨著刺激字複雜度的增加而增加，亦即其意義與斜率 0 是一樣的，因此本研究在斜率部分以兩種資料處理方式進行分析。其中一種是以原始計算所得的斜率（含正、負值，以下均以「原始斜率」稱之）進行信效度分析，另一種方式是將負值的斜率均轉換為 0，正值的斜率維持不變（以下均以「轉換斜率」稱之），然後再進行信效度分析。

分析結果如下：在重測信度的部分，複本 A 「平均反應時間」的重測信度係數為 .760 ( $p < .001$ )，「原始斜率」部分的重測信度係數為 .054 ( $p > .05$ )，「轉換斜率」部分的重測信度係數為 .286 ( $p < .001$ )；複本 B 「平均反應時間」的重測信度係數為 .742 ( $p < .001$ )，「原始斜率」部分的重測信度係數為 .069 ( $p > .05$ )，「轉換斜率」部分的重測信度係數為 .131 ( $p > .05$ )。至於複本信度的部分，複本 A 與複本 B 的複本信度係數，在「平均反應時間」部分為 .864 ( $p < .001$ )，在「原始斜率」部分為 .073 ( $p > .05$ )，「轉換斜率」部分則為 .273 ( $p < .001$ )。

根據上述分析結果，字彙辨認作業中的「平均反應時間」不論是複本 A 的重測信度、複本 B 的重測信度、或是複本 A 與複本 B 的複本信度考驗均達 .001 的顯著水準，信度係數介於 .742 ~ .864 之間。相反地，在「原始斜率」部分則不論是複本 A、複本 B 的重測信度，或是複本信度則均未達顯著水準。「轉換斜率」雖然在複本 A 的重測信度、及複本信度考驗均達 .001 的顯著水準，但是信度係數均很低，分別是 .286 與 .273；至於複本 B 的重測信度則未達顯著水準。

### 三、認字自動化指標的效度考驗

在效度考驗部分，本研究以受試者在第一次受測時的測驗資料進行效標關聯效度分析。本研究所採用的兩個效標為：黃秀霜（民 90）所編製的「中文年級認字量表」，及柯華葳（民 88）所編製的「閱讀理解困難篩選測驗」。如前所述，受試者在字彙辨認作業中正確率低於 .70 的資料在本研究中不納入分析，因此，複本 A 效度分析的有效樣本為 170 人，複本 B 效度分析的有效樣本為 168 人。

複本 A 的效度考驗結果為：以「中文年級認字量表」為效標時，「平均反應時間」與「中文年級認字量表」的效標關聯效度係數為  $-.543$  ( $p < .001$ )，「原始斜率」與「中文年級認字量表」的效標關聯效度係數為  $-.142$  ( $p > .05$ )，「轉換斜率」與「中文年級認字量表」的效標關聯效度係數為  $-.193$  ( $p < .05$ )。以「閱讀理解困難篩選測驗」做為效標時，複本 A 「平均反應時間」與「閱讀理解困難篩選測驗」的效標關聯效度係數為  $-.534$  ( $p < .001$ )，「原始斜率」與「閱讀理解困難篩選測驗」的

效標關聯效度係數為 $-.136$  ( $p > .05$ )，「轉換斜率」與「閱讀理解困難篩選測驗」的效標關聯效度係數為 $-.205$  ( $p < .01$ )。

複本 B 的效度考驗結果為：以「中文年級認字量表」為效標時，「平均反應時間」與「中文年級認字量表」的效標關聯效度係數為 $-.461$  ( $p < .001$ )，「原始斜率」與「中文年級認字量表」的效標關聯效度係數為 $-.083$  ( $p > .05$ )，「轉換斜率」與「中文年級認字量表」的效標關聯效度係數為 $-.182$  ( $p < .05$ )。以「閱讀理解困難篩選測驗」做為效標時，複本 B 「平均反應時間」與「閱讀理解困難篩選測驗」的效標關聯效度係數為 $-.438$  ( $p < .001$ )，「原始斜率」與「閱讀理解困難篩選測驗」的效標關聯效度係數為 $-.040$  ( $p > .05$ )，「轉換斜率」與「閱讀理解困難篩選測驗」的效標關聯效度係數為 $-.083$  ( $p > .05$ )。

根據上述分析結果，字彙辨認作業中不論複本 A 或複本 B 的「平均反應時間」與「中文年級認字量表」及「閱讀理解困難篩選測驗」的效標關聯效度考驗均達 $.001$ 的顯著水準，效度係數介於 $-.543 \sim -.438$ 之間。「原始斜率」部分則不論是複本 A、複本 B 的效標關聯效度考驗均未達顯著水準。「轉換斜率」部分，雖然複本 A 與「中文年級認字量表」及「閱讀理解困難篩選測驗」的效標關聯效度，以及複本 B 與「中文年級認字量表」的效標關聯效度達 $.05$ 顯著水準，但效度係數均很低，介於 $-.205 \sim -.182$ 之間；至於複本 B 與「閱讀理解困難篩選測驗」的效標關聯效度則未達顯著水準。

## 討 論

本研究以 Su 與 Samuels (2007) 的研究為基礎，嘗試探討以字彙辨認作業中的兩項測量資料做為認字自動化指標的可行性，其中一個指標為平均反應時間，另一個指標為以刺激字複雜度為預測變項，以反應時間為效標變項之迴歸線的斜率。在 Su 與 Samuels 的實驗一中，受試者在字彙辨認作業中的平均反應時間不僅隨著年級的增高而遞減，而且也呈現年級與刺激字複雜度的交互作用，亦即小學二年級學童的認字反應時間隨著刺激字筆劃數的增加而遞增；但小學四年級、六年級的學童及大學生的認字反應時間則不隨著刺激字筆劃數的增加而改變。本研究的複本 A 採用 Su 與 Samuels 實驗一的刺激字，根據受試者在第一次受測時所得資料之分析亦重覆驗證了 Su 與 Samuels 實驗一的結果。此結果也與 Samuels 等人 (1978) 及 Su (1997) 在美國以英文字為實驗材料，以美國學生為受試者所觀察到的現象一致。這幾個研究的結果均顯示：隨著認字技能的提升（亦即自動化程度的增加），認字的平均反應時間會逐漸減低；另外，文字的初學者傾向於以較小的單位來進行認字的處理，但是隨著認字能力的提高或自動化程度的增加，讀者會越來越傾向於以整個字元做為認字的處理單位。

至於以字彙辨認作業中的平均反應時間及斜率做為認字自動化的指標是否具有可行性，本研究則以這兩項指標的信度與效度考驗進行探討。在信度考驗部分，字彙辨認作業中的「平均反應時間」不論是複本 A、複本 B 的重測信度，或是複本 A 與複本 B 的複本信度考驗均達 $.001$ 的顯著水準，信度係數介於 $.742 \sim .864$ 之間。反之，在「原始斜率」部分則不論是複本 A、複本 B 的重測信度，或是複本信度則均未達顯著水準。「轉換斜率」雖然在複本 A 的重測信度、及複本信度考驗均達 $.001$ 的顯著水準，但是信度係數均很低，分別是 $.286$ 與 $.273$ ；至於複本 B 的重測信度則未達顯著水準。根據此研究結果，字彙辨認作業中的「平均反應時間」具有較佳的穩定度，但是「斜率」則否。相較於以答對題數計分的能力測驗，以毫秒為單位的反應時間測量更容易受到不確定因素的影響而變動，因此不容易像以正確題數計分的能力測驗可達 $.85$ 甚至 $.90$ 以上的信度。但即使如此，字彙辨認作業中「平均反應時間」的信度係數仍可達 $.742 \sim .864$ 之間，因此就關心辨識速度的認字自動化程度之評量，字彙辨認作業的「平均反應時間」應屬穩定的指標。



在效度考驗部分，不論複本 A 或複本 B 的「平均反應時間」與「中文年級認字量表」及「閱讀理解困難篩選測驗」的效標關聯效度考驗均達 .001 的顯著水準，效度係數介於  $-.543 \sim -.438$  之間。相反地，「原始斜率」部分則不論是複本 A、複本 B 的效標關聯效度考驗均未達顯著水準。

「轉換斜率」部分，雖然複本 A 與「中文年級認字量表」及「閱讀理解困難篩選測驗」的效標關聯效度，以及複本 B 與「中文年級認字量表」的效標關聯效度達 .05 顯著水準，但效度係數均很低，介於  $-.205 \sim -.182$  之間；至於複本 B 與「閱讀理解困難篩選測驗」的效標關聯效度則未達顯著水準。因此字彙辨認作業的「平均反應時間」具有較好的效度，但是「斜率」則否。

關於「斜率」在信、效度考驗的結果均不理想，原因可能與其變異量低有關。就 Su (1997) 的「轉換斜率」而言，二、四、六年級共 87 位受試者的轉換斜率標準差為 .13。而本研究中，以複本 A 第一次施測資料所得的轉換斜率標準差為 .07，複本 A 第二次施測資料所得的轉換斜率標準差為 .05，複本 B 第一次施測資料所得的轉換斜率標準差為 .08，複本 B 第二次施測資料所得的轉換斜率標準差為 .05。由於其變異極低，所以可能使得此指標不如「平均反應時間」對認字自動化程度來得敏感與穩定，因此其信、效度分析結果均不理想。

至於字彙辨認作業的「平均反應時間」與「中文年級認字量表」及「閱讀理解困難篩選測驗」的效標關聯效度係數只具中等相關，則可從以下三方面加以討論。首先，由於「中文年級認字量表」乃是測量識字量的工具，「閱讀理解困難篩選測驗」為評估閱讀理解能力的測驗，因此兩者雖與認字自動化程度有關，但並非直接測量認字自動化程度的工具，因此字彙辨認作業的「平均反應時間」與「中文年級認字量表」及「閱讀理解困難篩選測驗」的效標關聯效度係數不高，亦具合理性。

其次，就自動化發展的理論而言，認字技能從初學到自動化水準的發展通常會經過三個階段，分別是：「錯誤期」(nonaccurate stage)，「正確而未達自動化期」(accurate but not automatic stage)，「正確且達自動化期」(accurate and automatic stage，亦即流暢期)(Samuels, 2002)。「中文年級認字量表」是以認字反應的正確性做為計分的依據，因此所測得的能力與「錯誤期」及「正確而未達自動化期」較有關，但無法區辨「正確而未達自動化期」與「正確且達自動化期」的差異。相反地，字彙辨認作業的「平均反應時間」就可區辨後兩個階段的差別。由於兩者所測量的認字能力層面不同，因此自然不易有高相關。

另外，就閱讀理解理論而言，固然讀者的認字效能是影響閱讀理解的因素之一，但除此之外，句法能力、推論能力、口語理解能力、後設認知能力、以及與文章主題有關的先備知識等，也都會影響閱讀理解的表現 (van den Broek & Kremer, 2000)。因此，字彙辨認作業「平均反應時間」所反映的認字效能與「閱讀理解困難篩選測驗」所測得的閱讀理解能力只具中等相關，也符合理論的預測。

第三，Jenkins、Fuchs、van den Broek、Espin 和 Deno (2003) 的研究中也有類似的發現。這個研究是以四年級學童為研究對象。Jenkins 等人將「父親、兒子、與他們的驢」這篇故事改寫為大約三年級程度的文章，然後將文章中的字隨機排列成一個字單，做為研究工具。受試者需以盡快的速度唸這個字單，研究者則紀錄受試者的正確率(唸對字數)及速度(一分鐘唸對的字數)。此外，Jenkins 等人也以愛荷華基本能力測驗 (Iowa Test of Basic Skills) 中的閱讀理解分測驗測量受試者的閱讀理解能力。分析結果為：認字速度與認字正確率的相關為 .66，認字速度與閱讀理解的相關為 .53，均屬中等相關。此外，Speece 與 Ritchey (2005) 以讀字效能測驗 (Test of Word Reading Efficiency) 測量認字速度，以基本閱讀技能測驗 (Basic Reading Skills cluster，是 Woodcock-Johnson Psychoeducational Battery-Revised 中的分測驗) 測量認字正確率，結果認字速度與認字正確率兩者的相關為 .49 ( $p < .05$ )，亦屬中等相關。

根據上述三部分的討論，可做以下歸納：認字自動化程度、認字正確度、與閱讀理解程度，三者彼此相關，但非完全等同。因此，認字自動化程度與認字量表得分及閱讀理解測驗的表現，通常具中

等相關。

綜合以上的分析，字彙辨認作業的「平均反應時間」為一穩定且具參考價值的認字自動化評量指標。但是，以刺激字複雜度為預測變項，以反應時間為效標變項之迴歸線的斜率則不具穩定度，也缺乏良好的效度，故不適合做為認字自動化的評量指標。而以字彙辨認作業的「平均反應時間」做為認字自動化的評量指標，正可提供認字自動化評量系統之建構的參考。雖然 Su (1997) 的研究指出朗讀速率也是可參考的認字自動化評量指標，但是朗讀速率的測量必須由老師一對一施測，就師資人力而言並不經濟 (Fuchs & Fuchs, 1992)。因此若能以字彙辨認作業的「平均反應時間」為指標，發展出可以在電腦上進行的認字自動化評量系統，則對老師與家長而言將是很有幫助的工具。

### 參 考 文 獻

- 柯華歲 (民 88)：閱讀理解困難篩選測驗。台北：行政院國家科學委員會特殊教育工作小組。
- 郭生玉 (民 74)：心理與教育測驗。台北：精華。
- 國立編譯館 (民 88)：國民小學常用字彙研究字頻總表。台北：國立編譯館。
- 黃秀霜 (民 90)：中文年級認字量表。台北：心理。
- 楊憲明 (民 87)：閱讀障礙學生文字辨識自動化處理之分析研究。特殊教育與復健學報，6 卷，15-37 頁。
- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: The MIT Press.
- Bertelson, P. (1986). The onset of literacy: Liminal remarks. *Cognition*, 24, 1-30.
- Biemiller, A. (1977-1978). Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Research Quarterly*, 13, 223-253.
- Buswell, G. T. (1937). *Fundamental reading habits: A study of their development* (Supplementary educational monographs No. 21). Chicago, IL: University of Chicago Press.
- Calfee, R. C., & Piontkowski, D. C. (1981). The reading diary: Acquisition of decoding. *Reading Research Quarterly*, 16, 346-373.
- Cheng, C.-M. (1982). Analysis of present day Mandarin. *Journal of Chinese Linguistics*, 10, 282-358.
- Curtis, M. (1980). Development of components of reading skill. *Journal of Educational Psychology*, 72, 656-669.
- Fuchs, L., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review*, 21(1), 45-58.
- Goldman, S. R., & Pellegrino, J. W. (1987). Information processing and educational microcomputer technology: Where do we go from here? *Journal of Learning Disabilities*, 20, 144-154.
- Herman, P. A. (1985). The effect of repeated readings on reading rate, speech pauses, and word recognition accuracy. *Reading Research Quarterly*, 20, 553-565.
- Hue, C. W., & Tzeng, A. K. (2000). Lexical recognition task: A new method for the study of Chinese character recognition. *Acta Psychologica Sinica*, 32, 60-65.
- 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.
- Jone, K. M., Torgesen, J. K., & Sexton, M. A. (1987). Using computer guided practice to increase decoding fluency in learning disabled children: A study using the Hint and Hunt I program. *Journal of Learning*

- Disabilities*, 20, 122-128.
- Juel, C., Griffith, P. L., Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, 78, 243-225.
- Laberge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Mason, J. M. (1976). Overgeneralization in learning to read. *Journal of Reading Behavior*, 8, 173-182.
- McCormick, C., & Samuels, S. J. (1980). Word recognition by second graders: The unit of perception and interrelationships among accuracy, latency, and comprehension. *Journal of Reading Behavior*, 11, 121-130.
- Paris, S. G., & Carpenter, R. D. (2003). FAQs about IRIs. *The Reading Teacher*, 56, 578-580.
- Perfetti, C. (1985). *Reading ability* (pp. 99-122). New York, NY: Oxford University Press.
- Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). *Listening to children read aloud: Data from NAEP's Integrated Reading Performance Record (IRPR) at grade 4*. Washington, DC: National Center for Education Statistics.
- Rasinski, T. V., & Padak, N. D. (1998). How elementary students referred for compensatory reading instruction perform on school-based measures of word recognition, fluency, and comprehension. *Reading Psychology: An International Quarterly*, 19, 185-216.
- Roth, S. F., & Beck, I. L. (1987). Theoretical and instructional implications of the assessment of two microcomputer word recognition programs. *Reading Research Quarterly*, 22, 197-218.
- Samuels, S. J. (1985). Word recognition. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 256-276). Newark, DE: International Reading Association.
- 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.
- Samuels, S. J. (2006). Toward a model of reading fluency. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 24-46). Newark, DE: International Reading Association.
- Samuels, S. J., LaBerge, D., & Bremer, C. D. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*, 17, 715-720.
- Santa, C. M. (1976-1977). Spelling patterns and the development of flexible word recognition strategies. *Reading Research Quarterly*, 12(2), 125-144.
- Speece, D. L., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities*, 38(5), 387-399.
- Stanovich, K. E. (1982). Individual differences in the cognitive processes of reading I: Word decoding. *Journal of Learning Disabilities*, 15, 485-493.
- Stanovich, K. E. (1985). Explaining the variance in reading ability in terms of psychological processes: What have we learned. *Annals of Dyslexia*, 35, 67-96.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-407.
- Su, Y. F. (1997). *Indicators of automaticity in word recognition*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities, MN.

- Su, Y. F., & Samuels, S. J. (2007). *Developmental changes in word-length effect when reading Chinese script*. Manuscript submitted for publication.
- Taylor, S., Frackenpohl, H., & Pettee, J. (1960). *Grade level norms for the components of the fundamental reading skill* (Bulletin No. 3). Huntington, NJ: Educational Developmental Laboratories.
- Torgesen, J. K., (1986). Computer and cognition in reading: A focus on decoding fluency. *Experimental Children*, 53, 157-162.
- van den Broek, P., & Kremer, K. E. (2000). The mind in action : What it means to comprehend during reading. In B. M. Taylor, M. F. Graves, & P. van den Broek (Eds.), *Reading for meaning : Fostering comprehension in the middle grades* (pp. 1-31). Newark, DE: International Reading Association.

收 稿 日 期：2006 年 09 月 18 日

一稿修訂日期：2006 年 12 月 02 日

接受刊登日期：2007 年 01 月 25 日

附錄一 複本 A 刺激字

低複雜度			中複雜度			高複雜度		
刺激字	筆劃數	頻次	刺激字	筆劃數	頻次	刺激字	筆劃數	頻次
加	5	1604	情	11	1593	體	23	1574
代	5	1153	連	11	1190	覺	20	1230
早	6	895	跑	12	890	歡	22	892
北	5	772	趕	14	749	讀	22	753
句	5	521	練	15	542	護	21	498
另	5	491	語	14	502	響	21	480
仁	4	462	需	14	486	鐵	21	477
切	4	416	隊	12	414	驗	23	397
冷	7	386	豬	15	378	顧	21	375
忍	7	339	靠	15	322	魔	21	336
似	7	325	盤	15	321	顯	23	326
孔	4	299	源	13	291	繼	20	295
志	7	273	換	12	283	鐘	20	278
均	7	206	輪	15	205	蘭	21	242
巧	5	199	碰	13	198	露	21	204
平均數	5.53	556.07	平均數	13.40	557.60	平均數	21.33	557.13
標準差	1.15	380.84	標準差	1.45	382.28	標準差	1.01	381.66

附錄二 複本 B 刺激字

低複雜度			中複雜度			高複雜度		
刺激字	筆劃數	頻次	刺激字	筆劃數	頻次	刺激字	筆劃數	頻次
各	7	1607	帶	11	1515	變	23	1599
合	6	1157	理	11	1185	難	19	924
件	6	880	電	13	886	關	19	912
半	5	791	解	13	752	離	19	692
守	6	507	單	12	545	寶	20	621
巴	4	492	蛇	11	498	驚	23	522
仍	5	471	歌	14	481	願	19	468
村	7	419	境	14	444	識	19	441
良	7	404	鄉	12	413	麗	19	433
左	5	348	適	15	336	壞	19	365
沙	7	306	愈	13	322	羅	19	363
互	4	305	賞	15	300	續	21	323
究	7	273	價	15	290	藥	19	312
司	5	199	隔	13	194	勸	20	188
壯	7	191	敵	15	192	騙	19	178
平均數	5.87	556.67	平均數	13.13	556.87	平均數	19.80	556.07
標準差	1.13	395.61	標準差	1.51	378.69	標準差	1.42	366.05

Bulletin of Educational Psychology, 2007, 38 (4), 501-514  
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

## **Indicators of Character Decoding Automaticity: Reliability and Validity**

Yi-Fen Su

Hsueh-Chih Chen

Department of Educational and Counseling Psychology  
National Taiwan Normal University

The purpose of this research was to evaluate two possible indicators of character decoding automaticity for Chinese learners in elementary schools. One of the possible indicators was mean latency of lexical recognition task. The other possible indicator was the slope predicting latency from character complexity on the lexical recognition task. The test-retest method and equivalent-forms method were adopted to evaluate the reliability of the two indicators. The criterion-related coefficient was used to evaluate the validity of the indicators. The results of reliability and validity indicated that the mean latency of lexical recognition task was a good indicator of character decoding automaticity, while the slope relating latency to character complexity on the lexical recognition task was not. This finding should be useful in establishing a computerized assessment system of decoding automaticity for elementary school students in Chinese character recognition.

**KEY WORDS:** character decoding, automaticity, indicators, elementary school students