

國立臺灣師範大學特殊教育學系、特殊教育中心
特殊教育研究學刊，民88，18期，151—172頁

以文章寫作和造句測驗評估 作文能力之效度研究

葉靖雲

國立彰化師範大學

本研究為以效度為核心，主旨在於發展適用於國小中、高年級學習障礙學生的形成性作文能力評量工具。所發展的測量方式包括由兩種文體（記敘文和說明文）和四種題目類型（主題句、故事開頭、單圖、四個格連環圖）所組成之文章寫作作文能力測驗，使用總數字數及相異字數兩種評分方法。所探究的效度包括以作文和國語文學期成績為效標的效標關聯效度，以及比較不同年級（四、五、六）、能力（學習障礙與一般能力）組別學生表現差異的建構效度。除上述之文章寫作作文能力測驗外，本研究也同時檢驗三種造句測驗分數（意義正確的句子數、成語數、非常用字數）與作文能力效標間的關係，並比較不同年級、作文能力學生在這些造句測驗上的表現。本研究之樣本包括建立效標關聯效度之普通班學生樣本，以及發展建構效度之學習障礙樣本。結果顯示上述文章寫作作文能力測驗的總字數和相異字數與效標間皆有顯著的連結關係，除四年級外，效度係數大致在中度相關的範圍內。各種文章寫作作文測驗的分數也大多能夠有效區分不同年級、作文能力學生。造句測驗中意義正確的句數和非常用字數的效標關聯效度大致在中到高度相關的範圍內，這兩種造句測驗的計分結果能夠清楚鑑別不同年級、能力的學生。

關鍵字：小學、作文評量、文體、題目類型、學習障礙、造句、效度、總字數、相異字數、不同用字數

緒論

一、研究動機與目的

作文與閱讀、數學並列為重要基本學科能力。相關研究指出國小中、高年級以後，隨著年級的升高，作文能力對於學業成就的影響愈

來愈大（Montague & Graves, 1993）。作文又與說話同為自我表達的溝通技能，在人際與社會適應上佔有重要地位。根據學生的表現來調整教學的內涵是有效教學的基本原則之一（Goodman, 1985; Hoffman & Rutherford, 1984; Rieth, Polsgrove & Semmel, 1981），基於作文

能力對於學業學習和語言發展的重要性，教學者宜透過作文能力的持續監控以儘早發現問題並及時進行補救。作文是語文科目中技巧最複雜的一種，對於學習障礙學生而言更是最感困難的基本能力之一，多位研究者指出寫作學習障礙學生之作文成品不論在質和量上都較一般學生落後，作文技巧的缺陷也較一般學生顯著（參見Isaacson, 1988的文獻回顧）。所幸Christenson、Thurlow、Ysseldyke和Vicar（1989）以及Mastropieri和Scruggs（1994）的研究都顯示學習障礙學生的作文能力可以透過有效教學和形成性監控而提昇。然而，監控作文能力的改變需要穩定、可靠的工具。缺乏有效的作文評量工具，教師無法獲得學生作文表現進步與否以及成長幅度方面的明確回饋，也難以客觀的判斷各種作文教學法對於團體或個別學生的有效性。

本研究的主旨在於發展穩定、有效性的形成性作文能力測量工具。發展出來的測驗如果能夠明確區分不同能力的寫作者，則亦可用於寫作學習障礙學生的篩選。作文能力的評量有直接與間接、質與量、主觀與客觀層面的分野。研究者的興趣在於形成性測驗的發展，由於此類測驗需經常性實施，在編製、施測、計分和結果的解釋上必須符合便利和經濟的原則，故本研究所探究的測量方式偏重在客觀、量化的評量層面上。近年來，作業評量（performance assessment）日漸受到重視，已成為教室內評量的主流（Wiggins, 1989, 1993）。此種評量系統強調瞭解學生技能發展的最佳方法就是讓學生從實作中表現此種技能，例如從學生的朗讀表現來評定其認字技巧、以學生的文章寫作成品來評定其作文能力。本研究所發展的形成性作文測驗也採取此種直接評量的方式。

研究者預期達成的目標有以下四項：一、發展五種文章寫作作文能力測驗，包括故事開頭、單圖、四格連環圖片之記敘文寫作，以及

比較性、問題解決性之說明文寫作，學生在這五種作文測驗上的表現以「總字數」和「相異字數」來計分；二、發展用以評估作文能力之造句測驗，學生之表現以三種方式來評定，分別是計算「意義正確的句數」、「成語數」和「非常用字數」；三、比較各種作文能力測驗所得結果鑑別學習障礙與一般能力學生的效果；四、比較不同造句測驗評分方式所得結果鑑別學習障礙與一般學生之效果。

上述目標中，目標一的部分已在先前的研究（葉靖雲，民88）中探究過，此次複製的用意在於確認各種作文能力測驗和計分方式效標關聯效度資料。葉靖雲（民88）的研究顯示，整體而言，文章寫作的測量結果可以有效反映五、六年級學生之作文能力，四年級部分的效度係數較為偏低，可能與文章寫作對於此一年級的學生難度較高有關。在目前的研究裡，研究者在待檢驗之作文能力指標中加入了造句測驗。造句為作文之先備技巧之一，雖然此種技巧僅限於句子組成的層次，複雜程度不若文章寫作高，但造句和文章寫作之表現皆與寫作三元素（認知、語言學、形式）有密切關係。研究者假設作文與造句測驗的結果皆能夠反映學生真實的作文能力，但兩者適用的年級或能力水準可能不同。對四年級的學生而言，造句可能較文章寫作測驗適合作為作文能力的指標。先前的研究以年級間和高、低作文能力學生間的比較來發展建構效度資料，本研究除持續年級間的比較外，又加入了學習障礙與一般能力學生的比較以建立此種效度資料。透過先前與本研究對於數種作文能力指標效度的探討，結果作為發展形成性作文能力評量工具與寫作學習障礙篩選工具的依據。

二、待答問題

根據前面之陳述，研究者提出下列之研究問題：

（一）各種自編文章寫作作文能力測驗之「總

字數」與效標之間有沒有顯著關係？

(二) 各種自編文章寫作作文能力測驗之「相異字數」與效標之間有沒有顯著相關？

(三) 各種自編文章寫作作文能力測驗之「總字數」與「相異字數」之效標關聯效度有沒有顯著差異？

(四) 任兩種自編文章寫作作文能力測驗之效標關聯效度間有沒有顯著差異？

(五) 各種造句測驗分數與效標之間有沒有顯著關係？

(六) 各種造句測驗分數之效標關聯效度有沒有顯著差異？

(七) 學生在文章寫作作文測驗上的表現是否因年級的不同而有差別？

(八) 學生在造句測驗上的表現是否因年級的不同而有差別？

(九) 學習障礙與一般能力學生在各種文章寫作作文能力測驗的表現上有沒有明顯不同？

(十) 學習障礙與一般能力學生在各種造句測驗的分數表現上有沒有明顯的不同？

三、名詞詮釋

(一) 記敘文

記敘文是指根據事情發展的時間順序來組織情節，並有明確的故事發展的一種文章體裁，它不但在日常生活中廣泛應用，也是國內外作文研究最常使用的文體之一（施錚懿，民86；Newcomer & Barenbaum, 1991）。本研究使用「故事開頭」和「圖片」（包括單幅和四格連環）兩種刺激方式來誘發記敘文的寫作品。

(二) 說明文

說明文是指客觀的呈現或分析事理的一種文章體裁，具有說服和傳遞知識的功用，與記敘文同為國內外作文研究最常使用的文體（施錚懿，民86；Newcomer, et.al., 1991）。本研究以「比較性」和「問題解決性」兩種題目類型為刺激來引發出說明文的寫作品。

(三) 總字數

計算文章中所寫出的字之總數，包括用注音符號所替代的字，標點符號則不計算在內。此種分數為作文流暢性指標之一。

(四) 相異字數

將文章中重覆出現的字刪掉，計算剩下的字數，所得結果即為相異字數。相同的字不論出現多少遍，皆以一來計數。此種分數為測量作文中字彙變化的指標之一。

(五) 非常用字數

句子寫作中出現不在「中文五百個常用字」上之字的總數。所採用之字表源自林寶貴、黃瑞珍、彭千紅（民86）的附錄中（pp120-123）。這些研究者所使用之字表由國立編譯館（民58）和柯華威、吳敏而等（民79）之研究統整而來。非常用字如果出現一次以上，僅以一次紀錄。此種分數為作文中字彙變化性的指標之一。

研究方法

本研究的主旨在於發展適用於寫作學習障礙學生篩選以及形成性評鑑之作文能力測量工具，以效度的建立與比較為核心。測量採用實作的方式，換言之，從學生直接寫作的表現來評定其作文能力。經過檢驗後所推薦的測驗皆須符合以下之條件：一、分數之穩定性高；二、與其它可接受的作文和國語文能力效標有中度或中度以上的相關；三、能區別不同作文能力的學生。研究方法將依研究設計、樣本選取、研究工具、實施步驟及資料處理等部份陳述於下。

一、研究設計

本研究中有待比較者，在文章寫作方面涵蓋故事開頭、單圖、四格連環圖片之記敘文與主題句、故事開頭之說明文，共計五種作文能力測驗。這些文章寫作作文測驗共同的特色是採取直接、客觀、量化的評量，且皆以總字數

相異字數兩項計分結果作為各項比較的依變項。在造句測驗方面，皆以三種評分方式的分數（計算所寫出意義正確的句數、成語數與非常用字數）作為各項比較之依變項，除意義正確的句數牽涉部分之主觀判斷外，都採取直接、客觀、量化之評量。研究的樣本為中部地區小學四、五、六年級學生。發展的重點在於建立文章寫作作文能力測驗及造句測驗之效標關聯效度和建構效度資料。相關的設計分別說明如下。

(一) 效標關聯效度的建立

效標為學期末之作文和國語科學期成績。在五種文章寫作作文能力測驗方面，除了分別建立各種作文能力測驗所得分數與上述效標之間的相關係數外，不同作文能力測驗的效度也兩兩比較以檢驗是否有差異存在。以造句測驗來預測作文能力之效標關聯效度方面，除建立三種評分方式之分數與效標的相關係數外，不同造句測驗分數之效度也雙雙進行比較，以瞭解其間是否有差異存在。

(二) 建構效度的呈現

目的在檢驗各種作文和造句測驗分數能否鑑別不同年級或寫作能力的學生。首先，比較不同年級學生在五種文章寫作作文能力測驗和造句測驗之表現。其次，分年級來比較學習障礙與一般能力學生在五種文章寫作作文能力測驗和造句測驗上的表現。

二、樣本選取

本研究包括預試和正式施測兩個階段的樣本。預試樣本為臺中市省三國小和北屯國小高年級普通班學生，研究者請班級導師推薦高、低作文能力學生各5人參加預試的活動，扣除流失與資料不全者，共計有158人，抽樣和測試的工作已於八十六年完成。正式施測的樣本包括兩個部分：一是中部地區國小四、五、六年級普通班學生，研究者先自彰化縣、臺中縣、臺中市、南投縣各選取一所學校，依序為南郭、

塗城、大鵬與敦和國小，接著從前述合作學校的四到六年級各隨機抽取一個班級進行正式的施測，樣本資料整理於表一，總人數為524人。四到六年級受試樣本接受研究者自編之文章寫作作文能力測驗時的平均年齡分別是10.16、11.17和12.11歲。另一為學習障礙樣本，這一部份之樣本蒐集由本研究所屬之國科會整合計畫「國小輕度障礙學生作文錯誤類型分析及有效學習策略之研究」之總主持人來主導，過程如下：首先請彰化縣民生、臺中縣大雅、臺中市忠孝、南投縣南投國小全體四、五、六年級級任老師推薦該班國語科表現落在全班最後三名者；接著集中被推薦的學生，實施「羅桑二氏非語文智力測驗」（黃國彥、鍾思嘉、傅粹馨，民66）與「國民小學國語文成就測驗」（周台傑，民81）；每一位學生之智力測驗分數皆代入八十六年總計畫根據中彰投四縣市936位四到六年級普通班學生常模所發展之迴歸公式中。四至六年級的迴歸公式依序為 $Y=.34X+19.55$ ， $Y=.47X+16.85$ ， $Y=.31X+17.55$ ，皆以智力為預測變項、成就為依變項（周台傑，民86）。受試學生中，國語文成就測驗分數的實際表現落在預測值1.5個測量標準誤以下者，即被納入本研究之學習障礙樣本中。為了對照學習障礙學生之表現，研究者敦請每一位學障樣本就讀班級之導師推薦同班級中另一名與樣本同性別、作文能力在班級平均表現以上的學生，作為一般能力學生樣本。由於學障與一般能力學生除班級和性別外未進行其他條件之配對，故兩者在依變項上的表現僅以獨立樣本t考驗來進行組別間的比較。學習障礙與一般能力學生樣本資料摘要整理於表二。學習障礙學生接受測驗之年齡依四、五、六年級順序分別為10.06、11.17、12.28歲；一般能力學生則為10.21、11.13、12.15歲。在每一年級中，學障與一般學生之年齡差異皆未達顯著水準。

表一 普通班樣本的地區、年級、性別和人數資料

縣市	四年級			五年級			六年級			總數
	男	女	小計	男	女	小計	男	女	小計	
彰化縣	26	23	49	22	24	46	23	25	48	144
臺中縣	23	25	48	21	21	42	20	22	42	131
臺中市	26	23	49	18	22	40	22	19	41	132
南投縣	19	23	42	21	17	38	18	21	39	117
總數	94	94	188	82	84	166	83	87	170	524

表二 學習障礙學生的地區、年級、性別和人數資料

縣市	四年級			五年級			六年級			總數
	男	女	小計	男	女	小計	男	女	小計	
彰化縣	4	1	5	4	1	5	17	4	21	31
臺中縣	3	2	5	4	1	5	7	1	8	18
臺中市	3	1	4	4	3	7	12	1	13	24
南投縣	2	1	3	4	2	6	6	1	7	16
總數	12	5	17	16	7	23	42	7	49	89

三、研究工具

本研究所使用的工具包含作為效標的評量資料以及研究者自編之文章寫作與造句測驗。效標包括國語科和作文學期成績，原始資料取自八十七學年度上學期末學生成績單上國語科及作文分數。研究者檢視每一個班級國語和作文分數的分佈狀況，發現老師間給分範圍以及平均數差異頗大，為使班級間的分數得以直接比較，研究者將每一位受試學生的國語及作文成績先轉換成其就讀班級內的 T 分數，再以此分數作為效標。自編測驗包括文章寫作和造句兩部分，皆採直接、客觀、量化的評量方式，將分別說明於下。

(一) 文章寫作作文能力測驗

本研究所發展之作文測驗涵蓋兩種文體，分別是記敘文和說明文寫作，記敘文之下包括三篇文章，另兩篇則屬於說明文寫作。作文題目類型採用文獻上最常用在記敘文和說明文的

方式：記敘文方面包括故事開頭、單幅圖片以及連環圖片；說明文則是問題解決性與比較性兩種性質的主題句。每一種題目類型之下有數個備選試題，經過31位三年以上教學經驗的國小教師及兩位大學語文科系教師審核，決定預試的題目。選題的主要標準為：題目的明確性、具有可發揮的空間、符合一般小學生的生活經驗、和能區別不同作文能力的學生。經大部分老師審核通過的題目，接著在臺中市省三和北屯兩所國小高年級的普通班進行預試，藉此修正試題與說明不夠明確的地方，調整施測時間，並以最能刺激流暢的寫作表現和區別高低作文能力學生的試題作為正式施測的題目。經過預試後，以班為單位團體實施五種文章寫作作文能力測驗，實施的方式如下：在說明指導語後，每篇文章先提供三分鐘的時間構思，接著給予十分鐘的時間寫作，以相同的方式完成五篇文章。為平衡文體和題目類型呈現順序

所造成的影響，試卷上不同類型文章出現的次序是隨機排列的。包括寫作指導語的說明以及試卷的收發時間在內，完成五種文章寫作作文能力測驗約需一小時二十分鐘的時間。計分方式選用代表寫作流暢性的「總字數」以及標示字彙變化性的「相異字數」。普通班學生樣本於八十七學年度上學期結束前一個多月開始實施作文測驗，寒假前完成施測的工作。學習障礙與一般能力學生的部分，於八十七學年度下學期開學後實施，一個月內完成施測的工作。

(二) 造句測驗

造句題目中的詞、成語和句型皆選自國立編譯館出版之國小四、五、六年級國語課本。原包括35個題目，經過與文章寫作測驗類似的專家審核程序後，保留了20題。專家包含27位三年以上教學經驗的國小教師及兩位大學語文系教師。審核過的題目再經過預試的淘汰後，選擇了14個正式的題目。正式的造句測驗中，有6題為雙字詞、5題為四字成語、3題為牽涉到關係連接詞使用之句型（因為……所以……、不是……就是……、雖然……但是……）。測量的時間為十五分鐘。造句測驗的評分方式有三種：意義正確的句數、成語數、非常用字數。「意義正確的句數」計算受試者適切的將詞彙或句型擴展成完整句子的題數，重點在意義的合理性，文法、標點、寫字的正確性等因素非主要的考量，此種評分方式乃參考 Hammill 和 Larsen 在 Test of Written Test-3 (1996) 中字彙分測驗 (Vocabulary) 的計分法。「成語數」為計算在所有的題目中，受試者一共使用多少個成語或俗語，題目中所給的成語不計算在內。「非常用字數」的算法為對照林寶貴等人 (民86) 所匯整之中文500個常用字，然後計算在十四題當中用了多少不在此常用字表中的字，重複出現的字只算一次。本測驗與作文測驗同時實施，兩種類型的測驗間給予學生十分鐘的休息時間。

五種文章寫作作文能力測驗和造句測驗皆由彰化師大特殊教育系碩士級研究生及大學部高年級學生來執行，國語科和作文學期成績則是由合作學校教務處傳真給研究者，再轉換成班級內的 T 分數。

四、實施步驟

本研究的施行步驟含括試題編製、預試、正式施測、資料分析等部分綜合說明如下。研究者從小學課本及坊間兒童文學及作文相關書籍中找出有關於說明文和故事開頭的作文題目十三則，圖片題目則是參考 Engelmann 和 Silbert (1983) 以及坊間漫畫圖書，選出十一幅單幅圖片以及七幅四格連環圖片。發展出來的試題先請彰化師大特教系三年級一個班的學生進行初步篩選的工作，然後由31位三年以上教學經驗的國小教師及兩位大學語文系教師審核這些題目，每種作文能力測驗方式各選出兩個題目。造句測驗的題目選取自國立編譯館之國小四、五、六年級國語課本，參酌專家們的意見後，初步選定了二十題。

藉由前述程序發展出來的試題接著在臺中市省三和北屯國小進行預試，參與的16個班級由導師推薦高、低作文能力學生各5名，以兩種能力學生在各題目上的整體表現以及相對差異的大小和方向來修正指導語、調整施測時間、選擇正式施測的題目和評分的方式。預試時文章寫作表現以四種客觀、量化的評分方式來計算，分別是「總字數」、「成熟字數」、「相異字數」和「校正後相異字比例」，其中「總字數」是流暢性的指標，後面三項標示字彙的表現。由於「成熟字數」、「相異字數」和「校正的種類 - 符號比率」三者對於高、低作文能力學生的區別效果很接近，故正式施測時，只以其中最容易計算的「相異字數」作為代表字彙變化性的計分方式。造句測驗則從原有之題目中選擇鑑別度較佳的十四個題目。八十七學年度上學期開學後開始進行正式的施

測，取樣的地區包括彰化縣、臺中縣市、和南投縣。文章寫作和造句測驗皆建立有評分者間一致性信度，評分者為三位彰化師大特教系高年級學生，經過兩小時的訓練和討論後，從正式樣本之測驗卷中抽出三十份作文和造句測驗，結果如下：文章寫作測驗中「總字數」和「相異字數」的一致性係數分別為1.00與.99，造句測驗之「意義正確的句數」、「成語數」和「非常用字數」的係數為.87、.96與.99。信度分析的結果顯示作文和造句測驗的結果相當穩定。

測驗與相關資料經過整理、編碼、校對後透過 Excel 7.0 軟體登錄在電腦上，統計分析的工作則是藉由 SPSS 7.0 英文視窗版套裝軟體來進行，大學四年級學生助理處理資料登錄的工作，統計分析與報告撰寫乃由研究者自行完成。

五、資料分析

本研究採用的統計方法除了以平均數和標準差為主的描述統計外，尚包括以下幾種推論統計的類型：

(一) 檢驗任何兩元素間的連結關係：由皮爾遜積差相關 (Pearson product-moment correlations) 的統計方式來考驗。

(二) 探討任何兩相關係數之間是否有差異存在：以非獨立樣本相關係數差異性考驗 (Testing the difference between two nonindependent rs) 的方法來驗證 (Howell, 1987)。

(三) 比較不同年級學生之表現：各以單因子變異數分析 (one-way ANOVA) 來進行。

(四) 比較學習障礙與一般能力學生在文章寫作和造句測驗表現上的差異性：以獨立樣本 t 考驗來進行。

以上的統計考驗都以 $p < .05$ 為達到顯著性的標準。

研究結果

本研究從測驗發展的技術層面切入，主要目標在於建立自編之文章寫作作文能力測驗與造句測驗的效標關連效度，以及判斷測量之結果是否能夠鑑別不同年級和作文能力的學生。本章將從以下五個部分來呈現客觀數據分析的結果：首先呈現的是五種文章寫作作文能力測驗得分與效標間相關性考驗，以及效度係數間差異比較的結果；其次檢驗造句測驗不同評分方式的結果與效標間的關係，並比較效度係數間的差異；接下來呈現不同年級學生在文章寫作和造句測驗表現上比較的結果；然後是學習障礙和一般能力學生在文章寫作和造句測驗上的比較結果；最後為研究結果摘要。

一、五種文章寫作作文能力測驗的效標關聯效度比較

研究者自編之文章寫作作文能力測驗包括故事開頭、單圖、四格連環圖片三種記敘文以及比較性和問題解決兩種說明文。表三、四分別陳列有各種測驗之「總字數」和「相異字數」的平均數和標準差。這兩張表上較顯著的趨勢為：學生在故事開頭之記敘文與單圖記敘文所寫之總字數與相異字數，皆隨年級增高而增加。

在推論統計方面，首先進行的是五種文章寫作作文能力測驗上的表現與效標之間的相關性考驗，本研究所採用的效標包括八十七學年度上學期作文和國語科學期成績，兩種學期成績皆先轉換成班及內的 T 分數，再進行相關性考驗以及其他統計分析的工作。文章寫作作文能力測驗與效標相關性考驗的結果陳列於表五到六。由這些表中可看出五種文章寫作作文能力測驗或其加總後的分數與效標間的連結關係全都具有統計上的顯著性 (除一項係數的顯著水準達 .05 外，其他所有的相關係數皆達到 .01 的顯著水準)，而且除了四年級以及其他少數

表三 四到六年級學生在五種文章寫作作文能力測驗上「總字數」的平均數與標準差

文章寫作作文能力測驗	四年級		五年級		六年級	
	平均數	標準差	平均數	標準差	平均數	標準差
比較性說明文	104.77	41.13	121.37	41.26	117.25	48.38
故事開頭之記敘文	105.93	45.14	133.86	47.33	137.85	55.20
單圖之記敘文	89.98	40.55	102.12	44.78	119.56	54.23
四格連環圖片記敘文	92.96	31.83	84.06	39.01	91.12	46.30
問題解決性說明文	147.71	53.24	138.12	50.33	165.27	55.17
五種測驗的總和	541.34	170.46	579.53	173.57	631.05	198.07

表四 四到六年級學生在五種文章寫作作文能力測驗上「相異字數」的平均數與標準差

文章寫作作文能力測驗	四年級		五年級		六年級	
	平均數	標準差	平均數	標準差	平均數	標準差
比較性說明文	63.03	23.38	74.13	25.40	75.99	30.54
故事開頭之記敘文	66.60	26.96	84.73	30.37	93.65	32.98
單圖之記敘文	60.09	24.21	70.73	29.37	82.26	34.87
四格連環圖片記敘文	57.20	19.87	54.61	25.43	62.71	28.86
問題解決性說明文	57.81	21.04	54.83	25.25	67.79	30.17
五種測驗的總和	304.72	97.78	339.05	108.60	382.39	140.57

表五 五種文章寫作作文能力測驗之總字數與效標間的相關係數摘要表

文章寫作作文 能力測驗	效 標					
	四年級		五年級		六年級	
	作文 學期成績	國語文 學期成績	作文 學期成績	國語文 學期成績	作文 學期成績	國語文 學期成績
比較性說明文	<u>.21**</u>	<u>.17*</u>	.40**	.41**	.54**	.55**
故事開頭之記敘文	.37**	<u>.32**</u>	<u>.38**</u>	<u>.35**</u>	.49**	.52**
單圖之記敘文	<u>.30**</u>	<u>.22**</u>	<u>.41**</u>	<u>.37**</u>	.56**	.57**
四格連環圖記敘文	<u>.22**</u>	<u>.21**</u>	<u>.38**</u>	<u>.33**</u>	<u>.52**</u>	<u>.50**</u>
問題解決性說明文	.25**	.31**	.46**	.43**	<u>.62**</u>	.56**
五種測驗的總和	<u>.32**</u>	<u>.30**</u>	<u>.50**</u>	<u>.47**</u>	.65**	.63**

*p<.05 **p<.01；四、五、六年級人數分別為188、166、170人

表六 五種文章寫作作文能力測驗相異字數與效標間的相關係數摘要表

文章寫作作文 能力測驗	效 標					
	四年級		五年級		六年級	
	作文 學期成績	國語文 學期成績	作文 學期成績	國語文 學期成績	作文 學期成績	國語文 學期成績
比較性說明文	.27**	.30*	.43**	.42**	.55**	.57**
故事開頭之記敘文	.40**	.39**	.47**	.46**	.53**	.55**
單圖之記敘文	.42**	.39**	.53**	.47**	.57**	.56**
四格連環圖記敘文	.29**	.30**	.50**	.44**	.58**	.54**
問題解決性說明文	.21**	.29**	.49**	.43**	.64**	.57**
五種測驗的總和	.38**	.40**	.61**	.55**	.66**	.64**

* $p < .05$ ** $p < .01$ ；四、五、六年級人數分別為188、166、170人

例外，係數多在中度相關的範圍內。各種文章寫作測驗總字數和相異字數的效標關聯效度有隨年級的提升而增高的趨勢。

爲了瞭解總字數和相異字數兩種分數類型的效度有沒有差別，接著進行兩非獨立樣本相關係數差異性考驗的程序。由表五和六的格子數可算出每一個年級需進行12組比較（兩項效標乘以六項文章寫作測驗分數，格子中相關係數下畫橫線者表示在此測驗項目上，總字數與相異字數之效標關聯效度間有顯著差異。四、五、六年級各12組比較中，兩種分數之效度係數間存有顯著差異（至少達到.05之水準）的組數依序爲9、8、3組，且這些差異的方向皆爲

相異字數的效度優於總字數。整體來說，在四、五年級階段，相異字數預測作文能力的效果要比總字數來的好，這種現象在六年級較不明顯。

爲瞭解五種文章寫作作文能力測驗與加總後的表現兩兩之間在效標關聯效度上是否有差異存在，研究者再次以兩非獨立相關係數差異性考驗之統計方法來作比較。由於之前以國語文和作文學期成績爲效標之統計分析結果差異不大，爲簡化比較的內容，此項分析僅以作文學期成績爲比較的標的，結果摘要於表七和八中。綜合這些資料可以歸納出一些趨勢：一、單一實施任何作文能力測驗之效度皆不及所有

表七 文章寫作作文能力測驗間的效標關聯效度係數差異性考驗結果摘要表--以總字數為依變項的比較¹

文章寫作作文能力測驗間的配對比較	效標：作文學期成績		
	四年級	五年級	六年級
比較性說明文 vs. 故事開頭記敘文	**		
比較性說明文 vs. 單圖記敘文	*		
比較性說明文 vs. 四格連環圖片記敘文			
比較性說明文 vs. 問題解決性說明文			*
故事開頭記敘文 vs. 單圖記敘文	*		

表七 文章寫作作文能力測驗間的效標關聯效度係數差異性考驗結果摘要表--以總字數為依變項的比較¹ (續)

文章寫作作文能力測驗間的配對比較	效標：作文學期成績		
	四年級	五年級	六年級
故事開頭記敘文 vs. 四格連環圖片記敘文	**		
故事開頭記敘文 vs. 問題解決性說明文	*		**
單圖記敘文 vs. 四格連環圖片記敘文			
單圖記敘文 vs. 問題解決性說明文			
四格連環圖記敘文 vs. 問題解決性說明文			**
比較性說明文 vs. 五種測驗的總和	**	**	**
故事開頭記敘文 vs. 五種測驗的總和	*	**	**
單圖記敘文 vs. 五種測驗的總和		**	**
四格連環圖記敘文 vs. 五種測驗的總和	**	**	**
問題解決性說明文 vs. 五種測驗的總和	*		

格內的符號標示「兩非獨立樣本相關係數差異性考驗」的結果，*表示t值的 $p < .05$ ，**表示t值的 $p < .01$ 。

表八 文章寫作作文能力測驗間的效標關聯效度係數差異性考驗結果摘要表--以相異字數為依變項的比較¹

文章寫作作文能力測驗間的配對比較	效標：作文學期成績		
	四年級	五年級	六年級
比較性說明文 vs. 故事開頭記敘文	**		
比較性說明文 vs. 單圖記敘文	**	*	
比較性說明文 vs. 四格連環圖片記敘文			
比較性說明文 vs. 問題解決性說明文			*
故事開頭記敘文 vs. 單圖記敘文			
故事開頭記敘文 vs. 四格連環圖記敘文	*		
故事開頭記敘文 vs. 問題解決性說明文	**		**
單圖記敘文 vs. 四格連環圖片記敘文	**		
單圖記敘文 vs. 問題解決性說明文	**		*
四格連環圖記敘文 vs. 問題解決性說明文			*
比較性說明文 vs. 五種測驗的總和	**	**	**
故事開頭記敘文 vs. 五種測驗的總和		**	**
單圖記敘文 vs. 五種測驗的總和		**	**
四格連環圖記敘文 vs. 五種測驗的總和	**	**	**
問題解決性說明文 vs. 五種測驗的總和	**	**	

格內的符號標示「兩非獨立樣本相關係數差異性考驗」的結果，*表示t值的 $p < .05$ ，**表示t值的 $p < .01$ 。

測驗同時實施後之效度；二、在四年級階段，故事開頭記和單圖記敘文之效度優於其他兩種說明文體以及四格連環圖片；三、在五年級階段，各種測驗的效度多無顯著差異；四、到了六年級階段，問題解決性說明文之效度顯著優於其他四種作文測驗的效度。

二、造句測驗不同評分方式的效標關聯效度比較

研究者以自編的造句測驗來預測學生作文能力的表現。普通班學生樣本所造出之意義正確的句子數、句子中所使用的成語數，以及寫出之非常用字的平均數與標準差陳列於表九，明顯的趨勢為各種分數之平均數都有隨年級升高而增加的狀況。從四到六年級成語數的使用次數來看，學生自發性的在句子中使用成語的狀況仍不普遍，即使到了六年級，在十四個句子中出現的成語數平均還不足0.3個。由於一般

人所說或所寫的句子中極少出現一個以上的成語，扣掉四題成語題，剩下的十個句子中成語出現的比例仍然很低，可能的原因成語的意涵較一般詞彙豐富，由於造句測驗將寫作的框架侷限在句子的層次，不利於誘發成語的表達。

三種造句分數與作文和國語文學期成績之相關性考驗整理陳列於表十。除了成語數之效度係數偏低外，意義正確之句數與非常用字數之效度係數皆相當顯著，令人意外的是比作文測驗之效標關聯效度還要高。接著進行各種分數類型的效度差異性比較，採用非獨立樣本相關係數差異性考驗的統計方法來進行，結果顯示成語數之效度係數皆不如其他兩種分數。依年級來看，意義正確的句子數與非常用字數適用的階段不盡相同，前者在四年級的效度明顯優於後者，非常用字的評分方法在六年級超越前者，兩者在五年級階段沒有差別。

表九 四到六年級學生在造句測驗三種分數上的平均數與標準差

造句測驗分數類型	四年級		五年級		六年級	
	平均數	標準差	平均數	標準差	平均數	標準差
意義正確的句子數	7.93	2.53	9.10	2.70	11.07	3.14
成語數	0.01	0.10	0.10	0.62	0.29	0.61
非常用字數	21.91	7.02	26.71	9.14	35.53	11.20

表十 造句測驗三種評分方式與效標間的相關係數摘要

造句測驗分數類型	效 標					
	四年級		五年級		六年級	
	作文 學期成績	國語 文學期成績	作文 學期成績	國語 文學期成績	作文 學期成績	國語 文學期成績
意義正確的句子數	.63**	.61*	.70**	.72**	.52**	.57**
成語數	.08	.01	.19	.13	.24**	.24**
非常用字數	.37**	.31**	.72**	.74**	.72**	.71**

* $p < .05$ ** $p < .01$ ；四、五、六年級人數分別為188、166、170人

三、不同年級學生在文章寫作作文能力測驗和造句測驗表現上的比較

為建立文章寫作和造句測驗之建構效度資料，以下將以年級為自變項，各種類型的測驗分數為依變項，進行單因子變異數分析，藉此來檢驗本研究發展的文章寫作作文能力測驗結果能否正確區別因年級所形成之作文能力差異群體。各種文章寫作與造句測驗分數之平均數與標準差已呈現於表三、四和九之中，表十一直接摘要單因子變異數分析的結果，如果變

異數分析之 F 值達到顯著水準，接著以 Scheffé 的事後考驗比較年級間的差異。

表十一的摘要顯示在本研究所發展之作文能力相關分數類型中，僅有計算總字數之四格連環圖片記敘文測驗未達到顯著水準。從事後分析的結果整體看來，以相異字數計分之文章寫作測驗與造句測驗鑑別各年級學生的能力較總字數計分之測驗來的好。造句測驗的各種分數多能明確的區分不同年級的學生。

表十一 年級與作文、造句測驗分數之變異數分析暨事後比較摘要表

分數類型	One-way ANOVA		事後比較摘要		
	F值	顯著性	四年級vs. 五年級	五年級vs. 六年級	四年級vs. 六年級
比較性說明文－總字數	7.04	.00	*		*
故事開頭記敘文－總字數	21.88	.00	*		*
單圖記敘文－總字數	17.45	.00		*	*
四格連環圖片記敘文－總字數	2.47	.09			
問題解決性說明文－總字數	6.03	.00		*	
五種作文測驗總和－總字數	7.89	.00			*
比較性說明文－相異字數	12.63	.00	*		*
故事開頭記敘文－相異字數	35.20	.00	*	*	*
單圖記敘文相異字數	23.14	.00	*	*	*
四格連環圖片記敘文－相異字數	4.27	.01		*	
問題解決性說明文－相異字數	9.39	.00		*	*
五種文章寫作測驗總和－相異字數	18.44	.00	*	*	*
意義正確的句數	56.48	.00	*	*	*
成語數	15.57	.00		*	*
非常用字數	98.36	.00	*	*	*

四、學習障礙與一般能力學生在文章寫作作文和能力測驗造句測驗表現上的比較

本單元與上一節的目的相似，皆為建立文章寫作和造句測驗之建構效度資料而進行相關的分析。所要鑑別者仍為不同能力學生，樣本包括學習障礙與一般作文能力學生。在文章寫

作測驗部分依變項為學生在各種測驗上所寫出之總字數和相異字數，造句測驗比較之標的則包括意義正確的句數、成語數與非常用字數。學障與一般學生在文章寫作作文能力測驗上的平均數和標準差依年級順序陳列於表十二、十三。綜合比較這兩張表和表三中的數字，可以觀察到以下現象：

表十二 四到六年級學習障礙與一般能力學生在「總字數」的平均數與標準差

文章寫作作文能力測驗	四年級		五年級		六年級	
	學障	一般	學障	一般	學障	一般
比較性說明文	64.12 (36.71)	100.94 (34.89)	70.91 (34.98)	163.78 (42.40)	91.04 (40.94)	164.24 (39.54)
故事開頭之記敘文	49.82 (18.92)	117.53 (36.30)	74.74 (45.32)	168.30 (43.61)	89.16 (42.21)	180.96 (39.61)
單圖之記敘文	35.18 (16.28)	101.71 (36.63)	62.61 (38.59)	142.91 (48.80)	66.94 (32.38)	154.67 (39.38)
四格連環圖記敘文	42.12 (11.96)	96.94 (32.95)	59.04 (28.73)	125.78 (41.48)	58.37 (23.78)	123.20 (36.97)
問題解決性說明文	42.88 (22.41)	86.41 (34.85)	52.61 (39.01)	135.39 (34.03)	46.12 (26.38)	139.33 (43.60)
	N=17	N=17	N=23	N=23	N=49	N=49

* 括弧中的數字為標準差

表十三 四到六年級學習障礙與一般能力學生在「相異字數」的平均數與標準差

文章寫作作文能力測驗	四年級		五年級		六年級	
	學障	一般	學障	一般	學障	一般
比較性說明文	31.24 (14.52)	61.53 (19.75)	44.96 (42.68)	85.70 (20.82)	46.80 (19.23)	94.04 (19.15)
故事開頭記敘文	29.53 (13.79)	75.06 (21.97)	40.22 (26.03)	93.83 (28.22)	52.08 (23.10)	112.84 (24.97)
單圖之記敘文	21.94 (11.01)	65.71 (19.16)	41.04 (21.76)	82.57 (29.21)	43.76 (16.21)	100.14 (27.29)
四格連環圖記敘文	22.88 (12.08)	57.47 (19.17)	36.30 (20.02)	71.87 (26.01)	36.00 (10.60)	78.71 (26.14)
問題解決說明文	21.65 (9.72)	58.94 (24.00)	30.83 (24.32)	78.35 (18.04)	28.22 (12.54)	85.22 (28.49)
	N=17	N=17	N=23	N=23	N=49	N=49

* 括弧中的數字為標準差

首先，學習障礙學生在各種文章寫作作文能力測驗上的分數皆遠不及一般能力學生或該年級的平均表現；其次，不同年級學習障礙學生在各種文章寫作作文能力測驗上的差距皆小於普

通學生年級間的差異，這項資料提供了間接的證據，顯示學習障礙學生在作文能力上的成長較一般學生緩慢；三、學習障礙學生作文能力發展落後的情形非常嚴重，即使六年級的學生

在各種文章寫作作文能力測驗所得分數遠遠不如四年級普通班學生的平均表現。另外值得一提的是學習障礙學生之老師所推薦的一般作文能力學生，其寫作表現與年級平均數相比，有偏高的現象，應稱之為高能力組較恰當。此外，表十二與十三資料顯示不論寫作者能力的高低，說明文和記敘文體中最能夠刺激寫作者

流暢反應的作文形式分別為比較性說明文和故事開頭記敘文。接著，以獨立樣本 t 考驗來比較學習障礙和一般能力學生在各種文章寫作作文測驗上的表現，表十四、十五呈現比較的結果，顯示五種文章寫作作文能力測驗之總字數和相異字數都能夠明確區辨學習障礙者與一般作文能力學生。

表十四 四到六年級學習障礙與一般能力學生在「總字數」t 考驗之摘要表

文章寫作作文能力測驗	四年級			五年級			六年級		
	t	自由度	p	t	自由度	p	t	自由度	p
比較性說明文	-3.00	32	.00	-8.10	44	.00	- 8.99	96	.00
故事開頭之記敘文	-6.82	32	.00	-7.14	44	.00	-11.10	96	.00
單圖之記敘文	-6.84	32	.00	-6.19	44	.00	-12.05	96	.00
四格連環圖記敘文	-6.45	32	.00	-6.34	44	.00	-10.33	96	.00
問題解決性說明文	-4.33	32	.00	-7.67	44	.00	-11.98	96	.00

表十五 四到六年級學習障礙與一般能力學生在「相異字數」t 考驗之摘要表

文章寫作作文能力測驗	四年級			五年級			六年級		
	t	自由度	p	T	自由度	p	t	自由度	p
比較性說明文	-5.10	32	.00	-4.12	44	.00	-12.19	96	.00
故事開頭之記敘文	-7.24	32	.00	-6.70	44	.00	-12.50	96	.00
單圖之記敘文	-8.16	32	.00	-5.47	44	.00	-12.44	96	.00
四格連環圖記敘文	-6.29	32	.00	-5.20	44	.00	-10.60	96	.00
問題解決性說明文	-5.94	32	.00	-7.53	44	.00	-12.82	96	.00

兩類學生在造句測驗不同評分方式上的平均數與標準差陳列於表十六，綜合表十六與先前的表九，可以觀察以下的趨勢（由於造句測驗中成語使用的頻率很低，此部份不納入以下的討論）：首先，各類學生在兩種造句測驗分數上的表現皆隨年級的增高而增加；其次，相鄰年級在測驗分數上的比較結果顯示非常用字數的計分方法較能靈敏反映出學生造句能力的成長，因為其年級與年級間的差異分數較意義正確的句數來得明顯；再者，學習障礙學生在兩種造句測驗上的分數皆明顯低於一般能力學生或該年級的平均表現；最後，檢驗六年級學

障學生表現的結果顯示，此類學生在造句能力上落後的情形不若文章寫作能力明顯，就讀年級與實際能力的差距約在一到二個年級數間。一般能力學生之表現與年級平均數比較的結果顯示“高能力”較“一般能力”更能代表這些學生的表現。學習障礙與一般能力學生在兩種造句測驗分數上的差異亦以t考驗來進行分析，結果呈現於表十七。分析的結果顯示除了四、五年級學生在成語數的使用上無差異外，其他之比較全部達到 .01 以上的顯著水準，這個現象代表意義正確的句數與非常用字數的造句測驗分數都能夠鑑別學習障礙與一般能力學生。

表十六 四到六年級學習障礙與一般能力學生在各種造句測驗評分方法上的平均數與標準差

造句測驗之評分方式	四年級		五年級		六年級	
	學障	一般	學障	一般	學障	一般
意義正確的句數	2.71 (1.90)	11.59 (1.50)	7.48 (6.30)	12.30 (1.29)	7.80 (3.22)	12.90 (.80)
成語數	0.00 (.00)	0.00 (.00)	0.17 (.49)	0.52 (.99)	0.16 (.43)	0.65 (1.10)
非常用字數	12.88 (8.64)	30.35 (3.86)	17.91 (11.69)	38.09 (6.20)	23.53 (9.56)	42.51 (7.49)
	N=17	N=17	N=23	N=23	N=49	N=49

* 括弧中的數字為標準差

表十七 四到六年級學習障礙與一般能力學生在各種造句測驗評分方法之 t 考驗之摘要表

造句測驗之評分方式	四年級			五年級			六年級		
	t	自由度	p	T	自由度	p	t	自由度	p
意義正確的句數	-15.14	32	.00	-3.60	44	.00	-10.76	96	.00
成語數*				-1.51	44	.14	-2.86	96	.01
非常用字數	-7.62	32	.00	-7.31	44	.00	-10.93	96	.00

* 因兩個組別的標準差皆為 0，無法計算 t

五、研究結果摘要

統計分析的結果簡要摘述於下：

(一) 在五種文章寫作作文能力測驗中，故事開頭和單圖記敘文之總字數和相異字數兩種分數皆出現隨年級的增高而穩定、持續成長的趨勢。

(二) 不論在那一個年級內，五種文章寫作作文能力測驗之總字數與相異字數全都與效標評量有顯著的連結關係；除四年級部分和少數例外，效度係數多在中度相關的範圍內。

(三) 各種文章寫作作文測驗總字數和相異字數的效標關聯效度有隨年級提升而增高的趨勢，皆以六年級的效度最高。

(四) 在四、五年級階段，文章寫作作文測驗相異字數之效標關聯效度整體而言優於總字數。

(五) 在每一個年級階段中，單一實施任何文章寫作作文能力測驗之效標關聯效度皆不及所有測驗總分之效度，例外僅出現在四年級的單圖和故事開頭記敘文，此兩種測驗分數與總分沒有顯著差異。

(六) 在五種文章寫作作文能力測驗中，四年級以故事開頭和單圖記敘文的效標關聯效度較佳，問題解決性說明文則是六年級效度最好的測驗方式。

(七) 造句測驗中的各種分數皆出現隨年級升高而增加的穩定、持續成長趨勢。

(八) 各年級學生在造句測驗中，使用成語的頻率非常低。

(九) 造句測驗中，意義正確的句數與非常用字數兩種分數與作文能力效標之間有中度到高度的相關性。

(十) 依年級來比較造句測驗之意義正確句數與非常用字數的效標關聯效度：四年級意義正確句數的效度明顯優於非常用字數；兩種分數在五年級階段的旗鼓相當，都與效標分數間有顯著的連結關係；六年級則以非常用字數的效度最好，不過兩種分數在效度上的差異不及四年級部分大。

(十一) 說明文和記敘文體中最能夠刺激寫作者流暢反應的作文形式分別為比較性說明文和故事開頭記敘文。

(十二) 各種文章寫作作文測驗的計分方式大體而言能夠區分不同年級學生，相異字數的鑑別能力優於總字數。

(十三) 造句測驗的各種分數多能明確的區分不同年級的學生。

(十四) 學習障礙學生作文和造句能力之發展皆較一般學生明顯低落，尤以文章寫作作文測驗各種分數的落後情形最為嚴重。

結論與討論

本研究之主旨在於有效之形成性作文能力測量工具的發展，藉此老師得以監控作文障礙學生進步狀況，並以監測的結果作為修正教學的依據。所檢驗的文章寫作作文能力測量方式皆以總字數和相異字數來計分，包括故事開頭、單圖和四格連環圖片的記敘文以及比較性、問題解決性說明文，另外也嘗試以造句測驗來評估學生之作文能力。透過不同年級和能力組別學生在測驗表現上的比較，具有良好鑑別能力之測驗亦可作為寫作學習障礙學生之篩選工具。

一、結論

以下將先針對資料分析的結果來回答本研究的問題。針對問題一和問題二，研究結果顯示：五種文章寫作作文能力測驗之總字數和相異字數，與代表作文和國語文成就的效標之

間，有顯著的連結關係，相關性的強度大致在中度的範圍內。問題三中，各種文章寫作作文能力測驗的兩種計分方式--總字數與相異字數--之效標關聯效度間有顯著差距存在。大體而言，相異字數的效標關聯效度優於總字數。在問題四裡文章寫作作文能力測驗間之效標關聯效度差異性考驗結果顯示：各年級皆有部分的文章寫作作文能力測驗之效標關聯效度與其他測驗不同，狀況又因年級而有異。在四年級，記敘文體測驗中故事開頭或單圖記敘文之效度較說明文為佳；到了六年級，以問題解決性說明文的效度最優。

針對問題五，除成語數外，造句測驗中意義正確的造句數、非常用字數與作文能力的關係不但達到統計上的顯著水準，效度係數多在中度到高度相關的範圍內。在四年級階段，意義正確的句數其效度優於非常用字數，到了六年級情況則剛好相反，此發現也提供了問題六的答案，造句測驗評分方式之效度間還是存在有顯著的差異。單因子變異數分析的結果乃用以回答問題七和問題八，簡言之，學生在文章寫作和造句測驗上的表現會因為年級的不同而有差異，年級越高表現越好。學習障礙與一般作文能力學生在文章寫作和造句測驗上的表現有顯著的差異，前者明顯落後，這個發現也回答了最後兩個研究問題。以下將總結本次研究的發現，並與相關文獻進行比較。

二、討論

本研究首先複製前一年的研究（葉靖雲，民88），謹慎檢驗各種文章寫作作文能力測驗分數與一般認可之作文能力效標間的關係。兩次研究取樣的地區皆為中部四縣市（臺中縣市、彰化、南投），取樣的學校雖不相同，但有考量相同縣市內所選的學校其地區特性盡量與前一年的學校近似。第一、二年的普通班樣本總數分別為829和524位四到六年級學生。兩次施測時所使用的文章寫作作文測驗以及施測

時間完全相同，計分方法也一樣。不同點在於第一年納入的作文能力效標比第二年多，除了學期成績外，還包括標準化的作文和國語文成就測驗（陳英豪、吳裕益、王萬清所編之「兒童寫作能力測驗」與周台傑之「國民小學國語文成就測驗」）。由於第一年的資料顯示每種文章寫作文測驗與各個效標間的關係相近，基於時間和經費上的考量，本年度的研究僅以作文和國語科學期成績為作文能力發展的效標。以下將說明今年和先前研究的資料中所呈現之顯著趨勢。

首先，本研究中各種客觀、量化文章寫作文能力測量方式和分數類型皆與作文能力的效標間有連結關係。然而統計上的顯著性不一定表示有實質上的顯著關係，因為部分係數還未達到中度相關的水準，此種現象以四年級最明顯。對此階段的學生而言，可能還有其他的測量方式更能反映其作文能力的發展狀況。到了五、六年級，文章寫作的效度係數有升高的趨勢，這種趨勢在本年度的研究中尤其明顯。雖然文獻中也有關於國小階段作文能力指標間在效度上的差異性比較（如 Deno, Mirkin & Marston, 1980; Marston, 1982，引自 Marston, 1989; Videen, Deno & Matson, 1982），但多以跨年級的方式進行資料的分析，看不出年級間的趨勢。國外這些跨年級且涵蓋一般與學障樣本的研究指出有效作文能力指標，除總字數外，常見者還有正確的字母順序（correct letter sequence）以及正確之字的順序（correct word sequence）。字母順序著眼在字的範圍裡，字的順序偏重在字與字的連接，這兩種指標的有效可能表示書寫語文之低階、技巧的發展，在國小時期，可以用來標示作文能力的發展概況。另從語言結構的角度來看，寫作能力的發展由字、詞，到詞的組合、造句、敘述性的短文，再進入不同文體文章的寫作（曹綺雯、周碧紅，民82），一般四年級學生可能處於句

子、短文到文章寫作的過渡期，此時以文章寫作的實作表現來評估學生的作文能力可能並不恰當。研究者推測，在國小中年級（甚至高年級）適用之作文能力指標可能隱身在字、詞、句的層次，這也是研究者在第二年施測時加入造句測驗的主要理由之一。

在這兩年的研究中，另一個非常明顯的趨勢是相異字數的效標關聯效度優於總字數。換言之，寫作者在文章中所用的字彙變化性愈大，其作文能力愈好；這個現象顯示文字表達的能力與字彙的豐富、變化性有緊密關係。字彙豐富的寫作者在書寫表達時較能夠以精確、合乎時宜的字眼來作陳述或推理；字彙的豐富與否又與常識的多寡有高度相關（Hammill & Bartel, 1995）。知識豐富之人其寫作上的優勢為有較多的素材供自己選擇、組織、表達。過去，相異字指標的研究多集中在大學階段，如 Gajar 等人（Gajar, 1989; Gajar & Harriman, 1987）的研究顯示文章中的相異字數是評估大學生整體表現最好的指標。目前的研究也顯示相異字數的效標關聯效度非常穩定的隨年級的升高而增強。

文體也是一個作文能力評量時的重要影響因素。去年和今年的資料皆顯示四年級學生在故事開頭和單圖記敘文上所寫出的總字數與相異字數具有最佳的效標關聯效度，到了六年級則以問題解決性說明文測驗上的分數最能有效反映學生的作文能力。在四年級，單獨實施故事開頭或單圖記敘文的寫作評量，其預測作文能力的效果與五種類型的文章寫作的總分差不多，可見得記敘文寫作在此階段作文能力發展上之顯著地位。到了六年級，一般學生記敘文寫作技巧逐漸熟稔，課程中內容科目的比重增加，需要以書寫來完成的作業比重也愈來愈高，此時說明文寫作的重要性逐漸展現。然而，作文能力評量僅考慮到文體的層次似嫌不足，因為本研究中屬於記敘文體的測驗有三個

(故事開頭、單圖、四格連環圖片)，其中四格連環圖片這種測驗類型的效度顯然不及其他兩種。在說明文方面也有類似的情形：兩年的研究中，比較性說明文的效度皆不及問題解決性說明文。所以評量作文能力時，除了文體的考量外，測驗類型也應在考慮之列。

五種跨文體和試題類型的文章寫作文測驗總和之效標關聯效度優於任何一種測驗單獨實施的效度，例外僅出現在前面提及之四年級單圖和故事開頭記敘文的部分，這是另外一個兩年研究中的共同趨勢。以單一文章寫作文測驗分數評估作文能力的效度最好也不過在接近高度相關的位置，完全依賴學生在此種測驗上的表現來作重要的教學決定的風險不小。面對這個問題，研究者去年所提出的建議是先形成性實施(如每週一次)一種相對而言效度最好的測驗，每隔一段固定的時間(如每月一次)施予次佳的測驗，第二個測驗實施的目的在驗證或輔助解釋第一個測驗所反映之作文能力變化情形。所謂最佳、次佳的建議是由迴歸分析的資料中，比較各個測驗分數對於效標分數變異量的解釋程度來決定的。最佳者乃解釋變異量最大的測驗，次佳則是超越最佳的測驗額外增加變異數解釋量最大者，最佳、次佳的兩組測驗依年級順序分別為故事開頭和單圖記敘文、單圖記敘文和比較性說明文、故事開頭記敘文和問題解決性說明文。然而，按照這種方法實施形成性評量只能讓老師知道學生能力變化的大概情形，根本的問題在於這些文章寫作作文測驗的效標關聯效度從實際應用的層次上來看還不夠理想。如果為了提高測驗資料的效度而每次都執行兩、三種測驗又違背了形成性評量工具“快又有效”的特性。解決的辦法之一是嘗試其他可能的有效測量方式。這是研究者在今年的試題中加入造句測驗的最主要理由。

造句是文章寫作的先備技巧之一，雖然它

與作文能力有關，但過去還未有研究者直接以造句測驗上的表現來反映學生作文能力的情況。造句測驗與作文能力效標間的效度資料著實令研究者驚訝，造句測驗中，意義正確的句數與非常用字數的效度係數，顯然較作文測驗為高。研究者已計畫在下一年度複製這部分的研究，以獲致更肯切的結論。單從今年的資料來看，造句測驗的選擇仍需要考量年齡與技能發展的因素：四年級階段，學生能夠將語詞或句型建構成意義完整的句子，是最好的作文能力評量方式；到了六年級，學生造句時，非常用字出現的數量為最好的作文能力評量方式。在信效度條件相似，測驗的實施也都符合經濟、便利的情況下，測驗的分數是否具有“靈敏性”成為形成性評量工具選擇的一個重要考量要素。根據 Deno (1985) 的解釋，愈靈敏的測驗愈能清楚的反映出學生在短期內能力變化的情形。測驗分數愈能儘快反映學生能力的微小變化，教師愈容易獲得教學有效性的回饋資料，以修訂出符合學生需要的教學內容。在本次研究裡，五年級階段造句測驗中，意義正確句數和非常用字數都是穩定、有效的作文能力指標，這時便可從靈敏性的角度來選擇其中之一為形成性評量工具。比較兩種分數跨越年級的平均數變化情形：在普通班樣本部分，意義正確句數每年成長幅度約為兩個單位，非常用字數約在五個單位以上；在學生能力真的有變化的前提下，非常用字數的作文能力監控資料較容易讓老師察覺到能力變化的情形。除非有其他特別的考量，否則在五年級階段，研究者建議採用的系統是非常用字數的監控資料。

在建構效度的比較方面，本研究的發現與先前回顧的文獻一致，顯示各種作文測驗大都能夠正確區分不同能力的寫作者，由於本次研究中直接比較總字數與相異字數在鑑別效果上的差異，因此研究者進一步建議以相異字數的計分為優先採用的方式。如果加入造句測驗方

面的資料，則意義正確句數或非常用字數計分的造句測驗，也是很好的鑑別工具。這一部份的結果顯示這些測驗可以作為篩選寫作學習障礙或低寫作能力學生的簡便、有利工具。

參考文獻

一、中文部份

- 向惠芳(民82)：寫出圖來—談看圖作文。國教月刊，39期，40-49頁。
- 杜淑貞(民75)：國小作文教學探究。臺北市：臺灣學生。
- 林寶貴、黃瑞珍、彭千紅(民86)：兒童書寫語言發展指標研究。國科會報告。
- 周台傑(民81)：國民小學國語文成就測驗。臺北：精華。
- 周台傑(民86)：國民小學閱讀障礙學生閱讀錯誤類型分析及有效學習策略之研究(I)。國科會報告。
- 柯華蕙、陳俊文(民81)：小學生說話與作文產品之比較。國立中正大學學報，3(1)，27-46頁。
- 施錚懿(民86)：國小六年級寫作學習障礙與普通學生在故事與說明文寫作成果之比較。國立彰化師範大學特殊教育研究所碩士論文。
- 曹綺雯、周碧紅(民82)：寫作基本法。臺北：書林。
- 黃國彥、鍾思嘉、傅粹馨(民66)：羅桑二氏非語文智力測驗指導手冊。臺北：正昇教育科學社。
- 葉靖雲(民87)：國小學習障礙學生作文錯誤類型分析及有效學習策略之研究(一)。國科會報告。
- 葉靖雲(民88)：五種作文能力測驗的效度研究。特殊教育學報，13期，331-366頁。

二、英文部份

- Barenbaum, E., Newcomer, P., & Nodine, B. (1987). Children's ability to write stories as a function of variation in task, age, and developmental level. *Learning Disability Quarterly*, 10, 175-188.
- Brand, A. G. (1991). *Construction tasks for direct writing assessment: A frontier revisited*. (ERIC Document Reproduction Service No. ED 340 037).
- Christenson, S. L., Thurlow, M. L., Ysseldyke, J. E., & McVicar, R. (1989). Written language instruction for students with mild handicaps: Is there enough quantity to ensure quality? *Learning Disability Quarterly*, 12, 219-229.
- Crowhurst, M. (1987). Cohension in argument and narration at three grade levels. *Research in the Teaching of English*, 21, 185-201.
- Deno, S., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children*, 48, 358-371.
- Deno, S. L., Mirkin, P. K., & Marston, D. (1980). *Relationships among simple measures of written expression and performance on standardized achievement tests* (Research Report No. 22). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.
- Espin, C. A., Scierka, B. J., Skare, S., & Halverson, N. (1999). *Reading and Writing Quarterly*, 15(1), 5-27.
- Finn, P. J. (1977). Computer-aided description of mature words choices in writing. In C. R. Cooper & L. Odell

- (Eds.), *Evaluating writing: Describing, measuring, judging* (pp69-90). Urbana, IL: National Council of Teachers of English.
- Gajar, A. H. (1989). A computer analysis of written language variables and a comparison of compositions written by university students with and without disabilities. *Journal of Learning Disabilities*, 22(2), 125-130.
- Gajar, A. H. & Harriman, N. (1987). Identifying data based procedures for written expression disabilities at the university level. *Education*, 107(3), 252-258.
- Goodman, L. (1985). The effective schools movement and special education. *Teaching Exceptional Children*, 17, 102-105.
- Graves, A., Semmel, M., & Gerber, M. (1994). The effects of story prompts on the narrative production of students with and without learning disabilities. *Learning Disability Quarterly*, 17, 154-164.
- Hammill, D. D. & Bartel, N. R. (1995). *Teaching students with learning and behavior problems: Managing mild to moderate difficulties in resource and inclusive settings* (6th ed.). Austin, TX: PRO-ED.
- Hammill, D. D. & Larsen, S. C. (1996). *Test of written language: examiner's manual*. Austin, TX: PRO-ED.
- Hoffman, J. V. & Rutherford, W. L. (1984). Effective reading programs: A critical review of outlier studies. *Reading Research Quarterly*, 20, 79-92.
- Isaacson, S. L. (1988). Assessing the written product: Qualitative and quantitative measures. *Exceptional Children*, 54, 528-534.
- Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp18-78). New York: Guilford.
- Mastropieri, M. A. & Scruggs, T. E. (1994). *Effective instruction for special education* (2nd ed.). Austin, TX: PRO-ED.
- Montague, M. & Graves, A. (1993). Improving students' story writing. *Teaching Exceptional Children*, 25(4), 36-37.
- Neilsen, L. & Piche, G. L. (1981). The influence of headed nominal complexity and lexical choice on teachers' evaluation of writing. *Research in the Teaching of English*, 15(1), 65-73.
- Newcomer, P. L. & Barenbaum, E. M. (1991). The written composing ability of children with learning disabilities: A review of the literature from 1980 to 1990. *Journal of Learning Disabilities*, 24(10), 578-593.
- Nodine, B. F., Barenbaum, E. M., & Newcomer, P. L. (1985). Story composition by learning disabled, reading disabled, and normal children. *Learning Disability Quarterly*, 8, 167-179.
- Parater, D. & Padia, W. (1983). Effects of modes of discourse on writing performance in grades four and six. *Research in the Teaching of English*, 17(2), 127-134.
- Parker, R. I., Tindal, G., & Hasbrouck, J. (1991). Progress monitoring with objective measures of writing performance

- for students with mild disabilities. *Exceptional Children*, *58*, 61-73.
- Poteet, J. A. (1980). Informal assessment of written expression. *Learning Disability Quarterly*, *3*(4), 88-98.
- Quellmalz, E. S., Capell, F. J., & Chou, C. P. (1982). Effects of discourse and response mode on the measurement of writing competence. *Journal of Educational Measurement*, *19*(4), 241-258.
- Rieth, H. J., Polsgrove, L., & Semmel, M. I. (1981). Instructional variables that make a difference: Attention to task and beyond. *Exceptional Education Quarterly*, *2*(3), 61-72.
- Shinn, M. & Marston, D. (1985). Differentiating mildly handicapped, low achieving, and regular education students: A curriculum based approach. *Remedial and Special Education*, *6*(2), 31-38.
- Thomas, D. & Donlan, D. (1980). *Correlations between holistic and quantitative methods of evaluating student writing, grades 4-12*. (ERIC Document Reproduction Service No. ED 221 976).
- Tindal, G. A. & Marston, D. B. (1990). *Classroom based assessment: Evaluating instructional outcomes*. New York: Merrill/Macmillan.
- Vacca, R. T. & Vacca, J. L. (1989). *Content area reading* (3rd ed.). Glenview, IL: Acott, Foresman and Company.
- Vallecorsa, A. & Garriss, C. (1990). Story composition skills of middle grade students with learning disabilities. *Exceptional Children*, *57*, 48-53.
- Videen, J., Deno, S. L., & Marston, D. (1982). *Correct word sequences: A valid indicator of proficiency in written expression* (Research Report No. 84). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, *70*, 703-713.
- Wiggins, G. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco: Jossey-Bass.
- Wong, B., Wong, R., & Blenkinsop, J. (1989). Cognitive and metacognitive aspects of learning disabled adolescents' composing problems. *Learning Disability Quarterly*, *12*, 300-322.

Bulletin of Special Education 1999, 18, 151 – 172

National Taiwan Normal University, Taipei, Taiwan, R.O.C.

THE VALIDITY OF INDICATORS OF WRITING COMPETENCE

Ching-Yun Yeh

National Changhwa University of Education]

The main purpose of the study was to establish the validity -- criterion-related as well as construct validity -- of five direct, objective and quantitative measures of written expression. The writing measures were composed of four kinds of test stimuli (title, story starter, single picture and four pictures) , two modes of writing (story and expository), totally five types of writing samples. Two types of scores--total number of words and the number of different words--were derived from these measures. Besides, validity of scores derived from sentence construction task were also examined. The criterion measures included class grades of written expression and language. The results showed that writing measures with the number of different words as dependent variable correlated moderately with the criterion measures. Moreover, number of different words derived from each measure could differentiate students from various grades and with high/low ability on written expression. Scores of sentence construction correlated strongly with the criterion measures and were able to discriminate students with different writing competence.

Key words: elementary, assessment, validity, mode of writing, writing stimulus, learning disabilities, sentence construction, validity, total number of words, number of different words, matures words