

作文評分標準研究

柯華蕓
中央大學

摘要

作文評分常因評分者的主觀因素造成爭議。本研究旨在整理敘述文和說明文的各類評分項目，期冀以最少項目但是具有信度與效度的評分指標供批改作文時使用。經蒐集各類評分指標，並採用大考中心公佈的作文為研究對象，經刪除重複性高、與其他項目相關低、與大考中心暨有分數不相關的項目，結果，說明文有四類九項，敘述文有四類六項評分標準。經試作，證明這些標準有其信度與效度。精簡過後的項目可使作文評分更客觀，學生也較清楚作文的要求。

關鍵詞：作文評分標準、說明文、敘述文

1. 前言

寫作是以文字提出意見、描述事實、寫出個人經驗、敘述內心情感和想像等等的語文表達方式之一。寫作能力可以說是語言能力的綜合表現，因此寫作評量是考察語文能力一個有效的方式。寫作能力評量的方式有許多，如命題作文、實驗報告、學期報告、做摘要等等。本文主要探討的是以寫作為試題時的評量方式，在此以作文的寫作為主。

寫作雖可以考察語文綜合能力，但是它的評分卻不簡單。評分者主觀因素在評分過程中所扮演的角色常有爭議（大考中心，民 85）。因此作文評分的具體化和標準化一直是關心語文能力評量研究者探討的問題。過去國內有研究者投注於作文評定量表的編製（例：陳英豪、簡楚瑛、王萬清，民 77），但是數量不多。倒是有不少論文以寫作過程和作文（寫作）教學為研究主題，其中都會提到評分標準（例如：王萬清，民 81；張新仁，民 81）。這些研究中投入時間最長的應屬由大學入學考試中心自民國 84 年以來關於語文表達能力測驗的研究（大學入學考試中心，民 86a, b）。大考中心除研發適合語文表達

能力的各種試題，也探討各種題型的評分指標，其中包括作文的評分指標。

2. 作文評分指標研究

2.1 說明文評定量表

陳英豪、簡楚瑛、王萬清（民 77）曾編製一份說明文評定量表。標準分為五大項目，其下有十三個子項目，以五等量表方式評分。項目與子項目如表一。

表一：說明文評定量表評分項目

項目	標點符號	字詞正確	思想內容	組織結構	修辭美化
子項目	使用正確	用字正確 用詞正確 句子通順	內容充實 取材切題 列舉實例 符合事實	條理清晰 段落分明 前後呼應	措辭精緻 說明生動

陳、簡、王的說明文評定量表有其測量上的信度和效度，在往後的研究中也曾被採用（如王萬清，民 81；鄭博真，民 85）。

2.2 大考中心語文表達能力研究

大考入學考試中心為開拓多元語文能力中的文字表達能力命題方式，曾進行多年的研究。研究目的在找出適當的評量可以評量學子是否具有「利用文字作摘述，或以適當的發展方法，寫出有意義、有細節的文章，來表達經驗、感受、想像或意見、意念、事實」的能力（大考中心，民 86a）。也就是我們通稱的作文能力。大考中心文字表達能力測驗目標如下（大考中心，民 86b）：

- (1) 構詞、造句是否合乎文法，使用標點是否恰當。
- (2) 是否以清晰的條理，組織意義完整、首尾連貫的段落或篇章。
- (3) 是否依照描寫、說明、議論、抒情等表達上的需求，控引內在思緒，斟酌寫作素材。
- (4) 是否運用修辭技巧，設計優美、生動的表意方式。
- (5) 是否掌握讀者的特性與需求。

依據上述測驗目標，中心定出一般評分標準和特殊評分標準。一般標準包括：

- (1) 組織清晰，段落分明(配合目標二)。
- (2) 文法正確，標點適當(配合目標一)。
- (3) 文筆順暢，修辭妥貼(配合目標四)。

配合目標三和五則是特殊標準，屬於內容部分的評分，依各試題的要求訂定（大考中心，民 86b）。

顯然的，在大考中心的研究中認為作文不論文體與內容，可以有一些共同的評分標準，稱之為一般標準。至於「內容」的評量標準，在大考中心的報告中未進一步說明。基於此，本研究接受一般評分標準，但是亦不排除因文體或內容可能造成評分上的不同。表二是研究者整理文獻上相關的作文評分標準，其中不分一般標準或是特殊標準。

至於文體的種類，在國民中小學九年一貫課程綱要中對於寫作的文體分為童詩、記敘文、說明文、議論文和應用文。在小學前三年以記敘文和說明文以及日常生活中應用文的寫作為主（寫作能力 1-3-4）。四年級以上則加入議論文以及學校活動中需要的應用文（寫作能力 2-3-4）。但是文獻中所討論的評分標準基本上以說明文和記敘文為主要對象。因此本研究針對這兩種文體的評分標準作分析。

基本上，文獻的評分標準可以分為語言（包括文字、修辭、文法），組織和內容三大類。在語言一項中的文字與文法似乎屬於不同的範疇，但卻都被列在一起（大考中心，民 86b; 陳英豪、簡楚瑛、王萬清民 77;張新仁，民 81）。或許如大考中心的說法：「構詞、造句是否合乎文法」。因此，研究者仍維持文獻上的分類。

表二：說明文與敘述文量標準整理

文字修辭		組織結構		內容	
說明文	敘述文	說明文	敘述文	說明文	敘述文
修辭	修辭	段落分明	段落分明	取材適切	取材適切
文法正確	文法正確	起始有力	起始有力	見解獨到	新穎獨到
詞句豐富	詞句豐富	承轉曲折	承轉曲折	旁徵博引	X
文筆生動	文筆生動	結論得體	結論得體	精簡扼要	X
X	字詞優美	首尾連貫	首尾連貫	X	觀事有據
--	--	闡述合理	闡述合理	X	精彩動人
--	--	設定讀者	設定讀者	X	文題相符
--	--	論點豐富	X	--	--

X 表個別文體不要求的標準。

由表二的整理看得出，說明文和敘述文在文字修辭文法和組織上的標準大同而小異。小異處如說明文在組織上需要「論點豐富」而敘述文在文辭上需要「字詞優美」，這是可以理解的要求。在內容上，兩種文體的要求就比較有差異，這也就是大考中心所謂的「特殊標準」。

2.3 評分方式

有了標準後怎麼評分？說明文評定量表以五等分方式，每一項分開評分（陳英豪、簡楚瑛、王萬清，民 77）。至於大考中心則研發五等十級和三等九級的評分方式。

根據大考中心的研究發現評分者給一篇文章的分數差距可以有相當的距離。中心曾請三位評分者就相同文章給分數，滿分是 50 分，根據中心的統計，評分差異在 7.5 分至 9.6 分之間。基於此，大考中心研發五等十級制。五等分別為優、良、中、差和劣，每一等再分兩級。每一等皆有文字敘述供評分者參考。例如在[修辭]上優等的描述是「語藏豐富」。良是「用字穩當，對於典故、成語引用算適切」。中是「語藏較貧乏，用詞淺白不夠精確，偶有語病」。差是「語藏貧乏，用字淺俗瑣碎，文句多重複有語病」。劣是「語藏貧乏，用詞生澀粗造，文句冗贅重覆」（大考中心，民 86，p.54）。

民國 88 年(p.30)，大考中心再度精簡評分，研發出三等九級的評分方式，針對語言、結構、內容三向度，每一向度分上、中、下三等，各等再分三級（表三）。

表三：大考中心作文三等九級評分標準

		語言	結構	內容
上	A+ A A-	文句通順流暢，用詞精確，能運用修辭技巧	理路清晰，首尾連貫	內容新穎，能恰當表現出所欲刻劃的主題
中	B+ B B-	文句大致通順，用詞尚稱貼切	理路大致清楚，段落大致分明	內容平實，能大致表現出所欲刻劃的主題
下	C+ C C-	文句不甚通順，語彙貧乏，語病較多	理路不清楚，缺乏連貫	內容較僵化，不表現出所欲刻劃的主題

由這些演變不難看出大考中心有簡化寫作評分的決心。但是在評分過程中，因中心面對全國考生，必須謹慎執行評分的動作。根據大考中心的報告（民 88），評分過程不止有說明，更有訓練。

評分前：

- (1) 部分評分委員舉行「標準訂定會議」。
 - a. 擬定一般標準；
 - b. 挑出標準卷；
 - c. 在標準卷上寫出給分的理由；
 - d. 挑出十份試卷準備開「試閱會議」。
- (2) 舉行評分訓練的「試閱會議」。
 - a. 對評閱委員解釋評分標準；
 - b. 委員評閱十份試卷；
 - c. 檢討。

評分時：

- (1) 每一份考卷經初閱和複閱。
- (2) 初閱和複閱差距超過規定，由召集人主閱。

這樣的評分過程不是一般寫作評分時需要有的步驟，但是一般老師仍希望有較客觀的標準來評分，也可以讓學生知道寫作的要求。本研究主要目的就是在探討可不可以為每一類找到幾個具體、具有分辨力並容易執行的評分指標作為文章批改的依據。研究者主要在探討三類評分中的項目，怎樣才會恰到好處又反應出各評分類別的精神。研究者以 300 字左右的文章為研究對

象，找已經評分過的文章，輔以研究者在表二整理出的評分標準重新給分，以刪除信度、效度較低的標準，找出新的評分標準。當有了較具體評分標準後，再經過試用以確定它們的可用與可靠程度。

3. 研究一：評分項目的訂定

由大考中心的研究歷程得知，即使寫作評分的類別確定，因者文章的要求或是內容不同，需要評分委員開會取得評分上的共識。換句話說，不同的人對各類別下項目的想法可能不同。項目基本上是每類別的定義。如何在類別確定下，以適當的項目來描述各類別且有一定的共識是研究者要探討的。

3.1 資料來源

研究者以大考中心公佈的作文為主。作文分為說明文和敘述文，每一文體採用兩篇，一為高分篇，一為低分篇（大考中心，民 86b，請見表四）。取高、低分兩篇是要確認哪些標準可以分辨不同程度的作文。

表四：選文分數表

	高分平均	低分平均
敘述文	37	26
說明文	39	25

3.2 研究方法

研究者採用評分標準乃綜合過去有關作文研究的評分標準，並沿用三大類：

- (1) 文字、修辭、文法。
- (2) 組織結構。
- (3) 內容。

但因有文獻強調標點符號和錯字的評分（如陳英豪、簡楚瑛、王萬清，民 77、張新仁，民 81），因此有第四類：

- (4) 「其他」如標點符號、錯別字等。

各類的指標如表二，再加上標點符號和錯別字。因文體不同，各類評分項目略有不同。

在大考中心的作文中，看不出標點符號或錯字這一類的問題。因此研究者只好在範文中做些變更，如將標點符號、用字、用詞改為錯誤的，使評分標準的特點鮮明，看看教師在評分時是否看重這些特點，目的也是在確認哪些標準可以分辨不同程度的作文。因此本研究共有八篇文章，文體 2*高分或低分 2*變更與否 2（表五）。如此設計目的在探討：

- (1) 若本研究所定出的指標有效，可由高分卷和低分卷（1 和 3 及 5 和 7）的比較看出差異。
- (2) 若因應第四項指標而作的更動的有效，可由原版和變更版（1 和 2、3 和 4、5 和 6、7 和 8）的比較看出差異。

每一篇文章依其文體各自附一份評分標準。評分方法為 5 等量表，研究者請老師依所讀文章是否符合此標準勾選。5 等為：非常符合，符合，尚可，不符合到非常不符合。敘述文每一項最高分五分，共八十五分，若第四類有任何打勾，勾一題扣一分。論說文每一項最高分五分，共八十分，若第四類有任何打勾，勾一題扣一分。

研究者將八份作文刻意安排，使每一位老師評兩份在程度和文體上不一樣的文章。每份作文請 10 位老師評分。參與的老師都是高級中學的國文老師。回收的作文評分經整理，可用的有 7 至 10 份。

表五：研究文章的設計

	敘述文		說明文	
	A 卷（高分）	B 卷（低分）	A 卷（高分）	B 卷（低分）
原版	1	3	5	7
變更版	2	4	6	8

3.3 結果

3.3.1 操弄的效果

研究者預期原版和修改版中，可看出兩者在「其他指標」上的差異，但結果卻不一致。敘述文 A 卷原版及修改版(1 和 2)的比較，兩者在「其他指標」上，只有「錯字」達統計邊緣上差異($p < .06$)。說明文 B 卷(7 和 8)的標點

符號操弄達邊緣顯著性($p<.053$)。至於其餘任何兩個版本，在錯字和標點符號方面沒有顯著差異。沒有差異的原因，將在下面有進一步分析。

3.3.2 信度：各類項目間一致性

爲了解研究者所設定的評分標準有內在一致性，也就是說，這些標準都是用來評量各類作文能力，因此每一類項目間應有內部一致性。

(1) 說明文

除第四類評分不是每一項目間皆達統計上顯著相關外，其他三類分項目間的相關都達 $p<.01$ 的顯著水準。第四類項目「錯字」與其他各項無關，也與第四分類總分無關。「標點符號」只和第四類總分有相關。換句話說，錯字和標點符號的使用不相關。

(2) 敘述文

第一至第三類每一分項間都有統計上顯著相關。第四類分項間的相關與說明文不同，「錯字」與各分項間無統計上的相關但是與第四類總分有 $p<.05$ 相關。「標點符號」也與第四類總分有關。

整體而言，第一類至第三類各子項目的內部一致性都是肯定的。第四類的內部一致性不佳，其中「錯字」和「標點符號」不相關。這可能表示這兩種表現不重疊。

3.3.3 效度：分類項目間的關係

研究者以各類分項間的相關程度，分析評分項目間的一致性。

(1) 說明文

1-1. 四類總分間的關係

第一、第二和第三類各類總分間以及各類總分和四類總分間達 $p<.01$ 的顯著水準。第四類和其他三類各個總分以及四類總分皆不相關。

1-2. 文字、修辭、文法與其他三類的關係

第一類各分項和第二和第三分類總分以及四類總分都有 $p<.01$ 的顯著相關。第一類「修辭」和第二類的「起始有力」、「結論得體」、「首尾連貫」、「合乎邏輯」無統計上顯著相關，但是和第三類以內容爲主的各分項以及第二類的「論點豐富」和「設定讀

者」有關，顯示修辭與「內容」方面的項目有關。「文法正確」則與第二和第三類各項都有統計上顯著相關。「語句豐富」和第三類各項都有關但是和第三類「取材適當」、「旁徵博引」以及「見解獨到」無關。這結果似乎指出修辭與內容的關係比與文法和結構的關係密切，而文法結構的評分與內容也有關係。

1-3. 組織結構與其他三類的關係

組織這一類和第一及第三分類總分以及四類總分皆有關。組織和第一類用字遣詞的關係在上一節已描述，此處只說明組織結構和內容的關係。「首尾連貫」是唯一與第三類每一項都不相關的指標，而「起始有力」、「承轉曲折」、「論點豐富」、「設定讀者」與第三類每一項皆有關。這結果呼應上面提到的，組織的評分與內容不易分離。

1-4. 內容與其他三類的關係

第三類各項和第一和第二分類總分以及四類總分有關。至於內容與其他兩類分項關係已在上一節描述。

1-5. 標點、錯字與他三類的關係

標點符號和第一、二、三類各項分數皆無關。錯字則與第二類的「論點豐富」和第三類的「見解獨到」以及第三類總分有負相關。兩者與四類總分也無統計上的相關。基本上，在說明文評分中，標點符號和錯字的分數與語言、組織或內容分數沒有統計上顯著關係。

(2) 敘述文

2-1. 四類總分間的關係

四類各個總分和四類總分皆達 $p < .01$ 的顯著水準。第一至第三類總分間有顯著相關。第四類只和第一類（文字、修辭、文法）有統計上顯著相關，但是和第二類（組織）、第三類（內容）則不相關。

2-2. 文字、修辭、文法與其他三類的關係

第一類各項和第二及第三分類總分以及四類總分都有 $p < .01$ 的顯著相關。第一類「修辭」和第二類的「段落分明」、「承轉曲折」、「首尾連貫」無關，但是和第三類各項以及第二類的「起始有力」、「結論得體」、「闡述合理」和「設定讀者」有關。「文法

正確」和第二類「段落分明」、「承轉曲折」、「首尾連貫」和「設定讀者」無關，和第三類的「取材適當」、「文題相符」無關。「詞句豐富」和「字詞優美」和第二類的段落、起始無關，但與第三類「內容」各項皆有關。「文筆生動」則與第二類「段落分明」、「首尾連貫」和「設定讀者」、第三類「文題相符」無關。結果似乎指出老師批改敘述文時，對修辭和結構分開考慮，也對「語句豐富」和「字詞優美」和內容的考慮較接近。

2-3. 組織結構與其他三類的關係

組織各分項和第一和第三分類總分以及四類總分都有關。組織與第一類用字遣詞的關係上一節已描述，不再贅述。組織和內容的關係，只有「段落分明」、「首尾連貫」、「闡述合理」和第三類每一分項都有關。這與說明文的結果相近，就是組織的評分與內容的評分有關。

2-4. 內容與其他三類的關係

內容各分項和第一和第二分類總分以及四類總分都有關。與第一和第二類分項的關係上面已敘述，不再贅述。

2-5. 標點符號、錯字與他三類的關係

標點符號與四類總分有關，錯字則皆無。標點符號與第一類除「語句豐富」外皆有關但是只與第二類的「段落分明」、「起始有力」以及第三的「新穎獨到」、「精彩動人」有關。

不論說明文或是敘述文，明顯的是，修辭與內容有關係，而修辭與組織的評分似乎是可以分開處理的。至於文法正確一項與組織和內容都有部分關係。這或許就是文獻中把文法、語句豐富等放在一類的原因。至於組織的評分不容易與內容分開。也就是說在評量組織時，內容與組織會互相影響。因此，當以組織為一般評分項目，容易受特殊評分項目，也就是內容的影響。這使得研究者不能輕易把組織當一般項目而忽略其亦可能是特殊項目。至於錯字或標點符號，只有標點符號在敘述文中與其他類分數有統計上的關係。

3.3.4 能分辨高低分作文的評分項目

為找出最有效的評分項目，研究者比較不同評分項目對於高低分兩篇作文的區辨力。上面提過因錯字和標點符號的操弄不顯著，因此將原版和變更版合計，以增加統計上的效力。表六和表七列出能分辨高低分兩篇作文的項

目。說明文有十一項分別屬於三大類。

表六：可以分辨說明文程度的指標

類別	項目	A 卷 平均數	標準差	B 卷 平均數	標準差	t 值	顯著水 準
文字	文法正確	4.30	.48	3.14	.69	4.09	.001
修辭	文筆生動	4.30	.67	3.14	.69	3.45	.004
文法	語句豐富	4.00	.67	3.00	.58	3.21	.006
	修辭	3.90	.74	3.14	.69	2.14	.05
組織	承轉曲折	4.20	.63	3.14	.38	3.94	.001
結構	起始有力	4.60	.52	3.57	.98	2.84	.01
	闡述合理	4.30	.67	3.29	.76	2.91	.01
	論點豐富	3.80	.92	2.86	.38	2.55	.02
	段落分明	4.40	.69	3.43	.79	2.68	.05
內容	見解獨到	4.40	.94	2.86	.38	3.02	.01
	精簡扼要	4.10	.57	3.29	.76	2.54	.02
總分		72.70	7.73	59.00	10.26	3.15	.01

敘述文有七個項目都可以分辨高低分敘述文的程度（表七）。值的注意得是「文法正確」在敘述文不具區辨力。因此研究者在敘述文的「文字，修辭」一類中刪去文法一項。文法一項與第三內「內容」每一項都達統計顯著相關，與修辭亦達顯著相關。這顯示在敘述文評分中，評分者看文法和用詞遣詞及內容有高度重疊。因此「文法」一項雖不具區辨力，但也被包括在「文字修辭」和「內容」中了。

表七：可以分辨敘述文程度的指標

類別	項目	A 卷 平均數	標準差	B 卷 平均數	標準差	t 值	顯著水準
文字	字詞優美	3.78	.97	2.80	.42	2.90	.01
修辭	詞句豐富	3.89	.60	3.20	.79	2.12	.05
組織	段落分明	4.22	.67	3.30	.48	3.48	.01
結構	首尾連貫	4.22	.44	3.30	.67	3.48	.01
	結論得體	3.89	.93	2.90	.88	2.39	.05
內容	文題相符	4.33	.50	3.50	.85	2.57	.02
	新穎獨到	4.22	.67	3.40	.84	2.34	.03
總分		71.78	12.49	59.70	7.32	2.61	.02

透過分析能否分辨高低分作文的項目，淘汰掉不能分辨高低分的項目，使評分指標驟時減少。

3.3.5 與大考中心分數的相關

研究者以大考中心各篇的分數為效標，分析本研究評分項目的效度。不論文體，除第四類總分，其餘三類總分以及四類總分都與大中心的分數達統計上顯著相關。至於各類分項，只有表八和表九中所列項目與大考中心的分數有統計上顯著相關。這一來也提供刪減評分項目的依據。

表八：說明文評分指標與大考中心成績相關

類別	項目	相關係數
文字	修辭	.48
修辭	文法	.73
文法	詞句豐富	.64
	文筆生動	.67
組織	段落分明	.57
結構	起始有力	.59
	承轉曲折	.71
	首尾連貫	.48
	闡述合理	.60
	論點豐富	.55
內容	見解獨到	.62
	精簡扼要	.55
其他	錯字	.62

表八項目除錯字一項,皆有內部一致性，也有分辨力。錯字在論說文中不是一個有分辨力的指標，但與大考中心的分數有顯著相關，表示它是一個與作文整體表現有關的指標，但因沒有分辨力，研究者懷疑是本研究的操弄數量不夠造成分辨力不足。

表九：敘述文評分指標與大考中心評分的相關

類別	項目	相關係數
文字	詞句豐富	.46
修辭	字詞優美	.58
組織	段落分明	.65
結構	結論得體	.50
	首尾連貫	.65
內容	新穎獨到	.49
	文題相符	.53
其他	標點錯誤	.47

表九項目中除標點符號，皆有內部一致性也有分辨力。標點符號在敘述文中不是一個有分辨力的指標，但與大考中心的分數有顯著相關，表示它是一個與敘述文整體表現有關的指標。

3.3.6 新評分項目

本研究目的在找出最少最有效的作文評分標準。經過上述的分析，研究者以有內部一致性、有分辨力及和大考中心原始分數的相關達顯著水準者為評量作文的標準。只是項目仍嫌稍多。研究者決定把分辨力只達.05 的項目先保留不用（請見表六、表七）。被刪除的項目與同一類的其他項目都有顯著相關。如說明文的「修辭」與同一類中的「文筆生動」或「語句豐富」都有 $r=.60$ 或 $r=.61$ 的顯著相關($p<.01$)。這表示被刪除的項目與同一類的項目在意義上有重疊處，因此刪去並不會影響對同一類的定義。其次關於錯字和標點符號，基於可能是本研究操弄不足，因此暫時不刪除。研究者依文體的特色把有信度且最能分辨的評分項目列出如下並進行下一階段的試用研究。

說明文：

1. 文字，文法：文法正確、文筆生動、語句豐富。
2. 組織結構：承轉曲折、起始有力、闡述合理、論點豐富。

3. 內容：見解獨到、精簡扼要。
4. 錯字。

敘述文：

1. 文字，修辭：字詞優美。
2. 組織結構：段落分明、首尾連貫。
3. 內容：文題相符、新穎獨到。
4. 標點符號

關於錯字和標點符號，如上面提過，在分辨高低分作文上，分辨力不夠，可能是研究者所操弄的「量」不足。但是因他們與大考中心的分數有關，也就是說他們仍足以影響作文的得分，研究者乃保留說明文中的錯字和敘述文中的標點符號評量。至於他們在作文評分中扮演的角色，未來研究中值得再探討。

4. 研究二、新評分標準試用

為了解經過層層篩選出來的作文評分標準是不是適當且可用，研究者接著進行試用研究。試用的考慮是：

- 甲、若小學高年級學生能夠使用這一些評分標準，則表示這些標準定義的歧異性不大。
- 乙、不同人若對同一篇文章使用相同標準評分，結果差距不大，表示標準適用。

基於以上考慮，我們以目前研究者經營的作文網站上公布的文章為主，請參與網站的學生以上述的評分標準為他人的作文評分。我們也請一位作文老師以同樣的標準評量這些文章。結果如表十，前兩次結果指出老師和小學生的評量在故事體（敘述文）或是說明文都沒有統計上差異。但是第二篇敘述文結果有統計上差異（ $p<.05$ ，表十）。

表十：老師與同儕使用新標準評分差異分析

說明文	人數	平均數	標準差	t test	顯著水準
老師	41	3.46	1.00	-.92	.37
同儕		3.70	1.41		
敘述（一）	人數	平均數	標準差	t test	顯著水準
老師	59	3.68	.51	.21	.83
同儕		3.64	1.13		
敘述（二）	人數	平均數	標準差	t test	顯著水準
老師	42	3.81	.55	2.03	.05
同儕		3.24	1.62		

由於第二篇敘述文是同儕給的分數低於老師給的分數，而且標準差(1.62)相當大，表示同儕之間的看法有歧異，研究者以為有別的因素參與其中。目前，研究者接受新評分標準的可用性，但會繼續考驗與修改，使其更具公信力。

5. 結語

本研究透過評分項目內部一致性，高、低分分辨力以及和大考中心成績為外在效標的分析，找出說明文 10 項和敘述文 6 項的評分指標，並經過試作，以 5 點量表方式評分，大抵指出這些項目是可以接受的。與大考中心不同的是，本研究將因內容而異的特殊標準融入評分項目中，為不同的文體分別找出各自的評分項目。有具體的評分項目不但幫助教師評分時更有依據，更重要的是學生透過這一些標準知道作文的要求，進而能依標準來改善其作文水準。雖研究者嚐試操弄所採用材料中的錯字與標點符號以彰顯這兩項評分標準的信度與效度。但是結果未如預期。這有兩種可能：

（一）研究所變更的錯誤量不夠，顯不出效果。

（二）評分老師不覺得錯字和標點符號影響全文水準的呈現。

就如國外研究指出，作文修改若只在錯字和局部細節，對增進全文品質的幫助不大（例如 Parr, 1995）。對於標點符號和錯字是否分辨作文品質的效度，研究者覺得仍需要進一步探討。而基於錯字和標點符號與三大類評分項目的相關不高，因此不會影響本研

究所篩選出來的評分項目。研究者將擴大試用目前所得到的評分標準，更確認其為簡單、可用但不失作文評分精神的標準。

參考文獻

- Parr, J. (1995) When pens are passé: Students reflect on written composition. *Journal of Research on Computing in Education*, 27 (2), 221-231.
- 大學入學考試中心 (民 85)，八十四年度基礎科目國文科試題研發工作計畫研究報告。台北：大學入學考試中心。
- 大學入學考試中心 (民 86a)，指定科目考試規畫研究報告 (國文科)，語文表達能力測驗研究 (一)。台北：大學入學考試中心。
- 大學入學考試中心 (民 86b)，語文表達能力測驗研究 (二)。台北：大學入學考試中心。
- 大學入學考試中心 (民 88)，語文表能力測驗研究。台北：大學入學考試中心。
- 王萬清 (民 81)，寫作教學初學者之寫作教學內容知識結構及其改變歷程研究。台北：國立台灣師範大學教育與輔導研究所博士論文。
- 陳英豪、吳裕益、王萬清 (民 77)，兒童寫作能力測驗編製報告。台南：台南師範學院初等教育系。
- 張新仁 (民 82)，寫作教學研究－認知心理學取向。高雄：復文書局。
- 鄭博真 (民 85)，寫作修改教學策略對國小學生寫作修改表現、寫作修改能力、寫作品質和寫作態度之影響研究。台南：國立台南師範學院國民教育研究所碩士論文。

柯華葳

hwawei@cc.ncu.edu.tw

The Evaluation Criteria of Expository and Narrative Writings

Hwawei Ko

National Central University

Abstract

Grading writing is not an easy job. How to reach a set of comparatively objective criteria is the concern of many related studies. In common practice, there are 3 categories used to consider the quality of writing, namely, mechanical, organization, and information. In each category, 5 or more items are used to define the category. In this study, we try to eliminate items by analyzing : 1) items that are highly correlated; 2) items that do not have good discriminating power; 3) items that do not have criterion-related validity. The 3 categories still upheld. A total of 9 items left for expository and 6 items left for narrative. 5 point scale is assigned to each item, 5 stands for the most fitted and 1, the least fitted. After a tryout by different group of elementary school students, the reliability and validity of these writing criteria are acceptable. It is suggested that teachers of school adopt these simplified criteria and teach their students to use them for self and peer assessment.

Key words: written composition, assessment criteria, narrative text, expository text