

國小四年級學童閱讀摘要實作表現 及自我評量之研究

張碧容

臺南市立文元國小

鄒慧英

國立臺南大學測驗統計研究所

摘 要

本研究旨在發展閱讀摘要實作評量，及瞭解自我評量融入國小語文領域閱讀教學的實施成效；藉由探討國小學童實施自我評量之學習後，其在摘要能力上的成長狀況，及對閱讀摘要、自我評量之學習環境的改進與建議。本研究以台南市YY國小之四年級二個班級的64位學生為研究對象，隨機分配一班為實驗組，一班為控制組，進行閱讀摘要教學及自我評量之實驗處理；以研究者自編之「閱讀摘要實作評量」、「閱讀摘要能力自我評量表」、「閱讀摘要學習單」為研究工具，主要研究結果如下：

- 一、本研究國小四年級學童的摘要實作評量前測表現不甚理想。
- 二、自我評量對國小學童的摘要實作後測表現有顯著效果。
- 三、自我評量有益於國學童的摘要學習單表現。
- 四、閱讀摘要實作評量具有良好的評分者信度。

關鍵詞：閱讀摘要能力、摘要教學、自我評量、實作評量、評分者信度

壹、緒 論

本研究主題為國小四年級閱讀摘要實作表現及自我評量，本章將分為研究動機、研究目的、研究問題、名詞釋義及研究限制來加以說明。

一、研究動機

(一) 自我評量融入教學

從八十二年的課程標準到現今的九年一貫課程綱要，學生聽說讀寫作能力之培養依舊是國語文教學不容忽視的重點。在九年一貫課程中，語文領域中的國語文仍較其它學科佔有較多教學時間，其中又因教育部的推動，語文閱讀備受重視。值此之際，e學童語文能力普遍低落之批評更是時有所聞，尤其是閱讀與寫作能力的下降更令人憂心不已，針對後者，教育部更決定將寫作納入國中基本學力測驗，希望能因此喚起學校對寫作教學的重視。然而欲改善現今教學、學習情況，讓學習者、教學者突破以往舊觀念，除利用正式的考試制度外，將自我評量融入教學或與教學相結合或許是另一條可行途徑，因為正式考試已是教學成果的驗收，如何在教學過程中有效提升學生的學習成效方是重點！教師藉由自我評量方式，使學童體認出自己是學習的中心者，需對自我學習負責，進而養成其對自我學習的監控機制，這正是自我評量融入教學最主要的訴求。

(二) 閱讀摘要能力之培養與評量

閱讀摘要是閱讀理解能力之重要過程之一，由一篇文章的要點摘要，進而得知其是否真正理解此篇文章內涵，因此，閱讀摘要表現的好壞或可視為學生閱讀理解能力的一項指標。一篇文章主要是由一些重要訊息架構起來，故學生可經由閱讀摘要活動增強其閱讀理解能力，進而提昇其閱讀理解的表現；基此觀點，學生經過持續性地閱讀摘要練習之後，定能在閱讀理解方面能有所進步。由於摘要能具體呈現學生對閱讀材料的理解狀況，故閱讀摘要能力是閱讀理解能力相當重要的課題。

雖然摘要學習有助於閱讀理解能力的提昇，但以往的語文教學活動受限於種種因素，使得「摘取大意」學習活動備受忽視。而閱讀理解評量常見客觀式題型，少有閱讀摘要能力的評量。雖然摘要能力的評量有批改、評分標準等困難，但若能在大量閱讀的生活情境中，有效摘取文章大意，學童必能在學習成效上事半功倍，所以閱讀摘要實作評量有其教學上的重要性。

(三) 四年級為學童閱讀轉折期

Chall (1983) 在其研究中將兒童由初學閱讀到精熟閱讀分為六個發展階段，並將四年級視為開始轉變為理解閱讀的新知學習閱讀期。就學童閱讀能力的發展來看，國小一、二、三年級是以解碼認字為主，到了四年級，其理解閱讀大大提昇，並開始能利用閱讀學習各項新知，故可知四年級學童是由「閱讀認字」轉而「閱讀理解」之重要關鍵期。

對於四年級的學生，應該協助他們在摘要能力上能突破以往，因為閱讀理解的基石就是大意摘要，倘若學生現在能對自己的摘要表現有所要求，自然可以預期學生的閱讀理解會因其在摘要的自我要求而成長提昇。本研究將以實作評量方式，結合國小兒童閱讀摘要活動與自我評量，探討學童練習閱讀摘要（填寫閱讀摘要學習單）及自我評量對增進摘要能力的有效性。

二、研究目的

據上述，研究者認為閱讀摘要能力實為學童閱讀理解能力的一部份，因此若欲有效提昇學生的閱讀理解能力，不妨可以從閱讀摘要的教學做起。然如何能讓閱讀摘要教學產生效果，研究者以為自我評量的介入不失為一個好方法，因此，本研究旨在探究透過自我評量的介入閱讀摘要教學活動設計，能否有效改善學童的摘要能力，進而提昇其閱讀理解能力。據此而言，本研究目的分述如下：

- 1.發展閱讀摘要實作評量及評分規範。
- 2.探討國小四年級學童在閱讀摘要實作評量上的表現。
- 3.編製閱讀摘要能力自我評量表。
- 4.探討學童在閱讀摘要能力自我評量表上的表現。
- 5.探究自我評量對促進學童閱讀摘要能力之有效性。
- 6.探討不同評分者間評分之一致性。

三、研究問題

依據前述研究目的，本研究主要待答問題有下列數項：

- 1.如何建立閱讀摘要實作評量及評分規範？
- 2.四年級學童在閱讀摘要能力上的表現如何？
- 3.如何編製閱讀摘要能力自我評量表？
- 4.四年級學童在閱讀摘要的自我評量表表現如何？
- 5.自我評量的有效性？亦即四年級學童的閱讀摘要能力是否因自我評量而有增進？
- 6.不同評分者間之評分等級相關程度？各評分差異情形如何？

四、名詞釋義

（一）閱讀摘要實作評量

此項評量指稱由研究者自行設計的閱讀摘要作業；實施方式係由學童在結構化情境中，先閱讀一段文章，再以文字書寫方式擷取摘要該文章內容，爾後由研究者依照事先擬定之評分規範進行摘要的評分。

（二）閱讀摘要能力

意指閱讀者進行閱讀時，擷取、記錄文章主要重點的能力，也就是國語文教學中所強調的摘取大意能力。本研究以學童在研究者自編之閱讀摘要實作評量的得分代表其閱讀摘要能力高低，分數愈高代表其閱讀摘要能力愈佳。

（三）自我評量

意指個人對其學習歷程進行自我反省、評鑑的行為。本研究的自我評量是指受試者依據研究者擬定的評分規範與評定量表進行自我檢視，以評定作業單（研究者所自編的閱讀摘要學習單）摘要表現有那些地方符合標準，那些地方需要加強、改進。

五、研究限制

(一) 研究對象方面

本研究僅限於台南市某國民小學兩班四年級學生為研究樣本，因此結果的解釋及推論宜考量研究對象的地區性及年級的限制。

(二) 實驗處理時間方面

整個研究實驗處理期間為九十二學年度第一學期，但期間因遭遇期中考、期末考及放假而有所中斷，教學次數為18次，所以有限實驗處理時間對本研究結果所造成的影響是另一限制。

(三) 研究方法方面

由於時間與能力的限制，研究者無法長時間觀察學童自我評量的情形，僅能利用檢核表對學童的自我評量進行觀察；學童在填寫檢核表時，可能會受到其認知、當時情緒或社會期許等主客觀因素的影響，而選擇與實際情形不符的自評結果。此外，由於人力的限制，閱讀摘要實作評量方面，只有二位評分者，所以測量結果會有某種程度上的誤差存在，故在推論研究結果時需謹慎。

(四) 實驗教學與教材方面

本研究選用的閱讀文章數量有限，加上研究者本身的教學經驗有限，這將侷限本研究發現的外在效度，換言之，閱讀摘要活動的教學效果或許會因為閱讀文章的性質、種類而異，也可能因教學者的能力、經驗產生不同的成效，這都是本研究的限制所在。

貳、文獻探討

本章共分二節，第一節是閱讀摘要方面之研究，第二節是自我評量方面之研究。

一、閱讀摘要

(一) 摘要的定義

摘要即是「摘取要點」(何容，民63)，亦即摘取文章中的重點一整篇文章中最主要內容的意思。國內外皆有學者認為文章中各要點的組合便形成「大意」，彷彿全文的「點」、「線」交錯而成的「面」，就是「全文大意」(小光，民64)；Pearson和Johnson(1978)定義文章中最重要思想即是大意；Winograd(1984)定義大意是文章中的要點；而Hare和Bingham(1986)更是定義文章中最重要敘述就是大意。

大意一詞歧義甚多，以現行的國語「教學指引」或「習作」來說，與大意相關的有主題及主旨(要旨)。大意是指將作者說了什麼簡短的表達出來，限定在文章本身，包含具體細節；主題是作者所說的有何意指，目的在使讀者對人生或人性有所啟示；至於主旨又有重點及要點之分，重點是以詞組型式表達全課內容，而要點則是以句子型式表達分段內容(吳英長，民87)。

(二) 摘要能力的理論

Kintsch和Van Dijk (1978)認為整篇文章是由一系列的命題所組成，這些命題彼此間是以階層關係相連存在。讀者在摘取說明文大意的過程，要將文句轉變為命題，並且應有「刪掉不相關命題」、「刪掉重覆命題」、「用類別名詞取代類別舉例」、「整合細節命題」等四項能力。

Brown和Day (1983)更指出在摘取說明文大意的過程中，讀者應有「刪掉不重要訊息」、「刪掉重覆訊息」、「用屬性名詞代替屬性舉例」、「挑出主題句當作大意」和「自己寫主題句作為大意」等五項能力。

以上兩者所述理論大同小異，皆為探討在摘取說明文時所需能力，主要差別在於第一步驟：Kintsch和Van Dijk所指不相關命題是與文章重點完全無關連的概念，但Brown和Day卻認為不重要訊息與文章內重點有所關連，但因屬於不重要地位，給予刪除。此外Brown和Day將Kintsch和Van Dijk所提出的整合細節命題，細分為直接在文中找出主題句，及文中若未有主題句改由自己寫出主題句兩種類型。在本研究中所挑選的閱讀摘要文章，大多數為說明文，故可參照兩者摘取說明文所需能力之理論來給予教學指導。

(三) 閱讀摘要能力的表現方式

有關摘要能力的表現有多種不同的方式，常見的表現主要有口述摘要、文字摘要及圖解摘要這三種方式，以下將分別探討各個表現方式，及其優劣之處（方金雅、鍾易達，民87）。

1. 口述摘要能力：指閱讀文章後重點的摘錄，以簡要的口語來表達文章內容重點的能力。使用口述摘要可避免受試者文字表達能力所造成的測量誤差，更能測量出學生閱讀理解能力；但其缺點則是不易保留摘要的原貌，且不易體驗到整篇文章的結構。
2. 文字摘要能力：指閱讀後，以精簡文字描寫文章內容重點之能力。此文字摘述要點的方式，具有清楚顯示文章或段落結構的優點，且可用文字來合宜地描述文章重點，化繁為簡，增進閱讀理解；惟其缺點是對年幼學童不甚適合，因其文字敘寫能力未臻成熟，測量結果易受文字表達能力影響。
3. 圖解摘要能力：讀者在閱讀後能以圖表方式一方塊、圓圈、線條、箭號等圖形符號來表示文章或故事的結構與大意之能力。圖解摘要雖能了解文章的結構、顯示全體與部分間的關係、作為前導組織協助學生順利將新知納入舊知識體系等優點，但此法顯示文章或段落的能力則較有限。

由上述清楚得知，口述摘要能降低書寫能力所造成的影響，更精確測出學童的理解能力，但卻因不易保留原先摘要，難以認知到全文架構；使用文字句子能馬上清楚顯現文章內容重點，但易因學童文字表達能力影響測驗結果；至於使用圖形摘要方式可以具體、生動地顯示其文章架構和關係，但所顯示內容較為有限。由於本研究教學時間有限，為能有效清楚了解學童摘取內容重點之能力，故採用文字摘要方式，且本研究對象為四年級學童，其文字敘寫能力已具有一定水準，因此文字表達能力較不易造成測量誤差。

(四) 閱讀摘要之相關研究

1. 摘取文章大意的成果與過程

探討個體摘取文章大意的研究主要分為「摘取大意的產出成果」與「摘取大意的過程」二類。

- (1) 摘取大意成果（寫出大意）：一般判斷是否具有寫出大意的方式是當個體能答出與所設定的大意相似答案時，就表示其具有寫出大意的能力（楊韻平，民82）。如Williams, Taylor和Ganger (1981)、Williams, Taylor和deCani (1984)及Hare, Rabinowitz和Schieble (1989)皆先設定正確大意，若受試者能用自己的話寫出相似設定的大意，即可算具有寫出大意的能

力。

- (2)摘取大意過程：一般判斷是否具有摘取大意過程之能力是設定摘取大意過程各項工作的正確答案，個體若能答出正確答案，即表示具有該項能力（楊韻平，民82）。如：Brown和Day（1983）是根據受試者寫出的大意有無不重要訊息、重覆訊息、屬性名詞及主題句來判斷；而Hare, Rabinowitz和Schieble（1989）則認為受試者若能辨認所設定主題句工作的正確答案，就算具有辨別主題句的能力。

由上所述，摘寫大意乃是閱讀與寫作兼具，讀者需要多項能力共同運作才能做到理解一篇文章並正確地摘要內容重點。本研究以摘取大意的產生成果來決定學童是否具有摘要能力；若其能寫出相似於研究者設定的文章主要概念，就代表該學童具備寫出大意之能力，擁有完整、極佳摘要能力者，則能用自己的話語寫出大意，並以文章中的細節支撐其主要概念。

2. 學童摘要能力之現況

Brown和Smiley（1977）要求三、五、七年級學生及大學生以四點量尺評定文章中每一概念單位對文章主題的重要程度，結果發現只有大學生能區分各個層次的概念單位。此外，Hare, Rabinowitz和Schieble（1989）要求四、六、十一年級學生直接畫出文章的主題句，結果指出此能力具有年級差異，較低年級無法正確辨認主題句。

Williams, Taylor和Ganger（1981）發現只有43%的四年級學生和57%的六年級學生能正確寫出文章大意。Brown, Day和Jones（1983）研究發現只有16%的五、七年級學生是用自己的話來敘述大意，多數人還是逐字抄寫或部分照抄整篇文章當作大意。

楊韻平（民82）以台灣國小四、六年級學童為對象來探討其摘要能力之表現，研究發現大多數表現不佳學生並不能分辨文章中的主題句、支持主題句的細節，因此在摘取大意的過程與成果時難呈現出正確答案。同樣方金雅、鍾易達（民87）在有關國小學童閱讀摘要能力的應用研究上，發現台灣區14個縣市的980位國小四年級學生，只有11%學童能準確而完整摘要到文章重點。

由以上所述，可發現國小學童在摘取大意上的表現仍需加強，教學者應協助其發展出「迅速瀏覽、閱讀完一篇文章後，了解大意、重點所在」的能力；同時，教師應有非短短十、二十幾分鐘教學時間即可促使學童擁有此能力的認知，而是需教師進行長期的大意教學。至於教師在測量學童的學習狀態時，也應突破以往總是使用選擇題方式來測得學生摘取重點之窘，改以提供實作作業形式來讓學生展示其摘要能力。

3. 摘取大意教學

Pearson的研究小組將整個教學流程分為示範、指導式練習、固定期、獨立式練習及應用五個階段。強調大意策略具有可教性，經由教師的示範與指導，逐漸將閱讀策略移植到學生身上，進而能獨立應用在真實情境（Pearson & Dole，1987）。

Ellis的研究小組將理解的教學分為：先將閱讀策略作為手段，用以了解文章內容，之後再把策略當做目的，學會如何應用它。強調撰寫大意作為一種閱讀策略在進行教學時，也需要不同的考慮（Ellis, et.al., 1991）。

Gagn'e根據訊息處理理論，發展出引起注意、告知目標、回憶舊經驗、呈現新刺激、提供學習輔導、誘發表現、提供回饋、評鑑表現、促進遷移這九個教學事件。強調除了一般性的目標導向教學流程，更應配合學習者的內在歷程來安排（吳英長，民87）。

雖然閱讀研究方面文獻已有眾多學者提出，但國內有關摘取大意實際教學仍有不足處，於是在本研究當中，乃參照以上所探討的理論來設計教學流程，將閱讀摘要教學主要分為引起動機、閱讀活動、師生討論、摘寫大意及評分規範的討論五個步驟，可達到引起注意、呈現新刺激、教師示範及提供更多的解釋、提供回饋讓學生自我檢視……等功效，希望藉由教師策略性

的教導來協助學童培養其摘要能力。

(五) 兒童閱讀能力的發展

依據林清山(民87)的說法,許多學者提出「學習閱讀」與「經閱讀而學習」之區別,前者是指學習如何將印刷文字轉換成其它形式,此種轉換歷程的自動化,即學習閱讀的精熟乃幼稚園至小學三年級閱讀教學的主要重點;後者是指將閱讀當作一種工具,以獲得某些科目領域之特定知識,為四年級及其後的閱讀教學重點。

可知在國小三年級時,識字能力變得更自動化了,而學生閱讀時的認知處理資源改以更高的片語、句子或段落篇章層次為主,使得在四年級時,許多學童開始慢慢出現閱讀上的困難;因此本研究將以開始接觸複雜、抽象概念的四年級學童為研究對象,探討其摘要表現程度為何。

二、自我評量

(一) 九年一貫的潮流

九年一貫課程強調學習以學生為本位,且學校的最終目標在於培養自我導向的學生,所以應協助學生對學習具有自我掌握的能力,藉由自我評量來培養對學習狀況的監控,同時也喚起新的學習欲望(柯啟瑤,民90)。

教育目的在於發展個人自治學習能力,使其能自我教導、自我組織學習,為自己的學習負起大部分責任。Boud(1988)認為自治之學習者可由經由「從事自我評量」方法,使學習更有效率;Rogers(1983)同樣認為對自我之學習若能加以評量,將引發自我學習,並為學習負責;而自我評量不但可以使學習者朝向自治學習,而且符合學習者中心之理念,並且最後幫助學習者自我發展成為自我實現者(陳如山,民82)。因此教師應盡力協助學生在學習過程中進行自我評量,使其為自己邁進教育目標之進展證明。

由此不難推論,九年一貫課程強調學習者能主動、了解、及監控自我學習狀況,所以教師可藉由自我評量方式,來使學童體認出自己即是學習的中心、主體者,進而培養其對自我之學習反省、負責。

(二) 自我評量之定義與意義

自我評量(self-assessment)乃是監控、評鑑自己知識、表現與理解度的行為(Cariaga-Lo, Richards & Rrye, 1992)。自我評量也是對自我成就、進步的情形進行反省和覺察(Paris, Ayres, 1994)。評價的終極目標,以及最持久的效果,就是自我評量(Costa & Kallick, 2000; 李弘善譯,民90)。綜合眾學者的看法,可知自我評量是個人對於學習有所反省,扮演主動察覺之角色。

Wiggins(1998)提出學習者能經由評量或自我評量所提供的回饋來確定是否達到老師所期望的表現水準;Orsmond, Merry和Reiling(2000)認為透過自我評量可以提供師生清楚界定學習結果之機會,讓評分標準變得更清晰,對學生的學習產生重要的影響。自我評量提供回顧自我考量的機會,學生可以依據評量的項目來進行反省的動作,一方面可以加強自我自信心的建立,另一面也可發現問題所在,並思考如何做出新決定以有利於學習(柯啟瑤,民90)。

自我評量具有內化優良品質及標準的功用,學生運用自我評量表,曉得自己哪裡符合標準、哪裡有待加強之後,開始有意識地重視評分標準,也明瞭一份高品質作業是什麼模樣,進而可以自動自發地要求自我完成作業,努力將作業達到高水準,而不需老師耳提面命的提醒。在本研究中,學童透過閱讀摘要自我評量活動,使教學者有效地與學生溝通標準、規範何在,並幫助其自

我監控學習歷程，充分達到自我反省。

三、自我評量之使用

Towler和Broadfoot（1992）將自我評量分為四個主要階段：

- 1.知識階段：學生回憶之前已有舊經驗，檢視自己的學習成果與過程，並針對目前所學到的進行記錄。
- 2.分析與理解階段：學生必須先了解自我學習表現之品質，並進一步分析表現好、壞的原因為何。
- 3.評鑑階段：學生能判斷學習情境，並評鑑自己所學到的內容，即其能就所處的學習情境，更具體地說出自己的優劣表現所在。
- 4.綜合階段：學生能把他所學到的學習方法，運用於整體情境中，並設定自己未來的學習目標；因此學生將自己目前所習得的方法，優處保留，劣處改進後，即可融入整體情境中，適時使用。

Van Kraayenoord（1993）建議國小階段所有年級的老師，都應鼓勵學生利用自我評量來檢視自己的作品，透過記錄來檢視自己的進步情況，使其成為更獨立的學習者。Airasian（1997）指出老師除了使用規範來評量學生表現外，亦可要求學生使用相同的評分規範進行自評，確定作業中的缺點（引自蕭雅萍，民91）。

Paris和Ayres（1994）指出，自我評量在高等教育是一種普遍的評量方式，除此之外，自我評量在小學階段也能幫助學生使用自我評量來提升學習，如Opitz和Glazer（1995）以國小二年級為研究對象，每天進行自我評鑑單和每日評鑑的活動，結果發現對學習進行反省思考可以讓學習更加有效果，Smith（1997）提及墨爾本市郊的St. Bernadette's Primary School進行學生自我評量與評鑑的研究與發展，學生定期評量目標是否達成，結果發現透過自我評量可以讓學生評估自己的能力並選擇合宜目標進行學習。此外，Costa和Kallick（2000）要求學生進行畫線條、圖形之學習單，結果發現大部分學生運用自我評量表之後，才能有效提昇作業品質（Costa & Kallick，2000；李弘善譯，民90）。

藉由自我評量能讓學生將評量與教學做密切的結合，並根據他們所了解的標準來判斷自己的表現情況。在本研究中，在進行自我評量之實驗處理時，研究者讓學生使用閱讀摘要能力自我評量表，來改善自己的閱讀摘要能力之學習，使其在每次的摘要學習單上，可以更得心應手。

參、研究方法

本研究以台南市國小學童為研究對象，進行教學實驗設計，以研究者自編「閱讀摘要實作評量」及「閱讀摘要自我評量檢核表」，蒐集客觀具體資料，並應用統計方法加以分析。以下擬從研究樣本、研究設計、教學設計、研究工具和資料分析五方面分述之。

一、研究樣本

（一）預試對象

為確立本研究所用工具之一「閱讀摘要實作評量」選用文章的適切性，在正式施測前需先進行預試，研究者選擇台南市XX國小的一個四年級班級，由班級教師挑選該班國語文低、中、高程度學童各兩名，共六名做為預試對象。

(二) 正式施測及實驗處理對象

正式施測階段，以台南市YY國小隨機選出兩班四年級學生作為施測對象；在實驗處理階段，隨機分派一班為實驗組，另一班為控制組，兩班學童基本資料如表1所示。

表1 正式施測及實驗處理對象

	實驗組	控制組
班級人數	32	32
男生人數	17	15
女生人數	15	17

二、研究設計與實施程序

如圖1架構圖所示，本研究的自變項是有無接受自我評量的實驗處理，依變項為閱讀摘要實作評量後測表現。為了能有效掌握學生實驗處理前的摘要能力，筆者於實驗處理前，以二篇閱讀摘要實作評量進行前測做為共變量的測量，俾便排除學生既有摘要能力可能對實驗結果造成的混淆效果。依研究目的，進入實驗處理階段時，受試者依所屬班級隨機分為兩組，控制組部分只進行閱讀摘要教學，實驗組除了進行摘要教學外，並讓學生就自己的摘要表現進行自我評量活動。

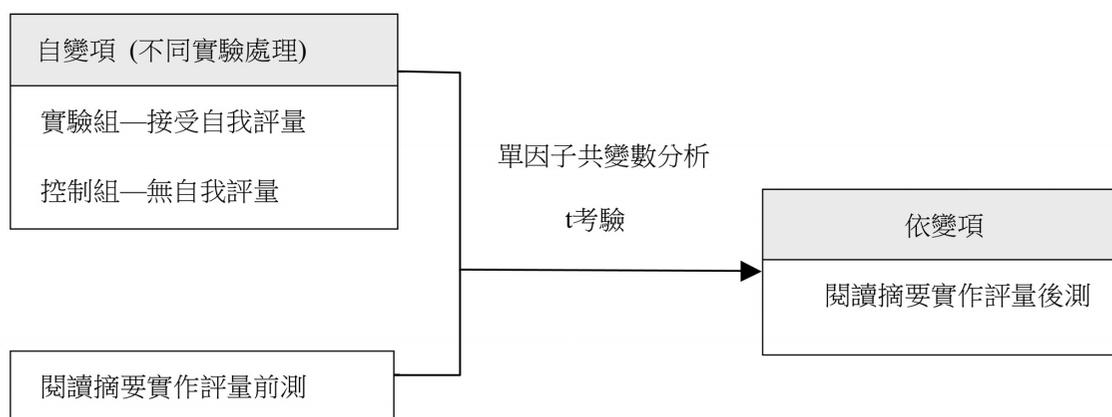


圖1 閱讀摘要實作及自我評量研究架構圖

本研究實施程序的步驟包括：

1. 蒐集相關文獻；
2. 編擬測驗工具；
3. 實驗設計；
4. 選擇及聯絡實驗學校；
5. 班級團體前測，包括閱讀理解篩選測驗及閱讀摘要實做評量；
6. 進行摘要學習，實驗組採用自我評量，控制組則無；

7. 蒐集研究資料；
8. 蒐集學生期末考資料；
9. 實施團體前測閱讀理解困難篩選測驗及閱讀摘要實作評量；
10. 整理資料及撰寫研究報告。

三、教學設計

本研究的實驗組與控制組均以同樣教學流程進行閱讀摘要教學活動，兩組主要差異在於教學活動結束後學生有無進行自我評量，兩組教學活動與實驗處理皆由第一位研究者親自進行，教學活動流程與實驗處理方式如下。

(一) 閱讀摘要教學活動流程

此部分為實驗組及控制組所共有部分，教學活動時間為37分鐘，流程主要分為引起動機、閱讀活動、師生討論、摘要活動及評分規範的討論五部分：

1. 引起動機：費時約五分鐘。在進行正式閱讀文章的教學前，研究者先以生活中的事例或故事引起學生學習動機及興趣。

2. 閱讀活動：費時約五分鐘。研究者發下閱讀摘要學習單後，先由學童自行閱讀文章內章，閱讀期間不得與同學進行討論。

3. 師生討論：費時約十五分鐘。在確定學童皆已閱讀完文章後，師生開始進行文章內容的討論，研究者採漸進方式，初期先由研究者說明文章的主要內容、大意，到後期改由學生主動說出所讀篇章的主旨內容，這段教學活動旨在讓學生從開始時的被動聆聽到最後的主動發言參與討論。

4. 摘要活動：費時約七分鐘。經過師生討論之後，學童開始進行閱讀篇章的摘要活動，即以書寫方式將重點摘要於學習單上。

5. 評分規範的討論：費時約五分鐘。當學童完成重點摘要後，研究者利用事先建立好的評分規範，與學童溝通討論一篇良好摘要所該具備的要件，此舉目的有二，一是讓學生瞭解後續教師評量其摘要的依據，二是讓學生熟悉這些評量規範做為省思其摘要寫作的準繩。

(二) 實驗處理

實驗組與控制組主要差異就在於有無進行閱讀摘要的自我評量。研究初期，實驗組於整個閱讀摘要教學結束後，再接著進行三分鐘的自評活動，即以閱讀摘要自我評量表對其剛完成的篇章摘要進行自我檢核。研究中期以後，研究者又增加十分鐘時間，針對前次摘要表現及自我評量表進行討論教學，即研究者於進行閱讀教學前，先就學童前一篇文章的摘要表現及自我評量表檢核狀況進行討論，以PowerPoint具體呈現學童的優劣表現，並再次與實驗組溝通如何利用評分規範省思文章摘要的品質。

四、研究工具

本研究共有三個研究工具，根據實作及摘要能力理論所設計的閱讀摘要實作評量，搭配閱讀課自編的閱讀摘要學習單，及參考各有關自我評量表而自編的閱讀摘要自我評量檢核表，分述如后。

(一) 閱讀摘要實作評量

1. 評量目的

本工具的設計旨在藉由真實評量測出學童的閱讀摘要能力。主要參考依據是民八十七年洪碧霞、邱上真等人所編製之閱讀理解測驗的摘要部分，及洪碧霞所設計的國小五年級長篇文章閱讀理解與摘要能力動態評量來設計，學童以文字摘要方式作答。

2. 評分規範

表2所示之閱讀摘要評分規範，係依據研究者與國小老師和專業人員討論後所擬定之評分指標和參考答案，俾便據此對學童的摘要表現進行公平計分。本規範將學童的閱讀摘要表現分為五個等級，從0分的不相關概念或誤解文章本意，到4分的主要概念與次要概念及支持性細節，顯現出學童在閱讀摘要表現上不同的層級能力。

表2 閱讀摘要評分規範

0—寫出文章中不相關的概念，誤解文章原本的意思。
1—(1)僅能寫出不重要的概念。 (2)雖能寫出次要概念，但描述不清楚或部分誤解文章中內容。
2—(1)能將次要概念描述清楚。 (2)僅寫出部分主要概念，但沒有寫到次要概念或部分誤解。
3—能清楚完整地寫出文章的主要概念。
4—能用自已的話清楚完整地將文章的主要概念寫出，且有次要概念/細節支持。

3. 預試文章挑選

研究者在準備研究所需閱讀篇章時，為求各篇文章的難度一致，分就下列幾點進行考量，首先，為避免學童對材料會有不同的熟悉度，在挑選素材時排除坊間出版的一般兒童讀本，改以教育部臺灣省國民學校教師研習會編定之「國語實驗教材」做為文章來源；其次，考量文章可讀性需適合四年級學童的閱讀水準，以及施測時間的限制，遂以四年級的實驗教材做為挑選預試文章的標的；第三，為顧及文章能有明確的中心主旨供學生進行閱讀摘要寫作，故在文體的選擇上以「說明文」為主，各篇名稱與字數如下。

- (1)測驗一：選自四年級下學期教材「猴子的記憶」，字數為542。
- (2)測驗二：選自四年級上學期教材「誰殺了大恐龍」，字數為696。
- (3)測驗三：選自四年級下學期教材「三多」，字數為710。
- (4)測驗四：選自四年級下學期教材「龍的傳說」，字數為792。
- (5)測驗五：選自四年級下學期教材「牆」，字數為606。
- (6)測驗六：選自四年級上學期教材「紅花綠葉變化多」，字數為891。
- (7)測驗七：選自四年級下學期教材「付出與感激」，字數為427。
- (8)測驗八：選自四年級上學期教材「談禮物」，字數為558。

上述八篇文章，除第七篇與第八篇屬於議論文外，其餘六篇均屬說明文體，作者挑選出這八篇文章後，又請一位曾經參與教師研習會國語文教材實驗教學的教師過目，以確認這八篇文章確實適合四年級學童的閱讀水準。

4. 預試實施方式

為檢視選用文章的適切性，研究者另外選擇有別於正式樣本（YY國小）之他所國小（XX國小）的一個四年級班級學生為預試對象，以避免預試文章的曝光。由於預試篇章多達八篇，在時間有限的壓力下，研究者商請該班教師推薦國語文能力高、中、低能力各二名做為預試對象，共計六名，以團測方式進行八篇文章的閱讀摘要預試寫作，一堂課四十分鐘寫兩篇文章，共施測四節課一百六十分鐘。

5. 預試結果說明

(1) 得分表現：表3是六位預試學童在各篇預試文章的摘要得分表現，結果顯示，無論學童國語文能力高低，都不會寫出與文章不相關的概念或誤解文章本意，顯示所選文章與學童的閱讀水準相當。

表3 預試學生在各篇測驗文章中的得分

編號	測驗一	測驗二	測驗三	測驗四	測驗五	測驗六	測驗七	測驗八
1	4	2	2	3	4	1	2	1
2	2	3	2	2	3	2	2	2
3	2	3	2	1	2	1	2	2
4	1	4	3	2	2	3	3	2
5	2	4	3	2	2	3	4	2
6	2	1	3	3	1	2	2	2

(2) 資料分析：表4是預試者在八篇文章摘要表現的敘述性統計，為瞭解這八篇文章的難度是否類似，研究者以重複量數的單因子變異數分析發現，八篇摘要的平均數差異未達顯著水準（表5， $F=0.83$ ， $p>.05$ ），顯示就這六名預試者而言，其在各篇章的摘要表現並無差異，進一步支持研究者選擇的八篇文章閱讀難度或可視為雷同或相似。

表4 閱讀摘要實作評量之預試得分之描述統計

	平均數	標準差	個數
測驗一	2.17	.98	6
測驗二	2.83	1.17	6
測驗三	2.50	.55	6
測驗四	2.17	.75	6
測驗五	2.33	1.03	6
測驗六	2.00	.89	6
測驗七	2.50	.84	6
測驗八	1.83	.41	6

表5 閱讀摘要實作評量之預試得分平均的變異數分析摘要表

變異來源	SS	DF	MS	F值
受試者間	4.17	5	.83	
受試者內	922.25	42		
文章	4.25	7	.61	.83
殘差	25.50	35	.73	

6.前、後測文章的挑選

雖然研究者挑選文章時已考量到篇章的閱讀水準，亦經國小教師確認其適切性，然為進一步確認各篇章間的難易度，研究者以六位高、中、低不同程度學生進行預試，預試結果顯示八篇文章之間的難易度相似，即便如此，基於文章字數與文體的考量，再參酌預試結果平均數與標準差等數據，研究者認為測驗一「猴子的記憶」與測驗五「牆」相近，測驗四「龍的傳說」與測驗六「紅花綠葉變化」相近，四篇文章均屬說明文體；其次考量正式前、後測的閱讀量應盡量相似，故以測驗一與測驗六、測驗四與測驗五組合成為本研究正式施測時前測與後測的閱讀素材，亦即讓前後測文章的閱讀量控制在一千四百個字左右。

7.正式施測實施方式

如所述，本研究對象為國小四年級學童，依前述取得的研究樣本，隨機分為實驗組及控制組，以團測方式進行，施測期間除了不認識的字可唸給學生聽之外，研究者不給予任何提示。測驗時間為一堂課40分鐘。

(二) 18份閱讀摘要學習單

1.內容

本系列學習單主要係參考民國八十八年宋文菊所編的閱讀理解實作評量作業單，及民八十七年洪碧霞、邱上真等人所編製之閱讀理解測驗摘要部分的設計模式，每張學習單安排有一篇文章及一個內容重點摘要的問題。文章來源為教育部臺灣省國民學校教師研習會「國語實驗教材」及康軒版本「國語課本」選出。本研究樣本學童使用非康軒版的國語教材。

2.目的

本工具的主要目的在學生經過閱讀教學實施後，透過閱讀摘要學習單的作業活動，可有效提昇其其閱讀摘要能力。

3.評分規範

與閱讀摘要實作評量之評分規範相同，共分為0、1、2、3及4五個等級，每次教學結束後，由研究者依據上述之「閱讀摘要評分規範」進行評分。

(三) 閱讀摘要自我評量檢核表

1.內容

主要參考來源為民國八十七年洪碧霞設計的國小四年級學童數學學習自我反省作業單、翰林文教雜誌及評量和記錄心智習性中的自我評量表，以檢核表型式讓學生對其閱讀摘要表現進行自我評量，評量項目如下：

- (1)我能將文章從頭到尾看完。
- (2)我能了解文章的意思。

- (3)我會先想過再寫摘要。
- (4)我的摘要包括文章中的次要概念。
- (5)我的摘要包括文章中的主要概念。
- (6)我能用自己的話寫出摘要。
- (7)我的摘要沒有不相關的概念。
- (8)我的摘要包括支持性細節。
- (9)我的摘要沒有錯別字。
- (10)我會將自己完成的摘要閱讀一遍。

2.目的

此自我檢核表的目的旨在讓學生有機會對其學習進行自我省思，評估其摘要優劣，同時據此內化閱讀摘要的評分規範（因為檢核項目多是評分規範所強調的評量要點），並且提供學童在九十二年十一月、十二月、一月、二月這四個月以來的學習歷程監控。

3.評分方式

學童在完成摘要後，依評量項目檢核自己完成的摘要內容，做到項目敘述內容者勾選「是」，若未做到者勾選「否」。

五、資料分析

本研究完成資料的蒐集後，利用統計套裝軟體（SPSS）進行資料的分析與處理。主要分析方法如下：

- 1.以描述統計來說明四年級學童閱讀摘要實作表現。
- 2.以學生閱讀摘要實作評量的前測成績作為共變數，使用共變數分析考驗兩組學生的閱讀摘要實作評量之後測表現。
- 3.以描述統計及獨立樣本t檢定分析兩組學生在閱讀摘要學習單上的表現。
- 4.以描述統計來說明自我評量表現情形。
- 5.以皮爾森積差相關來求閱讀摘要實作評量後測與自我評量之相關。
- 6.以次數、百分比來呈現各評分差異的人數分配比例。
- 7.以斯皮爾曼等級相關來分析評分者之一致性信度。

肆、結果分析與討論

本章主要目的在於呈現資料分析的結果，並加以討論，共分為三節。第一節探討國小四年級學童在閱讀摘要實作評量前測上的表現，並且比較實驗組與控制組在後測表現上的差異，以了解實驗處理對摘要表現的影響。第二節探究閱讀摘要教學中，有無自我評量在摘要學習單表現上是否有差異存在，及實驗組學童自我評量的情形。第三節則是進行不同評分者間評分一致性之檢驗，以了解本研究工具之一「閱讀摘要實作評量」的計分，是否會受到不同評分者的主觀判斷而影響得分結果，

一、閱讀摘要實作評量表現之分析與討論

（一）閱讀摘要實作評量之前測表現

本研究目的之一在於探討目前四年級學童的閱讀摘要能力，因此研究者利用受試學童於前測文章1「猴子的記憶」與前測文章2「紅花綠葉變化多」上的表現情形來進行分析與討論。

由表6的描述統計發現，前測文章1「猴子的記憶」的平均數為1.48，前測文章2「紅花綠葉變化多」的平均數為1.30，對照閱讀摘要實作評量之評分規範，可以得知受試學童在2篇前測文章的摘要表現，大都只能摘取到不重要概念及次要概念，而無法察覺或寫出文章的主要概念。

表6 閱讀摘要實作評量之2篇前測文章得分之描述統計

前測文章	平均數	標準差	個數
猴子的記憶	1.48	.91	64
紅花綠葉變化多	1.30	.92	64

從表7的前測文章1「猴子的記憶」得分分配表可以看出，受試學童摘要得分為0分者有11位（17.19%），其摘要顯示無法摘取到相關的概念或誤解文章的本意。摘要得分為1分者有18位（28.13%），顯示他們能摘取到相關但不重要的概念，或能摘取到次要概念，但描述不清楚或部分誤解文章內容。摘要得2分者有28位（43.75%），表示他們能將次要概念描述清楚或僅摘取到部分主要概念。至於得3分者有7位（10.94%），其摘要能完整而清楚的呈現文章主要概念。摘要得分為4分者為0人，表示此篇摘要作業沒有學童能以自己的話語將主、次要概念完整寫出，又能使用文章範例或細節予以支持。

表7 摘要實作前測文章1「猴子的記憶」之次數、百分比分配表

	0分	1分	2分	3分	4分
人數	11	18	28	7	0
百分比(%)	17.19	28.13	43.75	10.94	0

再看表8前測文章2「紅花綠葉變化多」的得分分配表，受試學童摘要得分為0、1、2、3、4分者分別為15位（23.44%）、20位（31.25%）、24位（37.50%）、5位（7.81%）、及0位，與前測文章1的摘要結果頗為雷同，綜合二篇前測文章的摘要表現，可以獲致以下結果，（1）無任何學童的前測摘要表現達到四分等級，（2）學童摘要表現多聚集在得分1和得分2，（3）得0分學生多於得3分者，顯示學生在閱讀完篇章後，多半未能察覺該篇文章的主要概念，甚至不少學生或出現不相關的概念與錯誤概念。顯然在學童的學習階段中，可給予更多閱讀教學與時間來改善學生之閱讀摘要表現。

表8 摘要實作前測文章2「紅花綠葉變化多」之次數、百分比分配表

	0分	1分	2分	3分	4分
人數	15	20	24	5	0
百分比(%)	23.44	31.25	37.50	7.81	0

(二) 兩組學童於閱讀摘要實作評量之後測差異表現

本研究在進行實驗處理時係以班級為單位，隨機分為實驗組與控制組，而非以學生為單位進行隨機分派，為顧及二班學生程度的差異，研究者將以摘要實作前測結果做為共變數，藉由統計控制方式，排除兩班學童既有之摘要能力差異對閱讀摘要實作後測的影響。表9所示為二組學童在閱讀摘要實作評量前、後測的表現，乍看之下兩組學童在閱讀摘要實作評量的後測得分均較前測有所進步，且實驗組表現似乎又優於控制組。為確知閱讀摘要教學是否具有成效，研究者先以成對樣本t檢定，考驗兩組學童在於前、後測表現上是否有顯著差異；結果發現實驗組的後測表現顯著優於前測表現 ($t=-7.23$, $p<.05$)，成績有明顯進步，控制組學童的前、後測成績未有顯著不同 ($t=-1.22$, $p<.05$)，即學童的閱讀摘要後測得分雖較前測得分為高，但卻未顯著優於前測。結果似乎顯示實驗組的教學成效優於控制組的教學成效。

然為進一步確認兩組學生的後測表現是否真有差異，將以共變數分析結果檢驗之。

表9 兩組學童在實作評量前、後測表現之描述統計

	平均數	標準差	人數	t值
實驗組-前測	2.72	1.53	32	-7.23*
實驗組-後測	4.91	1.84	32	
控制組-前測	2.91	1.59	32	-1.22
控制組-後測	3.31	2.15	32	

*達.05顯著水準

※摘要實作前測表現：二篇前測文章的得分加總之平均數。

※摘要實作後測表現：二篇後測文章的得分加總之平均數。

共變數分析前先進行變異數的同質性檢定，經過Levene法的F考驗結果顯示，兩組變異數的差異未達顯著水準 ($F=-1.60$, $p>.05$)，符合變異數同質性的基本假設，因此可進行共變數分析。此外，就「組內迴歸係數同質性」的基本假定而言，亦發現實驗組與控制組的組內迴歸線符合斜率同質的假定 ($F=0.20$, $p>.05$)。由表10可知，將共變項（閱讀摘要實作評量前測）對依變項（閱讀摘要實作評量後測）的影響剔除後，不同自變項（實驗處理）水準間的表現差異達顯著水準 ($F=15.69$, $p<.05$)，表示學童在實作評量後測得分的高低會因其所接受的實驗處理不同而有所差異，參照表9的平均數可知，實施自我評量的實驗組在閱讀摘要的後測表現優於未實施自我評量的控制組。此結果顯示，在進行閱讀摘要教學之餘，若能鼓勵學童透過自我評量進行省思，亦即建立學童檢視自己所寫摘要內容的機制，一段時間之後，將可有效提昇其摘取文章大意、重點的能力。

表10 兩組學童在實作評量後測得分的共變數分析摘要表

變異來源	SS	DF	MS	F值
組間(實驗處理)	46.98	1	46.98	15.69*
組內(誤差)	182.64	61	2.99	

*達.05顯著水準

二、自我評量表現之分析與討論

為瞭解受試學童在整個閱讀教學的自我評量歷程，及其與閱讀摘要表現的關連性，本節將先就兩組學童在各篇閱讀摘要學習單的表現進行比較分析，其次瞭解兩組學童在自我評量檢核表的表現，接著討論學習單表現與自我檢核表的關連性，最後再述及自我評量與閱讀摘要後測表現的關係。

表11 兩組學童在摘要學習單表現之描述統計

組別	平均數	標準差	個數	t值	
文章1	1	1.93	1.05	30	0.52
	2	1.77	1.33	31	
文章2	1	2.35	1.05	31	1.30
	2	2.00	1.08	30	
文章3	1	2.72	0.89	32	1.80
	2	2.23	1.22	30	
文章4	1	2.50	0.88	32	1.53
	2	2.08	1.19	25	
文章5	1	1.88	1.24	32	0.24
	2	2.26	1.34	31	
文章6	1	2.25	1.05	32	0.27
	2	2.17	1.20	29	
文章7	1	1.97	1.03	32	0.25
	2	1.91	0.96	32	
文章8	1	2.31	1.09	32	0.34
	2	2.22	1.13	32	
文章9	1	2.47	1.08	32	1.17
	2	2.13	1.23	31	
文章10	1	2.03	1.25	30	1.01
	2	1.74	1.00	31	
文章11	1	2.65	1.02	31	2.58*
	2	1.97	1.05	31	
文章12	1	2.03	1.12	29	4.00*
	2	0.97	0.91	29	
文章13	1	1.77	0.86	26	3.43*
	2	1.00	0.80	29	
文章14	1	2.63	0.87	32	2.70*
	2	1.93	1.14	30	
文章15	1	2.47	0.78	30	3.47*
	2	1.66	1.04	32	
文章16	1	2.68	0.91	31	5.01*
	2	1.25	1.32	32	
文章17	1	1.93	1.31	30	2.35*
	2	1.21	1.05	29	
文章18	1	2.26	1.06	31	2.25*
	2	1.60	1.22	30	

*達.05顯著水準

※組別1—實驗組 組別2—控制組

(一) 兩組學童在閱讀摘要學習單的表現

表11所示為兩組受試在各篇摘要學習單表現的描述統計及獨立樣本t檢定結果，很明顯地，以第11篇學習單為界，在這之前，實驗組與控制組的表現並無明顯差異，從第11篇起，實驗組與控制組的表現有顯著差異，前者優於後者。表示在閱讀摘要教學後半段，兩組學童的摘要表現開始有所不同，由於前節結果顯示實驗組的閱讀摘要後測表現優於控制組，即自我評量有助於提昇學童的摘要能力，此結果或可進一步說明自我評量的實施須待一段時間方能發揮預期效能，亦即學童需對評分規範內容有某種程度的內化，方能逐漸展現出自我評量對實驗組學童閱讀摘要表現的影響。

(二) 實驗組學童在自我評量檢核表的表現

表12呈現的是實驗組學童在十八次的閱讀摘要學習單表現和自我評量檢核表的相關係數，顯然在實驗的前半段，二者的相關不高，多未達顯著水準，但到了實驗的後半段，二者間的相關多達顯著水準，顯示二者間的關連性隨著實驗時間愈趨密切，惟就實際層面考量，二者間的統計關連性似乎大於實質的關連性。

表12 實驗組學童摘要學習單表現與自我評量之相關

	相關係數	個數		相關係數	個數
文章1	.03	32	文章10	.38*	30
文章2	.22	31	文章11	.35	29
文章3	.35	32	文章12	.46*	27
文章4	.26	30	文章13	.50*	26
文章5	.32	32	文章14	.41*	29
文章6	.45*	30	文章15	.46*	28
文章7	.38*	32	文章16	.08	31
文章8	.32	32	文章17	.62*	29
文章9	.66*	31	文章18	.36	29

(三) 閱讀摘要學習單表現優劣不同者在自我評量檢核表的表現

本研究目的之一在於探討閱讀摘要表現與自我評量的關連性，因此研究者利用閱讀教學時，實驗組學童於閱讀摘要學習單的得分表現及閱讀摘要自我評量表勾選情況進行分析、討論。

研究者將閱讀摘要學習單得分為3、4分者視為摘要表現優者；得分為0、1、2分者視為摘要表現劣者，以此探討摘要表現優劣不同者在閱讀摘要自我評量檢核表勾選「是」、「否」的數量與項目有何不同。結果發現摘要表現優者在自我評量檢核表的主要差異出現在以下三個評量項目的勾選。此群學童在評量項目5皆勾選「是」，但得3分者對於評量項目6和8未能完全勾選「是」，得4分者往往認為自己完全做到。

- (1) 評量項目5：我的摘要包括文章中的主要概念。
- (2) 評量項目6：我能用自己的話寫出摘要。
- (3) 評量項目8：我的摘要包括支持性細節。

至於摘要表現劣的受試者，其在自我評量檢核表的表現，可由以下四個屬於良好摘要基本條

件之評量項目略窺一二，部分學生在這些項目勾選「否」。

- (1)評量項目1：我能將文章從頭到尾看完。
- (2)評量項目2：我能了解文章的意思。
- (3)評量項目3：我會先想過再寫摘要。
- (4)評量項目10：我會將自己的完成的摘要閱讀一遍。

綜合上述，摘要表現優者較劣者更能認為自己摘要內容做到包含文章中的主要概念、用自己的話寫出、以及包括支持性細節。在閱讀摘要學習單上得4分者，除了認為自己不僅能「摘取到主要概念」之外，還能做到3分者所不能做到的「能用自己的話寫出」，即非直接從文章中抄錄下來，以及能「使用文章中的例子」來支持其所提出的主要概念。至於摘要表現劣者當中，常見的情形是部分學童認為自己未能完全具備良好摘要的基本條件。

此外，為瞭解在自我評量檢核表勾選「是」數量多寡者，其在閱讀摘要學習單的表現有無不同，研究者進一步以學童在閱讀摘要自我評量檢核表勾選「是」的數量多寡與在閱讀摘要學習單的得分進行探究。結果發現如下：

- (1)自我評量檢核表勾選「是」0-3個者，其摘要得分多為0、1分。
- (2)自我評量檢核表勾選「是」4-6個者，其摘要得分多為2分。
- (3)自我評量檢核表勾選「是」8-10個者，其摘要得分多為3、4分。
- (4)自我評量檢核表勾選「是」7個者，其摘要得分為2或3分。

由上述結果可知學童在自我評量檢核表勾選「是」數量的多寡與其在摘要學習單得分的高低具有關連性，勾選「是」的數量愈多者，其在閱讀摘要學習單的表現愈好，顯示自我評量檢核表有助於學童檢視其摘取的內容大意及重點。

(四) 閱讀摘要實作評量後測表現與自我評量表表現之相關

為了得知摘要表現是否與自我評量表表現有所相關，研究者在整個實驗教學結束後，於實驗組學童進行閱讀摘要實作評量後測時，每完成一篇後測文章，就給予填寫閱讀摘要自我評量檢核表，二篇後測共計二張自我評量檢核表。將學童在自我評量表上勾選「是」的數目總計起來，並與其在二篇後測文章上的得分，進行皮爾森積差相關考驗；由表13發現後測文章1「牆」的摘要得分與自我評量表的表現相關係數為.53， $p < .05$ ，達顯著相關，而後測文章2「恐龍的傳說」的摘要得分與自我評量表的表現相關係數為.52， $p < .05$ ，亦達顯著相關。此結果進一步顯示實驗組學童在摘要實作後測得分與自我評量勾選情形具有顯著相關，且是正相關，即摘要表現佳者，其在自我評量上，勾選「是」的數量往往多於摘要表現較差的學生。

表13 摘要實作後測表現與自我評量表表現之相關

後測文章1「牆」	.53*
後測文章2「恐龍的傳說」	.52*

*達.05顯著水準

三、評分者一致性之分析

為瞭解本研究工具之一「閱讀摘要實作評量」的計分，是否會受到不同評分者的主觀判斷而影響得分結果，因此本研究以另一位與研究者學習訓練背景相似者擔任第二位評分者，進行不同

評分者間評分一致性之檢驗。此部分分「預試樣本」與「正試施測樣本」兩階段進行。

(一) 預試樣本階段

此階段在於評分規範的練習與不同評分者間評分差異上的溝通，因此兩位評分者利用6名預試樣本來進行閱讀摘要評分規範中各個等級分數的了解；並且抽出彼此間不同給分的摘要表現，針對差異處進行討論，以期不同評分者經過此階段後，因預試樣本的練習而能增進評分上的一致性。

預試樣本階段共進行兩次練習，第一次練習文章為「猴子的記憶」及「龍的傳說」，第二次練習文章為「牆」及「紅花綠葉變化多」。由表14可看出，第一次練習時，兩位評分者在「猴子的記憶」及「龍的傳說」文章上，評分完全一致者有4份摘要，評分相差1分者有2份；不同評分者間評分之等級相關係數分別是0.87（達.05顯著水準）與.70（未達.05顯著水準），顯示二位評分者間的一致性仍有改善空間。第二次練習時，由表14可看出，兩位評分者在「牆」及「紅花綠葉變化多」文章上，評分一致者有5份，相差1分者有1份；其評分相關係數分別是0.91與0.90，皆達.05顯著水準，顯示二位評分者間的一致性已較第一次練習有所提昇，對評分規範的認知漸趨一致。

經此階段之練習，本研究的兩位評分者已熟練閱讀摘要評分規範中各個等級分數之評分內涵，並且在兩次練習的溝通與討論過程中，不同評分者間的等級相關最後達.05顯著相關，表示二位評分者可對正式樣本進行評分。

表14 預試樣本評分者間一致性與相關

兩位評者間評分	猴子的記憶	龍的傳說	牆	紅花綠葉變化多
完全一致	4	4	5	5
相差1分	2	2	1	1
等級相關	.87*	.70	.91*	.90*

*達.05顯著水準

(二) 正式施測樣本階段

透過第一階段預試樣本的練習，評分者於此階段對64位正式施測樣本進行閱讀摘要表現的評分工作。由表15可看出，兩位評分者在前、後測文章摘要表現的評分差異情形，前測部分二位評分者的差異均在一分之內，但後測的「龍的傳說」二位評分者在二份摘要的評分差異超過一分。再由評分者間評分之等級相關程度觀之，兩位評分者在前測文章的評分相關係數分別是0.85與

表15 正式施測樣本評分者間一致性與相關

兩位評者間評分	猴子的記憶	紅花綠葉變化多	牆	龍的傳說
完全一致	48 (75%)	55 (85.94%)	60 (93.75%)	53 (82.81%)
相差1分	16 (25%)	9 (14.06%)	4 (6.25%)	9 (14.06%)
累積總和	64 (100%)	64 (100%)	64 (100%)	62 (96.87%)
等級相關	.85*	.93*	.97*	.93*

*達.05顯著水準

0.93，皆達.05顯著水準；後測文章的評分相關係數分別是0.97與0.93，亦達.05顯著水準，顯示本研究二位評分者對評分規範之認知有一定程度的共識，使得本研究閱讀摘要實作評量具有良好的評分者信度。

伍、結論與建議

本研究旨在探究透過自我評量的介入閱讀摘要教學活動設計，能否有效改善學童的摘要能力，進而提昇其閱讀理解能力。為達此目的，研究者發展出適合國小四年級學童採用的閱讀摘要實作評量及閱讀摘要自我評量檢核表，並進行為期四個月的閱讀摘要教學，以實作評量評估有無自我評量對學童摘要表現之影響。研究者以台南市YY國小64位學童為研究樣本，依據上述研究結果，主要結論歸納如下：

一、研究結論

(一) 國小四年級學童的摘要實作評量前測表現不甚理想

由文獻中所提及四年級是由解碼認字轉變為理解閱讀的新知學習閱讀期，而由研究結果顯示：在文章「猴子的記憶」只有10%受試者，「恐龍的傳說」只有8%的受試者能念清楚而完整的寫出文章主要概念，卻分別有17.19%和23.44%的受試者無法寫出與文章相關的概念，或誤解文章的本意，由此可知在本研究開始進行前，大多數的受試學童在閱讀完一篇文章後，所摘取到的往往是不相關的概念、不重要的概念或文中的次要概念，對於主要概念的覺知極為有限，或是受試學童的閱讀理解能力不佳，或是受試學童的摘要寫作能力不佳，前者因理解能力有限，故無法寫出篇章的主要概念，後者或能理解篇章大意，但不擅於以文字表達其理解內容。總而言之，依據受試學童在前測的表現，國小四年級學童的閱讀摘要能力尚有許多改善空間。

(二) 自我評量對國小學童的摘要實作後測表現有顯著影響

實驗組學生於教學實驗處理結束後，利用研究者自編的自我評量檢核表對後測摘要進行自我檢核，結果顯示有進行自我評量的實驗組學童，其摘要實作後測的表現明顯高於無自我評量的控制組學生；由於兩組學生皆由研究者親自進行閱讀摘要教學，故就教學活動而言，兩組的處理幾乎完全相同，亦無受試者與教學者交互作用的問題，故兩組學生在摘要後測的差異情形，可以有無自我評量活動解釋之。換言之，依據本研究結果，自我評量之實驗處理能有效提升國小學童的閱讀摘要能力。

(三) 自我評量有益於國小學童的摘要學習單表現

為瞭解國小學童在進行自我評量活動時，對其摘要表現有何助益，所以研究者在閱讀摘要教學之餘，讓實驗組學童依據研究者事先設計的閱讀摘要能力自我評量表，對自己的摘要內容進行自我檢視；此外在整個教學實驗處理結束後，讓學生進行自編「閱讀摘要實作評量」之後測時，也給予填寫自我評量表。結果顯示：實驗組學童的摘要學習單得分與自我評量檢核表勾選「是」數量多寡具有一定程度的正相關，實驗中後期以後，二者間的相關係數多達.05顯著水準。透過此自我評量檢核活動，能幫助摘要程度佳者，除清楚完整地摘取出文章主要概念外，更有助於其「用自己的語言寫出」和「使用例子、細節來支持主要概念」；同時也能讓摘要程度差者，察覺到自己在良好摘要的四項基本條件—「能將文章從頭到尾看完」、「能了解文章的意思」、「會先想過再寫摘要」、「會將自己的完成的摘要閱讀一遍」尚未完全做到，藉此提醒學童摘要寫作

的省思重點。

(四) 閱讀摘要實作評量具有良好的評分者信度

兩位評分者在摘要前、測文章評分之完全一致性百分比高達75%以上，且評分者間之等級相關亦在0.85以上，達顯著相關，顯示運用本研究所研擬的摘要評分規範，確實能為閱讀摘要實作評量提供良好的評分者信度，只要評分者能熟悉評分規範的內涵與評分重點，再加上適當的評分者訓練與練習，便可為實作評量提供可信的測量結果。

二、建 議

(一) 對未來研究的建議

1. 班級教師擔任評分者的可行性

本研究為瞭解評分者間的一致性，使用兩位評分者進行摘要的評分工作，結果發現二位評分者間的一致性相當高，此結果顯示，只要能發展出適當的評分規範，事前給予適當的評分者訓練，便能提供可信的評量結果。然本研究的教學者並非班級任課教師，若將實作評量的評分工作交由班級教師擔任時，是否仍能獲致類似的結果，值得進一步探究，若班級教師亦能獲得高的評分者信度，當可提高實作評量在班級的可行性。

2. 另行增加一組對照組以瞭解閱讀摘要的自然成長

本研究實驗設計僅採用一組實驗組（含自我評量）與控制組（無自我評量）進行閱讀摘要教學，然卻無法窺出未進行閱讀教學情況下，學童閱讀摘要能力的自然成長發展，因此建議可另行增加一組對照組，不給予任何額外的閱讀摘要教學，俾便瞭解閱讀摘要教學對學童閱讀摘要能力的有無助益。

3. 檢視自我評量及評分規範的內化程度

本研究的自我評量設計採檢核表方式進行，對於學生判斷自我表現的敏銳度幫助有限，而評分規範的介紹與練習也僅以口頭方式進行，無法得知學生是否真正將評分規範內化；因此建議可與日誌結合，寫下自己平時練習自我評量及評分規範時的感受及個人看法，或與學童進行個別訪談，利用質與量並重方式來深入了解學童是否內化評分規範。

4. 師生可共同擬定實作評量之評分規範

由於本研究之受試者為剛升上四年級的學童，閱讀摘要實作對其而言是陌生的，所以研究者並未與學生共同研擬評分規範，而由研究者自行擬定；若未來研究對象是高年級，或對於摘要實作方式有一定程度的熟悉，建議可讓學生共同參與評分規範的擬定，使其經由與教師的互動討論，達成對優質表現的師生共識。

(二) 在教育應用及教學上的建議

1. 延長單元教學活動時間

由於本研究教學設計中的文章內容討論部分，時間只有十五分鐘，易導致學生無法充分理解文章所要傳達之意涵，因此建議一次完整的教學活動應延長為二節課，讓學生有更多的討論分享時間，促使其有效理解。

2. 文字摘要為主，口述摘要為輔

對於一些語文書寫表達能力較差者，建議先以口述方式輔導其摘要出重點，之後再協助其將重點書寫下來，或部分摘要可改由口述方式進行，俾便學童有更多元化、彈性的方式來呈現其理解閱讀篇章的能力；或者亦可同時採用書寫、口述方式來瞭解學童不同的閱讀理解組

型，惟此方式較為耗時，但卻是深入瞭解學生閱讀摘要能力的方法。

3. 教學文章應包含多種文體，並以多元方式進行教學

本研究挑選的篇章多屬說明文，實際教學可選擇各種不同類型之文章，讓學生感受到閱讀的多變性，經由閱讀可以獲得知識、抒發情感、發揮想像力等等。此外，在進行教學時，除了板書的使用，更需使用圖片、相關資料、視訊媒體來刺激、引起學生學習興趣，亦可透過有聲書方式讓學生體驗文字活化後的感受，換言之，閱讀教學方式不再限於書面的傳播，亦可採用聽覺傳播增加閱讀的趣味，以激發學童對閱讀的喜愛。

4. 討論過程中，多鼓勵學童發言與小組討論

在實驗教學過程中，因學生對於摘要實作與自我評量均感陌生，所以一開始研究者採用的方式是以教師講授為主，並提供書面資料、具體範例協助學生清楚、明白知曉自己的表現如何；之後逐漸改以引導方式，讓學生經由發言與小組討論方式來把握住摘要表現優劣之原則，鼓勵學生能舉出具體事證做為正確判斷其自我表現的依據。

5. 教師應清楚說明實作評量之規範

透過本研究的資料分析與討論，得知學童若能將自我評量與評分規範有效結合，將會提升其摘要實作表現；因此教師在提供實作作業時，應立即、清楚地告知學童評分規準，使其明白表現好壞的依據所在，必要時，教師應與學童一同開發評分規範，並發展出適用之自我評量工具，一旦學生能內化這些評分規範，當能善用規範進行學習的自我省思，因而提昇其表現。以國小四年級學童而言，評分規範愈具體愈好，初次使用的教師，最好能舉實徵範例讓學生瞭解各評分規準的內涵。

陸、參考文獻

一、中文部分

- 小光（民64）。國語段落大意的探討。中國語文，37卷6期，頁17-21。
- 王萬清（民86）。國語科教學理論與實際。台北：師大書苑。
- 方金雅、鍾易達（民87）。國小學童閱讀摘要能力評定規範之發展。載於臺南師範學院測驗發展中心主編國小教學評量的反省與前瞻，頁123-137。
- 何容主編（民63）。國語日報辭典。台北：國語日報社。
- 李詠吟（民82）。學習輔導。台北：心理。
- 李坤崇（民88）。多元評量。台北：心理。
- 吳英長（民87）。國民小學國語故事體課文摘寫大意的教學過程之分析。台東師院學報，第九期，頁149-184。
- 宋文菊（民88）。國小學童在閱讀理解實作評量上的表現分析。國立台南師範學院國民教育研究所碩士論文。未出版。
- 余民寧（民91）。教育測驗與評量。台北：心理。
- 林重新（民90）。心理與教育測驗。台北：揚智文化。
- 洪碧霞、邱上真、葉千綺、林素微（民89）。國小學童國語文成長組型之探討。測驗年刊，第四十七輯，第一期，頁1-26。
- 柯啟瑤（民90）。自我評量和交互評量的意義與功能。翰林文教雜誌，18期，頁9-13。
- 陳如山（2000）。淺談自我評量。教育研究，31期，頁56-60。
- 教育部（民91）。國民中小學九年一貫課程綱要。台北：教育部。
- 教育部（民89）。國民中小學九年一貫課程暫行綱要。台北：教育部。

- 教育部（民87）。國民教育階段九年一貫課程綱要。台北：教育部。
- 曹郁玲、黃淑娟（民89）。三年級學生閱讀理解與解題溝通能力的評量設計。載於國立臺南師院主編新世紀優質學習的經營研討會論文集。
- 彭森明（民85）。實作評量理論與實際。教育資料與研究，9期，頁44-48。
- 楊韻平（民82）。兒童摘取文章大意的能力。政治大學教育研究所碩士論文。
- 鄒慧英（民87）。實作評量的研發－以國小說話課為例。測驗與輔導，149期，頁3082-3087。
- 張嘉真（民89）。書蟲讀書會。台北：富春文化。
- 曾雅瑛（民87）。國民小學中文詞彙閱讀測驗之編製。國立台南師範學院國民教育研究所碩士論文。未出版。
- 蕭雅萍（民91）。評分準範與國小六年級學童英語口語作表現及自我評量之研究。國立台南師範學院國民教育研究所碩士論文。未出版。
- Costa, A. L., & Kallick, B. (2000)。李弘善譯（民90）。評量和記錄心智習性。台北市：遠流。
- Martin-Kniep, G. O. (2000)。陳佩正譯（民91）。成為更好的教師。台北市：遠流。
- Robert L. & Gronlund, N. E. (1995)。鄒慧英譯（2003）。測驗與評量。台北：洪葉。

二、英文部分

- Airasian, P. W. (1997) . *Classroom assessment*. New York: McGraw-Hill Companies, Inc.
- Boud, D. (ed) . (1988) . *Developing student autonomy in learning*. New York: Kogan Page, London/Nichols Publishing.
- Brown, A. L., Day, J. D., & Jones, R. D. (1983) . The development of plans for summarizing texts. *Child Development*, 54 (4) , 968-979.
- Brown, A. L., & Day, J. D. (1983) . Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, 22 (1) , 1-14.
- Brown, A. L., & Smiley, S. S. (1977) . Rating the importance of structure units of prose passage: A problem of metacognitive development. *Child Development*, 48 (1) , 1-8.
- Cariaga-Lo, L. D., Richards, B. F, & Frye, A.W. (1992) . *Understanding learning and performance in context: A proposed model of self-assessment*. (ERIC Document Reproduction Service No. ED 347 979)
- Ellis, E. S. et al., (1991) . An instructional model for teaching learning strategies. Focus on *Exceptional Children*, 23 (6) , 1-24.
- Hare, V. C., & Bingham, A. B. (1996) . Teaching students main idea comprehension: Alternatives to repeated exposures. In J. F. Baumann (Ed.) , *Teaching main idea comprehension* (PP.179-194) . Newark, DE: International Reading Association.
- Hare, V. C., Rabinowitz, M., & Schieble, K. M. (1989) . Text effects on main idea comprehension. *Reading Research Quarterly*, 24 (1) , 72-88.
- Kintsch, W., & Van Dijk, T. A. (1978) . Toward a model of text comprehension and production. *Psychological Review*, 85 (5) , 363-394.
- Marzano, Pickering, & JcTighe (1993) . *Assessing student outcomes: Performance assessment using the dimensions of learning model*. Alexandria, VA: Association for Supervision and Curriculum Development (ASCD) . 1-36.
- Orsmond, P., Merry, S., & Reiling, K. (2000) . The use of student derived marking criteria in peer

- and self-assessment. *Assessment & Evaluation in Higher Education*, 25 (1) , 23-28.
- Opitz, M., & Glazer, S. M. (1995) . Self-assessment and learning centers: Do they go together? *Teaching Pre K-8*, 25 (4) , 104-106.
- Paris, S. G., & Ayres, L. R. (1994) . *Becoming reflective students and teachers with portfolios and authentic assessment*. Washington: American Psychological Association.
- Pearson, P. D., & Dole, J. A. (1987) . Explicit comprehension instruction: A review of research and a new conceptualization of instruction. *Elementary School Journal*, 88 (2) , 151-165.
- Pearson, P. D., & Johnson, D. D. (1978) . *Teaching reading comprehension*. New York: Holt, Rinehart & Winston.
- Rogers, C. R. (1983) . *Freedom to learn in the 80s'* Columbus, Ohio: Charles E Merrill.
- Smith, C. (1997) . Student self-assessment at St Bernard's primary school. *Primary Educator*, 3 (4) , 7-9.
- Stiggins, R. J. (1987) . Design and development of performance assessment. *Educational Measurement: Issues and Practice*, 6 (3) , 33-42.
- Towler, L., & Broadfoot, P. (1992) . Self-assessment on the primary school. *Educational Review*, 44 (2) , 137-144.
- Wiggins, G. (1998) . *Educative assessment*. San Francisco, CA: Jossey-Bass.
- Williams, J. P., Taylor, M. B., & deCani, J. S. (1984) . Constructing macrostructure for expository text. *Journal of Educational psychology* 76 (6) , 1065-1075.
- Williams, J. P., Taylor, M. B., & Ganger, S. (1981) . Text variations at the level of the individual sentence and the comprehension of simple expository paragraphs. *Journal of Educational Psychology*, 73 (6) , 851-865.
- Winograd, P. N. (1984) . Strategic difficulties in summarizing texts. *Reading Research Quarterly*, 19 (4) , 404-425.

投稿日期：93年11月9日

修正日期：94年2月2日

接受日期：94年2月10日

The Research of Fourth Graders on Reading-Summarizing Performance Assessment and Self-Assessment

P. Z. Chang

Tainan Municipal Wunyuan Elementary School

H. I. Tzou

Graduate Institute of Measurement and Statistics, National University of Tainan

Abstract

The purpose of this study is to integrate self-assessment into reading-summarizing instruction to see if it improves fourth grades' performance on reading-summarizing assessment.

The main results are as follows :

- 1.The 4th-grades in primary schools performed not so well on the pre-test of summarizing assessment.
- 2.The self-assessment had significant effects on 4th-grades' performance on the summarizing assessment.
- 3.The self-assessment was useful to improve students ' summarizing performance.
- 4.An appropriate reliability estimate for the reading-summarizing performance assessment was derived using interrater consistency.

Key words : reading-summarizing, self-assessment, performance assessment, interrater reliability