

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89，19期，279—308頁

資優兒童問題解決能力* 實作評量之建構研究

呂金燮

國立臺北師範學院

本研究的主要目的在配合國小資優兒童鑑定的觀察期，建構以語文及數學領域為主的問題解決能力實作評量活動，觀察兒童解題歷程的認知複雜度類型。再以絕對和專家模式的方式，依據兒童的實際反應建立明確具體的觀察指標與評分系統。評量活動依結構度的特性分為結構清楚、弱結構與無結構等三類。研究對象為臺北市興隆國小以及臺北縣秀朗、中和和埔墘國小等四所設有資源式資優班學校的二年級與三年級的普通與資優兒童，共五十七位。主要研究結果有：(一)評量活動難度偏高，但能區辨解題能力之質的差異；(二)兒童於解題過程中表現出解題歷程的多元性，並呈現專家與生手的特質差異；(三)表現指標的層次分明，利於明確評分表的建立；(四)基於領域與專家生手的特質，評量活動的結構度對表現結果的影響不一。對研究結果的應用建議：(一)資優兒童的思考特質適合實作評量，但須有專業訓練的教師配合觀察與評分，方能發揮鑑定的功能；(二)當題目難度偏高時，兒童對題目情緒反應與堅持度成為解題的另一個關鍵；(三)以實作評量表現結果作為資優兒童鑑定依據，仍有待題庫的建立，方能作為大量施測的鑑定依據。

關鍵詞：資優兒童、問題解決能力、實作評量、認知複雜度

隨著認知心理學與認知科學的發展，學者對智力的看法有其新的詮釋，過去以兒童在智力發展階段的表現作為智力指標，而今學者所強調的是，兒童在未接受直接教學情況下，其學習與解決問題的能力 (Resnick & Glaser,

1980; Sternberg, 1982)。這些學者特別強調問題解決能力在資優兒童的教育與鑑定上的角色，也一再呼籲不應再以標準的智力或成就測驗來鑑定或評量資優兒童的學習能力，更提出以多元或實作評量的方式來符合認知心理學

*本研究獲國科會補助 (NSC 88-2614-H-152-001)，研究助理林妮芙、陳友琦和黃楨芬小姐，以及興隆、秀朗、中和和埔墘國小等四所學校十位資優班教師之協助，在此一併致謝；並感謝二位審查者提供寶貴的修改意見，使得研究結果的呈現更有意義。

上的研究發現 (Gardner, 1999; Sternberg, 1982; Wiggins, 1989)。目前國內資優教育的鑑定方式也力求以多元的指標來評量學生，如臺北市國小資優兒童的鑑定，在標準成就測驗與智力測驗之後，實施八週的教學活動觀察期，期能以更實作多元的資料評估學生的學習能力。但源於觀察期的活動缺乏系統化，評分標準不明確，目前觀察期的資料僅供參考，而較少實質的鑑定功能 (陳美芳, 民85)。費時費力的觀察期評量資料因缺乏系統與客觀的分析，而無法發揮其應有的功能，實屬可惜。有鑑於此，本研究乃配合國小資優兒童鑑定的觀察期，設計以國語及數學領域為主的問題解決能力實作評量活動，觀察兒童解題歷程的認知複雜度，並依據兒童的實際反應建立明確具體的評分標準，希冀提供給國小資優兒童鑑定之觀察期使用。

問題解決能力評量在資優教育上的重要性

以下將先從問題解決能力與智力的關係，再談鑑定過程中觀察期的落實、適用多元教育的多元評量需求以及鑑定與教學的聯結等，來討論問題解決能力評量在資優教育上的重要性。

一、問題解決能力與智力的關係

我們如何鑑定個人是否具有特殊才能或是資賦優異深受智力理論發展研究的影響，過去二十幾年的認知心理學研究，對特定領域的資優或特殊才能的表現優異有了不同的發現，對領域專長與智力發展的了解，逐漸改變我們對資優生的定義及其鑑定與評量的方式。智力理論從傳統的 IQ 理論到現今多元的理論，一直都以『問題解決能力』為中心概念，茲將其概念發展分為四個階段來討論：(1)認為智力是整合資料與經驗或做抽象思考的能力 (Spearman,

1904; Terman, 1925)。因此標準智力測驗如：比西和魏氏智力量表，多以抽象思考能力為評量的指標，測驗結果則以 IQ 的高低為依據；(2)認為智力是“調適”或“對環境的適應與調整的問題解決能力” (Piaget, 1976)；(3)問題解決能力是對文化及生理環境的調適 (Neisser, 1976; Gardner, 1983)。強調文化環境對兒童思考的影響，外在的刺激內化過程的能力，會因環境而有所差異，因而強調動態評量，或情境化的解題能力，而非去脈絡的抽象符號思考能力。Gardner (1983) 更提出多重智力理論，強調智力的多元化，且有學科領域的特定性，非一般性的思考能力；(4)以訊息處理理論為出發點的學者，則強調問題解決能力的行動力 (Resnick & Nelson-Le Gall, 1996; Sternberg, 1982)，其中 Resnick 和 Nelson-Le Gall 提出社會化的智力，強調在社會環境中實作與操作的重要性，Sternberg (1981) 提出智力三元學說強調除了成份智力 (componential)，更強調經驗 (experiential) 與情境 (contextual) 的重要性，這類學者也將問題解決和後設認知劃上了等號 (Glaser, 1991; Yussen, 1985)。

從抽象思考以 IQ 為指標到強調應用先備的經驗與知識到新的情境，並能監控自己學習的重要性。智力的概念不再是對抽象思考能力的評量，而是以解決問題的心智思考內容和過程等多元性，來區辨出資優與一般學生的內在差異 (Sternberg, 1981)。

二、鑑定程序的落實

為落實資優特質的多元性指標，臺北市各國民小學的資優兒童鑑定，從民國七十九年起即在團體智力測驗和個別智力測驗之後，實施了八週的觀察期，目的在評量兒童於實際情境下創新與解決問題能力的多元表現，以作為正式入班與教學之依據。臺北縣的資源式資優班學校也覺得有此必要，有意效法，但卻躊躇不前。因為目前實施的觀察期方式缺乏客觀的或

統整的資訊，評量目標不明確，各校活動不一，評分標準不明，僅用來作參考，而無實際鑑定效用。八週觀察期的立意在多元實作評量的基礎，卻因缺乏有系統的能力分析而無法彰顯其效果，是人力與資源的損失。

三、多元教育下多元化評量的需求

目前國小資優兒童的鑑定分為兩個階段，第一階段為團體智力測驗，第二階段為個別智力測驗。第一階段的團體測驗之一為成就測驗，一般學術性向資優兒童的成就測驗以國語及數學領域為主，以往這兩科的成就測驗都以國編版的教科書內容作為命題基礎，然而，近幾年來，開放教育與多元智慧教學等另類的教育方式，在教育人士有心推動下蓬勃發展，成為目前教育的潮流趨勢。教育朝多元化的目標發展，制式的教學內容與方式漸不合時宜，評量方式的多元化也成為發展的目標。再者，自民國八十五學年度起，教育部開放教科書的編寫，由國編版改為審定版，使得國民小學各科的教學更多元，教材也更具有地域特色。現有的成就測驗以舊有國編版教材為基礎，已不再適用目前教育開放的環境。學科領域的高層次問題解決能力的評量更形重要了。

四、鑑定與教學的銜接

資優教育的目的在培養學生問題解決能力 (Renzulli, 1977; Treffinger, 1986; Sternberg, 1997)，目前國內資優班的教學強調問題解決能力培養的課程（如主題本位、未來學課程及創造性問題解決課程等），朝向以統整各學科內容為主的教學方式，重視兒童的獨立思考、創造力與問題解決能力的培養。然而，以團體智力測驗、個別智力測驗及成就測驗等標準測驗工具的鑑定方式並無法提供學生在問題解決能力方面的潛能表現，再者，許多高智商的學生，並不表示其具有問題解決潛能，因此入班的部份高智商的學生，卻往往無法適應問題解決或創造力的教學方式 (Vantassel-Baska, 1994;

Sternberg, Ferrari, Clinkenbeard & Grigorenko, 1996)。問題解決能力是資優發展的目標，因此所需要的鑑定工具亦應能反應其問題解決潛能。能了解兒童真正的思考，方能掌握其真實的學習能力。以免造成鑑定與教學的無法銜接。

由以上幾點可見問題解決能力評量需求的迫切性，但是評量或鑑定工具的建立並非易事，本研究僅從認知複雜度的角度，探討資優兒童問題解決能力的特質，建構實作評量活動與評分表，希望提供資優教育的研究者與教學者另一種選擇。

問題解決能力的關鍵要素 與評量設計

一般而言，問題解決是指當沒有明顯解決方法時，解題者達成目標的認知歷程 (Mayer & Wittrock, 1996)。問題解決能力指的是高層次或複雜的認知活動，在解題過程中，解題者的思考過程會有某些程度的不確定與搜尋，需要應用到之前不同類型的學習結果與策略，而且解題之後會有新的學習 (Gagne & Medsker, 1996; Schoenfeld, 1989)。以下乃從問題解決能力的關鍵要素及其評量要件討論。

一、問題解決能力的關鍵要素

從對專家與生手的分析研究中發現，對優秀的問題解決者而言，領域特定的知識、認知複雜度以及動機和努力是影響解題表現的關鍵因素。

(一) 領域特定的知識

豐富的領域特定的知識 (domain-specific knowledge) 是首要影響其問題解決表現的重要因素 (Chi, Glaser & Farr, 1988)。領域特定的知識也就是對該領域相關的經驗與知識，對領域知識內容了解的豐富或淺薄，影響解題者對題目的了解及知識的應用 (Glaser & Baxter,

1997)。

優秀的問題解決專家在領域知識上有兩點主要特質：(1)選擇性的優秀：專家主要在自己的領域非常傑出 (Chi, 1987)，專家並非樣樣精通，而是選擇性的優秀；(2)有組織的知識架構與經驗：專家在短期記憶和長期記憶上都十分優秀，主要來自有組織的領域知識與經驗結構，也就是所謂的『基模』 (Chase & Ericsson, 1982)。

(二)問題解決認知複雜度

問題解決能力不在強調結果的對與錯，而更重要的是其解題結果背後複雜的認知思考歷程。問題解決的認知複雜度 (cognitive complexity) (Glaser, 1991)，又稱為認知歷程的多元性 (multiplicity) (Spiro, Vispoel, Schmitz, Samarapungavan & Boerger, 1987)，指的是解題過程中思考歷程多面向 (multi-facet) 的認知歷程以及其歷程的多層次 (multi-level) 深度。問題解決的過程中，若以認知歷程出現的時間先後為順序，首先就是了解問題的核心，再決定解決方式，然後執行。問題解決認知歷程的研究中，Mayer和Witrock (1996) 的研究發現問題的表徵、計畫和執行為研究問題解決的重要認知歷程向度，而Sternberg和Davidson (1992) 則以界定問題、形成內在表徵、編碼、計畫、策略選擇、結果監控為主要向度；Glaser (1991) 則強調表徵問題和尋求解決方法的認知歷程對問題解決品質的重要性。

專家的研究中發現，優秀的解題者在這些認知歷程的向度上有幾點主要特質：(1)深層的問題表徵：專家看他們領域中的問題深入，且以深層的訊息表徵問題，可以一針見血的找出解題的關鍵；生手傾向於以表面的訊息表徵問題，忽略了解題的重要關鍵，而在不相關的訊息上花時間 (Chi, Feltovich, & Glaser, 1981)；(2)對問題做質的分析：專家會花大量時間對問題做質的分析之後，再進行解題；相對於專家

的三思而後行，生手往往看到一些可行的方式就貿然行動，未能考量問題的整體性 (Voss & Post, 1988)；(3)解題快錯誤少：專家解題快速錯誤少 (Chase, 1983)；(4)強烈的自我監控能力：專家在解題時有強烈的自我監控能力，常有檢查、反思與自我調整速度的行為，生手則往往缺乏這些察覺能力 (Chi, 1987; Glaser, 1991)。

就資優生問題解決認知複雜度的研究發現，資優生在問題解決上比普通學生(1)有較大的彈性及精確性；(2)使用較少的時間；(3)較多不同的策略；(4)花較多的時間在思考理解問題與計畫目標；(5)同時也具有較高的後設認知能力；(6)能自發並成功的將教導的策略應用在不同的領域上 (Dover & Shore, 1991; Rowe, 1985; Swanson, 1992)。

(三)動機與努力

影響問題解決結果的表現，除了知識內容和複雜的思考認知歷程等要素外，仍有賴個人的動機和積極的態度，以及努力的程度 (Schoenfeld, 1989)。個人對所面對的問題的動機，對成就的影響不亞於能力與知識的多寡，甚至是影響問題解決結果的主要因素之一 (Deweck, 1989)。

二、問題解決能力實作評量的要件

要能明確清楚的區分出學生在問題解決能力上的差異，並非任何的問題都適合，過於簡單或難度過高都無法有效的區分學生的能力，因此能夠讓學生充分發揮其能力的問題難度與情境就成為問題解決研究中重要的條件。近來研究智力與資優的學者，從文化與情境認知的角度，批評標準化測驗的抽象思考，無法代表學生真正的能力，提出問題解決必須在實際應用的情境中評量 (Sternberg & Lubart, 1995; Wiggins, 1989)。了解兒童思考不再是單純的一般抽象思考能力，資優兒童的鑑定與評量，也應以新的且更多元實作 (performance assessment) 或卷宗評量 (portfolio assessment) 的角

度來進行（陳美芳，民85；郭靜姿，民85；Gardner, 1983）。綜合學者的論點，問題解決能力實作評量情境佈題上，應具以下條件：

（一）領域知識的應用與經驗判斷兼之：實作評量或卷宗評量等多元評量方式，在問題解決能力的內容設計上須兼顧領域或學科知識應用（formal knowledge）以及非知識方面的經驗判斷（informal knowledge）（Sternberg & Lubart, 1995）。知識是智力過程所運用的原始材料，指的當然就是領域中的理論概念，是可以教導的。知識的基礎使個人能注意並利用隨即而來的機會，以達成功的表現；非知識性的經驗判斷則是在一個領域或工作中所花的時間中，而獲得對該領域的常識基礎，亦即經驗的累積，這種經驗的累積需要興趣與毅力的累積方能獲得。

（二）弱結構（ill-structured）的情境：評量問題的結構度影響解題時的思考複雜度與層次，而問題的結構度都以解決問題必要條件的提供與否、方式的多元性、以及解決之結果的已知程度等為決定結構層次的依據，而這些要件的已知程度與解題者的能力有密切的相對關係（Getzels & Csikszentimihalyi, 1967; Yussen, 1985）。問題解決的結構度層次可視為一條連續性線上的許多點，從題目呈現的敘述方式、已知條件的多寡、可接受的方法或詮釋的多寡、以及問題、方法及結果的彈性來決定在線上的位置（也就是層次）。線的一端是最有結構（well-structured）、定義最清楚的層次，這個層次的問題敘述清楚，解題條件充分，問題解決的方法很明顯，出題者與解題者都知道，解題的結果出題者知道，解題者須經由一些已知的知識或方法中去發現，是解題者最熟悉的問題情境。而線的另一端則是最無結構（un-structured）、定義最模糊的問題層次，這個層次的問題中，問題定義不清（問題本身也許還含有許多不清的問題），解決問題的方法未

知，解題條件十分不足，有許多不同的方式，也許需要嘗試新的有創意的方式來解題，因此，出題者和解題者都不知道需要用哪個方式或哪個方式最恰當，解題結果出題者和解題者也無法預知，是解題者最陌生的問題情境。在這兩端之間則是其他層次的弱結構問題。

為了研究設計及分析上的限制，一般都將問題分為三或五個層次。結構愈清楚，條件愈充分，所需的高層次思考愈少；反之，結構愈模糊不清，問題情境愈新奇，條件愈少，難度愈高，也就愈需要高層次的計畫與策略監控等的思考活動（林妮芙，民88）。評量中若使用學生熟悉或結構定義清楚的情境，測到的能力是近遷移的能力，也就是成就測驗；若是新奇或無結構的情境，方能測到遠遷移的能力，也就是性向和能力（Lohman, 1993; Yussen, 1985）；雖然學者一再強調真正的問題解決能力要在弱結構或新奇的情境下才會發揮，然而問題的結構度與解題者的能力是屬相對性的。因此，多層次結構的評量設計在了解不同解題者的思考歷程仍是必需的（Resnick & Glaser, 1980; Spiro, Vispoel, Schmitz, Samarapungavan & Boerger, 1987）。

綜合上述，智力理論的發展一直強調應將其視為資優兒童鑑定的指標之一，問題解決能力在資優教育上的重要性是不置可否的。但是目前國內的資優兒童鑑定與評量仍以傳統的IQ及成就的百分等級為主要依據，偏重在抽象思考能力的測驗，在解釋測驗結果時，與教學配合的空間有限；另外，一般教師自編的學門領域的情境特定性觀察表或評量表又缺乏系統，難以客觀的評估學生的問題解決能力。如能針對國內學生問題解決思考歷程進行深入分析探討，必能有效編擬資優班兒童問題解決能力的鑑定評量與課程或教學策略。

基於上述的分析，本研究的主要目的在建構一個適用於國小二年級下學期資優兒童鑑定

觀察期使用的實作評量活動，不以傳統的成就測驗只重視結果的對與錯，更強調解題思考歷程的多向度與層次的分析與描述，因此研究的焦點不在增加評量的預測效度（predictive validity），而在思考歷程的分析與構念的效度（construct validity）。

本研究的主要目的有：

1. 發展可以推論兒童對領域知識了解層次的問題解決能力評量，而非單指測驗知識的『有』或『無』。
2. 發展可以讓兒童表現其思考歷程認知複雜度的評量活動。
3. 確立“優秀”表現的標準。
4. 發展評量學科領域問題解決能力層次的順序與評分系統。
5. 發展適用於觀察二年級兒童問題解決能力評量的觀察系統。

研究方法

本研究為配合國小兒童的學習經驗，乃以語文和數學為研究的學科領域。以不同結構層次的活動內容，觀察兒童在解題過程中的認知歷程和表現結果，以建構系統化的觀察向度與評分表。其中最大的挑戰就是表現的指標與評分表的建構。一般表現的指標常以：(1) 絕對的方式—課程標準或既有文獻中的相關標準來決定；(2) 專家的訪談—訪談該領域的專家，歸納其意見；(3) 領域分析—以概念或內容分析該領域的重點；(4) 專家模式—找出專家或學生在該項上的表現為標準；(5) 統計量表技巧—多向度量表（multidimensional scale），以及串連分析（cluster analysis）等方式來確立。本研究採以絕對的方式從文獻理論中歸納出該學科領域的重要知識或過程策略的代表性表現，建構一般的評分表（general scoring rubric）；為了使得評分的指標具實用及接近的特質，乃再以專家

的模式從兒童的實際表現中找出代表的表現特質，作為評分指標的具體例證，以建構明確的具體評分表（specific scoring rubric）。為了建構多層次的評分系統，研究的對象涵括二年級和三年級資優以及普通兒童。三年級資優兒童對評量活動的形式十分熟悉，且練習的機會最多，本次研究中相對於二年級普通兒童而言，將之視為專家的解題者，其次為三年級普通兒童和二年級資優兒童，而二年級的普通兒童接觸這類的題目機會最少，練習機會也最少，視為生手的層次。

一、理論架構

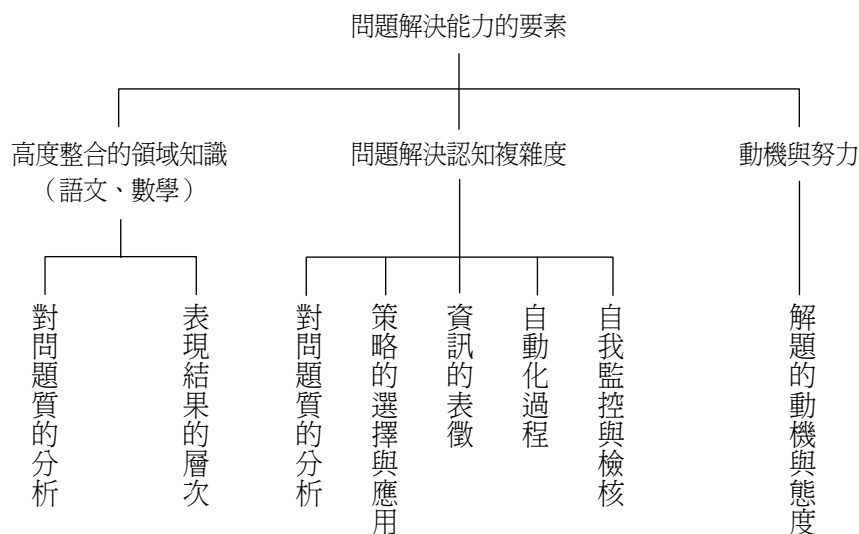
研究架構主要依據上述問題解決能力的三個關鍵要素，如圖一所示，以高度整合的領域知識、問題解決認知複雜度與動機和努力為三個主要向度。

(一) 高度整合的領域知識

雖然本研究評量工具的主要對象是國小二年級的兒童，但源於資優兒童對學科領域知識的理解高於一般兒童的考量（如兩個標準差以上），數學與語文的知識概念乃加深到三年級課程標準的深度與廣度。語文領域的能力強調對句子結構的知識。接寫故事，以複雜與合適的語言傳達想法與感覺等。數學能力則強調能依據特定的標準操弄具體事物的能力，了解因素間的因果關係，表現出標準數學運算與計算的能力。問題解決中連結不同數學運算的能力，以及從事歸納與演繹思考歷程的能力等。整合課程標準的能力指標與多元智力的評量標的（Lazear, 1995）的評量目標請詳見表二和表三。

(二) 問題解決的認知複雜度

問題解決認知歷程的向度則以對問題質的分析（界定問題）、資訊表徵、策略形成、自動化的過程、以及監控和結果檢核等。『界定問題』或『對問題質的分析』指的是了解問題的要求，知道問題的重點；『資訊的表徵』則



圖一 問題解決能力評量的要素

為解決問題必須將資訊以內在或外在的方式表徵，重新組合這些資訊間的關係或意義；『自動化的過程』則是如時間的分配等；『結果的監控與檢核』指的是隨時了解自己的進度，哪些未解，哪些需要修正等，也就是一般所說的後設認知，對結果檢核後將回饋轉換成行動計畫，而使解題更正確有效。這些認知歷程的層次則由研究對象的思考歷程中整合。

(三)動機與努力

基於人力時間的因素，此向度將不在此研究中討論，是為後續研究的重點。

二、評量活動

為配合國小資優兒童鑑定的觀察期使

用，乃以目前臺北市國小資優班觀察期的活動內容為主要類型，評量活動類型的選擇與改編原則，主要在能顯示兒童的解題策略和多樣的思考為參考。為了詳細分析了解不同解題者的思考歷程，以二年級兒童的能力為依據將問題依其結構度的特性分為結構定義清楚、弱結構和無結構等三類型活動。三類結構度之間的差異就在問題敘述、解題條件、解決方法和解題結果的已知程度和情境化的程度，詳細分類如表一所列。

語文的領域雖以聽、說、讀、寫為主要分野，但是基於語文的應用性，評量的內容大致都是四種能力的交互應用。閱讀的部份從圖、短詩到散文皆有，而寫的部份則從換

表一 評量活動的問題結構度分析

結構度	特質				
	問題標準敘述	解題條件	解決策略與方法	解題結果	情境化
結構定義清楚	清楚	充分	已知單一	已知	無
弱結構	不清楚	不足	已知多元	已知	模擬情境
無結構	含有更多問題	十分不足	未知多元	未知	生活真實

句話說、選詞造句、童詩仿作、實物的描述、到真實活動的敘寫等，而聽與說則配合讀與寫的活動進行。結構度清楚的題目，如童詩填寫，題目中清楚列出題目的要求標準，如這首詩有個規矩，就是每一句話的最後一個字都要有“一”的注音，如地（ㄉㄨㄛˋ），力（ㄌㄧˋ），目的是要兒童在此條件之下進行解題。而弱結構的題目，如古詩三則的解釋和看圖寫

故事，看圖寫故事的脈絡來自一連串的圖像，但兒童可發揮的則更多了，圖和古詩的文字給了暗示，但並不明示，端賴兒童的想像力與理解力，解題的向度就會有所不同。而沒有結構的如：小小說和實物體驗的描述，則由兒童對題目的理解，自定問題的向度，解題的方向更多元，答案更是設計者無法預知的。各評量活動與能力目標詳如表二。

表二 資優兒童語文問題解決能力評量內容與目標分析

評量活動	散文 龍龍 (閱讀與接寫)	新詩 樂隊、一寸高 (填詞)	古詩 山中、對花 白頭一老子 (解釋文意)	看圖寫故事 犀牛 (故事創作)	小小說 小狗 (故事分析與重 編)	影片描述 藍色星球 (影片內容描述)	活動報導 障礙路線 (活動敘述)
評量目標							
• 擴展字彙，以新的詞彙或新的用法應用在說與寫上	✓	✓			✓		
• 能了解用各種不同的語文方式呈現的幽默及含義	✓	✓	✓	✓	✓	✓	
• 理解閱讀的資料並以說或寫的方式表達對閱讀資料的看法	✓		✓			✓	
• 應用複雜且適當的語言溝通想法和需求及感覺	✓	✓	✓	✓	✓	✓	✓
• 以不同正式口語表達型式（辯論、論說、演講等）表達想法					✓		
• 能以已有的故事為架構，創新的故事	✓	✓	✓				
• 以不同的寫作形式來自我表達或溝通想法、意見、感覺等		✓	✓				
• 對活動或解決問題的過程，寫或報告活動內容						✓	✓

註：（）中為語文形式

數學領域的內容一般分為數、量與形，而過程推理能力則重在加減的概念推算乘除的運算以及空間形的推理，應用的範疇則以兒童熟悉的金錢、空間的測量、時間的計算、統計圖表以及情境的推理等。評量活動的情境從超級市場的空間規劃、彩繪玻璃的四則推理思考、撈魚比賽的統計記錄的應用等。有生活中的情境也有教室中常用的應用題的設計。超級市場的活動一和三，以及算算看的活動二為結構定

義清楚的題目，解題條件充分且問題的呈現十分直接，可從題目中找到解題線索，而彩繪玻璃、想一想和走路、拼圖等題目，則因解題條件不充分，且未提供解題策略，兒童的解題難度相對提高，而撈魚比賽和找規則則完全依賴兒童的組織分析能力，方能決定題目的要求而順利解題，解題結果多元且未知。各評量活動與能力目標詳如表三。

表三 資優兒童數學問題解決能力評量內容與目標分析

	超級市場		彩繪玻璃		撈魚比賽		想一想			算算看				邏輯推理			
	活動一	畫位置	禮盒	2色	3色	活動一	活動二	盒子	積木	七巧板	一	二	找規則	貼貼紙	擺豆子	走路	拼圖
1.計算或測量											✓	✓			✓		✓
2.數學運算過程或演算			1、2														
3.在複雜的問題中應用各種數學運算	5		3			✓								✓			
4.掌握不同問題解決技巧，及可能的方法	2、3、4	✓		✓	✓	✓				✓			✓		✓		✓
5.根據概念化的資訊從事抽象思考		✓			✓		✓	✓	✓	✓				✓			✓
6.數學型態或等式	3、4	✓											✓				✓
7.用圖表或其他等號表達		✓		✓		✓											
8.地圖、藍圖、流程圖的閱讀與轉換	1、2、3							✓	✓	✓				✓			
9.歸納或演繹的思考		✓	4		✓	✓	✓							✓		✓	✓
數學領域知識架構 (domain knowledge)																	
領域	數量		✓			✓	✓				✓	✓	✓		✓	✓	✓
	形	✓	✓		✓	✓		✓	✓	✓				✓			
過程	加或減										✓			✓	✓		✓
	乘或除											✓					
	四則混合			✓		✓											
	推理	✓	✓		✓	✓	✓	✓	✓	✓			✓				✓
應用範疇	金錢			✓													
	測量													✓			
	時間										✓						
	統計圖表	✓				✓	✓	✓	✓	✓							
	情境推理		✓		✓	✓							✓		✓	✓	✓

三、研究對象

研究的對象為民國八十八年臺北市興隆國小、臺北縣秀朗、中和與埔墘國小等四所學校中，二年級及三年級的資優班與普通班兒童。這四所國小的資優班都屬分散式的實施方式。因為評量活動的使用對象為資優生甄選時經過教師推薦後的兒童，因此研究對象中，普通兒童由任課教師推薦在語文與數學領域屬中等以上程度的兒童，而資優兒童則由教師抽選。這些兒童經家長簽名同意方參與研究。參與語文

評量的二年級兒童有十九位，資優生十位，普通生九位，其中男生九位，女生十位。而三年級有十二位，資優生與普通生各六位，其中男生女生各六位，共三十一名兒童參加語文的預試。至於數學的評量，二年級有十三位，其中資優生八位，普通生五位，男女生各為六與七位；三年級有十三位，其中資優生七位，普通生六位，男生六位，女生七位。

四、施測方式與時間

雖然是觀察期要使用的團體評量活動，為

了詳細記錄兒童評量過程中的所有反應以建立觀察期明確的觀察表與評分標準，乃採個別施測的方式進行，施測者為研究者訓練的四校資優班教師。施測過程中，兒童可以提問題澄清題目，也可以尋找解題的資源，施測者再予以回覆與提供資源。施測過程中，教師依據研究者所提供的觀察記錄表與觀察重點，詳細記錄兒童的所有反應。除了施測者詳細的記錄外，並同時錄音和錄影。施測時間則由各資優班教師選擇朝會、午休或其他可利用之時間進行，兒童的作答不限時間，直到做完為止，可以分次實施。

五、觀察指標與一般評分表

解題結果的評分指標與認知複雜度的觀察，主要以Glaser (1991) 與Sternberg和Spear-

Swerling (1996) 的問題解決構念指標為主要依據，表四詳細分列各向度的觀察指標。而表現結果的一般評分表中，語文領域則依據美國QUASAR (1993) 研究小組的實作評量構念架構與奧瑞岡州州教育局所頒佈的語文實作評量評分向度指引 (1993)，將評分向度分為正確性、相關性、複雜度和創意等四個指標，而每個指標再分為高度、適度 (大部分)、部份、缺乏等四個分數層次，而數學領域則依據Lane (1993) 對建構數學實作評量一般評分表的研究架構建議以及美國奧瑞岡州州教育局所頒佈的數學實作評量評分向度指引 (1993)，將評分向度分為數學的知識、策略的知識和溝通的能力等指標，而各個指標再細分為完整、主要、部份、和缺乏等三至四個分數層次。

表四 問題解決能力的觀察與評分指標

問題解決的要素	觀察指標
該領域高度整合組織的知識	問題理解的層次，對問題深層或表面的意義提出問題
對問題質的分析	重述問題、指出相關或關鍵的字詞，不相關的訊息，提問問題
策略的選擇與應用	在解題過程中於關鍵時能從不同的策略中，選擇適當的配合情境的策略，察覺解題中的錯誤，使用適當的策略
自動的過程性知識	速度、時間，整體和部份的時間分配
自我監控與檢核	在解題之間或之後，檢查自己的工作、重做、對自己的表現提問、自語後修正、頓悟阿哈的現象
動機與努力	面對問題正、負面的態度，遇到困難的堅持度與嘗試的過程

六、評分者

兒童反應類型的分類與解題結果的評分由研究者和兩位研究助理依據兒童的原始反應分類並依據一般評分表評分。兩位研究助理，一位為心理系碩士，另一位為國民教育研究所碩士專研評量與認知複雜度的研究。認知複雜度反應類型的分類由研究者編碼分類，找出大致的類型之後，再由研究者與心理系的研究助理將所有兒童的反應分類。語文解題結果的評分由研究者與心理系的研究助理評分，資料的分

析則以研究者的評分為主；數學解題結果的評分則由國教所的研究助理與一位資優班數學教師負責，資料分析以研究助理的評分為主。

七、資料的處理

兒童的解題過程由研究助理將教師的記錄再以錄影帶補充，將兒童的反應與解題結果轉錄為逐字稿。問題解決歷程的認知複雜度，由兒童的原始反應分類編碼後，計算反應次數，但因各組的人數不一，資料分析以各組每人平均次數的分析為主；而各評量活動的解題結果

得分由評分者評分後，因各分項的總分不一，而兩個領域的總分亦非以100分為滿分，為討論方便乃將原始分數化為百分數後再進行資料的統計分析。

結果分析與討論

本研究的目的重在建構評量活動的觀察與評分系統，結果的分析將從理論的構念分析與兒童反應原型（prototype）的向度與層次分析歸類，以建立觀察與評分層次的具體類型。分別從以下四點分析：(1)分析評量活動的相關度與難度；(2)分析兒童解題歷程的認知複雜度；(3)依據兒童反應的分析確定評量成就層次的順

序類別，建立特定的明確評分表；(4)討論活動結構對表現結果的影響。

一、評量活動的相關度與難度分析

(一)評量活動間的相關

評分者依據評分表對兒童表現結果的評分，語文的評分者相關係數為.91，而數學為.82。而各評量活動的得分相關可於表五和表六中看出。語文活動中，結構定義清楚的活動如散文閱讀、童詩填寫和弱結構的活動如古詩解釋、看圖作文的得分相關都達顯著水準，而無結構的評量活動，除了影片描述和實際活動經驗描述的相關達顯著水準外，這三個活動和其他二類結構度的活動相關皆十分低。

表五 語文評量活動得分相關

評量活動	散文 閱讀	童詩 填寫一	童詩 填寫二	古詩 解釋	看圖 作文	小小說分析 與演說	影片 描述	實際經 驗描述
童詩填寫一	.46*	1.00						
童詩填寫二	.47**	.76**	1.00					
古詩解釋	.53**	.53**	.39*	1.00				
看圖作文	.48**	.53**	.42*	.59**	1.00			
小小說分析與演說	.36	.37*	.39*	.55**	.52**	1.00		
影片描述	.48**	.33	.32	.10	.15	.16	1.00	
實際經驗描述	.65**	.22	.28	.36	.12	.35	.41*	1.00

* : $p < .05$ ** : $p < .01$

表六 數學評量活動得分相關

評量活動	積木	擺豆子	彩繪玻璃	七巧板	貼貼紙	走路	找規則	撈魚
擺豆子	.01	1.00						
彩繪玻璃	.29	.31	1.00					
七巧板	.01	.30	.30	1.00				
貼貼紙	.37	.56**	.57**	.40*	1.00			
走路	.23	.37*	.33	.14	.52**	1.00		
找規則	.02	.13	.22	.03	.48*	.26	1.00	
撈魚	.11	.38*	.45*	.33	.71**	.40*	.13	1.00

* : $p < .05$ ** : $p < .01$

而從表六中可發現，數學評量活動中，貼紙除了和積木活動的相關低外，和其他活動的相關皆達顯著水準，擺豆子和走路以及撈魚的相關亦達顯著水準，而彩繪玻璃和撈魚間的相關亦達顯著水準，其他的相關皆未達顯著水準。

(二)評量活動難度分析

評量活動的難度分析將由題目中所出現的難字、兒童要求澄清題意與解釋題目文字的問題平均次數、兒童得分的分布與平均得分的高低來分析。

語文領域。兒童對語文領域評量中的文字敘述，所提出的難字詞（不會讀或不認得的字）的總數整理如表七。由表七中可見，就閱讀文字的理解而言，對所有兒童而言，弱結構的活動難字最多；而就四組兒童的反應來看，所有題目對二年級普通兒童而言，難字最多（sum=51），其次為二年級資優兒童（sum=23）和三年級普通兒童（sum=16），而對三年級的資優兒童而言，幾乎沒有文字閱讀上的困難（sum=4）。由此可見，難字的總字數因年級的增加與資優與否而遞減。

表七 語文評量活動難字分析

活動結構度	小二普通生	小二資優生	小三普通生	小三資優生	總和
結構清楚	19	7	6	1	33
弱結構	26	12	7	2	47
無結構	6	4	3	1	14
總和	51	23	16	4	91

* 欄中的數字為難字的總和

兒童對評量活動中要求解釋文字與澄清題意的提問次數為難度的第二項指標。由表八中可以發現，對所有兒童而言，弱結構的評量活動平均提問次數最高（M=1.8）；比較四組兒童的平均次數，二年級普通兒童的問題平均次數最高（M=2.8），而三年級資優兒童的平均

次數最低（M=0.4），其中二年級資優兒童的平均次數也少於三年級普通兒童的平均次數（M=0.7 vs. M=1.1）。總之，提問的平均次數因年級的增加而遞減，而資優兒童的平均次數也低於普通兒童。

表八 語文評量活動提問平均次數比較

結構度	小二普通生	小二資優生	小三普通生	小三資優生	總平均
結構清楚	2	0.9	1.7	0.2	1.2
弱結構	4.5	0.8	1.1	0.6	1.8
無結構	2	0.5	0.5	0.3	0.8
總平均	2.8	0.7	1.1	0.4	1.3

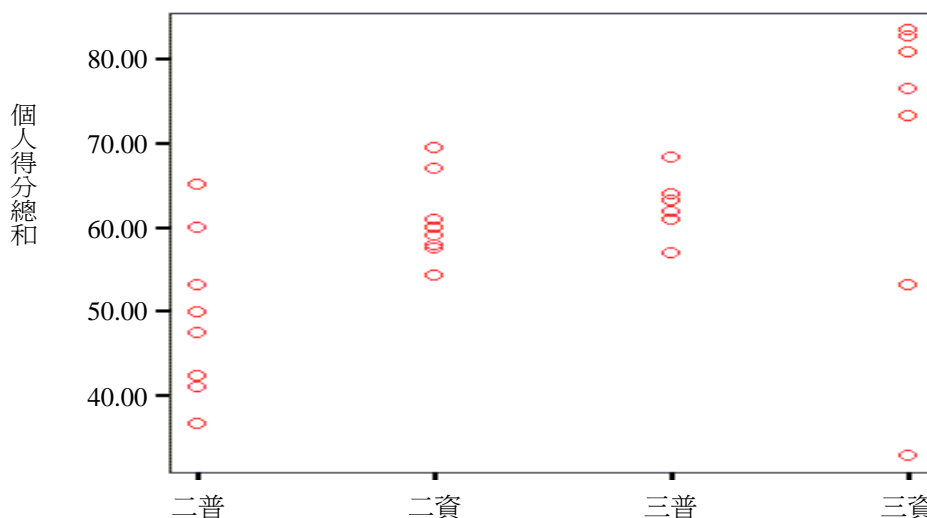
* 欄中為提問的平均次數

另外，由圖二中可看出，四組兒童的語文得分分數分布，二年級普通兒童的得分分布在

30~65分之間，而二年級資優兒童集中在40~70分之間，三年級普通兒童則集中在50~70分

之間，而三年級資優兒童的得分則落在55~90分的間距。四組兒童的語文平均得分依序是二年級普通兒童最低（ $M=48$ ），其次為二年級資優兒童（ $M=56$ ）和三年級普通兒童（ $M=$

63），而三年級資優兒童則是四組中平均得分最高的（ $M=75$ ），得分的分布因年級的增加與資優與否而遞增。



圖二 語文得分分布

數學領域。受試兒童對數學評量活動的文字敘述都能認得沒有難字詞。因此數學活動的難度分析就由每組的提問平均次數、得分分布和平均得分來看。由表九中可見，就所有兒童而言，評量活動的結構度對兒童的題意理解並無明顯的差異；但四組兒童中，可以發現二年

級普通兒童提問的平均次數最高（ $M=0.4$ ），其次為二年級資優兒童（ $M=0.2$ ）和三年級普通兒童（ $M=0.2$ ），而三年級資優兒童的提問最低（ $M=0.1$ ）。就對題目敘述的提問平均次數而言，對評量問題的理解因年級的增加與資優與否而遞增。

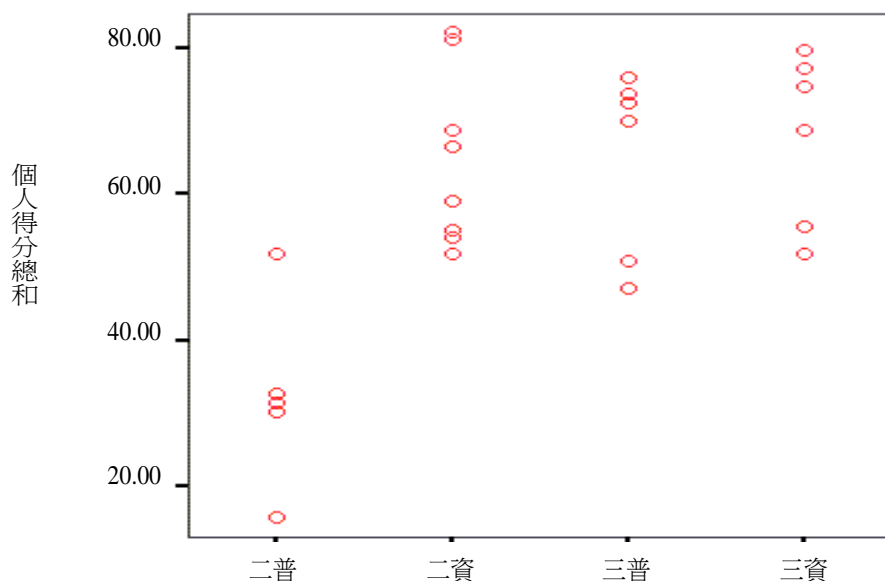
表九 數學活動提問平均次數

結構度	小二普通生	小二資優生	小三普通生	小三資優生	總平均
結構清楚	0.3	0.3	0.0	0.0	0.2
弱結構	0.4	0.1	0.1	0.1	0.2
無結構	0.6	0.1	0.5	0.1	0.3
總平均	0.4	0.2	0.2	0.1	0.2

* 欄中為提問的平均次數

由圖三中，可發現二年級普通兒童的得分都在20~50分之間，二年級資優兒童則在50~80分之間，三年級普通兒童都在45~80分之

間，而三年級資優兒童的得分則落在50~80分的間距。而四組兒童的平均得分中，二年級普通兒童最低（ $M=31$ ），其次為二年級資優兒



圖三 數學得分分布

童 (M=55) 和三年級普通兒童 (M=62)，而三年級資優兒童的得分最高 (M=66)，平均分因年級的增加與資優與否而遞增。

二、問題解決歷程的認知複雜度分析

問題解決的認知複雜度向度分別以解題的速度、對問題質的分析、資訊的表徵、策略的形成、以及過程監控與結果檢核為主要分析向度，每個向度中再依兒童的反應歸納類型層次。

(一) 解題速度的比較

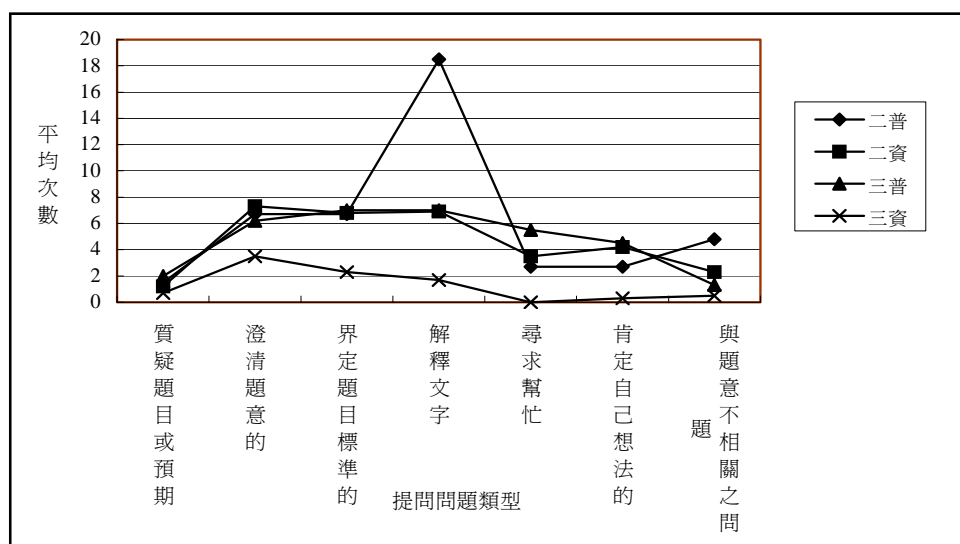
各組語文評量活動所需平均時間中，二年級普通兒童平均每題需要20分鐘，是最長的一組，而三級資優兒童的時間則是13分鐘，是速度最快的；而二年級資優兒童和三年級普通兒童則分別是18和16分鐘。而數學活動，二年級普通和資優兒童平均每題皆需要16分鐘，所需時間並無差異，三年級普通兒童需13分鐘，而三年級資優兒童則平均只需9分鐘，是解題速度最快的一組。

(二) 對問題質的分析

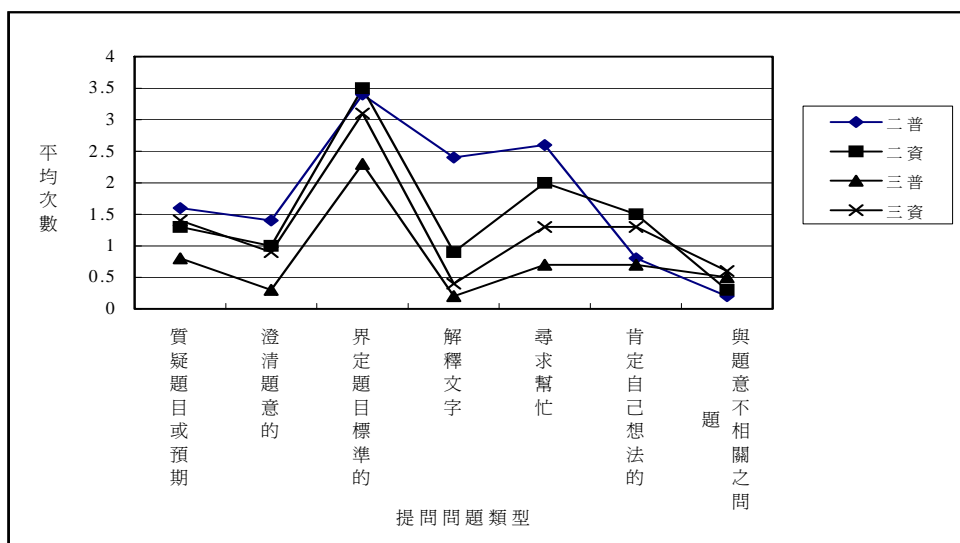
歸納兒童對所有評量活動所提出的問題類型分析之後，大致可將這些問題分為七類，依序為質疑問題或預期結果、澄清題意、界定題目標準、解釋文字、尋求幫忙、肯定自己的想法和與題意不相關之問題等，詳細的反應類型可見表十。從圖四中可以看出，在語文評量活動中，澄清題意 (M=5.9) 和界定題目的標準 (M=5.7) 的問題類型是所有兒童最常問的問題。而相對於其他組，三年級的資優生是最少提問的，而二年級普通和資優兒童以及三年級普通兒童則無明顯差異。在提問的問題類型中，解釋文字是二年級普通兒童最常問的問題，而界定題目的標準則是二年級資優兒童與三年級普通兒童最常問的類型，而三年級資優兒童最常問的則是澄清題意的類型。在數學方面，圖五中可見所有兒童都在界定問題的標準的提問平均次數明顯高於其他類型，而二年級普通兒童在文字解釋和尋求幫忙方面也明顯高於其他組；其中相對於其他組，三年級普通兒童則是在各類型提問中最少提問的。

表十 兒童提問問題的類型與反應舉例

提問問題類型	兒童反應舉例
質疑題目或預期	語文 真的有小矮人嗎？這是真的還是假的？這是我們拍的還是別人拍的？會不會很恐怖？
	數學 怎麼只有一顆（豆子）？沒說一個玩具區有多大？可不可能有誤差？怎麼題目第二回合成績只有2人？
澄清題意的	語文 這些名詞、形容詞選到的詞要變成句子嗎？
	數學 先用石子，再把它畫進去？用這七張紙排嗎？
界定題目標準的	語文 例子用過的詞，可不可以再用？每一個是不是一個就好？可不可以用注音？
	數學 用畫的可不可以？要寫做法嗎？可以全部塗一種顏色嗎？都要等於9嗎？
解釋文字	語文 「現代化」是什麼？什麼是「空翠」？
	數學 什麼叫「規則」？什麼叫「組合」？
肯定自己想法的	語文 如果想出來的字跟原來一樣可以嗎？只有一個字，可以嗎？
	數學 可以這樣塗嗎？可以畫兩個7嗎？可以排成這樣嗎？
與題意不相關之問題	語文 我可以走了嗎？寫完這個就好了嗎？是不是今天一定要寫完？
	數學 我要做幾張啊？可不可以休息？



圖四 語文提問問題類型分析比較

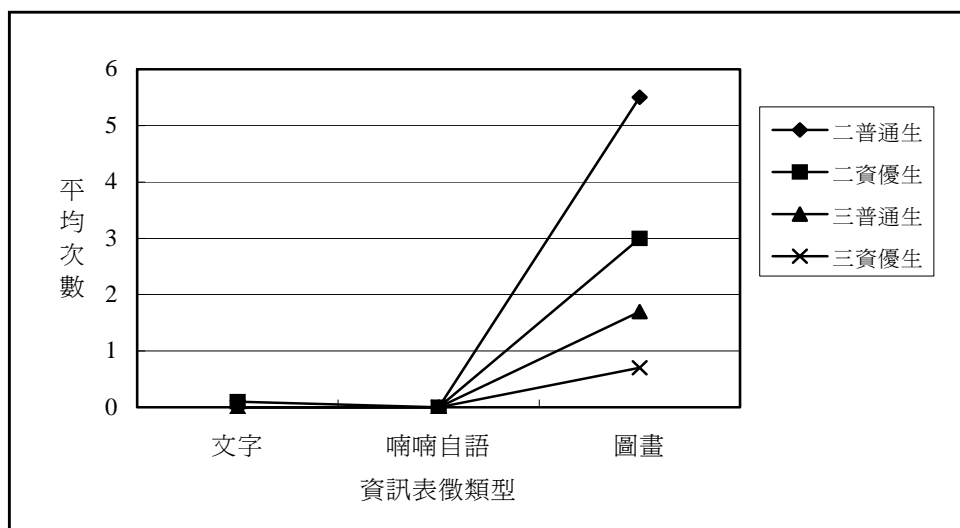


圖五 數學提問題類型平均次數比較

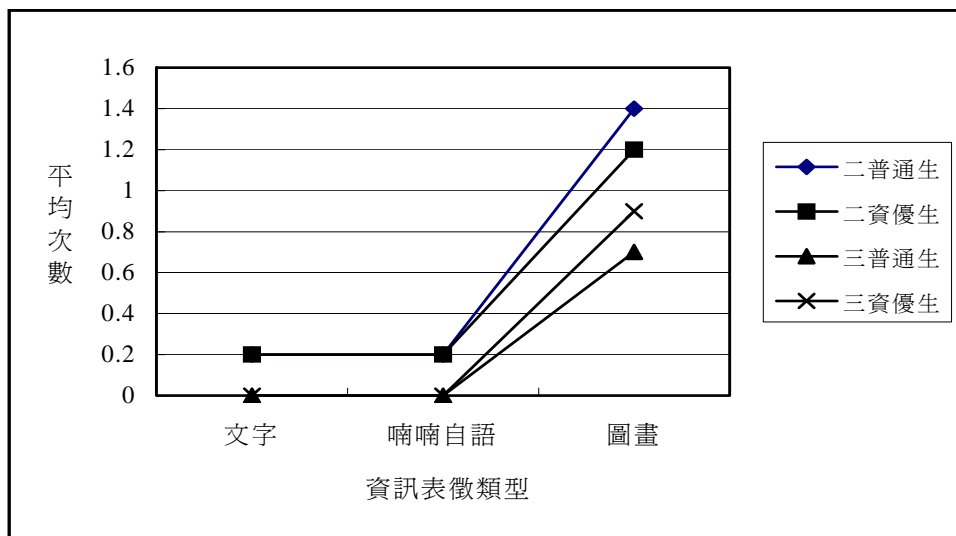
(三)資訊的表徵

在評量活動中觀察到兒童思考的資訊表徵方式，歸類後大致有三類：文字、喃喃自語和圖畫思考的方式。而從圖六和圖七中可以清楚的看出，所有兒童最常用的資訊表徵方式是圖

畫，而其中二年級普通兒童是在二個領域的評量中觀察到用得最多的一組；在語文評量中，三年級資優兒童則是用得最少的一組。而數學評量中，三年級普通兒童則是觀察到最少使用思考表徵方式的一組。



圖六 語文資訊表徵類型比較

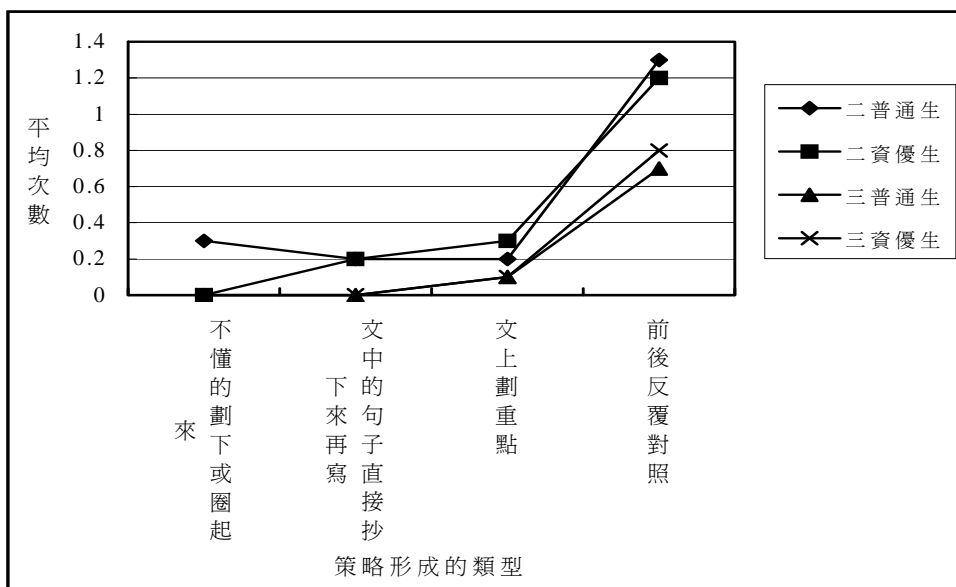


圖七 數學資訊表徵類型比較

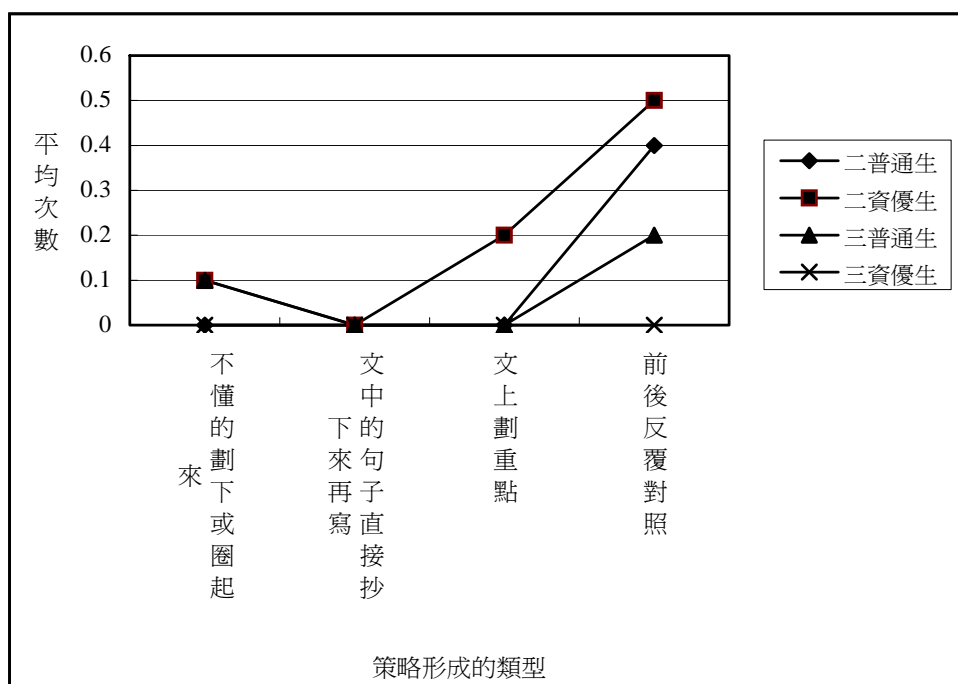
(四)策略的形成

於解題過程中觀察到兒童所使用的解題策略大致有：將不懂的劃下或圈起來、文中的句子直接抄下來、文上畫重點、前後反覆對照等四種類型。從圖八和圖九中可看出在語文或數

學的解題過程中，兒童最常使用的策略都是前後反覆對照；四組在各策略類型的平均次數的反應類型十分一致，而其中三年級的資優兒童都是觀察到的策略使用最少一組。



圖八 語文策略形成類型比較



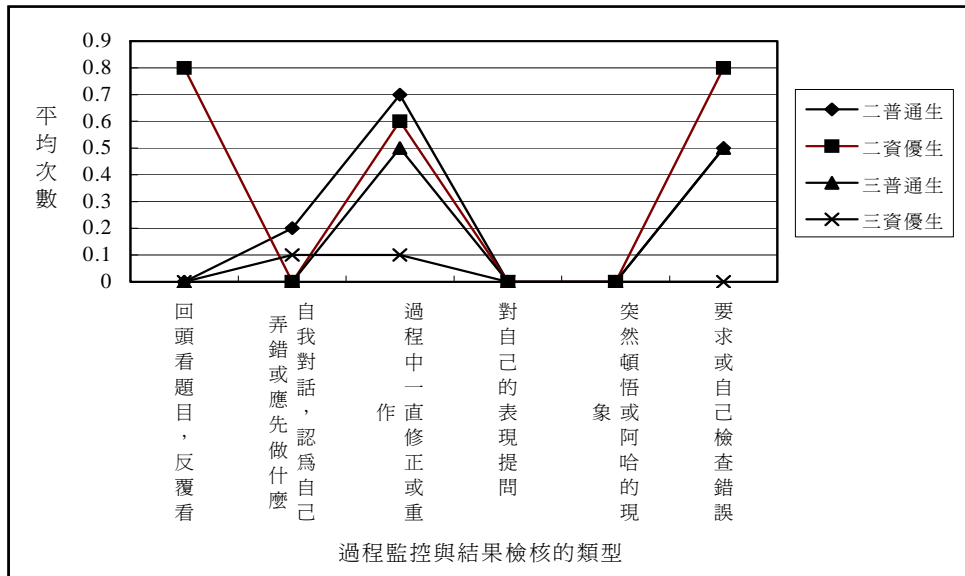
圖九 數學策略形成類型的比較

(五)過程監控與結果的檢核。兒童在過程的監控與結果的檢核中，觀察到的方式大致有：回答時回頭看題目或反覆看、自我對話（認為自己錯了或應先做什麼）、對自己的表現提出疑問、突然頓悟或阿哈的現象、過程中一直修正或重作（不包含自我對話和提問）、要求檢查錯誤等六種類型。從圖十和圖十一中可以看出，就語文評量方面，二年級和三年級普通兒童最常出現的方式是一直修改或重作，而二年級資優兒童則是回頭反覆看題目和檢查錯誤，而三年級資優兒童最常觀察到的是自我對話的方式；就數學方面，二年級普通和資優兒童以及三年級資優兒童都是要求檢查錯誤的方式最多，其中又以二年級資優兒童最高；而三年級資優兒童則在自我對話和對自己表現提問方面都明顯高於其他組；二年級普通兒童則在過程

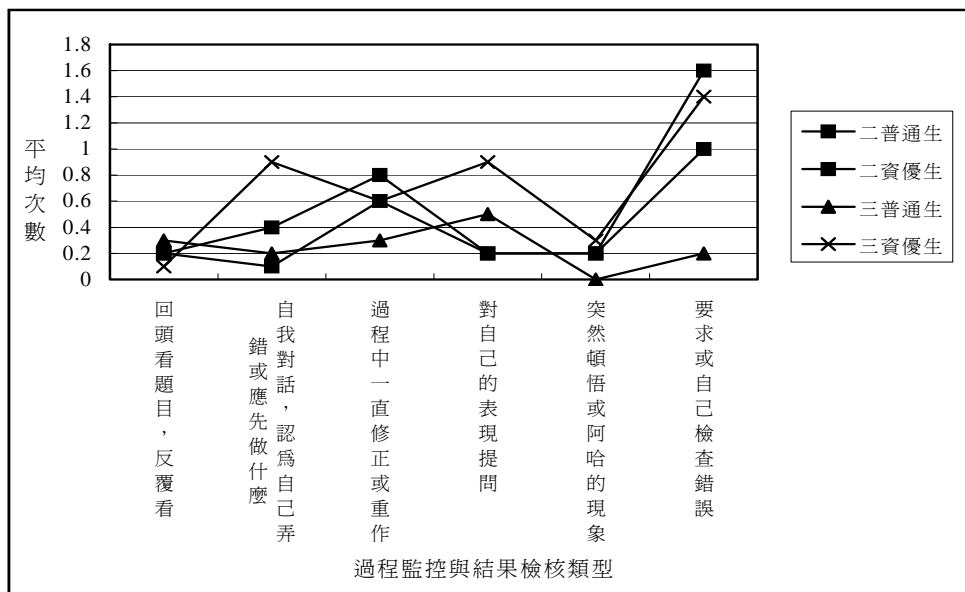
中一直修正或重作的方面明顯高於其他組。

三、表現結果的層次與具體評分表

兒童的表現結果依據上述語文與數學的評分層次評分歸類之後，發現就所有評量活動中每題得分層次中的人數比例而言，三年級資優兒童在高層次的分數（如4分）的人數比例最高，次為二年級資優生或三年級普通生，而二年級普通兒童在高層次的分數人數比例最低，而低層次的分數（如0分）的人數比例則是最多的。依據一般評分表所歸類出兒童在各個層次的反應類型和各組人數比例詳如表十一與表十二，分數的層次加上各個評量活動的特定反應類型將成為建構特定明確評分表的依據。限於篇幅，僅以語文和數學各一個評量活動為例。



圖十 語文解題過程監控與結果檢核類型比較



圖十一 數學解題過程監控與結果檢核類型比較

表十一 語文童詩填寫活動的分數層次與反應類型舉例

(1)沒有鑼就_____					
分數	反應類型舉例	小二普通生	小二資優生	小三普通生	小三資優生
4	敲鍋底，敲桌椅，用腳踢	17	33	33	50
3	敲自己	0	17	33	33
2	拿起雙手來拍，用盤子打	33	33	17	17
1	打鐵，彈鋼琴	33	17	0	0
0	沒有反應	17	0	17	0

(2)用手_____					
分數	反應類型舉例	小二普通生	小二資優生	小三普通生	小三資優生
4	拍東又拍西，打肚皮	17	17	33	67
3	打自己，就更容易	0	25	17	17
2	拍打樂器呀	33	42	33	17
1	當指揮棒，彈鋼琴用手敲鼓	33	17	0	0
0	沒有反應	17	0	17	0

* 欄中的數字為人數百分比

表十二 數學彩繪玻璃評量活動的分數層次與反應類型舉例

(1)選取解題條件					
分數	反應類型舉例	小二普通生	小二資優生	小三普通生	小三資優生
3	以深層結構選取解題條件：兩種顏色，每一扇窗戶都不一樣；位置改變就算不一樣。	0	62	67	67
2	以表面特徵選取解題條件：四格全塗一種顏色；找出規則，才不會重複或遺漏	20	15	17	17
1	其他無關聯的解題條件	0	0	0	0
0	沒有選出解題條件（沒有反應）	80	33	16	16

(2)策略表現					
分數	反應類型舉例	小二普通生	小二資優生	小三普通生	小三資優生
4	精確的專家解法，一次操弄一個變項：有系統的固定顏色以及顏色數目變項，改變位置變項	0	48	17	67
3	有規則的著色，有時候操弄兩個變項	0	0	33	16
2	有計畫，但不能貫徹（中途跳脫，又修正回來）	20	16	17	16
1	沒有明確的規則，想到什麼，就畫什麼	60	38	33	0
0	沒有反應，空白	20	8	0	0

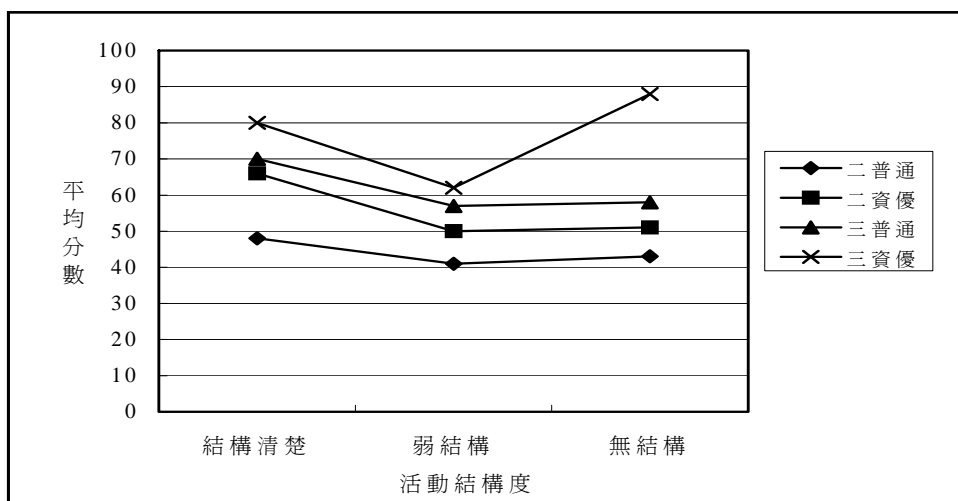
(3)答案的正確性					
分數	反應類型舉例	小二普通生	小二資優生	小三普通生	小三資優生
3	正確的塗了16扇窗戶	0	46	50	60
2	修改後為16扇	0	28	30	20
1	正確塗了10扇以上但16扇以下	40	10	20	20
0	塗了10扇以下	60	16	0	0

* 欄中的數字為人數百分比

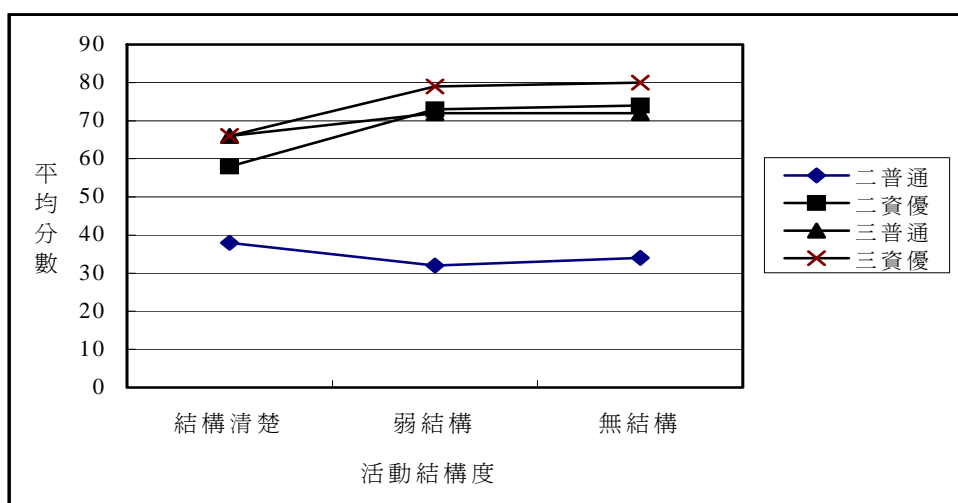
四、評量活動結構度對得分表現的影響

分析結構度與得分平均的關係發現，在語文評量活動中，兒童在結構清楚的題目上平均得分最高（ $M=66$ ），次為無結構的題目（ $M=60$ ），而弱結構的題目平均得分最低（ $M=54$ ），與難度分析的結果一致；而數學評量中，兒童在結構清楚的題目上平均得分則是明顯最低（ $M=58$ ），其次為弱結構（ $M=67$ ），而無結構的題目平均得分則是最高

（ $M=68$ ）。而從圖十二和圖十三中可以發現，語文評量中，二年級普通兒童和資優兒童以及三年級普通兒童的得分情形都是結構清楚最高，但三年級資優兒童則在無結構的題目上平均得分最高，而數學評量中，二年級普通兒童的平均得分以結構清楚的活動平均得分較其他兩類題目高，而二年級資優、三年級普通及三年級資優兒童的平均得分都隨題目弱結構和無結構而得分增加。



圖十二 活動結構度對各組語文平均分數的影響



圖十三 活動結構度對各組數學平均分數的影響

結論與建議

資優兒童問題解決能力的實作評量活動與評分系統的發展需有一系列的研究方可成就，本研究以兒童反應原型建構適用於二年級資優兒童鑑定觀察期使用的評量活動與觀察評分系統為主要目的，是這個系列研究的開端。以開放的角度觀察分析後，發現兒童思考的複雜度對結果分析是最大的挑戰，而為了統整性的呈現，也免不了將其複雜度概化與簡化，是後續研究必須突破的瓶頸。在此從問題解決能力的領域知識與認知複雜度的二個主要構念向度與能力的區別性來討論評量活動、觀察指標與評分系統的建構效度，並從此次研究的困難與限制提出未來研究方向的建議。

一、研究結果

(一)領域知識的構念相關

語文活動間的相關高，而數學活動的相關偏低。實作評量活動，以開放的情境設計，藉由不同評量活動的向度呈現兒童能力的不同面向，因此得分相關性不高，也是實作評量的特質之一（呂金燮，民88）。以此次的評量活動語文相關高，但數學相關分歧的現象，主要原因在於語文評量目標與評分表強調語文整合應用的能力，因此得分相關都相當高；而數學相關分歧主要在於一則因活動設計的目標向度分化，如空間推理的彩繪玻璃與數學運算的相關不高；二則因數學評分系統，除了解題結果的正確性外，並依個別評量活動的解題策略而區分分數的層次。因此各評量活動各有不同的策略與解題特質，而非結果的正確與否，更增加了相關不高的可能性。

(二)認知複雜度的多元向度與深度

1. 兒童於解題過程中表現出解題思考歷程向度的多元性與層次性。

從對問題質的分析到結果的監控檢核等認知複雜度的向度，兒童表現出多元化的類型。

有些類型對幫助解題的質的方面有其關鍵層次之分，如對問題質的分析，其依序是質疑題目或預期、澄清題意、界定題目標準、解釋文字、肯定自己的想法、以及與題意不相關之問題；而結果的監控檢核的六個類型的解題關鍵性，依序是自我對話、頓悟、自我懷疑、要求檢查、反覆看和一直修正等（Sternberg & Spear-Swerling, 1996）；而如資訊表徵則有三個類型，策略的形成四個類型等，表現出個人解題的思考特質之外，類型間的差異對解題的結果並無直接性的影響（Sternberg & Spear-Swerling, 1996）。

2. 呈現出專家與生手的特質差異。

從二年級普通兒童到三年級資優兒童的解題表現差異，可以窺見專家與生手間的特質差異。不論是數學或語文評量，三年級資優兒童不但解題正確性高、速度快、且所表現出的外顯策略和思考的資訊表徵較少，可見其解題已趨自動化；而其較常提問的是澄清題意與界定題目標準的問題，監控檢核過程在自我對話和自我提問甚至有突然頓悟的表現多，都顯示出專家優秀的解題監控特質。二年級資優生或三年級普通兒童的表現則為次之。而二年級普通兒童解題正確性最低，速度最慢，知識方面的組織不夠尚須分析解題，因此所表現的外顯策略與資訊表徵方式也就明顯較多。而因仍侷限於對題目的閱讀理解，因此提問的類型都集中在解釋文字與尋求幫忙等，未能考量到問題的整體性；反覆嘗試修正的過程也顯示了生手的解題特質。

(三)解題能力的區別性

1. 評量活動的難度偏高，但能區辨二～三年級普通兒童與資優兒童解題結果之質的差異。

整體而言，兒童在這二個領域的評量表現結果得分都偏低，凸顯了一般實作評量的特性之一（呂金燮，民88）。除了兒童不習慣這類

評量的活動之外，實作的成就評量活動必須具備高層次思考組織一些學習過的概念，而非習過概念的複製，也使得難度增加。雖然對兒童而言，題目可能較複雜困難，可是評量中每位兒童都能回答問題，而分數的差異是在解題質的層次而非答對與否。而從難度分析、各組得分平均與分數層次的人數比例來看，難度都因年級的增加與資優與否而遞減，其中除了語文平均分外，二年級普通兒童與資優兒童間的差異度都遠大於三年級普通兒童和資優兒童間的差異。再經過後續研究的修正與效度考驗，必更能適用於二年級資優兒童鑑定。

2. 兒童的表現原型分布各分數層次，評分表的分數層次具有解題能力質的差異。

實作評量依賴主觀的評分，因此明確具體的評分系統愈顯其重要性。評分系統必須容易使用，並能充分呈現能力的多向度，能將兒童的解題歷程的複雜度以不同分數層次代表，方能為鑑定之依據。就解題結果的分數層次分布與反應類型看，例如二年級普通兒童的回答都十分直覺而缺乏精細的描述，得分偏向低層次的分數（如1分或2分），而三年級的資優兒童的回答除了正確性外，其組織與邏輯性都相當高。高層次（如4或3分）的得分人數比例最高，這種質的差異是好的實作評量的評分表標的之一。層次分明的表現指標加上具體兒童反應類型建構了明確評分表，即能提供具體的評分依據。

3. 兒童在不同結構度的評量活動中的解題表現，因領域特性與兒童解題能力而有所差異。

由於評量活動結構度與兒童的能力是屬相對性的特質，對二年級兒童屬弱結構的問題，相對三年級而言可能就是結構定義清楚的問題，因此，不同結構度的活動對各組兒童的難度十分分歧。大體而言，源於語文的擴散性與沒有標準答案的限制下，結構度越不清楚，兒

童對評量標準與重點的掌握愈分歧，也最容易以直覺的方式答題，因此兒童在條件標準清楚的結構定義清楚的活動中表現優於其他兩類活動；而數學活動雖以開放的情境設計，仍屬聚斂性的思考活動，解題最終仍有正確答案的限制下，無結構活動富情境化的設計下，更能引發兒童的自我思考而表現優異。另外，二年級普通兒童不論是語文或數學評量，都十分依賴結構清楚的題目中，清楚的敘述、充分的解題條件、已知的解題方式與結果的協助解題，因此，在結構定義清楚的評量活動中表現較其他兩類的活動優秀；反之，專家型的三年級資優兒童在無結構定義的評量中，愈能發揮自我的思考與組織能力。兒童在不同結構度評量活動中的解題表現，因領域的特性與解題能力而有所差異。

二、研究的困難與限制

(一) 評量方式依賴文字敘述，而解題的思考觀察過度依賴外顯行為，有待突破。

不論是題目的敘述，或兒童解題的結果，常依賴文字的敘述。是實作評量無可避免的缺失。文字的描述過多時，語文能力對解題的結果相對重要之外，兒童的耐性也因大量的閱讀與書寫而耐性遞減。而解題歷程的複雜度除了解題時間可以計時，對問題質的分析可以請兒童表達外，其他的認知複雜度向度均需兒童有明顯的外顯行為方可觀察到，沒有表現出的解題歷程，並不意謂兒童沒有這樣的思考歷程，而是無法觀察到。而解題後，為更了解兒童的想法，進一步的追問，也未必是兒童原來的想法。因此，其準確度有限，也降低了觀察結果的可靠性。

(二) 評量活動強調解題思考深度的了解，但領域知識內容的範圍廣度不足。

由於實作評量強調兒童解題思考深度的了解，開放性的情境設計，讓兒童充分思考表達，因此施測費時。如此則犧牲了範圍的廣

度。每個向度的題型在時間的考量之下僅能有一至二題，在知識範圍廣度的代表性上仍有待商榷。

(三)難度偏高時，兒童本身的積極度影響解題結果甚鉅。

研究中最大的困難是部份兒童一遇難題就放棄，需要不斷的鼓勵，雖然最終都能順利解題，但這類兒童若於正式的施測中將無法有如此的協助，其表現結果也可想而知。有些兒童的語文能力並不佳，卻能積極爭取幫忙；有的兒童十分樂觀，解題過程不斷嘗試，有些兒童則較急躁，反而不容易找出規則。因此，實作評量的實施，兒童對難度高題目的情緒反應與堅持度成爲另一個解題的關鍵要素。

(四)評量費時，容易流失研究對象。

本研究是八週觀察期評量活動的題型預試，爲了讓兒童有充分發揮問題解決能力的機會，評量過程沒有時間限制，加上施測者與兒童間的互動，評量觀察的進行相當費時。有的兒童用了相當久的時間，有的兒童因爲題目難度高，很快就放棄，施測的時間並不一致。平均做完所有數學領域題目需要三個半個小時，而做完語文領域的題目需要 4.6 個小時。爲了分散兒童的壓力，以及不影響兒童的正常上課，以每天施測一個分項評量爲原則，平均每位兒童需要一週的時間方能完成一個領域的題目。因此原則上，每位兒童不同時參與兩個領域的評量。也因所需時間相當長，施測過程中遭遇部份家長拒絕，而有七位兒童因沒有完成所有活動，而無法列入資料分析。於實際應用時，在將所有活動分散於八週的觀察期之後，希望這樣的現象不會出現。

(五)樣本過於侷限。

爲了詳細觀察記錄兒童的解題思考歷程，人力時間的因素之下，無法廣泛抽樣。研究樣本的代表性有待商榷。因此，本次研究結果尚須經過一次正式的施測實驗，分析修正後，本

研究的評量活動與評分表方能有具體的實用價值。

三、研究結果之應用與未來研究之建議

(一)資優兒童的鑑定方面

1. 資優兒童的思考深度與解題能力都較一般兒童優秀，適合實作評量，但實作評量的應用費時費力，更須有專業訓練的教師配合。才能發揮鑑定的功能。

實作評量的應用可以深入了解兒童思考歷程，彌補一般標準化測驗的不足，增加對兒童解題能力的了解，給予鑑定時決策的依據。但施測費時，必須考量時間人力的因素之外，觀察與評分更須有專業訓練的教師參與方能發揮該有的效能。

2. 實作評量解題後的表現結果較適合作爲鑑定依據，而認知複雜度的層次作爲教學設計參考。

由於解題後的結果是兒童的了解與知識應用的高低層次，適合用來比較兒童的表現；而解題歷程中的認知複雜度各向度的表現類型，並無實質的高低之分，並不適用於比較。而適合依個別特質作爲教學設計的依據。

3. 須多次反覆評量與評分表的修正過程，實作評量的活動複雜度與評分系統才能配合資優兒童的特質，提供有效的依據。

一旦鑑定評量的目標確立之後，一般的評分表就具雛形了，評量活動的設計就有所依據，但是決定評分系統的複雜度與適度之前，必須先有兒童的實際反應作爲評分系統的回饋；其中最爲困難的是得分順序的決定，例如如何決定數學評量活動中不同解題策略的得分順序。因此，評量活動和評分系統應同時建立，而同時進行則必須經過多次的反覆過程後，方能更臻完備。因爲國內對二、三年級資優兒童學科領域問題解決能力與思考歷程的分析尚缺乏，仍有待後續研究的支持。

4. 基於實作評量強調深度而犧牲廣度的特

質，鑑定效度的提高有賴題庫的建立。

美國實作評量用於教學安置決策的實施已朝大量施測並建立題庫的方向發展，或許是我們未來研究的努力方向。若有題庫，則各校可依據其資優教學目標組合選擇使用評量活動，而不受有限的評量活動，更能切合實際需要。

(二)教學方面

1. 高層次的思考歷程列為教學目標。

兒童的問題解決歷程並沒有有一致穩定的表現，尚未成為思考的路徑，或許源於缺乏教學情境的支持，教學設計應可朝高層次的思考歷程設計，如對問題質的分析以質疑或澄清題意，而監控檢核以自我對話質疑等為主，依序列為教學目標。

2. 多元的解題情境，以適合領域與兒童的思考特性。

兒童在不同結構的評量情境中的表現會因領域的特性有所差異，而生手型的兒童依賴活動的結構度解題，而專家型的兒童則在無結構的情境下，最能發揮其優秀的自我思考與組織能力。因此，教師在設計活動情境，應兼顧這二個因素，以更了解適合兒童個別差異的解題思考教學。

3. 加強兒童自我思考的建構與傳達能力。

實作評量強調兒童對自我思考的建構與呈現，但此研究結果發現大部分兒童得分偏低是因為缺乏組織性的思考傳達。因此，教室中的學習應可對兒童學習概念的建構與傳達方面多予表現的機會。

4. 兒童面對難題堅持度的培養。

因篇幅所限，兒童的動機與努力在解題過程中的差異和結構度對解題歷程認知複雜度的影響，都未能討論。但從觀察中可以清楚發現，兒童的積極度影響解題結果甚鉅。因此，教學方面除了解題技巧，思考等認知方面的訓練外，兒童的堅持度與對待難題的態度，應該要設計情境積極有目標的培養。

總之，問題解決能力牽涉兒童整合許多知識經驗以及解題過程的努力，且受活動的結構度影響極大。因此，準確的反應分析對了解成功與不成功的解題思考十分重要。本研究僅從兒童的反應原型歸納分析，初步建立了適用於二年級資優兒童問題解決能力評量的表現結果評分表與解題歷程的複雜度觀察表，但鑑定用的表現結果評分表，則仍須待更多樣本研究的修正方能提供有信度與效度的鑑定依據。

參考書目

一、中文部分

呂金燮，民88年：實作評量－理論。載於王文中、呂金燮、吳毓瑩、張郁雯與張淑慧著，**教育測驗與評量－教室學習觀點**，169-203頁。臺北：五南出版社。

林妮芙，民88年：**實作評量的題目結構度對認知複雜度的影響**。國立臺北師範學院，碩士論文。

陳美芳，民85年：資優學生身心特質與評量。**教育資料集刊**，21輯，13-27頁。

郭靜姿，民85年：資賦優異學生的鑑定與教育安置。**教育資料集刊**，21輯，28-54頁。

二、英文部分

Chase, W. G. & Ericsson, K. A. (1982). Skill and working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, 2-58). NY: Academic Press.

Chi, M. T. H. (1987). Representing knowledge and metaknowledge: Implications for interpreting metamemory research. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp 239-266). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive science*, 5, 121-125.
- Deweck, C. S. (1989). Motivation. In A. Lesgold & R. Glaser (Eds.), *Foundations for a psychology of education*. Hillsdale, NJ: Lawrence Erlbaum.
- Dover, A. C., & Shore, B. M. (1991). Giftedness and flexibility on a mathematical set-breaking task. *Gifted child Quarterly*, 35, 99-105.
- Flavell, J. (1987). Speculations about the nature and development of metacognition. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp.1-30). Hillsdale, NJ: Erlbaum.
- Gagne, R. M. & Medsker, K. I. (1996). *The conditions of learning: Training applications*. Fort Worth, TX: Harcourt Brace College Publishers.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligence*. NY: Basic Books, Inc.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. NY: Basic Books.
- Getzels, J. W. & Csikszentmihalyi, M. (1967). Scientific creativity. *Science Journal*, 3(9), 80-84.
- Glaser, R. (1991). Expertise and assessment. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition* (pp. 17-30). Englewood Cliffs, NJ: Prentice Hall.
- Glaser, R. & Baxter, G. (1997, February). *Improving the theory and practice of performance based assessment*. Paper presented to the Board of Testing and Assessment. National Research Council/National Academy of Sciences Education Standards Conference: "The Assessment of Science Meets the Science of Assessments." Orlando, FL.
- Lane, S. (1993). The conceptual framework for the development of a mathematics performance assessment instrument. *Educational Measurement: Issues and Practice*, 12(2), pp. 16-23.
- Lazear, D. (1995). *Multiple intelligence approaches to assessment: Solving the assessment conundrum*. Tucson, AZ: Zephyr Press.
- Lohman, D. F. (1993). Teaching and testing to develop fluid ability. *Educational Researcher*, 22(7), 12-23.
- Mayer, R. & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47-62). New York: Simon & Schuster Macmillan.
- Neisser, U. (1976). "General, academic, and artificial intelligence." In L. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.
- Oregon Department of Education. (1993). *Oregon dimensions of problem solving*. Salem, OR: Author.
- Piaget, J. (1976). *The Psychology of Intelligence*. Totowa, NJ: Littlefield, Adams.
- QUASAR group. (1993). *Quantitative and qualitative understanding: amplifying students achievement and reasoning*. Paper presentation at the annual meeting of American Educational Research Association.

- Renzulli, J. S. (1977). *The enrichment triad model: A guide for developing defensible programs for gifted and talented*. Mansfield Center, CN: Creative Learning Press.
- Resnick, R. B. & Glaser, R. (1980). Problem solving and intelligence. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Lawrence.
- Resnick, R. B. & Nelson-Le, G. S. (1996). *Socializing intelligence*. A speech at National Taiwan Normal University.
- Rowe, H. A. (1985). *Problem solving and intelligence*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schoenfeld, A. H. (1989). Teaching mathematical thinking and problem solving. In L. B. Resnick & L. E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 83-103). Alexandria, VA: Association for Supervision and Curriculum Development.
- Spearman, C. (1904). "General intelligence objectively determined and measured." *American Journal of Psychology*, 15, 206-221.
- Spiro, R. J., Vispoel, W. P., Schmitz, J. G., Samarapungavan, A., & Boerger, A. E. (1987). Knowledge acquisition for application: Cognitive flexibility and transfer in complex content domains. In B. K. Britton & S. M. Glynn (Eds.), *Executive control processes in reading* (pp. 177-199). Hillsdale, NJ: Lawrence Erlbaum.
- Sternberg, R. (1981). A componential theory of intellectual giftedness. *Gifted Child Quarterly*, 25, 86-92. (Beyond IQ)
- Sternberg, R. (1982). Nonentrenchment in the assessment of intellectual giftedness. *Gifted Child Quarterly*, 26, 63-67.
- Sternberg, R. (1997). *Successful intelligence*. NU: Simon & Schustives.
- Sternberg, R. J. & Davidson, J. E. (1992). Problem solving. In M. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1037-1045). New York: Macmillan.
- Sternberg, Ferrari, Clinkenbeard & Grigorenko. (1996). Identification, instruction, and assessment of gifted children. *Gifted Child Quarterly*, 40(3), 129-137.
- Sternberg, R. J. & Lubart, T. (1995). *Defying the crowd*. NY: Free Press.
- Sternberg, R. J. & Spear-Swerling, L. (1996). *Teaching for thinking*. Washington, DC: American Psychological Association.
- Swanson, H. L. (1992). The relationship between metacognition and problem solving in gifted children. *Roeper Review*, 15(1), 43-48.
- Terman, L. M. (1925). *Genetic studies of genius: Vol. 1 Mental and physical traits of a thousand gifted children*. Stanford, CA: Stanford University.
- Treffinger, D. J. (1986). Fostering effective, independent learning through individualized programming. In J. S. Renzulli (Ed.), *Systems and models for developing programs for the gifted and talented* (pp. 429-460). Mansfield Center, CT: Creative Learning Press.
- VanTassel-Baska, J. (1994). *Comprehensive curriculum for gifted learners*. Needham Heights, MA: Allyn and Bacon.

- Voss, J. E. & Post, I. A. (1988). On the solving of ill-structured problems. In M. T. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp 261-287). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappa*, 27-32.
- Yussen, S. R. (1985). The role of metacognition in contemporary theories of cognitive development. *In Metacognition, cognition, and human performance, Vol 1*, pp. 253-283. Hillsdale, NJ: Lawrence Erlbaum.

Bulletin of Special Education 2000, 19, 279—308
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

MEASURING THE GIFTED CHILDREN'S PROBLEM- SOLVING ABILITIES: A MODEL OF PERFORMANCE ASSESSMENT

Chin-hsieh Lu

National Taipei Teachers College

ABSTRACT

The purpose of this study was to construct a model of performance assessment for measuring gifted children's problem-solving abilities. Through solving three kinds of performance assessment activities (well, ill and un-structured), children's problem-solving processes were observed and analyzed. Based on the absolute and expert models, children's performances on these activities were analyzed and categorized into different levels for constructing the specific scoring rubric and observing indicators. The data were collected from 57 gifted and no-gifted children enrolled in four elementary schools with pull-out gifted programs in November-December, 1998. The results demonstrated four main characteristics of the performance assessment model: (1) Though the difficulty level of the performance assessment was higher than expected, it was able to detect the discrepancy between the intuitive and organized solutions; (2) Children's problem-solving processes were very diversity and multiple. The discrepancies of the performances between expert and novice children were detected; (3) The children's responses were able to be ordered into vary levels of the indicators of problem-solving abilities based on the generic scoring rubrics; (4) The influences of the activity structure were varied based on the characteristics of academic areas and children's problem-solving abilities. Based on these results, three suggestions for implementation were proposed: (1) The results of performance assessment is appropriate to use as an identification indicator of gifted children with well-trained teacher for observing and scaling. (2) Children's attitude and commitment become the other keys factor for good performance on performance assessment when the difficulty level of the assessment activity

is high. (3) For large-scale assessment, the activity pool is need for valid use of the performance assessment as an indicator of the identification of gifted children.

Key words: Gifted children, Problem-solving Abilities, Performance Assessment, Cognitive Complexity