

國小四年級數學資優生數學成就測驗鑑定工具之編製

呂玉琴

國立臺北教育大學數學暨資訊教育學系教授

侯成龍

新北市樹林區武林國民小學教師

摘 要

本研究在發展鑑定國小四年級數學資優生的鑑定工具，編製「數學成就測驗」45 題，以國小四年級數學高成就學童為研究對象，施測樣本 235 名，施測時間為 80 分鐘。在信度考驗方面，Cronbach α 係數為.90，Rulon 折半信度為.91，顯示有良好的內部一致性及穩定性。在效度考驗方面，以內容證據、反應過程證據、內在結構證據、與其他變項的相關證據四方面加以驗證，顯示擁有良好的效度考驗。最後另以北部地區 10 所國小，進行常模參照截斷分數之建立，在 318 位學童中平均答對題數為 13.64 題，標準差為 5.95，以正二個標準差訂出國小四年級「數學成就測驗」截斷分數為 25.54 題。

關鍵字：數學成就測驗、數學資優生、鑑定工具

通訊作者：呂玉琴 leu@tea.ntue.edu.tw

壹、前言

資賦優異教育的主要目的是希望對於具學習潛能優異而無法在普通課程中受益的學生，提供適性教育的機會及引導，使其發揮潛能。因此，在資優教育中，如何鑑定需要接受特殊服務的學生，便成爲首要面對的問題。

民國 95 年中部縣市舉辦資優班聯合招考，遭教育部制止，地方縣市政府、家長、教育部各自堅持立場互不退讓，引起媒體輿論沸沸揚揚。爲正本清源，教育部舉行「全國資優教育發展研討會」並修正「特殊教育設施及人員設置標準」，其中第十條條文確立國民教育階段資優班以分散式辦理爲限，96 學年度起資優生不得集中編班。另外，教育部（2006）修正「身心障礙及資賦優異學生鑑定標準」部分條文，其中第十四條至第十八條條文同時將各種資優類別鑑定的標準提高，由原先正「一點五」個標準差，提高到正「兩」個標準差以上；並在第二條修文中規定除了一般智能及學術性向優異學生的鑑定之外，其他各類資優學生，不得以學科成就測驗鑑定。

長久以來，社會上許多人認爲資賦優異教育只是爲少數人做分類及做錦上添花的工作，違反教育公平的原則，故提出廢除資優教育回歸普通教育（McDaniel，1993）。其實鑑定資優學生的目的除在發掘普通教育情境中無法充分發展的學生外，更重要是在發現具有發展潛能的學生。林幸台（1996）指出，潛能應不止智力一種，舉凡人類所擁有之能力與表現，均屬於可能發展的範圍。因此資優鑑定乃是以學生的權利、潛能之發展爲目標導向所進

行的必要措施。

資優學生的資優特質鑑定有特殊測量工具。目前國內各中小學在鑑定資優學生時，採用的正式評量工具包含智力測驗、特殊性向測驗、成就測驗及創造力測驗，當然亦包含非正式的評量工具或資料，如學習行爲特質觀察量表、教師推薦資料、口試及觀察成績等。亦即，國內在鑑定資優生時，不管在工具及步驟方面，均具有多元的觀點。然而郭靜姿（1996）指出，國內對於資優概念的探討、測驗工具的鑑定效度，以及教育安置方式的合適性，都是值得再深入探討的課題。

郭靜姿（1995a）亦指出，國內以往在鑑定資優生頗看重智力測驗的結果，但智力測驗較適合鑑別年幼而未顯現特殊性向的孩子。相較於智力測驗，成就測驗因測驗的內容本身就與學校學習內容同質性較高，較能預測學業表現，可鑑定學術性向優異資源方案所欲服務的學科專長資優生。

在國內，依據「身心障礙及資賦優異學生鑑定標準」第十五至二十條，將資賦優異分成 6 大類，分別爲一般智能優異、學術性向優異、藝術才能優異、創造能力優異、領導才能優異、及其他特殊才能優異（教育部，2012a）。國小階段則以一般智能優異爲主，國高中大都採學術性向優異。教育部（2012b）修正「特殊教育學生調整入學年齡及修業年限實施辦法」，規定資賦優異學生可以免修該科、逐科加速、逐科跳級、各科同時加速、全部學科跳級等 5 種多元彈性方案。雖然如此，但在國小階段多屬一般智能優異類，且多採分散式安置措施，課程設計採外加方式（即每週可有 2 至 6 節的資優充實課程，抽離時數視學生需要調

整，或利用濃縮課程所節省的時數進行加速或加廣教學），因此國內對國小學術性向優異學生的發掘與鑑定較不重視。

以近年來由新加坡華僑中學所主辦的亞太地區國際性數學競賽為例，參加資格為國小五、六年級的學生，參加的國家包含臺灣、新加坡、南韓、澳洲...等十餘國，數千位參賽選手。我國自 2002 年參賽以來，歷年來表現優異屢奪佳績，更多次獲得團體冠軍。由此看來，雖然在國內國小階段並不強調學術性向優異學生的發掘與鑑定，但實際在國小校園中除一般智能優異類資優生外，實際上亦包含許多擁有特殊學術性向的學生值得去發掘並培育，尤其在數學資優方面，深具數學潛力的資賦優異兒童，更是國家民族的瑰寶，也是人類社會進步的泉源（陳龍安，1989）。以臺灣有限的人力、土地、物質資源，實在更應著重人力資源的開發，透過教育的機制，挑選各類智能優異的學生，給予適性的學習方案，培養成為對國家建設有貢獻的人才（吳克典，1999）。

朱中梧（2002）指出，在其擔任一般能力資優資源班數理科教師多年期間，發現一般能力資優生雖然都是經過正式甄試而錄取，但仍存在著許多個別差異，每位學童所專精的科目並不相同，尤其到了五、六年級，個別間的差異更為明顯。新加坡認為五至六年級為學童之定向階段（orientation stage）（翁婉珣，2005），在小學四年級結束前，學校會依據學生的學習能力實施“分流課程”分別施以不同難度的課程。因此四年級為適合鑑定數學資優生之階段，以利高年級階段的安置及教育。

綜觀國內，鑑定學術性向資優生有其迫切性，也因此需要優良之鑑定工具。郭靜姿、吳

淑敏、侯雅齡、蔡桂芳（2006）在全國資優教育發展研討會，有關「資優生鑑定與安置」的議題中的引言報告指出，教育部近年來雖已積極委託資優教育學者專家編製評量工具，然而資優鑑定工具在部分類別上仍有欠缺，如學術性向資優班、提早入學或縮短修業年限所需之個別智力評量工具。即使目前已有的各種評量工具也常因使用多年、保密性不再、仍需加速重編，以應學校鑑定之需。由此可知，在國小階段數學資優生的鑑定工具仍相當缺乏，實有待特教學者與數學教育學者攜手努力共同編製。因此，編製一有效的數學學術性向優異鑑定工具，可說是刻不容緩之課題。

探究資優生學術性向為何方面？除教師平時課堂觀察學生表現，與觀察學習特質之外，依據教育部（2006）「身心障礙及資賦優異學生鑑定標準」規定有三個途徑，其中之一為“學生在某領域學術性向或成就測驗得分在平均數正二個標準差或百分等級九十七以上，經專家學者、指導教師或家長觀察推薦，並檢附專長學科學習特質與表現等具體資料者”。因此可得知成就測驗與學術性向測驗在資優鑑定的過程中佔有舉足輕重的重要性。

除此之外，數學成就測驗亦可用來預測數學資優生未來之表現，如郭靜姿（1995b）以臺灣區學術性向優異班 21 所學生為樣本，做資優生鑑定工具得分間之相關及對於學業成就之預測分析的研究中指出，國中數學資優生的鑑定工具瑞文氏測驗、羅桑語文測驗分數、及數學性向測驗分數與學校自編各種段考測驗得分間的相關幾乎都為負相關，但未達顯著水準。數學段考對數學學業成就預測能力較高，而數學段考為教師自編成就測驗，為數學

成就測驗之一種。

楊麗華（1997）以“柯氏國民小學數學測驗”來探討充實制教學對資優兒童數學能力的影響，其第一年前測的結果，三年級資優班與普通班數學成就在應用與總分達顯著性差異，第二年後測，四、五、六年級資優班與普通班的數學成就無論在概念、計算、應用皆達顯著性差異。

國內目前已出版之數學成就測驗，如“柯氏國民小學數學科成就測驗”（1994）、國民小學三年級數學科成就測驗（1993）、國民小學低年級數學學習障礙兒童篩選測驗（1992）、國民小學數學能力發展測驗—初級、中級、高級（1987）、國小中年級數學成就測驗（1983）……等，年代大都久遠，大都針對 64 年版課程綱要所編製，並非針對現行九年一貫數學課程正式綱要所編製，實有重新編製之需要。

以成就測驗鑑定學術性向優異學生，Lohman（2006）指出大多數在能力成就測驗上居前面百分比的孩子卻常常無法保持這樣的狀態超過一、二年，即使在高信度的測驗亦是如此。Lohman 指出孩子退步主要是因個體成長上的差異、測量上的錯誤、測量假設上的錯誤、分數量尺（score scale）內容發展的改變及年級及級數人口標準數的改變。因此，資優鑑定必須避免發生類似的錯誤，用來辨認特殊學生的程序，必須避免“回歸效應”（regression effect）的產生。

其中，在測量假設上的錯誤方面，測驗分數最頂端與最低通常會產生較大的誤差，但以前的研究總是忽略最頂端與最低的誤差。尤其是測驗的編製只要忽略掉一個測驗細目，就會

帶給成績分級很大的影響，也因此如果一個班級大都使用相同等級的測驗且適合大多數學生程度的測驗，對資優生而言通常都太容易了。以使用試題反應理論（IRT）來分級的測驗的 Otis-Linnon 學校能力測驗和 CogAT 來說，也都顯示出此假設性測驗誤差。但根據 Lohman 和 Hagen（2002）的研究指出，在 CogAT 語文分級成績得 221 分，但測驗 Level A 中的測驗標準誤（standard error of measurement, SEM）為 14.8，可是在 Level D 卻只有 7.4，也因此給予程度較高的測驗可以減半預期中的測驗標準誤，所以假設性測驗標準誤可因為給予資優生較高程度的測驗而大幅減少。

Martin（1985）指出最大的回歸（regression）大多由於測驗標準誤，主要發生在第一、二年，三年級綜合成績在前 3% 的學生，只有 40% 可以維持到四年級，然而 ITBS 綜合成績擁有很高的信度（0.98），並且每年表現高度穩定性（3 到 4 年級 $r=.91$ ），但是這樣的情形依然會發生，就像所預料的，回歸效應在閱讀、語言、數學分科成績上會比在綜合以上所得的成績上有較大的影響，第一年，每個分數內容跌落了將近 50%，然而，每個年級的回歸是以較慢的比例進行著。

綜合上述，研究者將鑑定數學資優生的評量內容範圍提高一至二個年級，發展國小四年級數學資優生之成就測驗鑑定工具，期望研究結果能提供國內資優教育人員參考，以增進鑑定之效度，或作為未來研究的參考。

本文將先介紹編製數學成就測驗鑑定工具的架構，接著依序介紹鑑定工具的編製過程與施測、鑑定工具的信、效度及北部地區常模

參照截斷分數之建立，最後提出本研究的結論與建議。

貳、編製數學成就測驗鑑定工具的架構

數學成就測驗鑑定工具的編製架構主要考量下列三個面向：測驗內容涵蓋的年級範圍、鑑定工具的內容向度以及試題的題型與題數。

本研究欲發展國小資優生數學科成就測驗鑑定工具，測驗的目的在測量學童對已學得的數學智能或對已學得數學知能的推論能力。因此須顧及學校數學課程內容的同質性，並參考九年一貫數學課程正式綱要，如此才能預測其學業表現，挑選出數學才能優異學生，並且也能將數學成就表現優異，而智力測驗表現結果在標準以外，而被排除於資優教育門外的學生納入資優教育服務對象。

除依據九年一貫數學課程正式綱要之外，為避免「天花板效應」及鑑定回歸效應的產生，本研究成就測驗將其試題提高至五年級課程範圍。

九年一貫數學課程正式綱要將數學分「數與量」、「幾何」、「代數」、「統計與機率」和「連結」五個主題。因「連結」指的是數學內在結構的連結，及數學在生活情境以及和其它學科的連結，無特定的數學學科知識，所以課程綱要依數學學科知識主要分成四大主題，其中「數與量」主題在國小數學課程中具有最重要的位置，其主要概念的形成以及演算能力的培養均奠基於國小階段。「幾何」方面則是參考幾何歷史發展的軌跡與學童認知發展階段，盡量讓學童發揮、拓展其幾何直覺，在操作中，

認識各種簡單幾何形體與其性質，再慢慢加入簡單的推理性質與彼此之間的關係。國小的「代數」部分則是題材較少，主要是運用未知數作數學表示式、認識變數的概念、理解等量公理等。「統計與機率」則以學生的生活經驗為主，從學生感興趣的主題出發，使其學會敘述統計所呈現出的數字和圖表的意義，強調圖表的表達和溝通，並了解抽樣、機率的初步概念，且能正確地運用各項統計資料於實際的生活中。

本研究參考數學專家王富祥教授及數學教育專家陳光勳教授、張淑怡教授、陳幸玫教授的建議，依據數學學科之本質及考量九年一貫數學課程正式綱要中，「數與量」占課程中極大比重，將「數與量」分成「數」、「量」二個向度的分測驗，另考量「代數」在國小階段所占比重不多，且有許多指標與「數與量」共用，因此將「代數」納入於「數」的分測驗，至於「幾何」則歸為「形」的分測驗，另外「統計與機率」於國小階段所占比重極少，且機率大多屬於國中課程，因此本研究不予納入。亦即本數學成就測驗含數、量、形三個內容向度的分測驗。

本研究成就測驗的題目類型都是單一選擇題，主要是因選擇題適合大量的試題，涵蓋課程的範圍較廣，內容的取樣也較具代表性，二來資優學生的鑑定有其標準，選擇型題型亦較適合大量施測並建立常模，因此本研究設計數學成就測驗三個分測驗各 15 題試題，合計 45 題。

數、量、形三個向度題數分配均等，主要依據特殊教育專家呂金燮教授的建議，其考量有二：

一、數學資優生之特質

每位資優生所展現之天份並不相同，有些資優生展現在「數」，如數字的歸納規律...等；有些資優展現在「形」，如空間的概念、圖形推理...等，需給予每一類型的資優生，都有平等機會展現其天分。因為如依照九年一貫數學課程正式綱要比重分配各向度試題題數，則會出現「數」向度題目較多，「量」及「形」向度題目較少。為了兼顧各種數學資優特質之學童，不致於因某一類的試題較少，而被排除在外。

二、內容代表性

考量各向度試題題數，如依照九年一貫數學課程正式綱要比重加以分配，各向度試題並不均等，則試題題數較少之向度，較不具內容代表性，不利測驗統計分析。

參、鑑定工具的編製過程與施測

研究者根據研擬之數、量、形三向度、九年一貫數學課程正式綱要並參考/修改國內國小數學相關研究之試題來編製本研究的測驗工具，編製過程包括決定參考/修改試題的原則及二次的預試，並根據預試結果修改鑑定工具後進行施測。

本研究數學成就測驗研究工具，參考/修改自(呂玉琴、譚寧君, 1996)、洪碧霞(1996)國小數學診斷測驗、劉曼麗(2001)國小學童的小數知識研究、林珮如(2002)國小學童因數解題與迷思概念之研究所發展出的筆測試題、詹婉華(2002)國小高年級學童分數概念之探究、何健誼(2002)直觀法則對 K-6 年級學童在體積概念學習上的影響、許嵐婷(2003)

國小五年級面積概念之教學研究，所參考/修改之試題原本其研究已對其研究樣本進行施測，並分析答對率及學童錯誤類型，具有相當的信度與效度。又因本研究之目的在鑑定數學資優生，因此試題難度需較高，才能測量出資優生與一般學生之差異，所以本研究參考/修改這些相關研究測驗題目之原則，是以在其原本研究中五年級答對率低於五成、四年級答對率低於四成爲主。以往這些測驗試題主要是診斷學生學習困難及學習迷思概念，研究者參考這些研究試題歸屬、信效度分析、困難度，認爲許多試題可以發展爲鑑定國小四年級數學資優生數學成就測驗試題。而發展過程中需對其原本試題加以修改，其修改依據主要如下：

一、因應不同版本數學課程結構的改變

國內數學課程歷經 64 年版、82 年版、九年一貫暫行綱要和九年一貫正式綱要的變革，許多數學概念在課程中的編排不盡相同，例如“概數”在 64 年版課程原爲第八冊第五單元(四年級)之教材，但在九年一貫正式綱要課程，「概數」爲第九冊(五年級)之教材，因此必須參酌目前課程編排做適度調整與修改。

二、數學文字表徵及專有名詞的修改

本研究成就測驗將試題範圍提高至五年級範圍，主要目的在評量學生是否可運用四年級所學習的數學知識與能力，來對其超出所在年級課程試題進行解題，因此本測驗題目對於數學專有名詞應盡量避免使用，期望檢測學生問題解決、應用已學數學知識及推理的能力，不希望學生被許多數學專有名詞所困惑。

三、試題呈現方式

本研究成就測驗的題型為單一選擇題，所參考試題在其原本研究中，有些是以數學文字題方式呈現，大都需要算式及說明想法，因此需要修改，以符合本研究需求。

四、情境與語意表達

本研究所參考之研究，如洪碧霞（1996）國小數學診斷測驗，因經歷時間變革，題目情境與語意之表達與現在學童認知或許會有些許差異，又如詹婉華（2002）之研究，因研究目的之不同，題目呈現具有引導操作之情境，因應本研究需要則予以修改。

五、選項的修改

本研究之單一選擇題均為四選項，有些試題在其原本研究中為五選項試題，且已對研究樣本進行學童錯誤類型分析，故本研究之選擇題選項均以原試題中答錯率較高的錯誤類型為優先考慮之選項。

本研究「數學成就測驗」為涵蓋九年一貫數學課程正式綱要五年級數學課程架構，發展出 45 題試題。為了解題目語意是否通順、題目中干擾學童作答之文字、圖形訊息，以及降低選擇題題型學生可能較易猜題而會降低研究之信效度之因素，研究共進行二次預試。分別取樣五年級及六年級各一班，二次預試取樣人數皆為 67 人，並以 SPSS for Windows 12.0 及 TESTER2.0 進行題目分析。其分析說明如下。

第一次預試的目的，主要是了解將各種題型題目修改成選擇題型式後學生作答情形，其中答對率（P）值介於.06 至.85 之間，低於.40 以下共有 25 題，整份試卷的平均答對率

為.39，難易度偏難。難度最高之試題為第 24 題的.03，最低為第 18 題的.84，整份測驗的平均難度為.40。在鑑別度方面，整份試卷的平均鑑別度為.38。鑑別度低於.30 的有 13 題，佔總題數 29%，其中在第 39 及 43 題更是出現鑑別度為負數的情形。研究者針對 13 道鑑別度低於.30 之試題，與高分組學童進行訪談，以了解施測工具題目語意的描述是否符合學童的語彙程度？參酌訪談內容並與數學教育專家討論，進行題目之修改或更換。

第二次預試主要進行選項分析及二次預試比較，比較题目的修改或更換是否合適。其中答對率（P）值介於.12 至.79 之間；答對率低於.40 以下共 27 題，整份試卷的平均答對率為.37。難易度偏難。難度最高之試題為第 26 題的.13，最低為第 17 題的.78，整份測驗的平均難度為.40。在鑑別度方面，鑑別度低於.30 的有 12 題，佔總題數 26.6%，鑑別度並未再出現負數的情形，整份數學成就測驗的平均鑑別度為.40，也因此僅進行題目之修改。

研究者將數學成就測驗第二次預試作答資料，以 TESTER 2.0 統計軟體進行每一道試題的選擇題選項有效性分析，郭生玉（1990）指出每題不正確選項至少有一個低分組的受試者選它，與不正確選項低分組比高分組更多人選擇為選項誘答力的兩個原則，在本研究中大部份的題目選項都符合「不正確選項至少有一個低分組選它」，及「不正確選項低分組比高分組更多人選擇」。

二次的預試過程，皆根據前次施測結果徵詢國立台北教育大學一位數學教授、三位數學教育教授，逐一審查測驗題目，以了解數學成就測驗試題是否符合九年一貫數學課程正式

綱要。研究者就四位教授的建議事項逐一修改試題，以得到更高的信度，建立正式「數學成就測驗」鑑定工具。

本研究編製之「數學成就測驗」其目的在鑑定數學資優生，其試題難度偏難，研究採團體施測，對象訂為國小四年級學童，條件限制在其三年級上、下學期數學學業成績皆為「優等」之數學高成就學童，共 235 名。其中男生 174 名，女生 61 名；依縣市區分：台北市 103 名、新北市 95 名、桃園縣市 13 名、新竹縣市 12 名，其他縣市共 12 名。施測時間為 80 分鐘。

本數學成就測驗的題目類型為選擇題，每答對一題得一分，得分越高表示學童有越高的數學成就。

表 1

數學成就測驗之測驗分析表

受試者總數 (N)	235.00	變異數 (樣本)	79.26
問題總數 (K)	45.00	標準差 (SD)	8.92
平均答對率 (P)	.45	標準差 (樣本)	8.90
變異數 (Variance)	79.60	內部一致性係數	.90

(二) Rulon 折半信度

本研究另將「數學成就測驗」的四十五題試題的難度依大小順序排列，再按排列後的奇偶數題分成前後兩半測驗，把全體受試者在前後兩半測驗答對題數相減得到差異分數 (D)，並求得這些全體受試者差異分數 (D)

肆、鑑定工具的信、效度

一、鑑定工具的信度

本研究信度分析，探測驗內部的一致性，因為本測驗題目為選擇題，為非對即錯的測驗類型，所以不論用庫李 20 號信度 (KR_{20}) 或 Cronbach α 係數，求得的信度係數都是相同的。因此，本研究使用的考驗方法為 Cronbach α 係數以及 Rulon 折半信度。

(一) Cronbach α 係數

研究對象在「數學成就測驗」的測驗結果以 TESTER 2.0 統計軟體分析，其內部一致性 Cronbach α 係數值為 .90，達 .01 顯著水準，如表 1 所示。顯示本測驗內部一致性良好及具穩定性。

的變異數 σE^2 ，再求全體受試者答對題數的變異數 σX^2 ；帶入 Rulon 公式計算，如表 2 所示，求得 Rulon 折半信度為 .91，亦顯示內部一致性良好及具穩定性。

$$\rho = 1 - \left(\frac{\sigma_E^2}{\sigma_X^2} \right)$$

表 2

數學成就測驗 Rulon 信度分析表

差異分數		答對題數	
人 數	235	人 數	235
差異的平均數	-0.40	平均答對題數	20.23
標準差	2.64	標準差	8.92
變異數 σE^2	6.96	變異數 σX^2	79.60

二、鑑定工具的效度

本鑑定工具的效度包括測驗內容證據、反應過程證據、內在結構證據及其他變項的相關證據（王文科、王智弘，2005）。

（一）鑑定工具的測驗內容證據

「數學成就測驗」在測驗內容代表性方面，依據數、量、形三個內容向度及 Anderson 和 Krathwohl（2001）所修訂 Bloom 認知領域教育目標之認知歷程向度為認知向度，編製雙向細目表，參考相關研究及文獻，與數學教育

專家一同研商，確認了測驗的內容證據，如表 3。

雙向細目表中「認知歷程」部份，研究者只採用 Bloom 認知歷程向度中之「瞭解、應用、分析、評鑑」四個向度，其主要原因為施測對象為數學資優生，目的在測其數學能力，而「記憶」是屬於較簡單之認知程度，故將「記憶」向度刪除，「創造」向度是屬於較複雜之認知歷程，包含需產生—提出假設、規劃—設計、製作—建造，亦不適合以選擇題方式出現，也予刪除。

表 3

數學成就測驗雙向細目表

數學內容	認知歷程				小計
	瞭 解	應 用	分 析	評 鑑	
數	2、8、31	1、7、14、19 、25、32	13、20、38、 43、37	26	15
量	4、21、39	9、15、27、28 、40、44	3、10、22、33、 34	16	15
形	11、23、24	5、30、35、 42	6、12、18、 29、41、45	17、36	15

(二) 鑑定工具的反應過程證據

本研究在第一次與第二次預試後，就鑑別度較低之試題，直接與高分組學生中答錯的學童進行訪談，了解題目中干擾學童之因素，研究者依據訪談的內容修改或更換測驗試題。

正式施測後，研究者依據學童作答分析，

將數學成就測驗題目中所欲測驗學童已學得的數學知能之應用與推論能力，與學童表現及反應做相互驗證，如表 4 所示，本研究「數學成就測驗」除第 6、18、29、31、38 題，合計 5 道試題鑑別度低於 0.3，其餘試題皆擁有良好之鑑別度。

表 4

數學成就測驗數學知能之應用與推論能力分析表

題號	數學知能之應用與推論能力	答對率 N=235	高分組 答對率	低分組 答對率	鑑別度
1	未學過異分母加法前，能運用原本的分數知識進行推理解題。	.36	.67	.20	.46
2	單位量不同，不能直接透過分數大小的比較來決定分數量的大小。	.47	.76	.14	.61
3	在錢幣的情境中，利用幣值解題。掌握題目訊息，使用倍數判斷法解題。	.78	.98	.51	.48
4	能將題目訊息，發展速率的意義概念—單位時間之內所走的距離，以進行解題。	.55	.76	.30	.45
5	測驗學童對於正立方體的面的垂直關係和平行關係，以及對面數、視圖的理解。	.60	.91	.25	.66
6	對於各種不同三角形性質的判斷。	.24	.42	.14	.28
7	小數的除法中，是否具有單位量轉換的能力，能夠將被除數與除數同時轉換單位。	.30	.70	.07	.63
8	對於全部是 1 的概念是否清晰。	.45	.80	.16	.64
9	測驗學生是否完全理解單位面積的意義而算出圖形面積。	.24	.48	.07	.41
10	學童能否發現圖形間的關係，在一個長、寬並不相等的長方形中，如何才能剪出一個面積最大的正方形。	.61	.94	.26	.68
11	等腰三角形的內角的求法。	.86	1.00	.65	.35

(續) 表 4

數學成就測驗數學知能之應用與推論能力分析表

題號	數學知能之應用與推論能力	答對率 N=235	高分組 答對率	低分組 答對率	鑑別度
12	學童是否可以查覺直角三角形的斜邊邊長大於另二個邊長或是長方形對角線大於兩邊邊長。	.56	.86	.28	.59
13	是否能由列式中，觀察乘法的分配律，將其中共同的乘數提出，將被乘數先做運算。	.55	.89	.25	.65
14	對於不同位數小數的大小比較，及是否運用小數的減法來進行解題。	.35	.74	.10	.64
15	計算空心水泥柱的體積。	.33	.64	.14	.49
16	學童曾學過體積是可測量的，並能在測量後以單位正立方體的幾倍來表示體積。	.23	.50	.10	.40
17	是否了解四邊形的性質。	.77	.95	.55	.40
18	是否可運用切割或移補的概念，來判斷及比較圖形的大小。	.58	.74	.45	.29
19	使用四捨五入法取概數到指定位數再計算的能力。	.40	.62	.23	.39
20	可否在四個包含不同未知數的等式中，推論出未知數的相對大小。	.50	.76	.20	.56
21	排他性體積求法，但又涉及保留概念、測量單位的概念。	.25	.53	.10	.43
22	用兩個天平來呈現三種物體間的重量關係，進行兩物體重量的比較，涉及了遞移比較的因素。	.53	.79	.29	.50
23	理解三角形的內角和為 180 度，而 180 度等於二個直角。	.65	.89	.35	.55
24	是否可瞭解圓周長與直徑的關係，來比較不同圓形的圓周長。	.74	.89	.58	.31
25	乘法分配律的應用，掌握題目訊息正確處理倍數關係。	.49	.86	.19	.68
26	學童對於小數除法的理解，能否真正了解除數與商的關係。	.33	.59	.13	.46

(續) 表 4

數學成就測驗數學知能之應用與推論能力分析表

題號	數學知能之應用與推論能力	答對率 N=235	高分組 答對率	低分組 答對率	鑑別度
27	時和分的分數、小數化聚。	.32	.61	.10	.51
28	是否瞭解圓柱體的側面展開圖為一長方形，其表面積為圓周長乘以高。	.58	.82	.41	.41
29	可否藉由題目的圖形提示，理解外角的意義，並運用圖形內角和及平角的關係來求出外角和。	.09	.17	.09	.08
30	是否可了解球體與長方體間的形體關係，發現長方體的箱子的邊長恰巧等於五個球的直徑。	.63	.91	.32	.59
31	理解題目「只能被 1 和自己整除」的訊息，並推理一套判斷的策略，來找出 10 至 40 以內質數的個數。	.26	.38	.16	.22
32	是否可以藉由平分的概念，找出整數的所有正因數個數。	.20	.41	.07	.34
33	速率單位的化聚。	.42	.76	.10	.66
34	學童是否可以由大小兩個立方體邊長的倍數比，來推理出這兩個大小立方體的體積倍數比。	.18	.42	.03	.40
35	可否由每個選項所給定的邊長，來判斷哪一個選項「不能」圍成三角形，主要的判斷依據為三角形任意兩邊邊長和大於第三邊。	.55	.80	.38	.43
36	是否由各小題給定的圖形構成要素，判斷何者「不能」使對摺後的圖形重合，也就是要找出何者不是對稱圖形。	.44	.61	.22	.39
37	學童在未學過分數與小數互換的課程前，是否可以運用完整的分數及小數概念，推理比較分數與小數的大小。	.65	.85	.54	.31
38	解公倍數文字題的能力。	.19	.30	.14	.16

(續) 表 4

數學成就測驗數學知能之應用與推論能力分析表

題號	數學知能之應用與推論能力	答對率 N=235	高分組 答對率	低分組 答對率	鑑別度
39	長方體的體積計算中，題目所給定的邊長單位並不相同時，是否會先行做長度單位的換算，再計算長方體的體積。	.33	.62	.16	.46
40	複合立體圖形體積求法的能力。	.24	.56	.04	.52
41	學童是否理解對稱軸兩側圖形全等，由題目所給定的對摺圖形中，發現展開的圖形為一等腰三角形。	.70	.89	.49	.40
42	是否認識圓心角及扇形，及圓心角與圓內接正多邊形的關係，由圓心角判斷出圓內接正多邊形的邊數。	.32	.74	.10	.64
43	學童對於分別使用結合律及交換律的不同算式中，查覺何者的運算結果與 $2\frac{1}{9} - \frac{1}{25} - \frac{1}{7}$ 有所不同。	.43	.82	.19	.63
44	時間的運算能力。	.72	.91	.48	.43
45	是否可利用切割的方法，由多邊形內切割出的三角形數目，判斷多邊形的內角和。	.27	.59	.06	.53

(三) 鑑定工具的內在結構證據

內在結構證據涉及評量工具和題目間的相關，以及同一工具各部分間的相關。本研究以內部一致性分析來加以驗證，得到以下結論。

1. 在試題內的一致性方面

將全體受試者中的高分組與低分組，在每一道試題上作答的答對率做一比較，以 t 考驗分析兩組答對率的差異性後，發現「數學成就測驗」除第 29 題之外，高分組、低分組差異皆達 .01 顯著水準，顯示「數學成就測驗」具有高度的內部一致性，如表 5。

表 5

數學成就測驗高低分組平均數差異考驗分析表

題號	高低分組	平均答對率	標準差	t 值
1	低分組	.20	.41	6.09**
	高分組	.67	.48	
2	低分組	.14	.35	8.99**
	高分組	.76	.43	
3	低分組	.51	.50	7.64**
	高分組	.98	.12	
4	低分組	.30	.46	5.87**
	高分組	.76	.43	
5	低分組	.25	.43	10.48**
	高分組	.91	.29	
6	低分組	.14	.35	3.74**
	高分組	.42	.50	
7	低分組	.07	.26	9.59**
	高分組	.70	.46	
8	低分組	.16	.37	9.72**
	高分組	.80	.40	
9	低分組	.07	.26	5.93**
	高分組	.48	.50	
10	低分組	.26	.44	11.14**
	高分組	.94	.24	
11	低分組	.65	.48	6.02**
	高分組	1.00	.00	
12	低分組	.28	.45	8.54**
	高分組	.86	.35	
13	低分組	.25	.43	10.01**
	高分組	.89	.31	
14	低分組	.10	.30	9.79**
	高分組	.74	.44	

(續) 表 5

數學成就測驗高低分組平均數差異考驗分析表

題號	高低分組	平均答對率	標準差	t 值
15	低分組	.14	.35	6.70**
	高分組	.64	.48	
16	低分組	.10	.30	5.53**
	高分組	.50	.50	
17	低分組	.55	.50	6.15**
	高分組	.95	.21	
18	低分組	.45	.50	3.61**
	高分組	.74	.44	
19	低分組	.23	.43	4.93**
	高分組	.62	.49	
20	低分組	.20	.41	7.70**
	高分組	.76	.43	
21	低分組	.10	.30	5.96**
	高分組	.53	.50	
22	低分組	.29	.46	6.66**
	高分組	.79	.41	
23	低分組	.35	.48	7.89**
	高分組	.89	.31	
24	低分組	.58	.50	4.43**
	高分組	.89	.31	
25	低分組	.19	.39	10.57**
	高分組	.86	.35	
26	低分組	.13	.34	6.27**
	高分組	.59	.50	
27	低分組	.10	.30	7.13**
	高分組	.61	.49	
28	低分組	.41	.49	5.40**
	高分組	.82	.39	

(續) 表 5

數學成就測驗高低分組平均數差異考驗分析表

題號	高低分組	平均答對率	標準差	t 值
29	低分組	.09	.28	1.39
	高分組	.17	.38	
30	低分組	.32	.47	8.83**
	高分組	.91	.29	
31	低分組	.16	.37	2.93**
	高分組	.38	.49	
32	低分組	.07	.26	4.91**
	高分組	.41	.50	
33	低分組	.10	.30	10.17**
	高分組	.76	.43	
34	低分組	.03	.17	6.12**
	高分組	.42	.50	
35	低分組	.38	.49	5.56**
	高分組	.80	.40	
36	低分組	.22	.42	4.95**
	高分組	.61	.49	
37	低分組	.54	.50	4.16**
	高分組	.85	.36	
38	低分組	.14	.35	2.22*
	高分組	.30	.46	
39	低分組	.16	.37	6.18**
	高分組	.62	.49	
40	低分組	.04	.21	7.79**
	高分組	.56	.50	
41	低分組	.49	.50	5.60**
	高分組	.89	.31	
42	低分組	.10	.30	9.79**
	高分組	.74	.44	

(續) 表 5

數學成就測驗高低分組平均數差異考驗分析表

題號	高低分組	平均答對率	標準差	t 值
43	低分組	.19	.39	9.35**
	高分組	.82	.39	
44	低分組	.48	.50	6.13**
	高分組	.91	.29	
45	低分組	.06	.24	7.93**
	高分組	.59	.50	

* $p < .05$

高分組人數為 66 人

** $p < .01$

低分組數人為 69 人

2.在試題間的一致性方面
將全體受試者在每一道試題的答題情形
(答對或答錯)與其答對題數,分別求其積差
相關,求出每一道題目的點二系列相關係數,

經過每一道題目的點二系列相關係數經過統
計考驗後,發現「數學成就測驗」每一道試題
均達顯著水準,顯示測驗總分與每一道試題的
得分皆成正相關,如表 6。

表 6

數學成就測驗各題與測驗相關係數

題號	點二系列相關	題號	點二系列相關	題號	點二系列相關
1	.445**	16	.498**	31	.235**
2	.487**	17	.385**	32	.427**
3	.430**	18	.261**	33	.546**
4	.372**	19	.314**	34	.545**
5	.498**	20	.459**	35	.345**
6	.288**	21	.472**	36	.347**
7	.590**	22	.417**	37	.302**
8	.537**	23	.452**	38	.175*

(續) 表 6

數學成就測驗各題與測驗相關係數

題號	點二系列相關	題號	點二系列相關	題號	點二系列相關
9	.453*	24	.323**	39	.382**
10	.535**	25	.527**	40	.531**
11	.361**	26	.447**	41	.314**
12	.470**	27	.482**	42	.604**
13	.503**	28	.373**	43	.538**
14	.569**	29	.228**	44	.342**
15	.483**	30	.460**	45	.520**

** $p < .001$

(四) 測驗工具與其他變項的相關證據

此部分在考驗某一測驗得到的分數與類似的或不同之特質相關的情形。若在某一測驗工具的得分與測量同一特質的另一種工具的得分高相關，即具有聚斂性證據 (convergent evidence)。當在某測量工具上的得分與測量不同東西的分數不具高度相關，即呈區辨性證據 (discriminant evidence) (王文科、王智弘，2005)。本研究在外在效度的聚斂性證據方面以二方面來加以驗證。

1. 智力測驗變項

本研究的目的是在發展國小數學資優生的鑑定工具，但目前國小的資優班並未特別成立數學資優班，而是屬於一般智能資優學生，不論從智能理論的觀點，或心理學及社會學的觀點來看，一般智能資優學生是經過層層鑑定及

篩選，其智力測驗分數優於普通班學生之平均水準，也因此一般智能資優學生與普通班學生在本研究數學成就測驗的表現，也應該會有所差異，一般智能資優學生的表現應該會優於普通班學生。

本研究「數學成就測驗」分別對新北市國小四年級一般智能資優學生 (17 人) 與普通班學生 (29 人) 予以施測，經以 t 檢定分析結果發現，一般智能資優學生平均答對題數皆高於普通班受試者，達到 .01 顯著水準，如表 7。藉由資優班與普通班之常態曲線 (如圖一) 及一般智能資優學生與普通班學生在數學成就測驗答題數的峰態與偏態，如表 8，可判斷一般智能資優學生優於普通班學生，亦即本研究之「數學成就測驗」擁有良好的聚斂性證據。

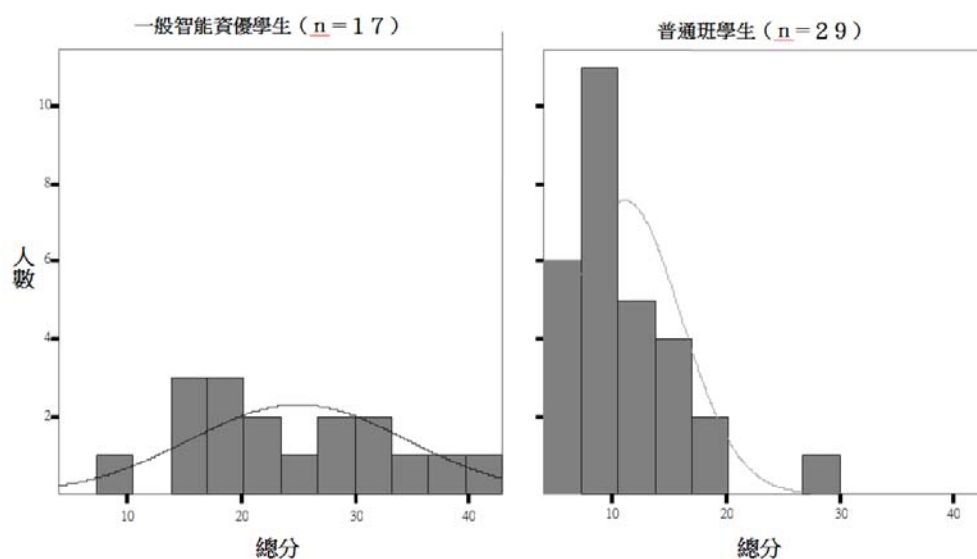
表 7

一般智能資優學生與普通班學生在數學成就測驗答題數的 t 檢定

	人數	平均 答題數	標準差	平均數估計標準誤	t 值	p 值
一般智能 資優學生	17	25.00	9.57	2.322	5.58	<.001**
普通班 學生	29	11.07	4.96	0.922		

總題數：45 題

** $p < .01$



圖一 數學成就測驗一般智能資優學生與普通班學生之常態分配曲線圖

表 8

一般智能資優學生與普通班學生在數學成就測驗答題數的峰態與偏態

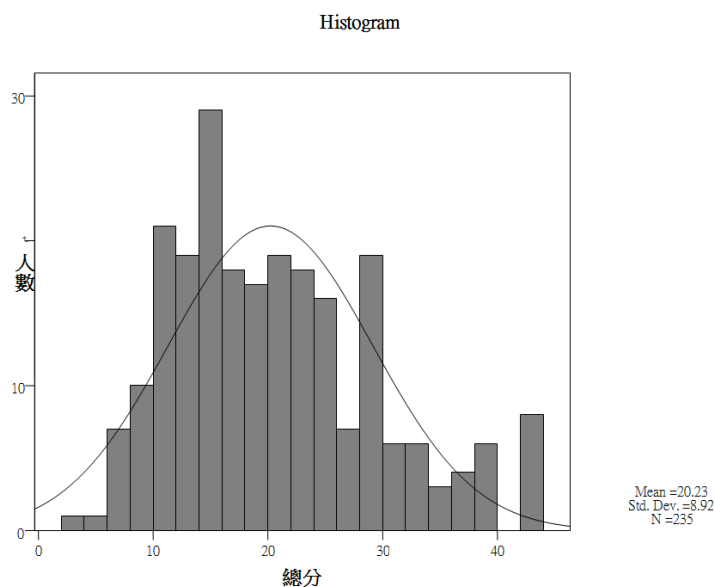
	峰態 係數	峰態標準誤	偏態 係數	偏態係數 標準誤
一般智能資優 學生	-0.743	1.063	0.209	0.550
普通班 學生	2.530	0.845	1.382	0.434

2. 受試者在鑑定工具表現

本研究的 235 位受試者為數學高成就學童，由學童的作答表現分析，發現平均答題數為 20.23 題（如表 9），由常態曲線圖（如圖 2）及答題分析（如表 10）可看出峰態係數為 -0.095 接近常態峰，偏態係數 0.667 接近正偏

態。代表分數偏低。

由此可知，數學高成就學童在接受高一年級的「數學成就測驗」時，大部份的學童得分偏向低分的一端，只有部份的學童真正具有高一年級的數學成就。



圖二 受試者在數學成就測驗之常態曲線圖

表 9

數學成就測驗答題分析表

	個數	總題數	平均答題數	標準差	平均數估計標準誤
數學成就測驗	235	45	20.23	8.92	0.58

表 10

數學成就答題數的峰態與偏態表

	峰態係數	峰態標準誤	偏態係數	偏態係數標準誤
數學成就測驗	-0.095	0.316	0.667	0.159

伍、北部地區常模參照截斷分數之建立

本研究的目的是在建立國小數學資優生鑑定工具，因此數學資優生在測驗中的表現需符合教育部（2006）修正「身心障礙及資賦優異學生鑑定標準」部分條文，其中第十四條至第十八條條文所規定資優類別鑑定的標準為正「兩」個標準差或百分等級百分之九十七以上，可界定為數學資優生。

因此，本研究另以北部地區十所學校，其中包含台北市大安區市立大安國小、私立復興小學、中山區中正國小、萬華區萬大國小、大同區日新國小，台北縣板橋區安和國小、三鶯區武林國小、七星區北峰國小、文山區深坑國小，桃園縣光明國小，每校四年級各一個班級學生，進行「數學成就測驗」施測，合計十個班級，共 318 位學童，其平均答對題數為 13.64 題，標準差為 5.95。得到「數學成就測驗」正二個標準差的截斷分數為 25.54 題。

陸、結論與建議

本測驗共編製國小「數學成就測驗」四十五題，測驗內容分「數」、「量」、「形」三個向度，根據初步的研究結果，本測驗之信度與效度頗佳，可作為國小四年級數學資優生鑑定之有效工具。

本測驗建立北部地區常模參照截斷分數，提供國小教育及國小數學資優生鑑定之參考。然而本測驗結果建議不宜單獨使用，可輔以其他的資訊，如智力測驗、性向測驗等，作

一綜合性之鑑定判斷。本研究礙於時間、財力等限制，只做北部地區常模參照截斷分數，因此各校使用本測驗時，宜參考城鄉差距、學校背景，來解釋測驗成績。

研究者根據研究結果提出下列建議

- 一、本研究所編製之「數學成就測驗」，因研究者時間、財力等限制，除針對國小四年級數學高成就學童施測，另針對北部地區十所學校進行施測以獲得北部地區常模參照，以定義出「數學成就測驗」正二個標準差之截斷分數。除此之外，應可更深入建立全國性常模參照以成為常模參照測驗，並建立不同年級全國性常模，以比較分析不同年級學生間的差異情形。
- 二、本研究考量施測人數、時間等因素，將題型設定為選擇題，因此不易知道學童解題的腦中思考策略。建議日後可朝向以不同的題型作深入的質性分析，以比較一般學童與數學資優生的解題差異，並觀察其解題特質，以做為鑑定之參考。

誌 謝

本文承行政院國科會補助，計畫編號 NSC 95-2521-S-152-009-MY3 之部分結果，感謝新北市六所國小：中和、北峰、安和、武林、深坑、樹林，台北市五所國小：大安、中正、日新、萬大、私立復興國小，桃園縣光明國小之全力支持，更感謝受試班級教師的熱心幫助，使本研究順利完成。

參考文獻

一、中文部分

- 王文科、王智弘 (2005)。教育研究法。台北：五南出版社。
- 朱中梧 (2002)。國小一般能力資優生之數學解題探究 (未出版之碩士論文)。國立台北師範學院，台北市。
- 何健誼 (2002)。直觀法則對K-6年級學童在體積概念學習上的影響 (未出版之碩士論文)。國立台北師範學院，台北市。
- 吳克典 (1999)。資優教育全方位發展策略的整合型研究。載於中華資優教育學會 (主編)，資優教育全方位發展 (1-5頁)。台北：心理出版社。
- 呂玉琴、譚寧君 (1996)。國小教師對學生數學認知知識了解之探討。行政院國家科學委員會專題研究成果報告 (編號：NSC 84-2513-S-152-002)，未出版。
- 林幸台 (1996)。資優學生鑑定與安置問題之探討與改進芻議。資優教育季刊，60，2-7。
- 林珮如 (2002)。國小學童因數解題與迷思概念之研究 (未出版之碩士論文)。屏東師範學院，屏東縣。
- 洪碧霞 (1996)。國小數學診斷測驗。臺北市：教育部。
- 翁婉珣 (2005)。台灣與新加坡之十二年數學課程比較 (未出版之碩士論文)。國立中央大學，桃園縣。
- 教育部 (2006 修訂公布)。身心障礙及資賦優異學生鑑定標準。臺北市：教育部。
- 教育部 (2012a 修訂公布)。身心障礙及資賦優異學生鑑定標準。臺北市：教育部。
- 教育部 (2012b 修訂公布)。特殊教育學生調整入學年齡及修業年限實施辦法。臺北市：教育部。
- 許嵐婷 (2003)。國小五年級面積概念之教學研究 (未出版之碩士論文)。臺中師範學院，台中市。
- 郭生玉 (1990)。心理與教育測驗。台北：精華書局。
- 郭靜姿 (1995a)。資優生鑑定效度研究的省思—再談測驗在資優鑑定的運用。資優教育季刊，56，4-12。
- 郭靜姿 (1995b)。資優學生鑑定工具得分間之相關及其對於學業成就之預測分析。特殊教育研究學刊，13，175-202。
- 郭靜姿 (1996)。資賦優異學生的鑑定與教育安置。教育資料集刊，21，27-54。
- 郭靜姿、吳淑敏、侯雅齡、蔡桂芳 (2006)。鑑定與安置。全國資優教育發展研討會手冊。臺北市：國立臺灣師範大學。
- 陳龍安 (1989)。創造思考教學的理論與實際。台北：心理出版社。
- 楊麗華 (1997)。充實制數學教學對台北市日新國小資優兒童數學能力影響之研究。資優教育季刊，65，30-35。
- 詹婉華 (2002)。國小高年級學童分數概念之探究 (未出版之碩士論文)。國立台北師範學院，台北市。
- 劉曼麗 (2001)。國小學童的小數知識研究。屏東師院學報，14，823-858。

一、英文部分

- Anderson, L.W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman.
- Lohman, D. F., & Hagen, E. P. (2002). *Cognitive Abilities Test. Research handbook*. Itasca, IL:Reversde.
- Lohman, D. F., (2006). Gifted Today but Not Tomorrow ? Longitudinal Changes in Ability and Achievement During Elementary School. *Journal for the Education of the Gifted*. 29, 451-484.
- Martin D. J. (1985). *The measurement of growth in educational achievement* (Unpublished doctoral dissertation). University of Iowa, Iowa City.
- McDaniel, T. R. (1993). Education of the gifted and the excellence equity debate: Lessons from history. In C. J. Maker (Ed). *Critical issues in gifted education* (pp.6-18). Austin, TX: PRO-ED.

The Compilation of Mathematics Achievement Test for Elementary Mathematics-Gifted 4th Students as the Appraisalment Tool

Yuh-Chyn Leu

Department of Mathematics and Information
Education, National Taipei University
of Education

Cheng-Lun Hou

Wulin Elementary School, Shulin District,
New Taipei City
Elementary Teacher

Abstract

The research is to develop the appraisalment tool for elementary mathematics-gifted students, namely the Mathematics Achievement Test (MAT). There are 45 questions in the MAT and the research samples are 235 fourth graders who demonstrate excellent performance in mathematics. The test time is 80 minutes. Regarding to the reliability, the values for Cronbach α are .90. The Rulo split-half reliability is .912. All of them indicate excellent internal consistency and stability. Regarding to the validity, the four aspects of content evidence, response process evidence, internal structure evidence and other variables all indicate satisfactory validity. The norm-reference cut-off score is established from ten elementary schools located in the Northern part and the average correct responses among these 318 normal students are 13.64 question. The standard deviation(SD) is 5.95 and choosing +2 SD as the cut-off score, students who score above 25.54 questions are qualified as the mathematics-gifted students.

Keywords: mathematics achievement test, mathematics-gifted students, appraisalment tool