

《測驗學刊》
第六十七輯第四期 2020 年 12 月 333~353 頁

小六數學素養表現分析之研究

曾建銘¹ 曾招文² 趙珮晴³

摘要

過去教學相當重視知識和能力，如今數學學習領域和 PISA 數學素養均強調數學知識與生活要能連結、實踐，讓「素養」落實教育現場，強化學生應用所學於日常生活中，因此本研究即設計一份國小六年級測試的數學素養試題，以了解小六學生數學素養表現、錯誤類型和相關背景表現差異。本文經專家會議從 PISA 2012 所公告之數學素養題本中，選取適合國小六年級的試題，並自編符合九年一貫課程綱要能力指標的 A、B 卷數學素養試題，採方便取樣 27 校 56 班國小六年級學生 1,258 名為正式施測對象，發現數學素養開放問答題的空白比例高，從錯誤類型可知計算不難，答題關鍵會落在題意的理解，對低分組學生易造成不了解、任意作答或猜測。背景變項方面，北區數學素養試題作答優於南區學生，而性別則是男性顯著優於女性學生。根據上述研究結果，提出結論和建議以供參考。

關鍵詞：地區、性別、國小六年級、國際學生能力評量計畫、數學素養

1. 曾建銘，國家教育研究院測驗及評量研究中心副研究員

2. 曾招文，桃園市觀音國民小學教師

3. 趙珮晴，大學入學考試中心統計研究員

收件日期：2019.04.13；完成修改：2020.07.29；正式接受：2020.11.03

通訊作者：趙珮晴；Email：99152513@nccu.edu.tw

地址：臺北市大安區舟山路 237 號 大學入學考試中心

A Study on the Mathematical Literacy of Grade 6 Students

Chien-Ming Cheng¹ Chao-Wen Tseng² Pei-Ching Chao³

Abstract

In the past, mathematics education focused on knowledge and skills. Now, both the Mathematical Curriculum and the PISA (the Programme for International Student Assessment) have hoped to promote a more practical usage of mathematical knowledge. The current study selected some items from the 2012 PISA mathematical literacy released items that were suitable for Grade 6 students as suggested by a group of experts. These were then combined together with two self-made assessments A and B, which were based on the Mathematical Curriculum Guidelines. A total of 1,258 grade 6 students from 27 schools and 56 classes were purposively sampled. Results showed that open-ended questions had higher ratio of blank answer. Further analysis of the error items showed that the assessment was actually not difficult to compute, but the problem lies with the understanding of the question itself. More so, with the low-scores group; students tend to have no idea about the question, which resulted in guessing and randomly selecting the answers. As for the students' background, students from the Northern region of Taiwan performed significantly better than the students from the Southern region. In addition, male students performed better than their female counterparts. According to these results, some conclusions and suggestions were proposed for practical application and future research.

Keywords: area, gender, 6th-grade, PISA (the Programme for International Student Assessment), mathematical literacy

1. Chien-Ming Cheng, Associate Research Fellow, Research Center for Testing and Assessment, National Academy for Educational Research

2. Chao-Wen Tseng, Teacher, Taoyuan Municipal Guan-Yin Elementary School

3. Pei-Ching Chao, Statistical Researcher, College Entrance Examination Center

Received: 2019.04.13; Revised: 2020.07.29; Accepted: 2020.11.03

Corresponding Author: Pei-Ching Chao; Email: 99152513@nccu.edu.tw

Address: No. 237, Zhoushan Rd., Da'an Dist., Taipei City 106, Taiwan

College Entrance Examination Center

壹、緒論

臺灣以往相當重視知識與能力，如今教學與評量不應再以學科知識技能為限，也要關注學習和生活的連結和實踐，所以課程要重視「素養」（教育部，2014；蔡清田，2010）。對於如何設計數學素養試題？鄭英豪（2013）指出，描述現實情境是必要的，亦可廣泛地鋪陳脈絡，才可強化「帶著走」的知識與能力；如此，數學就不只是數學符號操作，而是可以具體實踐應用於生活中。

國內外有許多大型測驗，依照測驗宗旨性質劃分，國內都屬於成就測驗，TIMSS（Trends in International Mathematics and Science Study）傾向能力測驗，至於 PISA（the Programme for International Student Assessment）與 PIRLS（Progress in International Reading Literacy Study）則為素養測驗，素養測驗又以 PISA 最廣為人知，本文以 PISA 數學試題為主軸，嘗試自編符合九年一貫能力指標的數學素養試題，一可藉此了解學生能力和錯誤類型，二來可作為數學素養命題參考。

目前國內 PISA 施測的學生都是以國三、高一學生為主，洪雪芬（2011）分析 PISA 樣本試題發現，若是以九年一貫課程綱要之數學能力指標來對照的話，幾乎有半數以上的題目是國小學生能力所及的。因此，本文嘗試透過專家選擇適合國小程度的試題，期許可「擷取」PISA 測驗評量的精神，讓國小學生體驗類似 PISA 測驗的數學素養試題。

另考量 PISA 數學素養可分成四大內容領域（臺灣 PISA 國家研究中心，2014，2017；Organization for Economic Cooperation and Development [OECD], 2005），而性別（臺灣 PISA 國家研究中心，2014；蔣治平、廖敏治、蘇懿，2012；龔心怡、李靜儀，2016；Voyer & Voyer, 2014）和城鄉差異（黃敏雄，2015；楊淑萍、林煥祥，2010；龔心怡、李靜儀，2016；Howley, 2003; Kifer, 2001）在數學的表現差異一直是研究學者和單位關心的。

據此，本文研究的目的是擬設計一份適合國小六年級測試的數學素養試題，一部分嘗試從 PISA 試題樣本中選取；另一部分參酌九年一貫能力指標配合 PISA 數學素養四大內容領域自編，以了解國小六年級學生數學素養表現和錯誤類型，並進一步分析在性別、地區和數學內容領域的表現差異，最後根據研究結果提出結論和建議，以茲教育單位、學者和教師參酌。

貳、文獻探討

以下將從素養與數學素養的定義開始，接著闡述九年一貫課程綱要與素養關係，最後列舉可能影響數學表現的相關因素。

一、素養與數學素養的定義

（一）何謂素養？

近年國際組織與評量常提到「素養」，尤其是國際學生能力評量計畫（PISA）（周玉秀，2006；單維彰，2016；蔡清田，2010）。臺灣自 2006 年加入 PISA 素養評比，每隔三年評比一次，每次均進行國際排名與檢討，讓臺灣教育相關單位不得不重視素養。

關於素養翻譯，單維彰（2016）文中第 4 頁指出「不論 Competence 還是 Literacy，都是指『能力』」；「Competence 較為綜合性與一般性（知識、技能與行為表現），而 Literacy 傾向於特定基礎能力（識、讀、寫、算）」。

素養可激發個體內部的認知、情意、技能的心智運作，協助個體發展成完整的個體，以因應生活情境的需求或複雜任務（蔡清田，2010，2011a，2011b），如同李國偉（2014）提到 DeSeCo 計畫以整體性觀點，將素養視為成功處理情境問題的能力，包括使用的認知知識、技能和態度動機等。蔡清田（2011b）進一步提出「課程改革中的素養模式」有四個重要構成要素：個體必須因應生活情境的複雜需求、個體內部情境的社會心智運作機制、個體的行動、「整體」互動體系。因此，「素養」可以彌補「知能」的不足（蔡清田，2011a），而「知識」是臺灣傳統教育重視的，強調素養可以更正重知識、忽略教育的缺失（蔡清田，2010）。

單維彰（2016）指出，華人的素養其實有「修養」意涵，蘊含有「知、行、識」要素，可依照學生生活經驗發展出完整脈絡；范信賢（2016）也有相同看法，認為素養是學以致用、做中學／學中做的展現。劉柏宏（2016）引《辭海》解釋素養為「平時的修養」，參酌聯合國教科文組織定義素養（literacy）已不侷限於讀寫能力，關係到社會溝通實踐以及知識、語言和文

化，所以素養定義已經隨社會產生動態演變；如同楊淑萍、林煥祥（2010）指出，OECD 所制訂的素養已經因時代環境變遷產生調整，不同於學校的學業成就評量。

綜合上述，素養是指一種個體要運用心智運作解決生活情境問題的能力，不再僅強調知識、讀寫能力而已；至此，素養已經帶著教育的期望，期許改善過去的傳統教育缺失，讓學生能將所學運用在現實生活中。本文「素養」即為強調個體認知、情意和技能的綜合性能力，並期待學生能應用於生活情境中，由於以數學特定領域為範圍，故以「Literacy」翻譯之。

（二）數學素養

劉柏宏（2016）認為，數學素養可以改善臺灣「高成就、低態度」的現象，文中以表格整理各國數學素養一詞，英國用 numeracy、美國用 quantitative literacy 或 mathematical literacy、北歐國家用 mathematical competence、荷蘭則是以 OECD 所採的 mathematical literacy，但是無論何種翻譯，其定義都隱含能夠處理使用日常生活脈絡中的數據，並理解數學表格或圖形，擁有描述、解釋、運用數學的能力。秦爾聰、劉致演、尤昭奇（2015）引美國國家研究會 Kilpatrick 等人的主張，數學素養是數學學習成功的表徵，由五種數學能力所構成：數學概念理解整合、執行數學程序過程流暢、善用解決數學問題策略、在情境中展現邏輯推理、成為數學知識實踐者。左台益、李健恆、潘亞衛、呂鳳琳（2018）指出，數學素養要素有內容知識、數學識能和建模歷程，內容知識像是數學的學習單元範圍，而數學識能是學生學習數學具備的能力，最後透過建模歷程讓個體在數學和現實世界中，運用數學來理解應用。

單維彰、鄭章華（2017）引用教育部說明「核心素養」是在九年一貫課程基礎上，藉由整合知識、技能和態度，使個人可以適應生活與面對挑戰，而十二年基本國民教育數學領綱將從「結果」與「過程」來建構核心素養，教學模組要營造現實生活或故事情境，認識到數學的實用性。李國偉（2014）指出，提升數學素養就是學習數學以便解決生活問題並具備反思能力。換句話說，個體能在不同情境下，將情境問題轉化成使用數學的能力（林福來，2011）。

測驗學刊，第 67 輯第 4 期

在臺灣 PISA 國家研究中心（2014）的《PISA 2012 臺灣精簡報告》第 8 頁中指出：「數學素養是個體在不同情境脈絡中，形成、應用以及詮釋數學的能力」，亦即透過數學思維行動來解決現實生活的問題，也在 PISA 2012 發展了數位數學素養評量，主要評量數學推理與歷程，無關乎電腦精熟程度；臺灣 PISA 表現，PISA 2012 數學素養平均分數 560 分較 PISA 2009 進步，但是學生表現優劣離散大，至於數位數學素養平均數為 537 分，相對於書面數學素養表現則略低。周玉秀（2006）引荷蘭數學教育大師 Freudenthal 在《數學結構的教學現象學》一書主張的「真實世界的問題，發展數學概念」，認為 PISA 採現象學取向明確與傳統學校課程區隔。

PISA 數學素養有數量（Quantity）、空間與形狀（Space and shape）、改變與關係（Change and relationships）、不確定性（Uncertainty）四大內容領域，臺灣 2012 年在這些分測驗的國際排名，除了空間與形狀是第 2 名外，其他都是第 4 名；PISA 數學歷程：「形成數學情境」將問題轉譯成數學可處理形式，對臺灣學生是相對容易的部分；「應用數學概念、事實、程序以及推理」要辨識自己的知識庫與問題的關係，並將知識有系統組織地應用解決，對臺灣學生而言相對困難；「詮釋、應用以及評鑑數學結果」須將結果與原情境連結，臺灣學生表現則稍弱（臺灣 PISA 國家研究中心，2014，2017；OECD, 2005）。

綜合上述，「數學素養」多指向個體要有能力在日常生活中，解讀、應用、詮釋數學相關問題，即除了具備有數學知識帶著走的能力外，並能因應於日常生活中的問題解決，這是一種異於傳統數學課程的現象學取向。據此本文「數學素養」為數學領域的綜合性能力，個體要有能力應用於生活情境中並解決數學相關問題。

二、九年一貫課程綱要與素養關係

「素養」近年來不但受到國際組織重視，也是國內教育研究單位及相關學者的期盼，其實國民中小學九年一貫課程綱要數學學習領域（教育部，2009）與 PISA 數學素養評量（臺灣 PISA 國家研究中心，2014，2017）宗旨其實相近，概述比較如下。

（一）數學內容

九年一貫課程綱要數學學習領域，將數學分為「數與量」、「幾何」、「代數」、「統計與機率」、「連結」等五大主題，而教師、教科書編輯者及相關審查單位除須注意「數學內部連結：解題能力的培養」及「數學外部連結：生活應用解題」外，也要與其他學科教材連結（教育部，2009）。至於 PISA 數學素養則有「數量」、「空間與形狀」、「改變與關係」、「不確定性」等四大內容領域（臺灣 PISA 國家研究中心，2014，2017；OECD，2005）。

可見九年一貫課程綱要和 PISA 數學素養內容領域相似，但是用詞不一，例如：九年一貫課程綱要，除了「連結」是察覺、轉化、解題、溝通，以及評析用以加強數學與日常生活的連結外，其他是數學專有名詞；PISA 數學素養內容領域，「數量」如同九年一貫「數與量」，而「空間與形狀」主要是空間、圖形結構，類似「幾何」概念，「改變與關係」與「代數」的部分概念是相同的，「不確定性」與「統計與機率」也是大部分相似，但「不確定性」更重視可能性與資料詮釋的推理（曾建銘、吳慧珉、趙珮晴，2019）。

（二）階段性表現

九年一貫課程綱要數學學習領域，將九年國民教育區分為國小到國中四個階段，每階段都有教學目標，以此階段可以逐步奠定高中數學基礎。PISA 將數學素養能力分成六個水準，並針對素養水準進行典型表現描述，尤以數學素養水準 2 被視為是數學精熟度基礎，未達此標準則表示要加強。

可知九年一貫課程綱要和 PISA 數學素養能力，都有階段性、層次性、階層性意涵，然而究其意義卻不一樣。九年一貫課程綱要是一種能力描述，其描述是一種教學目標，期許學生未來可以達成；至於 PISA 數學素養能力，則是將所有應具備能力統整後，以分數劃分，主要用於表現回饋。

整體來說，九年一貫課程綱要和 PISA 數學素養內容領域相似，均期盼教學要能融入生活情境，讓數學知識與生活得以連結、實踐；而九年一貫課程綱要和 PISA 數學素養能力都是階段性概念呈現定義、詮釋，但是究其意義則因目的不同而有異。

三、影響數學表現的相關因素

（一）性別

關於數學表現是否受性別影響？議論紛紛，莫衷一是。普遍來說，數理學業成就以男性優於女性，隨著社會進步開放而逐漸縮小差距（Else-Quest, Hyde, & Linn, 2010; Hedges & Nowell, 1995; Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Lindberg, Hyde, Petersen, & Linn, 2010）。但是，有些研究認為女生表現較佳，龔心怡、李靜儀（2016）採縱貫研究設計以七至九年級三波共 1,025 位國中學生為研究樣本，發現男生具有較高的數學自我概念和自信；女生則具有較佳的數學成就表現。Voyer 與 Voyer（2014）以後設分析發現女性在語文科最具優勢，而數學科表現也不差，但是女生優勢會在進入高中以後漸漸消失，此結果會受學校層級、國家和種族的調節影響。

有些研究則認為男女各有優勢，整體來說並沒有顯著差異。蔣治平等人（2012）編擬數學素養測驗共 44 題，其題目從認知、理解以及應用著手，並滿足生活化和盡量貼近真實情境，以 487 位正修科技大學五系二至三年級學生為測驗對象，發現 16 項數學素養能力在學系間產生差異，至於性別和年級並無明顯差異，僅有「具極限值推理能力」、「具解讀抽樣過程能力」呈現女生優於男生；而「具敏感性分析能力」是男生優於女生。在臺灣 PISA 國家研究中心（2014）的報告中，臺灣數學素養平均 560，名列全球第 4 名（僅次於上海、新加坡和香港）名列前茅，性別方面，男學生雖然優於女學生，但亦未達統計顯著水準。

由以上研究可知不同性別對於數學的表現，並非呈現一致的結果，因為隨著年齡與學習階段會有差異，一般是國小女生略優於男生，國中階段差不多，至高中階段以上則呈現男生優於女生，至於臺灣小六數學素養表現的性別差異為何？乃本文欲探討目的之一。

（二）地區

許多研究證實城鄉差距一直存在（黃敏雄，2015；楊淑萍、林煥祥，2010；龔心怡、李靜儀，2016；Alordiah, Akpadaka, & Oviogbodu, 2015; Butt & Dogar, 2014; Li, 2007）。龔心怡、李靜儀（2016）的三波縱貫研究設計分

析，發現無論是數學自我概念與數學學業成就皆達到城鄉的差異，以城市地區的學生表現較好。楊淑萍、林煥祥（2010）以臺灣PISA 2006 年的 8,815 位學生進行分析，認為城鄉差異的原因在於大城市的經濟資源與文化資源持有率最高，隨人口減少而遞減外，而經濟資源和文化資源在數學素養上交互作用和主要效果都達顯著，以有愈多資源則素養表現愈好，所以素養不佳，恐非能力有限，而是資源不足的問題。

黃敏雄（2015）以臺灣教育長期追蹤資料庫（Taiwan Education Panel Survey, TEPS）和國際數學與科學教育成就趨勢調查（TIMSS）進行城鄉差異分析，有許多符合世俗認知，例如：臺北市與鄉村國中學生的平均數學表現差距確實很大；大都市、城鎮、一般鄉村、偏遠鄉村，學生的平均學習表現依序遞減；臺北市國中學生隨年級提高，數學表現進步幅度較大，所以從國一升到國三與鄉村地區差距也有擴大現象；此外學生家庭社經背景、之前的學習表現也是導致城鄉差異的原因。

僅有少數研究指出城鄉未有顯著差異，也認為應該更細部剖析。Howley（2003）從 1996 年和 2000 年的國際教育進步評量（National Assessment of Educational Progress, NAEP）的數學成績，發現偏鄉和小型學校的數學成就與全國平均未達顯著差異，歷年來都只有微小差距而已。Kifer（2001）對於教學現場的城鄉差異，認為「變異」重於「平均」，「平均」是整體大致情況，但是生活方式、地區文化和才能成就不可忽略，顯然 Kifer 建議多細部深入探討。

綜合上述可知城鄉學生數學成就差異一直存在，且平均都是城市大於鄉鎮，但變異大小與影響因素都值得持續關注，因此本文將以臺灣小六數學素養表現進一步分析。

參、研究方法

先說明研究題本的設計並以預試修正題目，然後研究對象的抽樣與背景描述，最後則是資料處理與分析方法。

一、研究題本

本文設計 2 份研究題本（甲、乙兩題本），每題本分三部分：第一部分是 PISA 試題 13 題和修改自 PISA 試題 1 題共 14 題；第二部分分成 A、B 卷，

測驗學刊，第 67 輯第 4 期

均有 12 題；第三部分則是調查性別、地區背景問卷資料等。2 份研究題本則以 PISA 之 14 題作為共通題（定錨題），以利後續兩卷等化。

PISA 試題主要取自 PISA 2012 數學素養題本，經專家會議選取適合國小六年級試題 13 題，另 1 題修改自林福來（2013）臺灣 2011 數學素養評量樣本試題並參酌 PISA 題目定義加以修改，合計共 14 題；另一部分取自九年一貫數學學習領域 103 學年度各版本教科書，選取適合國小六年級學生施測的單元，並考量不同版本的公平性，再請 5 位國小資深教師依照 PISA 數學內容領域編擬 30 題，後經專家會議挑出 24 題，分成 A、B 兩卷，都有二元計分與多元計分題目。

本文專家會議的組成，是由 5 位國小教學年資超過 10 年以上具數理背景或教育測驗統計研究所畢業之教師與 1 位測驗評量博士組成，根據雙向細目表命題，除了數學內容領域外，認知層次在 PISA 部分分為形成數學情境（20%）、應用數學概念、事實、程序與推理（20%）、詮釋、應用與評估數學結果（10%），自編部分則是參照 TASA 分為概念理解（15%）、程序執行（20%）、解題與推理（20%），分別有其定義，命題時就要思考其歸類，亦要考慮試題情境的合適性，修題除了試題本身外，也再次討論、檢核歸類是否正確，最後才組卷預試、再次修審、正式施測。

表 1 研究題本試題設計

PISA 數學內容領域	PISA 試題			自編試題	
	形成數學情境	應用數學概念、 事實、程序與推理	詮釋、應用與 評估數學結果	A 卷	B 卷
數量	2012 P46Q1 (1) 2012 P57Q2 (3)	2012 P48Q3 (2) 2011P-RQ-006 (12)	2012 P3Q1 (9)	A2 A3 A4 A5	B3 B4 B5
空間與形狀		2012 P42Q1 (4) 2009 P11Q1 (13)	2012 P44Q1 (10) 2012 P65Q1 (11)	A6 A10 A11 A12	B8 B10 B11 B12
改變與關係	2012 P68Q1 (5)		2012 P8Q2 (14)	A1	B1 B2
不確定性與關係	2012 P7Q1 (6)	2012 P27Q5 (8)	2012 P25Q1 (7)	A7 A8 A9	B6 B7 B9

註：PISA 試題格內代表「PISA 試題代號（題本題號）」。

二、預試

預試對象係方便取樣，選取臺中市兩校共 4 班國小六年級 113 位學生，施測甲題本 56 人、乙題本 57 人。依據預試分析結果，修改說明如下：

1. 二元計分題目：PISA 試題，問題 7、8 同一題組，問題 8 較簡易，本著試題由易而難的原則，調整順序。A 卷維持原樣，B 卷問題 12 為避免正方體與長方體概念干擾，題幹從原題「長方體的最少須由幾組相同的面所組成？」修改成「正方體是由一組相同的六個面所組成的，則長方體的面最少須由幾組相同的面所組成？」。
2. 多元計分題目：PISA 試題，問題 3 空白占 16.8%、得 0 分占 71.7%，合計高達 88.5%顯示需再修題；討論後，在題幹表格內多增一列會員費，使題目更清楚易懂；問題 12 空白占 21.2%、得 0 分占 66.4%，合計高達 87.6%顯示需再修題；討論後，將 1 歐元 = 34.5 臺幣修改為 1 澳幣 = 25 臺幣使能整除，再將四捨五入取至小數第 2 位刪除。A、B 卷答對率都在一半以上，維持原樣。

三、研究對象

本文採方便取樣 27 校 56 班國小六年級學生 1,258 名（有效樣本）為正式施測對象，施測時間為二節，第一節施測共同的部分——PISA 試題，第二節施測時是按照學生坐位，採 A、B 卷（含背景問卷）輪流發放給學生作答，研究對象以中部學生居多占 35.9%、其次南部占 27.8%、接著是北部占 21.5%、最後離島有 14.7%；無論是性別或卷別，都各占研究對象約一半的比例。

四、信度

受試者分隔信度係數（Separation Reliability）為受試者變異與真分數變異（即 Rasch 單向度模式可解釋之變異）之比值，係數愈高表示試題愈能穩定區隔受試者差異，受試者的排序愈穩定（Smith, 2001）。

本文是以學生之原始作答再加上背景問卷資料，採 EAP 法來估算學生能力的可能值（Plausible Value, PV），得到受試者分隔信度係數為 0.99，表示試題愈能穩定區隔受試者差異，受試者的排序穩定；而根據 Conquest 所得到的 EAP/PV 信度估計值則為 0.82。

一般而言，一份優良的教育測驗至少應該具有 .80 以上的信度係數值，才比較具有使用的教育價值（Camines & Zeller, 1979）。而根據 Henson（2001）的觀點認為，如果是以發展測量工具為目的時，信度係數應在 .70 以上。此表示本測驗具有不錯的信度。

五、資料處理與分析

二元計分試題以試題反應理論 (Item Response Theory, IRT) 一參數邏輯式模式 (one-parameter logistic model) 以及古典測驗理論 (classical test theory, CTT) 的難度和鑑別度進行分析；多元計分試題則是以 IRT 的部分計分模式 (partial credit model) 的 1 分和 2 分閾限參數 (threshold parameter)。接著，再以不同題本不同題型的得分與空白人數百分比了解試題得分情形，並列舉試題錯誤類型。

為了進行性別、地區及其與數學內容領域的表現差異分析，本文是結合學生的試題作答反應與背景問卷資料，利用一參數邏輯式模式 ConQuest 軟體做資料分析得出學生的期望後驗能力估計值 (EAP)，然後採 PISA 以平均數為 500 和標準差為 100 進行線性轉換；自編 A 卷和 B 卷能力值採用 PISA 定錨題與所有試題同時估計，然後固定試題參數後再分別算出 PISA、A 卷、B 卷的學生能力估計值。差異分析方法包含：單因子變異數 (one-way ANOVA)、獨立樣本平均數 t 檢定和單因子多變量變異數 (multivariate analysis of variance, MANOVA) 分析。

肆、研究結果

將先呈現二元計分與多元計分得分情形，再列舉數學素養試題錯誤類型，然後進行地區、性別差異分析，最後參酌文獻提出綜合討論。

一、試題得分情形

無論是 PISA 試題、自編 A 卷或自編 B 卷，開放問答的空白比例均高於選擇題和封閉問答，PISA 試題第 3 題空白的人數甚至高達 42.4%，PISA 試題第 12 題占 19.2% 次之，接著則是自編 B 卷第 8 題有 19%，可能是題幹文字內容長度過長、不懂題意而自動放棄；至於選擇題和封閉問答空白的人數比例則明顯較少。

二、數學內容領域難易度分析

二元計分試題分析，IRT 難度參數愈大則表示試題愈難，若從 IRT 難度參數平均來看，PISA 試題和自編 A 卷的數學內容領域由難到易依序為數量、

不確定性和空間與形狀，而自編 B 卷則是空間與形狀、數量和不確定性，另改變與關係僅 1 題在 PISA 試題，不列入排序。除此之外，本文另以 CTT 的難度和鑑別度進行分析，CTT 難度愈高表示試題愈簡單，試卷平均難度由難到易排序同 IRT 難度參數排序，而 CTT 鑑別度如果高於 .40 表示試題非常優良（郭生玉，2004），本文除了 PISA6-2 之填充題、自編 A 卷第 4 題外，其餘試題鑑別度都達優良標準。

多元計分試題分析，1 分和 2 分閾限參數分別以 IRT_b 1 和 IRT_b 2 表示，其差異為閾限差異。如果從 1 分閾限參數來看，可知 PISA 試題第 3 題數量要得 1 分最難，而 PISA 試題第 10 題空間與形狀要得 1 分最易；如果從 2 分閾限參數來看，可知自編 B 卷第 3 題數量要得 2 分最難，而 PISA 試題第 14 題改變與關係要得 2 分最易；如果用 2 分閾限參數減去 1 分的閾限參數，可知得 1 分要跨到 2 分，最難的是自編 B 卷第 3 題數量，最易則是 PISA 試題第 14 題改變與關係。

依據 Linacre 與 Wright（1994）建議 IRT 的適配度 MNSQ 介於 0.80~1.20 之間為佳。本文 PISA 試題的 MNSQ 介於 0.95~1.12 之間；A 卷介於 0.90~1.18 之間；B 卷介於 0.92~1.06 之間，可知本文 IRT 適配度還不錯。

三、試題錯誤類型說明

本文依照 PISA 數學內容領域挑選試題錯誤類型，並進行說明如下。

（一）數量錯誤類型

1. PISA 第 3 題開放問答題

本題全部答對占 15.3%；部分答對占 2.1%；有作答但答錯占 40.2%；不會作答且空白者占 42.4%。不會和做錯合起來 1,039 人，占 82.6%，高達 8 成以上不會此題。究其原因，發現國小六年級學生租借 DVD 的生活經驗不足，沒有將會員費分攤至租借費用的概念，大都以 $10 \div 2.5 = 4$ ，答：4 片為錯誤答案。由此可知，學生對於無經驗之生活情境題，將造成答題困難，所以選擇合適的生活情境是命題時需要考量的，否則有可能變成試題差異功能題，導致獨厚有經驗的學生。

測驗學刊，第 67 輯第 4 期

2. PISA 第 12 題開放問答題

本題全部答對占 25.9%；有作答但答錯占 54.9%；不會作答且空白者占 19.2%，不會和做錯合起來占 74.1%，高達 7 成以上不會此題。究其原因，發現國小六年級學生的小數乘法 $10000 \div 25 = 400$ ， $400 \times 0.3\% = 1.2$ 時，少除 10 倍計算錯誤答案為 12，或是 $0.3\% = 0.003$ 誤當成 0.0003，多除了 10 倍；所以 0.3% 對多數國小六年級學生造成認知上負荷，應該更改為 3% 較適合。

（二）空間與形狀錯誤類型

1. PISA 第 4 題選擇題

本題答對率 48.25%，屆近一半，高分組作答情形正確率高達 80%，相較低分組學生而言，錯誤選項（A）22%、（B）27% 和（D）25% 與正確選項（C）24% 差不多，代表低分組學生有可能猜測，不是對於題意不了解，就是在空間與形狀的應用數學概念、事實程序與推理上概念不足。究其原因，試題中除了圓的圓周長觀念外，還加上了時間、速度的概念相關單元，如果將圓當成時鐘，半小時剛好在錯誤選項（A）R 的位置，但忽略了題幹中轉一圈恰好是 40 分鐘；如果將摩天輪當成時鐘，直接判斷 40 分鐘，剛好位在錯誤選項（B）在點 R 和點 S 之間；如果誤將登艙點 P 點進入摩天輪當成 R，直接判斷 40 分鐘，剛好位在錯誤選項（D）在點 S 和點 P 之間，建議教師可以針對學生在時間、速度與距離關係做概念澄清。

2. PISA 第 11 題選擇題

本題答對率 60.33%，顯示有一半以上學生答對，高分組作答情形正確率高達 95%，相較低分組學生而言，錯誤選項（B）19% 和（D）41% 加起來約 60%，代表低分組學生在空間與形狀應用數學詮釋、應用與評估上概念仍須加強。究其原因，除了車庫的立體形體觀念外，還包含了一個窗戶和一道門的相對位置，選擇錯誤選項（B）和（D）的學生忽略了面向車庫時，門在左側，而窗在右側。或許未來可透過實際操作，幫助學生建立立體空間的概念。

（三）改變與關係錯誤類型：自編 A 卷第 1 題開放問答題

本題全部答對占 36.2%；部分答對占 7.5%；有作答但答錯占 42.6%；不

會作答空白者占 13.7%，不會和做錯合起來占 56.3%，超過半數答錯。究其原因，本題除了文字敘述外，還需運用表格數據計算，有三個不同方案，再搭配網內和網外，以秒計費的時間換算，在題意的理解上思緒要清晰，且需應用計算兩位小數乘法合併加法，學生容易產生題意不清、粗心或者計算錯誤的情形。此外，手機雖然是生活中的一部分，但是對小六學生而言，幾乎都是父母幫忙申辦與繳費，學生對於費率並不熟稔，再加上解題需讀懂表格的意義，也是答對率偏低的原因之一。

（四）不確定性錯誤類型：自編 A 卷第 7 題選擇題

本題答對率 28.3%，顯示僅有約 3 成學生答對，高分組作答情形正確率高達 60%，相較低分組學生而言，錯誤選項（A）21%、（B）25%和（C）41%都高於正確選項（D）9%；其中錯誤選項（C）的選項誘答力明顯，代表多數學生的迷思概念是 $20 \times \frac{40}{100} = 8$ ，而忽略必須涵蓋之前的 100 球進球率 38% 部分，正確計算該是 $120 \times \frac{40}{100} = 48$ ， $48 - 38 = 10$ ，高於 10 選（D），因此應用「比和比值」解決問題時，利用比率解決由總量求部分量概念的澄清。

綜合素養試題錯誤類型，可知學生在 PISA 或自編 A、B 卷要能夠正確解題，所需要的計算過程並不繁雜，僅有 PISA 第 12 題要從 0.3% 轉化為 0.003 稍感困難，答題關鍵會落在題意的理解，畢竟數學素養講求學生要有能力在日常生活中，解讀、應用、詮釋數學相關問題，對低分組學生易造成題意不了解，而任意作答或猜測，如：PISA 第 4 題選擇題將摩天輪當成時鐘、PISA 第 11 題選擇題忽略了面向車庫時的門窗位置、PISA 第 3 題開放問答題無租借 DVD 的生活經驗、自編 A 卷第 1 題開放問答題無手機費率概念等。

四、性別、地區及其與數學內容領域的表現差異分析

（一）地區的表現差異分析

本文為了解不同地區表現差異情形，以單因子變異數進行分析，採 Scheffé 法進行事後比較，可知地區在能力值均達到統計 .05 顯著水準，整體 $F(3, 1254) = 3.62^*$ 、PISA 試題 $F(3, 1254) = 2.90^*$ ，或自編試題 $F(3, 1254) = 3.21^*$ ，而整體和 PISA 試題事後比較，以北區優於南區學生。

測驗學刊，第 67 輯第 4 期

（二）性別的表現差異分析

本文為了解不同性別表現差異情形，以獨立樣本平均數 t 檢定進行分析，可知性別在 PISA 試題 t 值 2.24* ($p = 0.025$) 達到統計 .05 顯著水準，以男性優於女性學生，至於整體 t 值 1.81 ($p = 0.07$) 和白編測驗 t 值 1.27 ($p = 0.204$) 則是沒有顯著差異。

（三）性別、地區與數學內容領域表現差異分析

本文以單因子多變量變異數分析，當 Wilks' Lambda 達統計顯著水準時，再以 Scheffé 進行事後比較。結果發現各區域的數學內容領域達統計顯著水準 (Wilks' Lambda = .97, $p < .001$)；以 Scheffé 事後比較，發現空間與形狀以北部顯著優於南部，而在改變與關係則是北部顯著優於離島，其餘內容領域各區域間則沒有顯著差異。至於性別的數學內容領域亦達統計顯著水準 (Wilks' Lambda = .99, $p < .04$)；以 Scheffé 事後比較，發現男學生僅有在數量顯著優於女生，其餘三個數學內容領域則沒有顯著差異。

五、綜合討論

素養受到國際組織關注，未來將持續重視並期待能落實於教育現場。

本文發現，無論是 PISA 或自編試題，開放問答的空白比例高，從錯誤類型可知數學素養試題計算不難，答題關鍵會落在題意的理解，對低分組學生易造成不了解、任意作答或猜測。「數學素養」異於傳統「數學能力」，如今的「數學素養」帶有現象學解釋，期待學生可解讀、應用、詮釋數學相關日常生活問題，讓數學知識得以活化應用；然而，「數學素養」卻有可能增加診斷學生錯誤作答的難度。舉例來說：「 $1+1=?$ 」，答錯可以明白知道學生對於「 $1+1$ 」有作答困難；「買 1 個棒球，再買 1 個棒球，總共有多少個棒球？」加入情境描述的題目，學生答錯可能是不懂題目意涵，不懂「再買」、「總共」用語，或者沒有棒球的生活經驗，也如本文分析錯誤類型結果：將摩天輪當成時鐘、忽略了面向車庫時的門窗位置、無租借 DVD 的生活經驗、無手機費率概念等。因此，重視「素養」將知識運用於生活的立意良好，但是在試題設計時，為了公平公正，不得不思考試題用語與情境的適切性。

接著，地區差異分析顯示北區數學素養試題作答優於南區學生；內容領

域來說，北部空間與形狀顯著優於南部，而改變與關係則是北部顯著優於離島。本文與多數研究結果一致，數學城鄉差異一直存在，即使數學素養亦是（黃敏雄，2015；楊淑萍、林煥祥，2010；龔心怡、李靜儀，2016；Alordiah et al., 2015; Butt & Dogar, 2014; Li, 2007）；究其原因有經濟資源與文化資源持有率（楊淑萍、林煥祥，2010）、學生家庭社經背景和之前學習表現（黃敏雄，2015）等。建議未來要考量各地區思考命題情境外，秦爾聰等人（2015）在國中七年級學生以臆測為中心的 5E 數學探究教學模式下，發現學生在非例行性問題情境時，數學臆測是數學素養的趨力，秦爾聰等人建議學生要經由計畫性的設計始能探究學習，教師教學可嘗試合作或引導式教學。

至於性別，本文結果只有在 PISA 素養題部分是男性顯著優於女性學生；內容領域來說，則僅有數量以男性顯著優於女性學生，其他未有顯著差異；臺灣 PISA 性別差異未達統計顯著差異（臺灣 PISA 國家研究中心，2014），建議未來持續深入探討釐清，畢竟導致性別差異的因素眾說紛紜，數理信心（龔心怡、李靜儀，2016）、學校層級、國家和種族（Voyer & Voyer, 2014）、時代進步開放（Else-Quest et al., 2010; Hedges & Nowell, 1995; Hyde et al., 2008; Lindberg et al., 2010）等都可能是影響的因素。

最後，由於本文是方便取樣，係根據認識之教師、主任或校長分屬北中南及離島選擇研究對象，並非分層隨機抽樣，且施測學校會因所處之地理位置和學生成就表現有所不同，不宜將研究結果推論至全臺灣，此乃本文之研究限制。

伍、結論與建議

以下依據研究結果提出結論與建議。

一、結論

素養是臺灣未來課程發展的重點之一，本文設計一份適合國小六年級測試的數學素養試題，用以了解學生數學素養表現和錯誤類型，並進一步分析地區、性別差異。本文發現數學素養講求學生要有能力在日常生活中，解讀、應用、詮釋數學相關問題，學生不需要繁雜的計算過程解題，答題關鍵是落在題意的理解，尤其是低分組學生易造成題意不了解而任意作答或猜

測驗學刊，第 67 輯第 4 期

測，如果是開放問答題的空白比例也會偏高。至於背景變項的差異分析，北區數學素養試題作答優於南區學生，顯示數學城鄉差異一直存在，即使數學素養亦是，可能受資源持有率、家庭社經背景和之前學習表現的影響；而性別則是 PISA 素養題與數學內容領域的數量部分呈現顯著差異，且都是男性優於女性學生，其他未有顯著差異，但是導致性別差異的因素莫衷一是，有待未來再深入探討剖析。

二、建議

根據研究結論，提出以下建議。

（一）數學素養試題設計要思考試題用語和情境適切性

本文發現數學素養開放問答的空白比例高，從錯誤類型可知計算不難，關鍵在題意理解，特別是低分組學生易造成不了解、任意作答或猜測。數學素養係活化數學知識並應用於日常生活，但是在數學素養試題設計時，務必思考試題用語與情境的適切性，避免獨厚特定學生，以維持試題公正公平，而數學教師除了教學概念要嘗試融入生活經驗外，也要加強素養試題的錯誤作答剖析能力，以協助低分組學生提升數學素養。

（二）關注、鼓勵、協助偏鄉和女性學生的數學素養學習

數學素養試題，北區作答優於南區學生，而男性顯著優於女性學生，顯然有城鄉差異和性別差異存在，導致城鄉差異的原因眾多，例如：經濟資源、文化資源、家庭社經地位等；性別差異亦是，如：數理信心、學校層級、國家、種族、時代進步開放等，期盼相關教育單位持續關注、鼓勵、協助偏鄉和女性學生的數學素養學習。

謝誌

本研究承蒙科技部專題研究計畫（MOST 103-2511-S-656-002-）提供經費補助，並感謝期刊與審查委員提供寶貴意見。

參考文獻

中文部分

- 左台益、李健恆、潘亞衛、呂鳳琳（2018）。臺灣、新加坡及巴西數學教科書中數學素養內涵之比較：以畢氏定理為例。**教科書研究**，**11**（3），33-62。
- 李國偉（2014）。教育部提升國民素養實施方案：數學素養研究計畫。取自 <https://reurl.cc/Q3GzL0>
- 周玉秀（2006）。從 PISA 看數學素養與中小學數學教育。**科學教育月刊**，**293**，2-21。
- 林福來（2011）。數學素養的教學思維。取自 <https://reurl.cc/py4zZx>
- 林福來（主編）（2013）。臺灣 2011 數學素養評量樣本試題（上、下）。臺北市：國立臺灣師範大學。
- 洪雪芬（2011）。運用 PISA 評量試題於國小六年級之數學解題初探。發表於 2011 年南臺灣教育論壇。日期：2011 年 7 月 25 日。
- 范信賢（2016）。核心素養與十二年國民基本教育課程綱要：導讀《國民核心素養：十二年國教課程改革的 DNA》。**教育脈動**，**5**，取自 <https://www.naer.edu.tw/ezfiles/0/1000/img/67/379316110.pdf>
- 秦爾聰、劉致演、尤昭奇（2015）。探討七年級學生在以臆測為中心的數學探究教學脈絡下其數學素養展現情形。**臺灣數學教師**，**36**（1），1-16。
- 教育部（2009）。國民中小學九年一貫課程綱要數學學習領域。取自 <https://cirn.moe.edu.tw/WebContent/index.aspx?sid=9&mid=248>
- 教育部（2014）。十二年國民基本教育課程綱要總綱。臺北市：作者。
- 郭生玉（2004）。教育測驗與評量。臺北市：精華。
- 單維彰（2016）。素養、課程與教材：以數學為例。**教育脈動**，**5**，1-19。
- 單維彰、鄭章華（主編）（2017）。十二年國教數學素養導向課程設計與教學案例。新北市：國家教育研究院。
- 曾建銘、吳慧珉、趙珮晴（2019）。臺灣學生數學學習表現之探討：來自跨資料庫比較的證據。**測驗學刊**，**66**（1），27-50。
- 黃敏雄（2015）。學生數學表現的城鄉差異。**教育研究集刊**，**61**（4），33-61。
- 楊淑萍、林煥祥（2010）。由家庭經濟資源及文化資源探討我國學生在 PISA 科學、數學素養的表現。**科學教育學刊**，**18**（6），547-562。

測驗學刊，第 67 輯第 4 期

- 臺灣PISA國家研究中心（2014）。**PISA 2012 臺灣精簡報告**。取自 http://pisa.nutn.edu.tw/dl_result_report.htm
- 臺灣PISA國家研究中心（2017）。**PISA 國際學生能力評量計畫**。取自 <http://pisa.nutn.edu.tw/index.html>
- 劉柏宏（2016）。從數學與文化的關係探討數學文化素養之內涵：理論與案例分析。**臺灣數學教育期刊**，3（1），55-83。
- 蔡清田（2010）。課程改革中的「素養」（competence）與「能力」（ability）。**教育研究月刊**，200（12），93-104。
- 蔡清田（2011a）。課程改革中的「素養」（competence）與「知能」（literacy）之差異。**教育研究月刊**，203（3），84-96。
- 蔡清田（2011b）。課程改革中的「素養」。**幼兒教保研究**，7，1-13。
- 蔣治平、廖敏治、蘇懿（2012）。數學素養能力的測驗命題設計與評量。**全球商業經營管理學報**，4，119-133。
- 鄭英豪（2013）。從PISA看國內數學命題趨勢。取自 <https://reurl.cc/R1Epy6>
- 龔心怡、李靜儀（2016）。國中學生數學自我概念與數學學業成就相互效果模式之縱貫研究：性別差異與城鄉差距之觀點。**科學教育學刊**，24，511-536。

英文部分

- Alordiah, C. O., Akpadaka, G., & Oviogbodu, C. O. (2015). The influence of gender, school location and socio-economic status on students' academic achievement in mathematics. *Journal of Education and Practice*, 6(17), 130-136.
- Butt, I. H., & Dogar, A. H. (2014). Gender disparity in mathematics achievement among the rural and urban high school students in Pakistan. *Pakistan Journal of Social Sciences*, 34(1), 93-100.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment* (Vol. 17). Thousand Oaks, CA: Sage.
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136, 103-127.
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269, 41-45.
- Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development*, 34(3), 177-189.
- Howley, C. (2003). Understanding mathematics education in rural context. *The Educa-*

- tional Forum*, 67(3), 215-224.
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321, 494-495.
- Kifer, E. (2001). Why research on science and mathematics education in rural schools is important or the mean is the wrong message. In S. A. Henderson (Ed.), *Understanding achievement in science and mathematics in rural schools: Conference proceedings* (pp. 44-48). Lexington, KY: Appalachian Rural Systemic Initiative.
- Li, Q. (2007). Mathematics, science, and technology in secondary schools: Do gender and region make a difference? *Canadian Journal of Learning and Technology*, 33(1), 41-57.
- Linacre, J. M., & Wright, B. D. (1994). Dichotomous meansquare infit and outfit chi-square fit statistics. *Rasch Measurement Transactions*, 8(2), 360.
- Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, 136, 1123-1135.
- Organization for Economic Cooperation and Development. [OECD] (2005). *PISA 2003 technical report*. Paris, France: Author.
- Smith, E. V. (2001). Evidence for the reliability of measurement and the validity of measure interpretation: A Rasch measurement perspective. *Journal of Applied Measurement*, 3, 205-231.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174-1204.

測驗學刊，第 67 輯第 4 期