

## 「國中生線上閱讀理解能力測驗」之發展及 相關因素之研究

黃秀霜<sup>1</sup> 曾雅瑛<sup>2</sup> 甘孟龍<sup>3</sup> 許力云<sup>4</sup> 黃品甄<sup>5</sup>

### 摘要

本研究旨在編製具有信效度之「國中生線上閱讀理解能力測驗」，透過閱讀理解策略與線上資訊之使用，以了解學生線上閱讀理解能力，並分析不同背景變項對於該測驗工具得分之影響及其差異表現。樣本為 1,107 位來自臺灣北、中、南三個區域的國中一年級學生，本研究使用 Rasch 部分給分模式進行測驗之心理計量檢核。研究結果顯示：

1. 「國中生線上閱讀理解能力測驗」各試題皆符合 Rasch 測量模式，多元計分試題的等級適切且難度成階層關係，且性別在所有題目均無明顯的差異試題功能存在。
2. 不同背景變項雖然在線上閱讀理解能力具有差異，惟影響程度並不高。

最後，本研究根據研究結果作出具體之相關研究建議，以供未來教學及研究之參考。

**關鍵詞：**Rasch 測量模式、線上閱讀理解能力

---

1. 黃秀霜，國立臺南大學教育學系教授  
2. 曾雅瑛，彰化縣後寮國小校長  
3. 甘孟龍，寧德師範學院副教授  
4. 許力云，雲林縣口湖國中教務主任  
5. 黃品甄，臺南市政府教育局機要秘書  
收件日期：2017.10.12；完成修改：2019.09.25；正式接受：2019.09.26  
通訊作者：曾雅瑛；Email：yyt9487@gmail.com  
地址：彰化縣芳苑鄉後寮村芳寮路 30 號 彰化縣後寮國小

## Research on the Development and Related Factors of The Online Reading Comprehension Assessment for Junior High School Students

Hsiu-Shuang Huang <sup>1</sup> Ya-Ying Tseng <sup>2</sup> Mon-Long Gan <sup>3</sup>  
Li-Yun Hsu <sup>4</sup> Pin-Chen Huang <sup>5</sup>

### Abstract

The main purpose of this study was to develop The Online Reading Comprehension Assessment for Junior High School Students. With the understanding of students' reading strategies and using online information, junior students' online reading ability was explored. Furthermore, the effect of students' background variables on reading ability were analyzed. 1,107 junior high school students from north, middle and south part of Taiwan were collected to examine the psychometrics characteristics of research tools and the relational model. The results were showed as following:

1. The Online Reading Comprehension Assessment for Junior High School Students fits Rasch Model, there is measurement invariance (gender DIF) existed.
2. Students' background variables affected online reading ability, yet the effect was not strong.

Finally, several suggestions were offered for further studies.

**Keywords:** on-line reading comprehension, Rasch measurement model

---

1. Hsiu-Shuang Huang, Professor, Department of Education, National University of Tainan

2. Ya-Ying Tseng, Principal, Houliiao Elementaty School of Changhua County

3. Mon-Long Gan, Associate Professor, Ningde Normal University

4. Li-Yun Hsu, Director of Academic Affairs, Yunlin County Kouhu Junior High School

5. Pin-Chen Huang, Secretary of Director-General, Bureau of Education, Tainan City Government

Received: 2017.10.12; Revised: 2019.09.25; Accepted: 2019.09.26

Corresponding Author: Ya-Ying Tseng; Email: yyt9487@gmail.com

Address: No.30, Fangliao Rd., Houliiao Vil., Fangyuan Township, Changhua County 52861, Taiwan Houliiao Elementaty School of Changhua County

## 壹、緒論

十二年國民教育課程綱要中之「科技資訊與媒體素養」能力中揭示：學生須具備科技與資訊應用的基本素養，並理解各類媒體內容，進行分析與批判（教育部，2014），其強調學生能運用科技資訊以判斷媒體資訊來源之正確性與適當性。因此，引導學生如何使用科技蒐集資訊，並且在網路龐大資料中找尋正確知識的技巧，已成為學生線上閱讀之必備能力，亦為教學與研究之重點。

Henry（2006）指出，學生倘無線上閱讀與搜尋能力，容易迷失在網路龐雜資訊中，引導學生學習並具備足夠線上閱讀理解素養將會影響學生線上閱讀表現，也連帶影響學生日後應用龐大網路資源能力與職場競爭力。

雖然線上閱讀的重要性已經廣泛受到肯定，但和傳統紙本閱讀相比，不論國內外都仍然只有少數具有信效度的評量工具（張貴琳，2011；曾雅瑛，2012；葉雅鏞，2009；Organization for Economic Co-operation and Development [OECD], 2011, 2013）。當前之教學強調績效責任（accountability），透過科學化的方法達成目標，其中重要關鍵就是需要有可信之工具以評量學生的發展階段及成就水準，透過工具才能從評量結果來改善教學（Thomas & Pring, 2004）。

線上閱讀素養議題日益重要，教學中如何有效探知學生線上閱讀素養亦為推動線上閱讀教學之關鍵，若能透過有信效度的評量，將能協助教師了解學生線上閱讀理解之情形，以做為改進教學之依據。

線上閱讀能力的表現受到許多因素影響，因此 OECD（2006）的閱讀架構考量這些因素，在評比閱讀理解表現時，也同時蒐集影響線上閱讀表現的相關變項，本研究亦從不同背景變項探討學生之線上閱讀理解表現。

綜合上述背景動機，本研究之目的如下：

- 1.編製「國中生線上閱讀理解能力測驗」。
- 2.分析不同背景變項學生之線上閱讀理解能力及其差異。

## 貳、文獻探討

### 一、線上閱讀素養之定義與歷程

線上閱讀情境通常都涉入一個以上的線上閱讀歷程，例如：幾乎所有的線上閱讀活動都需要以訊息提取為基礎，像是評鑑討論區文章前，必須要進行訊息的提取才能有適當的內容來評鑑，若評鑑是針對跨文本的主題內容，就需要整合解釋的歷程，才能歸納出評鑑所需的具體資訊。因此，在設計線上閱讀測驗的時候，考量歷程的連貫性，通常會將關聯的歷程整合成一整體的構念，以利設計題目。國際學生能力評量計畫（Programme for International Student Assessment, PISA）的測驗就將五種歷程整合為「擷取與檢索」、「統整與解釋」及「省思與評鑑」三個構面設計題目（OECD, 2006, 2011, 2013）。因此，以下就此三個構面分述探討其與線上閱讀的關係。

#### （一）「擷取與檢索」是線上閱讀素養的基本歷程

擷取與檢索是線上閱讀最基礎的歷程，若細分其內涵可以包含直接提取、直接推論與線上搜尋。PISA 的閱讀素養主要分為擷取與推論兩大技能，其中之直接提取與直接推論都同樣也是紙本閱讀的基礎歷程，而搜尋則是線上閱讀獨特而重要的閱讀歷程（OECD, 2011, 2013）。

Coiro 與 Dobler（2007）指出，由於線上可以進行搜尋，以協助資訊的提取與統整，因此在閱讀的先備知識上，必須具備網站結構的先備知識以及搜尋引擎的先備知識。

綜上所述，可知「線上搜尋歷程」為理解線上閱讀之重要成分，因此相關評量都將「線上搜尋」納入考量。

#### （二）「統整與解釋」與線上閱讀素養的關係

在 PISA 的閱讀架構中，閱讀歷程可以分為兩類：一類是使用文本內資訊，另一類是利用外部知識；前者除了上述的直接擷取外，另一個就是統整與解釋。透過統整與解釋可以達成「形成廣義的了解」以及「建構詮釋內容」（OECD, 2006, 2011, 2013）。

統整在線上閱讀中占有重要的地位，因為線上閱讀有較高的比例會在跨文本以及非連續文本的情境下進行，例如：在 PISA 2012 中，線上情境中的跨文本題目就占了 81%，比紙本情境中的 2% 高出許多（OECD, 2013）。

### （三）線上閱讀結合文本以外的知識形成高層次的「省思與評鑑」

省思與評鑑是 PISA 閱讀架構中需要連結外在知識的閱讀歷程，兩者都必須在原來文本之外，連結其他的知識才能進行（OECD, 2006, 2011, 2013）。

針對文本內容的省思與評鑑，舉例來說，像是針對議論文中的不同觀點進行異同的比較，決定何者較為可信或是正確。若是針對文本形式進行省思與評鑑，則是對於文本的形式本身與所知的其他文本形式做比較或是給予評價，如對於論述文本的推論或是歸納的段落安排進行評議。

文本的省思與評鑑屬於較高層次的閱讀表現，因為需要對文本達到基本的了解，才有辦法進行此類的閱讀歷程。以所有參加評比的國家來看，普遍來說，「省思與評鑑」的表現不如「統整與解釋」（臺灣 PISA 國家研究中心，2012），可知提升學生線上閱讀理解高階能力的重要性。

## 二、線上閱讀測驗的測量情境

Leu（2007）將線上閱讀的測驗工具根據施測情境分為三類：一種是在虛擬的網路上進行，一種是在真實的網路情境下進行，最後一種以呈現特定的線上瀏覽畫面讓受試者作答。張貴琳（2012）則是將線上閱讀依評量情境的真實性分為三種不同的測量取向：第一為在真實的網路情境中進行小規模的質性研究，採用認知歷程分析蒐集學生的線上閱讀理解策略和行為組型；第二為在有限的資訊空間評量學生的問題解決表現，採用虛擬的網際網路，但其無法充分代表線上閱讀的全觀複雜度和挑戰性，例如：美國教育測驗中心進行的大規模 PISA 資訊與溝通科技（Information and Communication Technologies, ICT）素養評量；第三為提供靜態的片段螢幕畫面，評量學生在面對某個特定的線上閱讀情境之選擇，但其無法真正測量學生擷取、分析、統整和溝通資訊的能力。

上述的三種測量方式各有不同的優缺點，真實的情境最符合現實的線上閱讀情形，但無法掌握各個網頁的變動情形，因此無法穩定控制測驗的各項變數；採用模擬的線上閱讀情境雖能有效地解決上述問題，但開發成本很

高，且若模擬不當就無法完整呈現真實情境，此類的測驗內容固定，題目類型相較之下就比較受限。本研究之線上閱讀理解測驗係在真實網路情境中測量，貼近學生日常的線上閱讀經驗。

### 三、影響線上閱讀理解能力之因素探究

閱讀能力的表現受到許多因素影響，本研究分析學校所在地區、性別、家庭網路環境與其線上閱讀理解的關係。

#### （一）性別與線上閱讀理解的關係

以臺灣學生來說，不論是在促進國際閱讀素養研究（Progress in International Reading Literacy Study, PIRLS）或 PISA 的閱讀表現都是女性優於男性（柯華蕓、詹益綾、張建妤、游婷雅，2009）；張貴琳（2012）的研究亦發現，中學生的線上閱讀素養表現存在顯著的性別差異；Wu（2014）分析 2009 年 PISA 的 19 個國家、34,104 個 15 歲受試者的資料，結果顯示大多數國家不同性別學生的線上閱讀活動並沒有顯著差異，女性在後設認知策略、導航技巧以及書面閱讀評量的表現較佳，但在線上閱讀評量並沒有差異。在紙本閱讀上，國外研究也普遍指出女性在語文的表現優於男性（Logan & Johnston, 2009; Marksa, 2008），但性別在線上閱讀理解的影響則結果不一，因此本研究欲探討性別與線上閱讀理解的關係。

#### （二）家中電腦、網路可及性（accessibility）與學校位置

Forzani、Leu、Kulikowich、Sedransk 與 Coiro（2014）的研究發現，低社經學區學生的線上閱讀理解能力明顯低於高社經學區學生，因來自低社經學區學生在家中可能較少有使用電腦的機會，臺灣的學校位置常與學校擁有的資源及家長社經背景有關，都會型學校在這些條件上一般優於鄉鎮型學校，而這些因素是否如國外研究一樣，會影響學生的線上閱讀理解能力，值得探討。

## 參、研究方法

### 一、研究架構

本研究主軸為編製「國中生線上閱讀理解能力測驗」，由於線上閱讀理



解能力測驗同時涵蓋「擷取與檢索訊息」、「統整與解釋訊息」以及「省思與評鑑訊息」三個層面試題，故模式中將以此三個層面分數，做為線上閱讀理解能力的測量指標。

## 二、研究樣本

### （一）預試樣本

本研究以中、南部地區的兩所國中，共計 200 名學生作為預試樣本，根據其作答反應修正相關試題，形成正式題本。

### （二）正式樣本

本研究採取立意取樣，自臺灣北、中、南地區抽出 9 所學校，共合計 1,188 名國中學生做為正式研究對象。扣除無效問卷資料後，共得有效樣本數 1,107 名，有效樣本率為 93.18%。在這些樣本資料中，男性共計 527 名，女性則計有 568 名，另有 12 名漏填性別選項。

表 1 正式抽樣人數之學校背景資料分布

學校類型	學校地區	樣本人數
鄉鎮型	高雄市 A 國中	114
都會型	臺南市 B 國中	110
都會型	臺南市 C 中學	187
都會型	臺南市 D 國中	115
鄉鎮型	雲林縣 E 國中	175
鄉鎮型	雲林縣 F 國中	129
鄉鎮型	新竹縣 G 國中	92
鄉鎮型	苗栗縣 H 國中	95
都會型	基隆市 I 國中	90
總人數		1,107

### 三、研究工具

#### （一）編製依據

針對「國中生線上閱讀理解能力測驗」的編製，本研究主要係依據 PISA 2012 (OECD, 2013) 的閱讀評量架構：「擷取與檢索訊息」、「統整與解釋訊息」以及「省思與評鑑訊息」做為試題編製的理論依據。在試題型式上則參考「國小高年級線上閱讀理解測驗」（曾雅瑛，2012）進行設計。本研究依此編擬 10 題涵蓋各種線上閱讀文本之試題，並將這些試題建置在 SurveyMonkey.com 線上施測平臺。

#### （二）計分方式

「國中生線上閱讀理解能力測驗」的題型包含選擇題 7 題與開放性試題 3 題，表 2 為試題雙向細目表。選擇題採取答錯給 0 分、答對給 1 分的二元計分；開放性試題採取 0 分、1 分、2 分、3 分的多元計分，測驗總分為 16 分。表 3 為經過預試後修改的開放性試題之評分規範與作答示例。

#### （三）施測方式

本測驗在 Survey Monkey 平臺上採線上團體施測，測驗時間為 40 分鐘。題目中提供連結，學生點選連結即進入資料查找畫面，關掉連結就回到施測平臺，無法連結到與試題無關網頁，以避免測驗干擾誤差。此外，施測人員亦巡視學生作答情形，以確保學生不會進行與測驗無關活動，提高測驗完成率。

#### （四）預試分析

由於本測驗同時包含二元及多元計分試題，故採用 Rasch 部分給分模式（partial credit model）進行預試題本的試題品質檢核。詳細分析結果如下所述。



表 2 「國中生線上閱讀理解能力測驗」雙向細目表

	記憶	理解	應用	分析	評鑑	創造
1. 從該網頁內容中的哪一個分類項目，可查詢到舞獅一詞訊息？		◎				
2. 你認為以下網頁連結中所提供的辭典，哪一部最適合大衛學習？			◎			
3. 請根據臺南市公共圖書館網的年度借閱排行榜，指出「科學類」的圖書中，哪一本書被借閱的次數最多？		◎				
4. 【短片：同理心】請判斷下列選項中何者最符合「同理心」的意義？				◎		
5. 【短片：僱傭人生】下列選項中，何者較能描述影片中之主角以外人物的象徵意涵？			◎			
6. 【短片：僱傭人生】下列選項中，何者是該片創作者比較想要傳達的意義？					◎	
7. 請判斷下列選項中何者最符合文天祥心目中對國家盡「忠」的意義？					◎	
8. 【短片：同理心】你認為大和最後有了解同理心嗎？請說明理由。		◎				
9. 判斷小玉說法是否正確，並從文章中找出資訊來支持你的說法。			◎			
10. 小華收到轉寄的電子郵件「何瑞君小弟病危」，請幫小華判斷是否該捐款給何瑞君小弟？					◎	

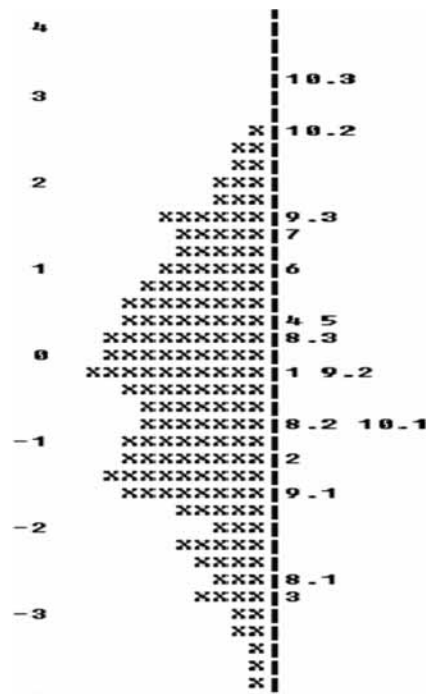
表 3 「國中生線上閱讀理解能力測驗」之開放性試題評分示例

試題	評分規範			
	計 0 分	計 1 分	計 2 分	計 3 分
【短片：同理心】你認為大和最後有了解同理心嗎？請說明理由。	未作答；學生認為大和最後並沒有了解同理心意義。	學生雖回答大和最後有了解同理心意義，但在作答反應中，沒有任何證據顯示其判斷說法是正確的。	學生認為大和最後有了解同理心意義。且在說明中提到「同學面質大和」的這段過程，但未能再進一步指出大和由同學面質過程中，也感同身受到他傷害欣好，進而同理欣好受到傷害的心情。	學生認為大和最後有了解同理心意義，並且能於作答中明確指出，大和在同學面質過程中，也感同身受他確實傷害了欣好，因而同理欣好受到傷害的心情。
判斷小玉說法是否正確，並從文章中找出資訊來支持你的說法。	未作答；學生指出小玉說法正確。	學生雖指出小玉說法不正確，但在作答中並沒有提出任何說明，或是無法清楚交代判斷小玉說法不正確的理由。	學生指出小玉說法不正確，但在作答中，僅回答「由文章中的實驗結果得知」，沒有進一步提及任何實驗細節（例如：使用何種溶液作測試）。	學生指出小玉說法不正確，於作答中，除了： 1. 可由文章中的實驗結果得知； 2. 並且，能夠具體交代實驗細節（例如：使用氫氧化鈉溶液測試水果的酸度；使用碘液為氧化劑，澱粉為指示劑來測試各種水果的維他命 C 含量）。
小華收到轉寄的電子郵件「何瑞君小弟病危」，請幫小華判斷是否該捐款給何瑞君小弟？	未作答；或作答反應中未寫出任何具體理由。	學生判斷不應捐款理由中，僅指出郵件內容中 1 個可疑之處。	學生判斷不應捐款理由中，能指出郵件內容中有 2 個可疑之處。	學生判斷不應捐款理由中，能指出郵件內容中 3 個以上可疑之處。

### 1. 試題契合度及難度分析

首先，為檢驗試題是否測量線上閱讀理解能力之構念，本研究以 *infit* (weighted) 及 *outfit* (unweighted) MNSQ 指標進行模式－資料的適配度分析。若題目適配度佳，表示題目所測量到的是相同構念；若題目適配度不佳，則須進行刪題。一般而言，多元計分題的合理 MNSQ 數值範圍介於 0.60~1.40 之間 (Bond & Fox, 2007)，超過此範圍表示該題目可能與理論模式不符。此外，依據 Chien (2006) 的建議，當未加權 MNSQ 和加權 MNSQ 互有高低時，以加權 MNSQ (*infit* MNSQ) 為決裁標準。

本研究結果顯示：所有測驗試題的 *infit* MNSQ 值介於 0.72~1.40 之間，因此可判斷所有試題的適配度皆能契合 Rasch 測量模式。其次，所有試題的難度值介於 -2.78~3.30 之間，圖 1 即為「試題－受試者分布的線性關係圖」(Wright Map)。



註：每個 X 代表 1.2 個受試者。數字 8、9、10 分別為本測驗開放性問題，其中數字 8.1 表示第 8 題的第 1 個閾值，8.2 表示第 8 題的第 2 個閾值，8.3 表示第 8 題的第 3 個閾值。

圖 1 預試分析之 Wright 圖

## 2. 試題等級適切性

該項分析主要檢定多元計分試題的閾值是否呈現合理的階層次序性。一般而言，低等級的閾值須低於高等級的閾值，表示當受試者能力值愈高時，達到較高等級的機會也愈高。此外，等級之間也應有適切的間距，應介於 1~1.4 個 logit，最多不得超過 5 個 logit。

本研究的開放題計分採 4 個等級（0~3 分），表 4 顯示每一開放題 4 個等級的閾值皆具有階層次序性，即得分愈高，所需要對應的能力值便愈高，惟第 8 題得 2 分跨到 3 分之間的間距過窄、第 10 題得 2 分跨到 3 分之間的間距也有過窄之虞。

## 3. 評分者一致性

兩位分別擔任中小學國文教師的研究團隊成員依據評分規範，隨機抽取 50 名預試樣本進行評分者間一致性的檢驗。兩位評分者針對三題開放性問題的評分者一致性信度介於 .88 ~ .90，評分者間的一致性甚佳。

綜合預試結果顯示：試題不僅符合 Rasch 模式契合度指標，試題難度值分布亦理想，惟在開放性問題的等級計分上似有改善空間，故研究者重新調整計分規準，修改後之計分規準如表 3 所示。

## 四、資料處理與分析

本研究針對各項研究目的，分別以 SPSS 12.0 版、LISREL 8.52 版，以及 CONQUEST 2.0 版進行資料分析與處理，詳細之各項檢定方式分別說明如下。

本研究旨在建立具有良好心理計量特性之「國中生線上閱讀理解能力測驗」。由於該測驗試題中，選擇題採二元計分（0、1），開放性試題則採多元計分（0~3）；因此，在資料處理與分析上，主要是確定作答資料是否符合 Rasch 部分給分模式，並同時針對相關的試題參數、受試者能力參數，以及性別之試題差異分析檢測（Differential Item Functioning, DIF）進行考驗。

## 肆、研究結果與討論

### 一、「國中生線上閱讀理解能力測驗」之心理計量分析

#### (一) 內在結構之憑證

##### 1. 等級適切性

本測驗中的開放性試題計有 3 題，採四等級計分，參考 Linacre (2002) 建議：(1)各等級的次數應大於 10，才能獲得較穩定的閾值估計值；(2)等級閾值應呈現次序階層性，即得分愈高，所需對應的能力值便愈高；(3)理想的閾值間距應介於 1.0~5.0 個 logit，以能夠反應等級適切性並提供足夠區辨受試者的能力（余民寧、陳柏霖、陳玉樺，2017；張麗麗，2015；Linacre, 2004）。

本研究結果如表 4 所示，在 3 題開放性試題中，每一試題之等級閾值皆能呈現階層次序性；然而，第 10 題最高等級次數不及 10，可能影響參數估計之

表 4 開放性多元計分試題各等級之次數及閾值

題號	修正前			修正後		
	等級	次數	閾值	等級	次數	閾值
8	0	259	---	0	259	---
	1	274	-0.40	1	504	-0.62
	2	230	0.38	2	344	1.16
	3	344	0.96			
9	0	297	---	0	297	---
	1	98	-0.24	1	692	-0.58
	2	594	0.01	2	118	2.52
	3	118	2.40			
10	0	389	---	0	389	---
	1	686	-0.22	1	686	-0.21
	2	25	3.02	2	32	3.85
	3	7	3.44			

穩定性；此外，試題 8、9、10 的閾值間距也皆有過窄情形（不及 1.0 個 logit）。因此，基於上述建議標準下，本研究改以三等級計分（0~2 分），修訂後結果即呈現較為理想的等級適切性。本研究也據此進行後續信、效度之分析。

## 2.單向度檢核

在個別試題契合度方面，本研究以 Bond 與 Fox（2007）提出合理的 MNSQ 數值範圍應介於 0.60~1.40 之間為標準，並再參照 Chien（2006）建議，當未加權 MNSQ（outfit MNSQ）和加權 MNSQ（infit MNSQ）互有高低時，以 infit MNSQ 為決裁標準。

由表 5 得知，所有試題的 outfit MNSQ 數值介於 0.87~1.14 之間、infit MNSQ 數值則介於 0.88~1.10。此一結果顯示，所有試題皆能契合 Rasch 單向度模式，即所有試題皆在測量相同的構念，故線上閱讀理解能力可以由單一構念解釋。

### （二）試題難度

由表 4 及 5 可知，就整份測驗而言，各試題平均難度介於 -1.26~1.82 之間，測驗難度之平均值則為 0，其中難度值最低的是第 4 題，難度值最高的是第 10 題；受試者能力值介於 -1.03~1.75，平均能力值為 0.44，此顯示受試者能力分布略高於題目的難度分布。

然而，若考慮多元計分試題的閾值難度，則試題難度最高依序出現在第 10 題的第二個閾值 3.85 及第 9 題的第二個閾值 2.52。換言之，受試者在這兩個開放性問題中要拿到高分是極為困難的。

其次，圖 2 顯示，不同能力受試者皆有相對應難度之試題，然而在難度值 1.00~2.00 間，似乎仍有填補試題的空間。

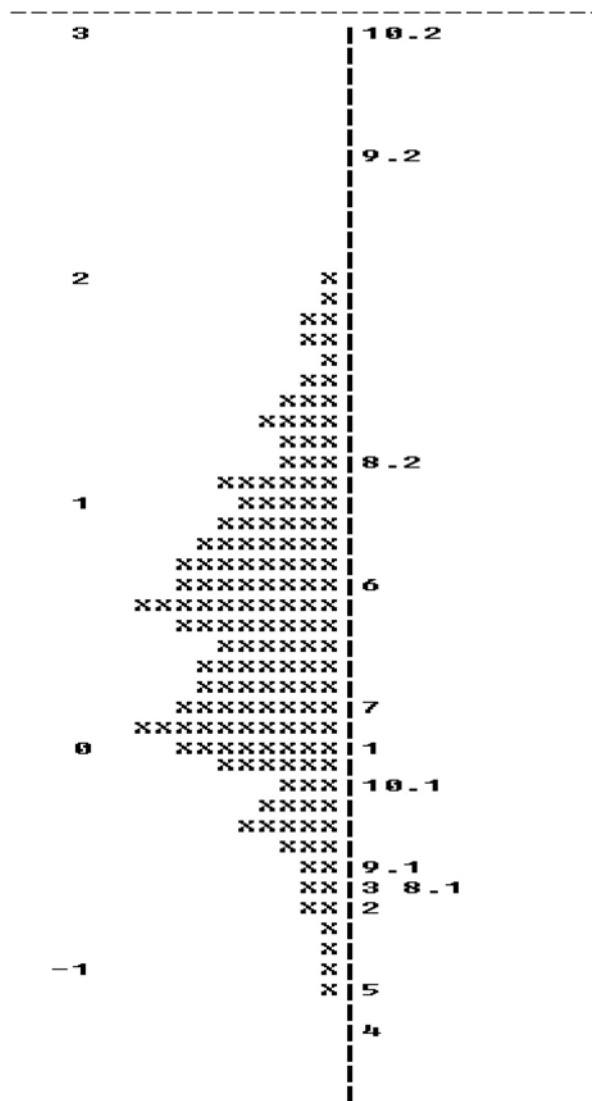
此外，整合不同的線上閱讀理解歷程發現，「統整與解釋」試題有時比「擷取與檢索」試題更容易，這樣的數據與 PISA 調查結果相同，也與張貴琳（2012）的研究發現一致。原因可能在於，文本中某些細部資訊不容易被察覺，而這些訊息卻是答題的關鍵資訊（OECD, 2011），且資訊導航複雜程度的差異性也可能是影響彼此難易度差別之因素。另外，省思與評鑑線上訊息之試題平均難度對全體受試者而言是最難的部分，大部分學生無法有效根據網站資訊作為判斷或推論的依據回答問題，而僅是根據自己的意見回答。整體而言，「國中生線上閱讀理解能力測驗」試題的難度屬於「適中偏易」。

表 5 試題難度及適合度指標

閱讀歷程	試題內容	題型	難度參數	標準誤	outfit MNSQ	infit MNSQ
擷取與檢索 訊息	1. 從該網頁內容中的哪一個分類項目，可查詢到舞獅一詞訊息？	選擇	0.04	.04	0.94	0.95
擷取與檢索 訊息	2. 你認為以下網頁連結中所提供的辭典，哪一部最適合大衛學習？	選擇	0.77	.05	1.09	1.06
擷取與檢索 訊息	3. 請根據臺南市公共圖書館網的年度借閱排行榜，指出「科學類」的圖書中，哪一本書被借閱的次數最多？	選擇	-0.62	.05	1.14	1.10
統整與解釋 訊息	4. 【短片：同理心】請判斷下列選項中何者最符合「同理心」的意義？	選擇	-1.26	.05	1.09	1.04
統整與解釋 訊息	5. 【短片：僱傭人生】下列選項中，何者較能描述影片中之主角以外人物的象徵意涵？	選擇	-1.14	.05	1.03	1.02
統整與解釋 訊息	6. 【短片：僱傭人生】下列選項中，何者是該片創作者比較想要傳達的意義？	選擇	0.67	.04	1.09	1.08
統整與解釋 訊息	7. 請判斷下列選項中何者最符合文天祥心目中對國家盡「忠」的意義？	選擇	0.10	.04	1.05	1.04
省思與評鑑 訊息	8. 【短片：同理心】你認為大和最後有了解同理心嗎？請說明理由。	開放	0.27	.04	0.93	0.94
省思與評鑑 訊息	9. 判斷小玉說法是否正確，並從文章中找出資訊來支持你的說法。	開放	0.97	.04	0.88	0.88
省思與評鑑 訊息	10. 小華收到轉寄的電子郵件「何瑞君小弟病危」，請幫小華判斷是否該捐款給何瑞君小弟？	開放	1.82*	--	0.87	0.88

註：\*為參數限制。





註：每個 X 代表 7.1 個受試者。

圖 2 正式施測分析之 Wright 圖

### （三）試題的 DIF 分析結果

為了避免學生因性別差異導致試題作答上產生有利或不利的情形，因此本研究進行性別 DIF 檢核，以探討各試題是否因性別影響而導致 DIF 存在。根據美國教育測驗服務社（Educational Testing Service, ETS）對於 DIF 效果的解釋，DIF 效果可以區分為 A、B、C 三個等級，將其換算為 logit 尺度而言，若難度差異值（即兩組 DIF 估計值之差距，取絕對值）低於 0.4 個 logit 以下，則歸於 A 類，表示幾乎沒有 DIF 存在；若難度差異值超過 0.6 個 logit 以上，則歸於 C 類，表示有顯著 DIF 效果存在；其餘則歸於 B 類。

表 6 分析結果顯示，本測驗所有試題難度估計值在性別上的差距皆小於 0.5 個 logit，換言之，男、女性在所有題目上並無明顯的 DIF。

表 6 試題與性別的交互作用

試題	DIF 參數估計值		全距
	男性	女性	
1	0.17	-0.17	0.34
2	0.07	-0.07	0.14
3	0.16	-0.16	0.32
4	0.20	-0.20	0.40
5	0.12	-0.12	0.24
6	-0.19	0.19	0.38
7	-0.01	0.01	0.02
8	-0.19	0.19	0.38
9	-0.13	0.13	0.26
10	-0.22	0.22	0.44

## 二、信度分析

本測驗在 Rasch 模式的試題分隔信度（item separation reliability）為 .99，表示受試者在分隔試題難度次序上極為穩定，即當試題給予另一組具備相同能力特質的受試者施測時，試題難度排序的一致性或穩定性極高。試題分隔指數（item separation index）為 9.95，高於 Linacre（2012）建議要大於 3 的標準，顯示測驗試題具有一定的難度階層，試題間的難易度有顯著不同。然而，本測驗的受試者分隔信度（person separation reliability）為 .54，則稍不

理想，表示試題較難以穩定分隔受試者能力，另外受試者分隔指數（person separation index）為 1.08，也意謂測驗對於兩極端受試者樣本的偵測較不敏感，可能原因包括：測驗題數不足、試題反應類別不足，或是受試者能力變異不足等因素影響導致（Boone, Staver, & Yale, 2014）。本測驗開放性問題之評分者原始計分的 Spearman 相關係數介於 .88~.89，顯示評分者一致性良好。

然而，相關線上閱讀理解測驗試題都有信度偏低的傾向（張貴琳，2012），顯示在考量受試者線上導航能力、線上施測時間，以及受試者在電腦教室使用線上網路的行為秩序等複雜因素，通常編製者對於此類測驗長度採取較保守的立場，為了要能更加以區分高、低能力受試者，本測驗未來仍有必要加以修訂。

### 三、不同背景變項國中生線上閱讀理解能力之表現及其差異分析

「國中生線上閱讀理解能力測驗」在經由 Rasch 分析確保具一定可信賴的信、效度基礎上，研究者進一步針對國中生的相關背景差異進行分析、比較，嘗試了解此背景因素和線上閱讀理解能力的關係。本測驗經先前開放性試題計分的修訂後，測驗原始得分範圍落在 0~13 分，若對應於 CONQUEST 軟體以最大概似法（maximum likelihood, ML）估計受試者能力參數值則落在 -4.03 至 5.26 個 logit。

表 7 即呈現不同背景變項受試者在「國中生線上閱讀理解能力測驗」的平均能力估計值，研究者據此進行各項平均能力差異檢定。其中，就背景變項所涉及家中是否具備電腦、家中是否能上網等兩個題項，研究者在考量家中即便有電腦，但能否上網對於線上閱讀的體驗和表現仍可能存在一定的效果差異下，故視資料處理情形評估二者間的交互作用效果。

整體而言，本研究結果顯示，除了「性別」、「就讀學校類型」以及「家中是否能上網」等因素外，其餘背景因素在線上閱讀理解能力上並無顯著差異。而就前述的性別變項而言，在 levene 檢定顯示  $F$  值為 0.42， $p > .05$  符合變異數同質性的假設下，本研究發現女性在整體線上閱讀理解能力的表現較男性為佳（ $t = -4.50$ ， $p < .05$ ），此結果與其他多數研究發現一致（柯華葳等人，2009；張貴琳，2012；Logan & Johnston, 2009；Marksa, 2008）。此外，性別對線上閱讀理解能力的  $\eta^2$  為 .018，依 Cohen（1988）標準判斷，兩者間屬於低度關聯程度。

表 7 不同背景變項之國中生線上閱讀理解能力表現

	背景變項	樣本數	平均能力值 (logit)	標準差
性別	1.男	527	0.35	1.04
	2.女	568	0.63	1.01
就讀學校類型	1.都會型學校	502	0.62	0.77
	2.鄉鎮型學校	605	0.39	1.20
家中是否有電腦	1.有	1,053	0.50	1.04
	2.否	54	0.35	1.12
家中是否能上網	1.能	1,069	0.50	1.05
	2.否	38	0.16	0.87
每週平均上網次數	1.不到 1 次	52	0.32	1.18
	2.每週 1 至 2 次	149	0.59	1.11
	3.每週 3 至 4 次	147	0.47	1.04
	4.每週 5 至 6 次	127	0.50	0.99
	5.天天上網	632	0.48	1.03
常使用電腦進行哪些閱讀活動？	1.上網閱讀電子書	73	0.54	0.98
	2.使用電腦看故事光碟	16	1.01	1.46
	3.上網蒐集相關資料	234	0.41	1.02
	4.上網閱讀部落格文章、臉書及 LINE	761	0.51	1.04
	5.下載電子書後離線閱讀	23	0.16	0.96

在就讀學校類型上，由於 *levene* 檢定  $F$  值為 71.47， $p < .05$  違反變異數同質性假設，故改採校正後的  $t$  值。研究結果發現，就讀都會型學校受試者的表現顯著優於鄉鎮型學校受試者（ $t = 3.60$ ， $p < .05$ ），而其  $\eta^2$  為 .012，顯示就讀學校類型與線上閱讀理解能力間屬於低度關聯程度。

此外，有關於家中是否能上網的 *levene* 檢定  $F$  值為 0.46， $p > .05$  同樣符合變異數同質性下，本研究發現在家中可以上網的受試者，其線上閱讀理解能力相對家中無法上網的同儕更具有優勢（ $t = 1.97$ ， $p < .05$ ），顯示過去一直存在的數位落差問題確實值得關注。然而，家中是否能上網和受試者線上閱讀理解能力的關聯性  $\eta^2$  為 .004，僅屬於微弱程度。研究者認為，兩者關聯性不如預期得高或許也隱含一定的合理性，因為根據相關研究發現，書面閱讀能力與線上閱讀理解能力具有相當程度之關聯性，顯示兩者在其中具有能力內涵上的重疊（曾雅瑛，2012；Coiro, 2003），而二者所共同需求的閱讀理解能力的培養不須具備電腦或網路設備。由此，某種程度或許也能同樣理

解為何天天頻繁上網的青少年，其線上閱讀理解能力的表現未必比一般未時常上網的同儕更佳，輔以目前青少年可以選擇上網的管道途徑眾多，這些因素或許皆降低此一背景性質的重要性。

最後，無論就受試者傾向應用電腦進行何種閱讀活動，其對線上閱讀理解能力亦無影響。推測原因，可能是在利用電腦進行各種閱讀活動的過程中，其實都需要具備一定程度線上閱讀理解能力，故無法由青少年所傾向的線上閱讀活動中去凸顯其線上閱讀理解能力的差異。

## 伍、研究結論與建議

### 一、結論

#### （一）「國中生線上閱讀理解能力測驗」符合 Rasch 測量模式，且所有題目在男、女性並無明顯的差異試題功能

本研究編製之「國中生線上閱讀理解能力測驗」，在構念效度上，無論個別試題的 Rasch 契合度指標、性別 DIF 檢定等結果，皆支持本測驗所測量的線上閱讀理解能力構念為單向度，受試者及試題可置於同一量尺上相互比較；在開放性問題的等級結構上，經閾值分析選擇適切的等級數；在試題難度適切性上，試題難度尚能與受試者能力相對應，開放性問題之難度明顯高於選擇題之難度，可分別測量中／高能力之受試者；在線上閱讀理解歷程的試題難度次序上，也呼應文獻上的發現；在受試者之線上閱讀理解能力表現上，本研究呈現女性優於男性的表現，其結果與文獻相同。

#### （二）不同背景變項雖然在線上閱讀理解能力具有差異，惟影響程度並不高

就線上閱讀理解能力而言，本研究所探討之各種背景變項中，除了性別因素外，包括：家中是否有電腦網路設備、上網頻率和線上閱讀活動等情形，這些背景變項在對比下，反應數位環境落差可能帶來的優劣勢區別，惟這些因素影響幅度皆不如預期得高。這原因可能是除了國中生使用電腦上網已是極為普及的趨勢下，某些學生即使無法在家中上網，仍可以透過許多途

徑獲得，因而縮短了彼此差距。

## 二、建議

### （一）「國中生線上閱讀理解能力測驗」在教學上的省思與改善

網際網路提供許多訊息，卻也併存許多零碎資訊、偽資訊或假資訊，這些有賴於使用者能針對不同訊息做出可靠評價。然而，就線上資訊的省思與評鑑能力而言，本研究發現國中生在此方面較薄弱。因此，未來可針對線上資訊的批判性作為閱讀教學關鍵，讓學生可以從中找出支持或否定某論點的有用資訊，更可從多元化線上文本評鑑資訊真偽，培養其良好的線上閱讀能力。

教學中，教師可關注如何協助學生在線上閱讀理解文本中，除理解文本外，更關鍵的是評鑑文本資訊之能力，透過多元文本舉例練習，協助學生有效率的評鑑線上資料。

### （二）「國中生線上閱讀理解能力測驗」之進一步修訂

本研究編製之「國中生線上閱讀理解能力測驗」，強調試題情境架構在真實網路瀏覽，以得知學生線上閱讀理解實況。雖然該測驗試題皆符合 Rasch 模式適合度指標，但其中建構省思與評鑑能力開放性試題，未來仍有進一步修訂之必要，例如：本測驗開放性試題其一描述：「小華收到一份轉貼寄來的電子郵件，信件標題是『何瑞君小弟病危』，信件內容如下述。請你閱讀信件後，根據信件提供的資訊，幫助小華判斷是否應捐款幫助何瑞君小弟？並說明你如此判斷的理由。」原先試題設計為避免影響受試者回答傾向，故未明確說明：「若能根據電子信件內容，指出愈多不可靠資訊，則得分愈高。」然可能因此之故，許多受試者雖回答不應捐款，但就判斷理由說明上多僅指出信件內容中一項不合理之處，這導致僅有極少數受試者能在該題得到高分，以致無法有效區隔受試者能力。此外，如何在有限作答時間下盡可能擴充測驗試題，進而提高信度，皆有待未來測驗的修訂。

測驗學刊，第 66 輯第 4 期

## 謝誌

本報告為科技部計畫編號 MOST103-2410-H-024-031 之計畫，感謝科技部支持，特此致謝。



## 參考文獻

### 中文部分

- Chien, S. (2006)。以 Rasch 分析協助測驗同分時之測量研究。取自 [http://raschsmile.blogspot.com/2006\\_09\\_01\\_archive.html](http://raschsmile.blogspot.com/2006_09_01_archive.html)
- 余民寧、陳柏霖、陳玉樺 (2017)。量表長度簡化研究：「簡式中小學教師主觀幸福感量表」修訂。**教育研究發展期刊**，13 (4)，27-56。
- 柯華蕓、詹益綾、張建好、游婷雅 (2009)。臺灣四年級學生閱讀素養：PIRLS 2006 報告 (第二版)。取自 <http://140.115.107.22/PIRLS%202006%20National%20Report%EF%BC%882nd%20Edition%EF%BC%89.pdf>
- 張貴琳 (2011)。國中學生線上閱讀素養構念之探討 (未出版之博士論文)。國立臺南大學，臺南市。
- 張貴琳 (2012)。國中學生線上閱讀素養發展現況調查。**教育研究與發展期刊**，8 (2)，87-118。
- 張麗麗 (2015)。Rasch 模式建置「國小四、五年級閱讀理解測驗」簡版。**屏東教育大學學報：教育類**，42，37-66。
- 教育部 (2014)。國民中小學十二年國民教育課程綱要。取自 <http://www.naer.edu.tw/files/15-1000-7944,c639-1.php?Lang=zh-tw>。
- 曾雅瑛 (2012)。國民小學高年級學生線上閱讀理解測驗編製與實驗教學驗證之分析 (未出版之博士論文)。國立彰化師範大學，彰化市。
- 葉雅鏞 (2009)。國中學生線上閱讀素養與策略調整特徵之探討 (未出版之碩士論文)。國立臺南大學，臺南市。
- 臺灣 PISA 國家研究中心 (2012)。臺灣 PISA 2009 結果報告。臺北市：心理。

### 英文部分

- Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Dordrecht, the Netherlands: Springer.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale,

- NJ: Lawrence Erlbaum Associates.
- Coiro, J. (2003). Reading comprehension on the Internet: Expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher*, 56(5), 458-464.
- Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, 42(2), 214-257.
- Forzani, E., Leu, D., Kulikowich, J., Sedransk, N., & Coiro, J. (2014). *The effects of teacher factors on students' ability to critically evaluate online information*. Storrs, CT: University of Connecticut. Retrieved from <http://www.orca.uconn.edu/research/research-reports/>
- Henry, L. A. (2006). Searching for an answer: The critical role of new literacies while reading on the Internet. *Reading Teacher*, 59(7), 614-627. doi:10.1598/rt.59.7.1
- Leu, D. J. (2007). *Expanding the reading literacy framework of PISA 2009 to include on-line reading comprehension*. A working paper commissioned by the PISA 2009 reading expert group. Princeton, NJ: Educational Testing Services.
- Linacre, J. M. (2002). Understanding Rasch measurement: Optimizing category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106.
- Linacre, J. M. (2004). Optimizing rating scale category effectiveness. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement: Theory, models and applications* (pp. 258-278). Maple Grove, Minnesota: JAM Press.
- Linacre, J. M. (2012). *Winsteps Rasch measurement computer program User's guide*. Beaverton, OR: Winsteps.com.
- Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading*, 32, 199-214.
- Marksa, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries. *Oxford Review of Education*, 34(1), 89-109.
- Organization for Economic Co-operation and Development. [OECD] (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris, France: Author.
- Organization for Economic Co-operation and Development. [OECD] (2011). *PISA 2009 results: Students on line: Reading and using digital information, explores students' use of information technologies to learn (Volume VI)*. Paris, France: Author.
- Organization for Economic Co-operation and Development. [OECD] (2013). *PISA 2012*

黃秀霜、曾雅瑛、甘孟龍、許力云、黃品甄 「國中生線上閱讀理解能力測驗」之發展及相關因素之研究

*assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris, France: Author. doi:10.1787/9789264190511-en

Thomas, G., & Pring, R. (2004). *Evidence-based practice in education*. NY: Open University Press.

Wu, J. Y. (2014). Gender differences in online reading engagement, metacognitive strategies, navigation skills and reading literacy, *Journal of Computer Assisted Learning*, 30 (3), 252-271.

