

國小高年級學童線上數位閱讀認知 負荷量表編製

陳新豐 屏東大學教育學系副教授

摘要

本研究旨在編制國小高年級學童數位閱讀認知負荷量表（Digital Reading Cognitive Load Scale, DRCLS），建立信度及效度資料。研究工具採用自編「國小高年級學童數位閱讀認知負荷量表」，命題內涵依任務／環境、學習者特性（認知能力、認知風格、先備知識與經驗）、以及環境與學習者特性的交互作用等 3 個向度，再區分為心智的努力（mental efforts）以及心理的負荷（mental load）等 2 個構面，採用李克特 5 點量尺的設計方式，總共 25 題。研究分二個階段進行，第一階段先編擬試題，並商請四位學科領域專家及二位國小教師來進行題項審查，具有良好的內容效度指數（Content Validity Index, CVI），繼而就 387 位預試樣本進行施測與蒐集資料分析，根據預試之量表信度與效度資料進行修正。第二階段，分層比率隨機抽取 35 個班級，針對 728 位國小高年級學童進行線上量表施測。測量信度方面，具有良好的內部一致性信度，在構念效度方面，以結構方程模式對二群隨機樣本群組分別進行理論模式和複核效度的驗證，結果具良好的模式適配度以及模式穩定度，在區辨和幅合效度方面，心智的努力與心理的負荷等二個構面分量表均可接受。綜上所述，DRCLS 在相關研究以及數位閱讀實務上具有應用價值。

關鍵詞：數位閱讀、認知負荷、結構方程模式、量表編制



Cognitive Load Scale of Digital Reading for Elementary School Students

Shin-Feng Chen

Associate Professor, Department of Education, National Pingtung University

Abstract

This study aims to establish a Digital Reading Cognitive Load Scale (DRCLS) for elementary school high-grade students in order to form a reliable and valid database. The DRCLS consists of three components: (1) task environment, (2) learner characteristic (cognitive capability, cognitive style, prior knowledge and experience) assessment, and (3) an interactive effect component to measure learner's engagement within the task environment. The interactive effect component measures a learner at two levels: mental effort and mental load. Twenty-five Likert-type scale items were produced for the DRCLS. The entire study was divided into two steps. Step 1 included developing the items and having them validated by four subject field experts and two elementary school teachers. Validated items were statistically analyzed using 387 pre-test samples. The 25 items were further refined according to the results from the pre-test data analysis. At the Step 2 data were collected from a stratified proportional sampling of 35 classes, which included 728 elementary school high-grade students. These students responded to the 25 items in a Web-based online environment. Analysis on data collected from these students revealed good internal consistency reliability. Structural equation modeling (SEM) was developed to measure two scale components: mental effort and mental load. The theory model was cross-validated using two random sample groups. Analysis of the SEM model to data revealed a good model fit to the data, good construct validity, and good divergent and convergent construct validity for both scales' components. The education value of using DRCLS to promote digital reading practice among students is also forwarded.

Keywords: cognitive load, digital reading, scale development, structural equation modeling



壹、緒論

認知負荷 (Cognitive Load, CL) 是認知心理學領域中的一項構念 (construct)，意指當個體在進行某項特定工作時，對個體認知系統的運作而產生的訊息負荷 (Sweller, van Merriënboer & Paas, 1998)。最早是由澳洲心理學家 Sweller 以認知心理學觀點進行人類認知資源運作的一系列研究後，彙集相關研究成果後所發展而成的理論 (Sweller, 1988)，對學習者而言，認知負荷的高低，存在許多影響因素，除了作業本身難易度與複雜性外，讀者個人對該項作業的專業知能與技能程度也會有所影響；換言之，當作業本身的性質越容易、越單純，則個體的認知負荷量就越少，個體本身具備此一作業的專業能力愈足夠，則個體的認知負荷量也就越少。此外，認知負荷的概念可以讓作業任務的設計者，考量工作過程中的認知訊息呈現的形式與強度，以避免強度過高的認知負荷讓作業任務執行者產生挫折感，進而降低了工作任務成效。認知負荷的相關研究在教育應用方面，國內有許多學者 (左台益、呂鳳琳、曾世綺、吳慧敏、陳明璋、譚寧君, 2011; 涂金堂, 2011)，研究如何降低學生在學習過程所承受的認知負荷，以提高學生的學習成效，這也是許多教育心理學者及第一線教師所關心的重點，因此本研究以認知負荷為主題以編制數位閱讀環境下的認知負荷量表。以下將以認知負荷的分類、測量策略、測量工具以及數位閱讀認知負荷等部分，說明如下。

一、認知負荷的分類

Gerjets、Scheiter 與 Cierniak (2009) 以及 Clark、Nguyen 與 Sweller (2006) 針對認知負荷區分為「內在認知負荷」(intrinsic cognitive load)、「外部認知負荷」(extraneous cognitive load) 及「增益型認知負荷」(germane cognitive load) 等三類，其中「內在認知負荷」(intrinsic cognitive load) 所指的是學習目標對學生所造成的心智負荷，這與學習者是否已具有學習內容、學習教材及學生對學習內容的先備知識有所關聯。其次「外部認知負荷」(extraneous cognitive load) 也被稱為「無益的認知負荷」(ineffective cognitive load)，外部認知負荷主要是學習者將認知資源，耗損在處理與學習要素之間沒有關聯的知識、技能與程序，並無助於學習要素的處理與學習基模的建構。另外「增益型認知負荷」(germane cognitive load)，與外部認知資源也有些類似，都是發生在學習者將認知資源投入

在處理學習要素以外的知識、技能與程序。但是，兩者不同的是：增益型認知負荷中額外處理的知識、技能與程序對學習內容的要素本身是有助益的，是有助於學習基模的建構，對於學習目標的達成也具有正向的效益。

Reid 與 Nygren (1988) 提出測量認知負荷可區分為時間負荷、心理負荷以及心理壓力負荷等 3 個向度，其中時間負荷包括時間壓力、時間可用性、時間需要性；心理負荷則包括所輸入的訊息內涵、互動性的活動、所耗費的能量、努力的程度、知覺任務的困難性以及任務的完整性；心理壓力負荷則是包括任務的成效、緊張的程度、失敗的可能性、不確定的程度、壓力、情緒上的壓力、動機、疲勞的程度等。另外 Hart 與 Staveland (1988) 則將測量認知負荷分為 3 個向度，分別是任務/環境、學習者特性（認知能力、認知風格、先備知識與經驗）、環境與學習者特性的交互作用等 3 個向度，再區分為心智的努力（mental efforts）以及心理的負荷（mental load）等 2 個構面。其中任務與環境的部分包括任務的目標、時間的因素、資源以及社會環境；學習者的特性則是包括選擇策略、認知能力、認知風格、先備知識與經驗等部分，本研究中所建構之國小高年級學童數位閱讀認知負荷量表，編制時考量認知負荷的完整性，主要參考 Reid 與 Nygren (1988) 之時間負荷、心理負荷以及心理壓力負荷等 3 個向度以及 Hart 與 Staveland (1988) 之任務/環境、學習者特性、環境與學習者特性的交互作用等 3 個向度為主要命題內涵，再區分為心智的努力以及心理的負荷等 2 個構面來加以編制。

二、認知負荷的測量策略

認知負荷的測量，主要是以學習者在學習作業時所知覺的困難度視為學習者的認知負荷量，並且再以學習者解題與遷移之能力來輔佐判斷，因此測量認知負荷的向度可以從任務（作業）（task-based dimension）、學習者（learner-based dimension）等 2 個向度來測量，其中作業的向度是因應任務需求而產生的，意即需要完成作業任務下所需要的努力，稱為心理負荷，包括內在認知負荷與外在認知負荷；學習者的向度是為學習者執行作業時所付出的認知能力與資源，稱為心智努力（Brunken, Plass & Leutner, 2003），本研究主要是採用心智努力與心理負荷這 2 個向度來測量認知負荷，而測量方式則是採用直接性與主觀的方式，在數位閱讀的環境下請受試者自陳認知負荷的情形。

在教育領域中，測量認知負荷可採用的方式有訪談、量表施測等。例如：對

學習者的壓力感受進行訪談，或是使用量表工具進行施測，都是在實務場域中常見的認知負荷測量方式，對於認知負荷的測量策略，一直以來也都是認知心理學家關注的一項議題（左台益等，2011）。以測量認知負荷實施的作法來區分，認知負荷的測量策略大致區分為「主觀測量法」、「生理測量法」及「任務表現測量法」等3種（DeLeeuw & Mayer, 2008）。其中「主觀測量法」指的是受試者反思及回顧自身在學習歷程中的認知及感受，自身評估自己的負荷感受之後，在量表工具上圈選出符合自己感受的對應尺度；其次「生理測量法」是假定在受試者產生認知負荷的情況下，會造成腦波、血壓或眼動等生理上的反應，透過相關儀器的檢測，來判斷受試者的認知負荷程度；至於「任務表現測量法」，則是從受試者在工作任務中所展現的績效表現與任務本身在設計的難易度與複雜度，反推個體在執行任務期間的認知負荷程度（Sweller, van Merriënboer & Paas, 1998）。其中，較不易實施的是「生理測量法」，根據 Paas 等人的看法，生理測量需藉助相關儀器方可進行，測量進行時，個體並無法持續執行工作任務，個體的注意力容易受到測量方式的影響而轉移到工作任務以外的地方，而且大多數的測量儀器所費不貲，因此，在教育現場的研究中較難以生理測量法來測量認知負荷；另一方面，主觀測量法在測量工具的發展上，都通過一定的工具檢驗程序，其測量的信度與效度具有一定的水準，測量的敏感度也比生理測量法來得更好（Paas & van Merriënboer, 1994）。目前，由受試者填答量表或問卷的主觀測量策略，仍是當前在教育相關研究中較適合做為認知負荷測量工具的可行方法（Sweller, van Merriënboer & Paas, 1998）。綜合以上討論，本研究測量認知負荷所採用測量策略擬以填答量表的主觀測量策略為主，在數位閱讀的環境中，設計測量認知負荷的線上問卷的方式來測量受試者的認知負荷情形。

三、認知負荷的測量工具

目前利用填答量表主觀測量的方式來測量認知負荷的工具上，Paas（1992）編制認知負荷量表1題，採9點量表主觀衡量法，讓受試者可以利用等級層次來表達自己認知負荷的程度。Paas 與 van Merriënboer（1994）以及 Paas、van Merriënboer 與 Adam（1994）中所編制之認知負荷量表主要有2題，屬於李克特7點量表，分別是（1）我覺得本篇的困難程度；（2）要了解本篇內容，對我而言困難程度。Paas、Tuovinen、Tabbers 與 van Gerven（2003）從西元1988至2002年

期間，歸納共 27 篇測量認知負荷的工具中，其中利用主觀測量的李克特量表有 24 篇，其餘 3 篇的測量分別是產出系統 (production system)、次要任務技術 (secondary task technique) 以及瞳孔變化 (眼動儀) (pupillary responses)。由此可見，主觀衡量方法還是大部分研究者採用來測量受試者認知負荷的方法，在 24 篇的主觀測量方法中，7 點量表有 8 篇，9 點量表有 16 篇。Hwang 與 Chang (2011) 編制之認知負荷量表共有 4 題 7 點量表，分為成「心理負荷」(內在認知負荷)和「心智努力」(外在認知負荷與增益認知負荷)兩個面向，兩個面向各兩題，信度方面總量表的 alpha 係數 0.92，心智努力 0.86，心理負荷 0.85。Hadie 與 Yusoff (2016) 利用 Leppink、Paas、van der Vleuten、van Gog 與 van Merriënboer (2013) 所發展的 10 題認知負荷量表，採 10 點李克特式量表，由 93 位醫學系的學生為樣本來驗證其信度與效度，信度大於 0.70，效度經由 CFA 的分析方法，三個因子的模式得到模式適配的結果。

本研究關注於國小高年級學童在數位閱讀環境中的認知負荷，基於利用自我評估的認知負荷量表能方便且快速進行認知負荷的測量，不過因考量國小學童區辨的能力不似成人的敏感性高，量表不以 7 點以上的李克特量表來設計，而是只採用 5 點的李克特量表，編制 25 題的認知負荷量表，所建立的量表是以數位閱讀環境下所蒐集資料來建立量表的信度以及效度等測驗特徵。

四、數位閱讀認知負荷

在本研究中，主要關心的是數位閱讀的認知負荷，Shapiro 與 Niederhauser (2004) 認為數位閱讀的環境會影響閱讀者的閱讀理解，Wiley 與 Schooler (2001) 即指出影響的因素是數位閱讀材料的類型以及所需的認知負荷，這數位閱讀材料即是由紙筆類型轉變至非線性多視窗的閱讀材料類型。Hill 與 Hannafin (2001) 的研究中發現，非線性多視窗的數位閱讀材料與紙筆類型的閱讀材料相較之下需要更多的認知負荷。一般來說，在數位閱讀的初期情境下，認知負荷在數位閱讀時也許是加重的，在閱讀數位文本的歷程中有較多的困難和阻力，因為閱讀的發展是從學習閱讀 (Learn to Read) 轉變為藉由閱讀來學習 (Read to Learn) 的階段 (Chall, 1996)，對國小階段的學生來說，前段的學習閱讀 (Learn to Read) 階段，閱讀時主要認知是在文字的辨認與解碼上面，因此，數位閱讀環境下多文本及超連結的情境，反而造成閱讀者分散其注意力，進而增加閱讀時的認知負荷；

至於到了由閱讀來學習（Read to Learn）階段，基本上，多元表徵的數位閱讀環境對閱讀理解的影響則不一致，從學習環境的相關理論來說，從全方位的學習環境（Universal Design for Learning）理論中，全方位的學習環境即是具有多元表徵（multiple representations）、多元表達（multiple expressions）以及多元參與（multiple engagements）的環境（Rose & Dolan, 2006），但要注意的是，大量的訊息並不代表學習者就能把所有資訊全盤消化吸收，相反的，也有可能造成認知負荷過重的結果，以至於造成更混亂或是無效的學習效果，因此，教學者需要進一步思考的是「學生的需求在那裡？」要依照學生的個別需求來提供適當的訊息，而且這個訊息對於學習者來說應該既是一種挑戰又是對於學習的一種支持。

至於如何降低數位閱讀的認知負荷呢？Moreno 及 Mayer（2007）提出步調原則，在數位閱讀中，所閱讀的材料能透過依速度和順序傳送資訊單元的控制方式，可以適當平衡工作記憶中的資訊處理容量和速度，並且減輕暫時性的分散注意。此外，也可以依據學習者的專門知識，來調節資訊傳送的速度及資訊量，對先備知識較不足的學習者，除了需要一步步慢慢的進行資訊傳遞外，也可將資訊分割成較小的單元，以利學習者進行學習。另外 Wouters、Tabbers 與 Paas（2007）利用選擇適當資訊呈現方式的控制方法，來平衡資訊量和降低空間分散注意力，並經由適當的視覺提示，引導學習者注意，以消除或降低注意力分散的效應；而在互動式動態的視覺呈現上，則提供追蹤動畫接續改變的工具，以降低使用者的外在認知負荷。

上述的說明可以得知，數位閱讀中需要許多的策略來降低國小學童在數位閱讀中所產生的阻力，而這也是本研究建置測量數位閱讀中國小高年級學童認知負荷量表的主要目的之一。

綜上所述，本研究之目的在於編制一份適用於國小高年級學童，以心智的努力與心智的負荷等二個構面的數位閱讀認知負荷量表，進而針對測量結果進行量表信度與效度的考驗。第一階段的研究目的在於澄清構念定義與建立題項的適切性與可理解性，分析預試樣本之信度與效度資料並且進行數位閱讀認知負荷量表修正。第二階段的研究目的，在於希望對於大樣本資料，建立量表題項測量結果與理論的適配性與穩定性，並探討量表題項間的一致性，其中的效度考驗包括量表因素結構以及穩定性分析，信度考驗則是以內部一致性係數為主。

貳、研究方法

本研究旨在建置國小高年級數位閱讀認知負荷量表，並探討其相關因素，以下依本研究的研究對象、研究工具、資料分析方法等說明如下。

一、研究對象

本研究以國小高年級學童為研究對象，第一階段的研究中，先以 387 位屏東縣、高雄市中的高年級學童為預試樣本，立意抽取 12 個班，其中男生 210 位、女生 177 位，有效問卷 387 份。

第二階段的研究對象中，先針對屏東縣與高雄市的所有國小進行編碼，利用分層比率隨機方式，先隨機抽取各分層的學校，再利用電話與 Email 等二個方式同時進行邀請，其中有 23 所學校同意，再隨機選取 35 個班級進行線上施測，作為信度與效度樣本，經剔除填答不完整的問卷，得有效問卷 728 份，其中男生 343 位，女生 385 位。

二、研究工具

本研究採自編之國小高年級學童數位閱讀認知負荷量表，共計 25 題，題項量尺部份採 5 點李克特量尺計分，由 1 到 5，依序為非常同意、有點同意、普通、有點不同意、非常不同意，編制架構與量表的發展程序說明如下。

(一) 編制架構

本研究所自編之數位閱讀認知負荷量表，乃基於 Sweller (1988) 將認知負荷分為「內在認知負荷」、「外部認知負荷」及「增益型認知負荷」，命題內涵參考 Reid 與 Nygren (1988) 之時間負荷、心理負荷以及心理壓力負荷等 3 個向度以及 Hart 與 Staveland (1988) 將負荷量表之主要的向度依任務/環境、學習者特性(認知能力、認知風格、先備知識與經驗)、以及環境與學習者特性的交互作用等 3 個向度，再區分為心智的努力以及心理的負荷等 2 個構面，其中心智的努力 12 題，心理的負荷 13 題，總共 25 題(如附錄所示)，採用李克特量尺 5 點的設計方式，得分愈高代表認知負荷愈大，反之則認知負荷愈低。

(二) 量表的發展

量表的發展程序主要分為擬定初稿題目、預試量表內容與計分方式以及預試

實施與修正量表等 3 個部分，依序說明如下。

1. 擬定初稿題目

根據量表編制架構與參考文獻，編擬初稿題目。邀請從事數位閱讀研究之專家學者四名以及資深國小高年級教師二名，針對認知負荷量表的構念與所有題項，進行審閱（分別為適用、修改後適用和不適用）與修改建議。此外，為提高量表題項的可閱讀性，邀請 10 位高年級學生預作試答，並依試答結果做為是否修改與文字修訂的依據。

2. 預試量表內容與記分方式

透過上述專家審閱與初步樣本預試過程，建立本量表內容效度。依據專家意見，修正認知負荷構念的相關定義，以及修改和潤飾心智的努力與心理的負荷等 2 個向度的部分題項。

3. 預試實施與修正量表

實施與篩選預試問卷後，以 PHP、HTML、Flash 等程式語言建置國小高年級學童數位閱讀認知負荷量表，並以 MySQL 為後端收集資料的平台，針對 728 份樣本進行項目分析和因素分析，參照理論和統計分析的結果，檢視樣本於各題項分配的情形，檢驗題項品質、區辨性以及題項成份，修正或刪修部份題目，形成量表正式題本。

三、資料分析

本研究調查之資料回收後，先行檢視問卷填答狀況，針對有效問卷進行資料編碼，再以 SPSS 統計軟體和 AMOS 統計軟體進行相關統計分析。採用之統計方法如下：

（一）進行內容效度指數（Content Validity Index, CVI）來驗證量表內容的專家效度（Polit & Beck, 2006; Polit, Beck & Owen, 2007）。

（二）進行鑑別度考驗與因素分析，以了解各題項的鑑別度、各題項和總分的相關，以及題項成份和因素負荷量，作為刪除或修正預試目的依據。

（三）以驗證性因素分析來驗證量表的結構以及跨樣本的穩定性，考驗其建構效度。

（四）模式整體適合度的評估指標，主要為模式適配度指標及顯著性考驗，並且以下述的適配度指標和標準檢視模式適合度，絕對適配度指標部分包括

Chi-Square 不顯著， $\text{Chi-Square} / \text{df} < 3.000$ (Hair, Black, Babin, Anderson & Tatham, 2010)，SRMR (Standardized Root Mean Square Residual) < 0.080 (Hair et al., 2010)，RMSEA (Root Mean Square Error of Approximation) < 0.080 (Hair et al., 2010)，GFI > 0.900 ，AGFI (Adjusted Goodness-of-Fit Index) > 0.900 ，增益適配度指標包括 IFI (Incremental Fit Index) > 0.900 ，CFI (Comparative Fit Index) > 0.900 (Hair et al., 2010)，精簡適配度指標則是包括 PRATIO (Parsimony Ratio) > 0.500 (Blunch, 2008)，PNFI (Parsimony-adjusted Normed Fit Index) > 0.500 (Blunch, 2008)，CN (Critical N) > 200 (Hoelter, 1983)。

(五) 以 Cronbach's α 來考驗量表測量結果的內部一致性信度。

四、因素分析

因素分析可分為探索性 (EFA) 以及驗證性 (CFA) 二種，量表編制的應用中，EFA 主要在於了解蒐集資料的潛在意義，旨在探討量表編制之理論或者假設之探尋，至於 CFA 則是在於考驗量表編制理論所提出的假設是否適切，亦即 CFA 是理論驗證的統計方法。量表工具發展初期，EFA 比 CFA 更為適切，CFA 適用於經過嚴謹理論依據的編制過程，驗證階段可利用預試樣本或者是另一組樣本來檢驗 EFA 所發展的模式，來建立量表的複核效度 (cross validation)。本研究在數位閱讀認知負荷量表的建置中，先以 EFA 中的主成份分析法檢探索量表的因素結構，第二階段則是利用 CFA 來進行理論模式與複核效度的驗證。

參、研究結果

本部分旨在說明國小高年級學童數位閱讀認知負荷量表之發展過程及其應用，其中包括題目建立與量表內容、預試量表項目分析、因素分析、信度分析及建構效度。茲分別說明如下。

一、量表內容

本量表是建置於線上平台，而線上平台乃採用 Like UNIX 的作業系統，網站伺服器是以 Apache 為主，其中包括 PHP 的編譯程式以及 MySQL 資料庫作為受試者填答者的記錄資料庫，系統介面是以 HTML、PHP、Flash 等程式語言混合設計

完成。

數位閱讀認知負荷量表總共有 25 題，主要編制的向度依任務 / 環境、學習者特性以及環境與學習者特性的交互作用等 3 方面，再區分為心智的努力以及心理的負荷等 2 個向度，心智的努力 12 題，心理的負荷 13 題，採李克特量表 5 點設計，分別是非常同意、有點同意、普通、有點不同意以及非常不同意等，請四名專家及二位國小高年級教師審閱預試題目，針對量表題目的構念、內容和文辭等提出修改的建議，作為編制預試量表的參考。6 位專家的意見中並未有不適用的題目，故保留所有的題目，2 個分量表的 CVI，I-CVI 其最小值為 0.83，最大值為 1.00，平均數為 0.97，校正後的 Kappa 值其最小值為 0.82，最大值為 1.00，所有題目之 I-CVI 皆大於學者建議 I-CVI 在 0.78 以上，而 S-CVI/UA 為 0.80，亦大於學者建議值以上 (Polit & Beck, 2006; Polit, Beck & Owen, 2007)。本量表經過逐題審查與修改，確保題項構念內涵的完整性和適切性。

二、量表信度與效度分析

(一) 量表項目分析、因素分析

第一階段，本研究針對 387 份預試問卷，以項目分析和探索性因素分析進行考驗，據此修正和編制國小高年級學童數位閱讀認知負荷量表的正式量表。

1. 鑑別度和同質性考驗

依預試樣本資料進行項目分析，針對各題項進行鑑別度和同質性考驗。鑑別度考驗在檢驗預試題目是否能有效地鑑別出特質或行為傾向上個別差異之程度，將各分量尺總得分之前後 27%，分為高分組與低分組，計算出高低分兩組受試者於各單題項之平均數是否有差異，若決斷值 (Critical Ratio, CR) 達顯著水準 ($p < 0.05$)，即表該題項有鑑別力，反之，該題項則可以考慮刪除。同質性考驗，乃考驗各題項與分量尺得分之相關，以及各題項之因素負荷水準 (當把本量表因素設定為一個主成份時)，若相關係數或因素負荷量低於 0.30 時，予以刪除。在預試樣本的項目分析裡，修改與刪題的判斷原則，乃依據高低分組差異檢定之 p 值未達 0.05 顯著水準的題項，與各構面總分相關小於 0.30 的題目，或刪題後 Cronbach's alpha 明顯提高之題項，分析結果，25 個題目的 CR 值 (4.971~20.048) 均達顯著 ($p < 0.05$)，刪題後的 Cronbach's alpha 無明顯提高之題項 (0.936~0.942)，題目與總分之相關皆大於 0.30 (0.299~0.742)，因此全數

保留，25 題 2 個構面的數位閱讀認知負荷量表繼續進行探索性的因素分析。

2. 探索性因素分析

研究者將利用探索性因素分析中的主成份分析法分別檢驗二個構面的題項，依各題項的抽取主成份後之共同性估計值，作為刪減題項的標準。若共同性估計值 < 0.40 ，則刪除之，本量表進行主成份分析法後，心智的努力構面的共同性估計值皆大於 0.40 ($0.414\sim 0.847$)，心理的負荷構面的共同性估計值亦皆大於 0.40 ($0.583\sim 0.724$)，因此經過探索性因素分析主成份分析後，所有的題項亦皆保留。

3. 內容效度

本研究所建置之數位閱讀認知負荷量表基於認知負荷相關理論，以聚焦於了解國小高年級學童數位閱讀後之認知負荷，整理文獻後形成認知負荷的理論架構，在量表的題項上，參考 Hart 與 Staveland (1988)、Reid 與 Nygren (1988) 等人實徵的研究編制而成，之後邀請從事數位閱讀研究之四位專家學者審查，作為修刪題的參考，以及二名資深國小高年級教師，針對認知負荷量表的構念與所有題項，並且為提高量表題項的可閱讀性，邀請 10 位高年級學生預作試答，並依試答結果做為是否修改文字修訂的依據，經由上述程序，確認觀察指標與構念定義的一致性，因此本量表具有良好的內容效度。

(二) 正式量表效度分析

第二階段，本研究檢驗國小高年級學童數位閱讀認知負荷量表之建構效度、跨樣本效度與信度分析，用以檢驗測量結果與理論架構的適配度、理論架構類化到不同樣本群的適配度與穩定性、量表題項間的一致性，分析結果如下所述。

1. 數位閱讀認知負荷量表內因素間相關

本研究正式樣本 ($N=728$ ，女生 $=343$ ，男生 $=385$) 之數位閱讀認知負荷量表之平均數、標準差與相關係數摘要表整理如表 1。

表 1 數位閱讀認知負荷量表各分量表之平均數、標準差與相關係數摘要表

相關係數	心智的努力	心理的負荷
心智的努力		0.523 / 女
心理的負荷	0.584 / 男	

(續下頁)

表 1 數位閱讀認知負荷量表各分量表之平均數、標準差與相關係數摘要表（續）

相關係數	心智的努力	心理的負荷
平均數 (Mean)		
男	9.93	10.00
女	9.74	9.66
標準差 (SD)		
男	3.738	4.048
女	3.276	3.545

註：相關係數細格右上角為女生，左下角為男生相關係數

由上述表 1 中可以得知，男生在 2 個向度間的相關為 0.584，女生則為 0.523，全體則是為 0.560，男生在心智的努力與心理的負荷等 2 個向度的平均數分別為 9.93 與 10.00，女生則為 9.74 與 9.66，至於標準差，男生在 2 個向度為 3.738 與 4.048，女生則為 3.276 與 3.545。

2. 建構效度與跨樣本效度

以驗證性因素分析檢驗國小高年級學童數位閱讀認知負荷量表的建構效度與跨群體效度。首先以 SPSS 20 檢驗本研究 25 個觀察變項的偏態和峰度，偏態介於 0.001 到 0.656 之間，而峰度介於 -0.923 至 0.022 之間，符合 Kline (2011) 所提出的單變量常態分配之檢驗標準，意即偏態的絕對值小於 3.0，而峰度的絕對值指小於 10.0，顯示適合採取最大似法，繼續利用 AMOS 20 軟體，以最大似法進行參數估計和檢驗。在結構方程模式中，多群組分析中利用同一母群中不同樣本群的比較，以檢驗同一模式在不同樣本裡是否仍然成立，測量工具是否能類化到不同的樣本群裡，符合本研究目的，故採取多群組分析的方式檢驗模式之穩定性。

本研究將受試者隨機分為兩群，為測試樣本 (N=370) 和效度樣本 (N=358)。首先進行模式適配前，先檢驗樣本資料的基本適配度，意即測試樣本和效度樣本的參數值和誤差變異都合理且顯著 ($p < 0.05$)，誤差變異皆為正值。第二步驟，以結構方程模式，對測試樣本進行二構面理論模式的驗證，檢驗模式的適配性。第三步驟，以結構方程模式，對測試和效度樣本進行寬鬆模式與嚴格模式之跨樣本效度檢驗，檢驗模式的穩定性，結果呈現如下。

由上表 2 可知，測試樣本和效度樣本於模式驗證之各指標都呈現良好適配結果的範圍。測試樣本與效度樣本的卡方值分別為 22.558 ($p=0.257>0.050$) 和 28.140 ($p=0.081>0.050$)，未達顯著水準，表示資料與模式適配，Chi-Square / df 分別為 1.187 和 1.4815，小於 3.000 為佳 (Hair et al., 2010)。RMSEA 為 0.023 和 0.037，而 PCLOSE 近似 RMSEA 檢定考驗機率，其值為 0.949 與 0.687，皆未達顯著水準，接受模式適配的虛無假設 (Hair et al., 2010)。SRMR 為 0.023 和 0.024，達配適程度，低於 0.080 為佳 (Hair et al., 2010)。AGFI 為 0.974 和 0.960，大於 0.900 為佳 (Hair et al., 2010)。CN 為 526 與 355，皆大於 200 (Hoelter, 1983)，綜合上述的適配度指標顯示本模式有良好的絕對適合度。在增益和精簡適配度方面 IFI 值為 0.998 與 0.992，大於 0.900 為佳，CFI 值為 0.998 和 0.992，皆大於 0.90 的適配規準，達適配程度，而 PRATIO 為 0.679，大於 0.500 為佳 (Blunch, 2008)，PNFI 為 0.668 與 0.664，大於 0.500 為佳 (Blunch, 2008)，顯示本模式具有良好的增益適配度以及精簡適配度。

此外，比較測試樣本和效度樣本於模式驗證的參數值，進行複核效度來釐清因素結構穩定性，採寬鬆複製模式 (loose replication)，隨機將測試樣本的迴歸權值設定為驗證樣本的參數數值，列述於表 2。兩模式的離差卡方值為 1.436 ($df=6$)， p 值為 0.965，大於 0.050，未達顯著水準，而測試樣本與驗證樣本的估計數值相似且 CFI 值之差為 0.006，小於 0.100，顯示出國小高年級學童數位閱讀認知負荷量表模式具有模式複核效度。施俊名、吳裕益 (2008) 提出複核效度之測量恆等性檢定，即檢視該量表從某一個樣本上所得到的參數值是否可以類推到其他的樣本上，若可以的話，即可宣稱該量表具有一定程度的模型穩定性 (model stability)，因此由上述模式的檢定結果兩模式的離差卡方值未達顯著水準，即表示國小高年級學童數位閱讀認知負荷量表具有模型的穩定性。

綜合上述分析結果顯示，國小高年級學童數位閱讀認知負荷量表的測驗結果與二構面的理論架構具良好的適配程度，且此理論架構對兩隨機分派的樣本群都有不錯的適配程度。

表 2 國小高年級學童數位閱讀認知負荷量表整體模式適配度結果摘要表

評鑑指標	適配規準	測試樣本 (N=370)	效度樣本 (N=358)	評鑑結果		
絕對適配度指標						
Chi-Square	$p > 0.050$	21.165 ($p=0.328$)	30.317 ($p=0.050$)	適配		
Chi-Square / df	< 3.000	1.114	1.596	適配		
SRMR	< 0.080	0.023	0.030	適配		
RMSEA	< 0.080	0.018	0.041	適配		
GFI	> 0.900	0.986	0.979	適配		
AGFI	> 0.900	0.974	0.960	適配		
增益適配度指標						
IFI	> 0.900	0.998	0.992	適配		
CFI	> 0.900	0.998	0.992	適配		
精簡適配度指標						
PRATIO	> 0.500	0.679	0.679	適配		
PNFI	> 0.500	0.668	0.664	適配		
CN	> 200	526	355	適配		
寬鬆複製模式比較						
	Chi-Square	Chi-Square/df	SRMR	RMSEA	CN	Δ Chi-Square
模式 A	51.482 (df=38)	1.355	0.023	0.022	754	
模式 B	52.918 (df=44)	1.203	0.022	0.017	831	1.436 (df=6, $p=0.965 > 0.050$)

註 1：模式 A= 測試樣本與驗證樣本的迴歸值允許不同

註 2：模式 B= 測試樣本與驗證樣本的迴歸值設定相同

在模式內在適合度部分，整理於表 3。測試樣本的標準化迴歸權值平方界於 0.167 與 0.697，大於 0.20 為佳，只有 P4 題項的標準化迴歸權值的平方 (0.167) 並未大於 0.200，但已是非常接近，其餘皆達到適配的範圍，參數估計值之 t 考驗皆達 0.050 的顯著水準。潛在變項的成份信度值為 0.806 與 0.866，心智的努力與心理的負荷二個構面的成份信度介於良好的範圍。潛在平均變異數抽取 (AVE)，意指潛在構念因素對題項變異的解釋百分比，為 0.524 與 0.619，心智的努力與心理的負荷等二個構念之潛在平均變異抽取大於其構念相關係數的平方，代表區辨效度佳。

效度樣本的標準化迴歸權值平方界於 0.211 與 0.702，大於 0.20 為佳，所有題

項皆達到適配的範圍，參數估計值之 t 考驗皆達 0.050 的顯著水準。潛在變項的成分信度值為 0.829 與 0.868，心智的努力與心理的負荷二個構面的成分信度介於良好的範圍。潛在平均變異數抽取為 0.559 與 0.623，心智的努力與心理的負荷等二個構念之潛在平均變異數抽取大於其構念相關係數的平方，代表區辨效度可被接受。

表 3 因素負荷量、信度係數、測量誤差、CR 與 AVE 摘要一覽表 (測試樣本 N=370 / 效度樣本 N=358)

測量指標	因素負荷量	信度係數	測量誤差	組合信度 CR	平均變異數抽取值 AVE
P1	0.751/0.812	0.564/0.659	0.436/0.341		
P2	0.835/0.838	0.697/0.702	0.303/0.298		
P3	0.816/0.815	0.666/0.664	0.334/0.336		
P4	0.409/0.459	0.167/0.211	0.833/0.786		
				0.806/0.829	0.524/0.559
Q1	0.822/0.790	0.676/0.624	0.324/0.376		
Q2	0.815/0.838	0.664/0.702	0.336/0.298		
Q3	0.834/0.834	0.696/0.696	0.304/0.304		
Q4	0.665/0.685	0.442/0.469	0.558/0.531		
				0.866/0.868	0.619/0.623

註：測試樣本 / 效度樣本

(三) 正式量表信度分析

本研究以 728 份樣本進行國小學童數位閱讀認知負荷量表的信度分析 (女生 =343, 男生 =385)，女生之內部一致性 α 值如下：心智的努力 0.866，心理的負荷 0.889，整體 0.922；男生之內部一致性 α 值如下：心智的努力 0.904，心理的負荷 0.931，整體 0.950。全部人數 728 份樣本下的內部一致性 α 值如下：心智的努力 0.939，心理的負荷 0.888，整體 0.915，顯示本量表的信度達到可接受的程度。

肆、討論與建議

根據上述研究發現與先前文獻對照，並進行討論，作為提出建議的依據。

一、國小高年級數位閱讀認知負荷量表之發展與量表信度與效度

本量表以 Hart 與 Staveland (1988)、Reid 與 Nygren (1988) 針對負荷量表中環境、學習者特性、以及環境與學習者特性的交互作用等 3 個向度，再區分為心理的努力與心理的負荷等 2 個構面，並且以線上施測環境收集資料，編制國小學童數位閱讀認知負荷量表，以符應線上施測環境及數位閱讀環境下收集認知負荷資料的需求。

經過項目分析與探索式因素分析後，正式量表包括心智的努力與心理的負荷二個構面，共計 25 題。本量表內部一致性信度考驗，心智的努力構面 0.939，心理的負荷構面 0.888，整體 0.915，顯示本量表的信度達到可接受的程度，與先前的測量工具相較之下具有較佳的內部一致性 (Hwang & Change, 2011; Hadie & Yusoff, 2016; Leppink et al., 2013)。此外，在效度考驗方面，本研究採用結構方程模式，檢驗與建立本測量工具的建構效度，乃為先前測量工具甚少建立的效度資料 (Hadie & Yusoff, 2016)。本研究將受試者隨機分為測試樣本 (N=370) 和效度樣本 (N=358)，分別進行驗證性因素分析，結果顯示認知負荷二個構念模式具適配性以及跨樣本的穩定性 (施俊名、吳裕益, 2008)，與 Hadie 和 Yusoff (2016) 的研究相較，雖然該研究也是採用 CFA，但除了本研究所收集的樣本人數符合 SEM 所建議的樣本人數外 (Schumacker & Lomax, 2016)，模式的適配程度也有較佳的情形。

此外，模式內在適合度部分，本模式潛在變項的成份信度皆為可接受的範圍。構念間相關為正值，符合理論上的期待，具理則效度。心智的努力與心理的負荷的幅合效度及區辨效度屬於可接受的程度，此與 Gerjets、Scheiter 與 Cierniak (2009) 的研究結果一致。此外本研究符應現在數位閱讀蓬勃發展的趨勢，行動載具與數位閱讀環境已趨成熟，國內國小學童數位閱讀的行為日趨頻繁 (林巧敏, 2011)，因此發展國小學童數位閱讀認知負荷量表有其必要性，並且可以提供數位閱讀相關研究收集認知負荷的資料。

二、研究限制與建議

綜合本研究成果，研究所發展的國小高年級數位閱讀認知負荷量表具有良好的構念效度及信度，為一有效、可靠的測量工具，可用以檢測國小高年級學童數

位閱讀的認知負荷狀況，未來在使用此一量表時，宜以此理論構念做為分析的基礎。其次，量表的發展並非單一研究便能克盡其功，而是持續累積證據的過程，對於量表的實際使用經驗與實證資料，均能有效提供量表的評鑑、改進及再發展（Marsh, Ellis, Parada, Richards & Heubeck, 2005）。因此，未來在國小高年級學童數位閱讀認知負荷的測量議題上，本研究嘗試提出以下建議：

（一）提供量表的重測信度或效標關聯效度等資訊

關於量表的重測信度及效標關聯效度等指標之提供，是量表發展相當重要的資訊。由於本研究屬試探性、前導性質的研究，僅針對南部地區高雄市與屏東縣所蒐集的樣本進行分析，未來，將可進一步擴大樣本收集的範圍及時間，比較該量表之重測信度等指標，以了解量表分析結果的穩定性與一致性。此外，尚可與量表中的其他心理構念，如數位閱讀理解能力，做進一步探討分析，以了解國小學童數位閱讀認知負荷狀況是否會影響其數位閱讀理解能力的發展。

（二）考量其他脈絡因素對量表建構之影響

其他相關脈絡因素，也可能對本量表的建構有某種程度的影響，例如國小學童的性別、社經地位、閱讀習慣、閱讀行為等都可能影響認知負荷的判讀結果。因此，不同的群體在此一模式上是否也具有相同的結果，亦有待未來進一步加以分析探討。

（三）運用長期縱貫性調查資料特性，進行跨年度比較

在本研究中，研究者企圖以另一組樣本，進行測量恆等性檢驗，希冀使此一模型能更具有穩定性與預測性，然而這只能說明某一時間點下的母群適用於此一模式，無法說明其他時期的群體也同樣適用。事實上，影響國小學童認知負荷的因素繁多，且會隨時間、社會環境變化而改變，因此，後續的研究可以嘗試比較不同年度下國小高年級學童身心狀況的變化趨勢，以提升整個模式的理論價值與應用層面。

謝誌

本研究感謝二位匿名評審提供專業的修改意見以及行政院科技部經費補助，研究計畫編號：MOST 103-2410-H-153-001-MY2、NSC 103-2420-H-153-001-MY2。

參考文獻

- 左台益、呂鳳琳、曾世綺、吳慧敏、陳明璋、譚寧君（2011）。以分段方式降低任務複雜度對專家與生手閱讀幾何證明的影響。*教育心理學報*，**43**，291-314。
- 林巧敏（2011）。臺灣國小學童數位閱讀興趣與行為之調查分析。*國家圖書館館刊*，**100**（2），30-59。
- 施俊名、吳裕益（2008）。「大學生身心健康量表」概念效度驗證之究。*教育研究與發展期刊*，**4**（4），201-230。
- 涂金堂（2012）。應用認知負荷理論的數學解題教學實驗。*屏東教育大學學報－教育類*，**38**，227-256。
- Blunch, N. J. (2008). *Introduction to structural equation modelling using SPSS and AMOS*. London, England: Sage.
- Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychology*, *38*(1), 53-61.
- Chall, J.S. (1996). *Stages of reading development* (2nd Ed.). Fort Worth, TX: Harcourt Brace.
- Clark, R.C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: evidence-based guidelines to manage cognitive load*. San Francisco, CA: Pfeiffer.
- DeLeeuw, K. E., & Mayer, R. E.(2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*(1), 223-234.
- Gerjets, P., Scheiter, K., & Cierniak, G. (2009). The scientific value of cognitive load theory: A research agenda based on the structuralist view on theories. *Educational Psychology Review*, *22*, 43-54.
- Hadie, S. N. H., & Yusoff, M. S. B., (2016).Assessing the validity of the cognitive load scale in a problem-based learning setting. *Journal of Taibah University Medical Sciences*, *11*(3), 194-202.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2010). *Multivariate data analysis (7th ed.)*. Upper Saddle River, NJ: Prentice Hall.

- Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam, Netherlands: Elsevier.
- Hill, J., & Hannafin, M. (2001). Teaching and learning in digital environments: The resurgence of resource-based learning. *Educational Technology Research and Development, 49*(3), 37-52.
- Hoelter, J. W. (1983). The analysis of covariance structures goodness-of-fit indices. *Sociological Methods & Research, 11*(3), 325-344.
- Hwang, G.-J., & Chang, H.-F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*(4), 1023-1031.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling (2nd ed.)*. New York, NY: Guilford.
- Leppink, J., Paas, F., van der Vleuten, C.P.M., van Gog, T., & van Merriënboer, J.G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods, 45*(4), 1058-1072.
- Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the self-description questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analysis. *Psychological Assessment, 17*(1), 81-102.
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309-326.
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429-434.
- Paas, F., Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71.
- Paas, F., van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load

- in instructional research. *Perceptual Motor and Skills*, 79, 419-430.
- Paas, F. G. W. C., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-Load Approach. *Journal of Educational Psychology*, 86(1), 122-133.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489-497.
- Polit, D.F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459-467.
- Reid, G., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in psychology*, 52, 185-218.
- Rose, D. H., & Dolan, R. P. (2006). Implications of universal design for learning for classroom assessment. In D. H. Rose & A. Meyer (Ed.), *A practical reader in universal design for learning* (pp.73-83). Cambridge, MA: Harvard Education Press.
- Schumacker, R. E., & Lomax, R. G. (2016). *A beginner's guide to structural equation modeling*. New York, NY: Routledge.
- Shapiro, A. M., & Niederhauser, D. S. (2004). Learning from hypertext: Research issues and findings. In D. H. Johassen (Ed.), *Handbook of Research on Educational Communications and Technology* (2nd ed.)(pp. 605-620). Mahwah, NJ: Lawrence Erlbaum Associates.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257-285.
- Sweller, J., van Merriënboer, J. J.G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10,251-296.
- Wiley, J., & Schooler, J. W. (2001). The mental web pedagogical and cognitive implications of the net. In C. R. Wolfe (Ed.), *Learning and Teaching on the World Wide Web* (pp. 243-257). San Diego, CA: Academic.
- Wouters, P., Tabbers, H. K., & Paas, F. (2007). Interactivity in video-based models. *Educational Psychology Review*, 19(3), 327-342.

附錄

國小高年級學童數位閱讀認知負荷量表

心智的努力 (mental efforts)

- 01 我覺得我花了很大的心力，才能利用這種方式進行閱讀學習
- 02 我覺得我花了很大的心力，才能理解這種閱讀活動的內容
- 03 在閱讀之前，我認為「數位閱讀」在閱讀上是容易的
- 04 在閱讀的過程中，我實際上用了很少的心力
- 05 看完作答說明後，我覺得「數位閱讀」是簡單的
- 06 在「數位閱讀」的閱讀過程中，找到相關訊息是容易的
- 07 在「數位閱讀」的閱讀過程中，我有足夠的時間思考
- 08 在「數位閱讀」的閱讀過程中，我覺得很有成就感
- 09 在「數位閱讀」的閱讀過程中，讓我非常認真地在閱讀
- 10 在「數位閱讀」的閱讀過程中，我非常投入
- 11 我覺得在「數位閱讀」的閱讀過程中，必須同時處理很多的訊息
- 12 我自己投入很多努力來準備「數位閱讀」

心理的負荷 (mental load)

- 01 在這樣的情況下利用這種閱讀呈現方式讓我無法專心學習
- 02 在這樣的情況下利用這種閱讀呈現方式造成我很大的壓力
- 03 在「數位閱讀」的閱讀過程中，我覺得是順暢的
- 04 在「數位閱讀」的閱讀過程中，我覺得沒有壓力
- 05 進行「數位閱讀」後，我對閱讀的內容理解有十足的把握
- 06 進行「數位閱讀」後，我更加喜歡閱讀的活動
- 07 進行「數位閱讀」，花費了我很多時間為了要完成閱讀活動
- 08 透過「數位閱讀」來進行閱讀，使我獲得鼓勵，讓我有信心繼續閱讀
- 09 透過「數位閱讀」來進行閱讀，我覺得很簡單，不再害怕閱讀
- 10 透過「數位閱讀」來進行閱讀，迅速增長了我的知識
- 11 我很有把握可以運用「數位閱讀」的方式來進行以後的閱讀活動
- 12 閱讀的內容利用「數位閱讀」的方式來呈現，對我來說是非常困難的
- 13 我沒有足夠的時間來進行「數位閱讀」