教育科學研究期刊　第六十六卷第一期
2021年，66（1），107-139
https://doi.org/10.6209/JORIES.202103_66(1).0004

從算式或文字閱讀數學科普文章：共變數分析與階層線性模式的比較

吳昭容
國立臺灣師範大學
教育心理與輔導學系
學習科學跨國頂尖研究中心

鄭鈐華
國立臺灣師範大學
教育心理與輔導學系

張凌嘉
國立臺灣師範大學
教育心理與輔導學系

摘要

理解數學推導是困難的，而文獻上對文本中數學推導段落採用算式或文字表徵何者較具優勢的結論不一。本研究以299位來自四所高中職的高一學生為對象，比較科普文章以算式或文字表達數學推導對理解的影響，以及對高、低能力學生是否有不同的效果。由於高中生能力是巢套於學校，因此以共變數分析與階層線性模式來比較結果。研究材料為三篇幾何的科普文本及其理解測驗，算式與文字兩個版本的內容只在關鍵句（分別僅1、5、5句）用了不同表徵，其餘內容、附圖及測驗都相同。受試者被隨機分派到算式版或文字版，閱讀後的理解測驗包含是非、計算及說明題。此外，受試者也接受閱讀推理篩選測驗與數學先備知識測驗。共變數分析結果顯示能力和版本不具交互作用，且高能力學生顯著優於低能力學生，算式版多數優於文字版。而階層線性模式分析在排除學校平均數學能力的顯著影響後，凸顯了版本的效果，算式版低能力學生的閱讀表現與文字版高能力學生無顯著差異。本研究從算式和文字的語言結構討論算式優勢，也透過文本主題領域、認知負荷、評量重點及測驗題型，和受試者特性與文獻進行比較。

關鍵詞：外在表徵、語言、數學科普、閱讀理解

通訊作者：張凌嘉，E-mail: alcc@ntnu.edu.tw
收稿日期：2019/05/03；修正日期：2020/01/13；接受日期：2020/03/02。

壹、緒論

數學閱讀是數學素養（literacy）的重要成分。能理解不同的數學表徵，對人們達成溝通互動、社會適應而言相當重要（左台益、李健恆，2018；Niss, 2003）。數學閱讀不單只針對數學教科書或作業，運用數學典故或圖畫書（以下統稱為數學科普文章）被認為具有數學教育的功能。人們認識數學的歷史與數學在生活上的意義，一方面可以跨越對數學概念的認知障礙、促進理解，另一方面能對數學學習產生正向態度（Jankvist, 2009; Whitin & Whitin, 2004），而實徵研究也支持這樣的主張（Lim & Chapman, 2015; Young-Loveridge, 2004）。要使科普文章中涉及數學推導的段落容易理解和利於轉換應用，必須選擇適當的表徵方式。文章中的數學推導是用算式還是文字來傳達，讀者可能會有不同程度的理解，而且此種表徵效果或許會因讀者能力的不同而有所差異。

多媒體研究有許多探討圖示相對於文字之表徵效果的文獻（Mayer, 2014），但比較算式和文字表徵的研究卻有限（Dee-Lucas & Larkin, 1988, 1991; Kolloffel et al., 2009; Leung et al., 1997; Mayer & Jackson, 2005; Österholm, 2006; Watkins, 1979）。雖然算式（equation，或方程式formula）或數學符號（symbol）被視為語文的一種型態，不像圖像（picture或diagram）與文字（word或文章text）的差異那麼顯著。然而，數學文本包含多樣的語符系統（semiotic systems），算式與數學符號能精簡且精確地表達數量關係，相當不同於一般文本，但對許多人而言極具挑戰性（Adams & Lowery, 2007; Schleppegrell, 2007）。數學科普文章的作者或大眾傳播的媒體工作者，在文章中應該選擇算式還是一般文字來表達數學推理，才能促進讀者的理解、有效地傳達概念呢？對不同能力的讀者，不同表徵的可讀性是否不同（陳昭珍等人，2020）呢？這是本研究的第一個目的。回答這個問題不只有助於數學閱讀的理解發展，也利於科普傳播與終身學習的實務。第二個目的是統計分析的議題，由於本研究的受試者來自不同高中職，資料具巢套性質，採取共變數分析（Analysis of Covariance, ANCOVA）或階層線性模式（Hierarchical Linear Modeling, HLM）所得的研究結論是否不同？比較傳統與晚近統計方法之結論的差異，可促進學界對資料分析方法的交流。

一、數學表徵系統

構成數學文本的基本單位，除了數字（例如5、2/5、0.25）、符號（例如＋、－、×、÷、√‾、＝、≥、≈）之外，也有文字組成的術語（technical word，例如函數、平行四邊形），以及與日常用語意義不同的準術語（subtechnical word，例如任意、相似、二次、根）（吳昭容等人，2018）。數學文本必須遵循數學特有的語法結構來組合上述各類的詞（O’Halloran, 1998; Schleppegrell, 2007）。例如，3＋（-4）＝7雖是個錯誤的算式，但符合算式的語法結構；而3＋（÷4）＝7則不合語法。基於語言學的分級，O’Halloran（1998）指出數學文本可以分成詞、

式（expression，例如$-16x^3$、$80x$、$-16x^3 + 80x$）、子句（clause，例如 $f(x) = -16x^3 + 80x$）、子句合成體（clause complex，例如 $f(x) = -16x^3 + 80x$、$\therefore f(x) = -16x(x^2 - 5)$）四個層級。

在詞的層次上，算式中的符號可以與文字相互替換，例如，運動學公式中的「初速」常以V_0表示、「加速度」以a表示、「重力加速度」以g代表。通常符號所占的篇幅較文字來得小，而這是一般人覺得數學精簡的理由之一。然而，精簡的符號在解碼上未必容易，例如，許多國中生對「數列與級數」單元的符號感到困擾：S_n、S_{15}、a_1、a_n、d、⋯⋯。另一個數學精簡性的來源，來自數學在語法層次有各種壓縮的方式（O’Halloran, 1998, 2008），例如，約定俗成的縮寫（$f(x)$代表以x為自變項的函數）、指數（x^3代表$x \times x \times x$）、省略（例如$80x$代表$80 \times x$、$3 + 5 \times 7$代表$3 + (5 \times 7)$）。O’Halloran（1998）以系統功能語言（systematic function language）觀點分析數學文本指出，這樣的壓縮經常產生級轉移（rank shift）的現象，也就是上一層級的語言單位以下一層級的型態表現。例如，$x \times x \times x$是由參與者（participant）x和過程（process）× 所組成的子句，而x^3降階成為式。當一個子句被名物化成為一個式之後，就可和另一個過程詞再組合成更複雜的子句，例如$-16 \times x^3$；類似地，$-16 \times x^3$這個子句名物化成$-16x^3$這個式之後，可以再加上別的式。$-16x^3 + 80x$這子句被展開後，是多達五個過程的子句合成體〔$(-16 \times x \times x \times x) + (80 \times x)$〕，也就是型態的精簡卻有概念複雜的代價。科學與數學的學習經常得面對這類語言的壓縮與展開（陳世文等人，2018；陳世文、楊文金，2006），而要展開被壓縮的算式需要熟練的解碼（decoding）能力，例如，在對學生的錯誤類型分析中常發現，學生錯把x^3當作$x \times 3$，或把$80x$誤作$80 + x$。

前述數學語言在語法上的壓縮與展開所帶來的理解困難，不只存在於算式表徵，文字表徵的數學也有同樣的現象。以「求過原點且與圓$x^2 + y^2 = 64$相切之所有圓的圓心軌跡方程式」為例，「求」是過程詞，而後續很長的文字是一個詞組，且這個詞組是由子句合成體級轉移而成的。再例如，本研究用了「兩地陽光斜射夾角θ_1與θ_2的差」，也是從子句「A地陽光斜射夾角 θ_1減 B地陽光斜射夾角θ_2」級轉移而成的名詞片語。因此，要比較算式與文句何者較易理解，並不是單從表徵是否用了算式或純粹用文字就能判斷的，至少還需從語言結構的層次探究表徵對理解的影響。

二、算式與文字表徵的閱讀理解

在比較算式與文字表徵效果的數學閱讀研究中，有些文獻指出閱讀含有數學符號的文本，其理解比閱讀單純的文字來得差（Dee-Lucas & Larkin, 1991; Mayer & Jackson, 2005; Österholm, 2006），但有些文獻則是認為沒有差異（Kolloffel et al., 2009; Watkins, 1979），或者認為是依閱讀文本的特性或讀者的能力而定（Leung et al., 1997; Watkins, 1979）。為了釐清這些發現得到表徵效果不一致的可能原因，本節透過回顧與比較文獻之操弄的語言結構層次、受試者特性、文本主題、測驗題型、統計方法，以及理論主張來指出可能的差異來源，並摘要於表1。

表 1

算式與文字表徵之相關文獻與本研究的特性對照摘要

結論	文獻	操弄層次	受試者	文本主題	測驗題型	統計方法
文 ＞ 算	Dee-Lucas & Larkin（1991）	子句	大學生	流體力學	直接應用、遷移應用、證明題	邏輯斯迴歸
	Mayer & Jackson（2005）	子句	大學生	海浪形成	問答題	t檢定、Cohen's d
無符號 ＞ 有符號	Österholm（2006）	詞	高中生與大學生	群論	問答題	Spearman's r、Mann-Whitney's U
文 ＝ 算	Kolloffel et al.（2009）	篇章	高中生	排列組合	概念題、程序題、分析與分類題	ANOVA
低能力：無符號 ＝ 有符號 高能力：無符號 ＜ 有符號	Watkins（1979）	詞	大學生	笛卡兒積&……	計算應用	ANOVA
低能力：文 ＞ 算 高能力：文 ＝ 算 冗長的文 ＜ 符號	Leung et al.（1997）	詞（Exp. 1-3）、篇章（Exp. 4）	九年級生	單利&……	計算應用	ANOVA
	本研究	子句	高中生	幾何	是非題、計算題、問答題	MANCOVA、HLM

Dee-Lucas與Larkin（1988, 1991）發現，生手大學生認為以算式呈現的內容比以文字呈現的來得重要，但閱讀文字呈現之科學文章的理解與解題表現優於閱讀算式版的讀者。Dee-Lucas與Larkin（1988）使用兩篇流體力學的文章，長度大約50句，其中各有9和11句的關鍵句有兩種表徵方式。同一篇文章的其中一版是奇數關鍵句為文字表徵、偶數關鍵句採算式表徵，另一個版本則相反。受試者被隨機分派到兩種版本，對文章的每一句進行重要性評定（但研究者只分析關鍵句）。結果生手大學生認為用算式表示的語句比用文字表達的來得重要，但專家對關鍵句的重要性評定不受表徵的影響。隨後，Dee-Lucas與Larkin（1991）以不熟悉文章知識的大學生、兩篇流體力學文章（長度各為21和25句）進行算式與文字對理解表現影響的探討。約有一半的句子在文字版全以文字描述概念，而算式版則採用代數符號（例如Δp、p_1、p_e、p_l等）、算式（例如$p_1 = p_e + p_l$），以及代數符號意義的界定來表示，例如以下是其中一關鍵句的文字版與算式版：

> The new pressure is equal to the sum of the original external pressure, the pressure due to the liquid above the point, and the increase in the external pressure (over the original external).

> The new pressure p_2 is equal to
> $p_2 = \Delta p + p_e + p_l$
> Where Δp is the increase in the external pressure (over the original external pressure p_e)

受試者被分派的文章是算式、文字各一篇。文章理解表現以直接應用、遷移應用以及證明題來檢測，每一類題目都有定性（例如只需掌握當變項x變大，變項y變小）與定量（求特定的數值）的試題。採用邏輯斯迴歸的結果顯示，不論定性或定量的問題，閱讀文字版的理解都比算式版好。Dee-Lucas與Larkin（1991）認為生手大學生缺乏領域知識，導致無法分辨實質內容的重要性，只能仰賴表層特徵，結果影響其注意力分配。算式版讓這些生手大學生花比較多的時間在算式上，但在未完全掌握算式意義下就放棄深究，同時花在其他文字部分的理解時間不足，整體的理解表現就比文字版差。由於該研究使用到的算式難度不高，因此Dee-Lucas與Larkin排除了大學生受試者讀不懂算式的可能解釋。

Mayer與Jackson（2005）的主張和Dee-Lucas與Larkin（1991）類似。Mayer與Jackson認為當學習目標是質性地理解科學系統的運作時，文本若增加數量資料，會將學習者引到與學習目的無關的量化細節上，而妨礙質性理解。該研究以大一學生為對象、海浪形成之因果的說明文為材料。受試者被隨機分派到兩種版本，即質性說明版與量化延伸版（增加了符號、等式、計算的內容）。六個開放式題組全部都不需要用到量化的資料，例如，其中一個題組是「什麼原因造成海浪？為何海浪會襲向海岸？為何在海岸會捲起浪花？」，結果發現，閱讀質性說

明版的受試者表現顯著優於量化延伸版。Mayer與Jackson以連貫性原則（coherence principle）描述這個現象——當量化資料與質性理解的學習目標不一致，會妨礙學習表現。不過，該研究的目的並不在比較同一概念由算式或文字表徵的優劣，因此並不能說明有算式或數學符號的文本就比較不利於理解。

Österholm（2006）也發現，數學新知以文字說明會比有數學符號的版本來得容易理解。然而，不同於Dee-Lucas與Larkin（1991）的版本差異在子句的層次，Österholm的版本差異只在詞的層次上。該研究以自然組的高三學生與主修工程和教育的大學生為對象，被隨機分派到「群論」文本的兩版之一，一版單純使用文字（無符號版），另一版則使用了數學符號（有符號版），如下：

> If two objects belong to the set, then the combination of these also belongs to the set. One says that the system is closed.

> If a and b belong to G, then a~b also belongs to G. One says that the system is closed.

另外，所有受試者還閱讀一篇歷史文章作為對照。除此之外，受試者都被測量了數學與歷史的先備知識，以及採開放性試題測量對文本的閱讀理解，再透過分類答案的層次來計分。結果顯示，閱讀無符號與有符號的兩群受試者的數學先備知識相當，但符號版學生的數學閱讀理解測驗的表現顯著地較無符號版差，且大學生與高中生閱讀符號版的理解狀況一樣差；而閱讀文字寫成之數學文本或歷史文本，則是大學生顯著優於高中生。此外，文字表徵組的受試者其數學和歷史的閱讀理解有顯著相關，但符號表徵組的受試者則其數學和歷史的閱讀理解無相關。Österholm依據前述結果主張，閱讀以符號表示的數學文本需要不同的閱讀技巧，且此種技巧的增長和一般文字閱讀能力的增長並不同步。

以上文字表徵優於算式表徵的結論，並未被完全複製。Watkins（1979）操弄了文本有無數學符號卻未顯示文字優勢，和Österholm（2006）的結論不同。Watkins除了操弄符號的有無外，還操弄了「語文結構」變項：口語式英文（Ordinal English, OE）與數學式英文（Mathematical English, ME），直交組合成四種版本的文本，包括有符號的ME、無符號的ME、有符號的OE、無符號的OE。數學內容包含笛卡兒積、整數行列式等四個主題，以93名大學生為對象，前測確認了受試者不具備學習內容的知識。三因子的實驗設計包括有無符號、語文結構，以及學生數學修課層級的高低，前兩個因子是受試者內變項，也就是每位受試者拿到的四個主題之文本分屬四種不同的表徵。每研讀了一個文本後，就各有五個題目測驗來反映學習效果。結果高階班優於基礎班、OE優於ME，但有無符號則沒有差異，不過與班別有交互作用：有無符號對基礎班學生沒有差異，不過有符號則利於高階班學生，這點和Österholm的發現不一致。對於此一不一致的發現，Österholm從測驗設計的差異來討論：Österholm的試題以開放式問題

瞭解學生閱讀後的心智模型，而Watkins測量了文本未直接提到的知識，可能會涉及較多背景知識的運用，且其試題測量了計算應用。因此，Watkins的發現可能代表符號有利於程序性的應用，也適合形成某些推論性的概念理解。

Kolloffel等人（2009）比較不同表徵對排列組合的學習效果，其結果顯示文字與算式兩種表徵對學習的影響沒有差異。該研究以123名高中生為對象，這些學生都尚未學過排列組合，在排列組合的學習上屬於低先備知識者。受試者被分派到五個組別，分別學習圖組、文組、算式組、文+算式、圖+算式五種數位化學習材料，其中有差異的部分，圖組是一張樹狀圖，文組是124個字的文字說明，算式組是一條與機率有關的等式。資料蒐集了測驗成績、認知負荷，以及學習時間。Kolloffel等人預期圖的表徵提供概念基模、有利於概念知識的建立；算式表徵有利於程序知識的掌握；文字表徵則有助於形成情境模式、善於分析與分類；不同表徵的組合則可能產生互補（complementary）（Ainsworth, 2014）或者冗餘（redundancy）（Kalyuga & Sweller, 2014）效果，也就是不同表徵可能提供獨特、相互支援的訊息而有正向效果，或者不同表徵提供了重複、增加整合負擔的訊息而有負面效果。就文組、算式組以及文+算式三組之間的表現而言，結果三組的測驗總分與分項（概念題、程序題、分析與分類題）都沒有差異，且整體認知負荷以及學習時間也無顯著差異，僅在內在認知負荷上，文+算式組感受的費力程度顯著比文組低。Kolloffel等人的研究顯示，採用一百多字的文字說明和用一條算式說明的效果（測驗表現、認知負荷及學習時間）類似，他們原先預期算式有助於程序性知識、文字利於建立情境模式的假設未被支持。

Leung等人（1997）以一系列的實驗結果主張，認知負荷是影響文字或算式教材學習效果的關鍵。他們採用計算單利的解題範例（worked example）為材料、九年級學生為對象，算式版、文字版分別如下，而整合版有八行，由文字版前四行加上算式版後四行組合而成：

P = \$500	Principal = \$500
R% = 7%	Interest Rate = 7%
I = \$245	Simple Interest = \$245
I = P × R% × T	Simple Interest = Principal × Interest Rate × Time
245 = 500 × 7% × T	245 = 500 × 7% × Time
245 = 5 × 7 × T	245 = 5 × 7 × Time
T = 245/(5 × 7) = 7	Time = 245/(5 × 7) = 7

實驗一的受試者被隨機分派到三組，並以練習階段的表現區分成高、低兩種能力狀態。結果低能力受試者的表現複製了以往文字版優於算式版的發現（得分較高、錯誤較少），但高能力受試者則三組都沒有差異。實驗二改採文字相較於算式長很多的學習材料，例如：

$n \times r = N \times R$　Number of turns of the front wheel × Radius of the front wheel = Number of turns of the rear wheel × Radius of the rear wheel

受試者經預試確認與實驗一的高能力者相當，但結果卻發現算式優於文字版。實驗三將學習時間（26分鐘和10分鐘）與表徵版本（文字和算式）組合成四組，結果發現在學習時間短的情況下，文字版優於算式版，而在學習時間長的情況下，文字版就不再有優勢。實驗四改編實驗一的材料，使文字版增加更多語句，且合併算式版和文字版兩版內容成為冗長版，結果算式版呈現明顯的優勢，有較高的正確率、較少的錯誤。該研究的測驗屬於計算應用，包括近遷移題和遠遷移題。Leung等人（1997）從認知負荷理論主張，算式或文字表徵孰優孰劣的問題沒有簡單的答案，必須視學生能力、材料以算式或文字表示的特徵，以及學習時間的長短而定。如果算式的意義過於複雜，就只有高能力的學生或者必須有更多的學習時間才能處理；然而，很多時候，算式的精簡性會比冗長的文字更有利於學習。

表1的摘要顯示，多數研究發現文字優勢，受試者多為高中以上的學生，材料的主題都是數學或者物理，測驗題型有些側重理解，有些則重計算，統計方法多半採用變異數分析，惟Dee-Lucas與Larkin（1991）以及Österholm（2006）因為資料不符合變異數分析的前提而改採相關分析。然而，各研究在材料操弄的語言結構層次上則有較大的差異。Österholm（2006）以及Watkins（1979）用來與文字版相比較的版本應該稱為符號版，其以數學符號表示的僅在「詞」和「式」的層次，並不在「子句」的層次，也就是在表達一個數學命題時混用了文字和符號，而沒有純以數學符號表示的數學子句，因此，兩種版本的特性都偏向文字。Leung等人（1997）操弄兩個版本的層次同樣只在「詞」，文字版用文字的全稱、算式版則用字首的字母作為代數符號，然而，兩種版本都偏向算式，因為都用了運算符號、顯現算式的形式。至於Kolloffel等人（2009）的算式版和文字版，雖然沒有前述偏頗其中一種表徵的問題，但卻有另一個缺點：兩個版本的篇幅差距太大（一條算式相對於124個字），未控制資訊量和研讀時間，很容易產生兩種版本傳達的訊息不等價的問題，該研究最後發現文+算式版在五種版本中效果最好，或許也反映了兩種表徵版本的訊息不等價的問題。Dee-Lucas與Larkin（1988, 1991）所設計的算式版是以完整的算式表達一個數學命題，該命題在文字版則以文字表示等價的語意，且兩種版本篇幅的落差小，而其餘非關鍵句都維持一模一樣，最符合本研究想檢驗兩種表徵在涉及數學推導之「子句」的層次效果。因此本研究在操弄層次上依循了Dee-Lucas與Larkin的設計，詳見以下「研究設計與假設」。

三、研究設計與假設

本研究選擇高中一年級學生，以與文獻的結果相互對照。材料從數學科普書籍選擇幾何典故進行改編，三篇文章都需要搭配圖示，圖上標示了文本說明所必要的代數符號。本研究採用Dee-Lucas與Larkin（1988）類似的設計，也就是文字版的內容也包含符號，且文字、算式兩個版本使用完全相同的圖示與測驗試題，僅在涉及數學推導的關鍵句有差異（參見後文「研究工具」有關「文本」的介紹及附錄一）。

本研究在關鍵句的設計上，文字版並非直譯算式，而是力求語句流暢、能聯繫其他語句，且篇幅能有所控制。例如，文本「地球半徑」第一個關鍵句的算式為 $d=\frac{\theta}{360}\times 2\pi r$ ，若採直譯可能是「弧長 d 等於圓心角 θ 除以360，再乘以2倍圓周率再乘以半徑」，一則非常不流暢，再則喪失文字的特色，無法反映數量推理的原意，比方 $\frac{\theta}{360}$ 要表達的是一種比例關係，而不只是除法運算。因此，該句被設計成「弧長可依圓心角占360°的比率乘以圓周長求得」。再如文本「島高」由平行線截角性質得「$\triangle PAC \sim \triangle CMF$，因此 $\frac{d_1}{d_2}=\frac{\overline{PA}}{\overline{CM}}$ 」，文字的表示方式為「三角形PAC與CMF相似，因此 d_2與 d_1的比值，等於 $\overline{PA}$ 與 $\overline{CM}$ 的比值」。

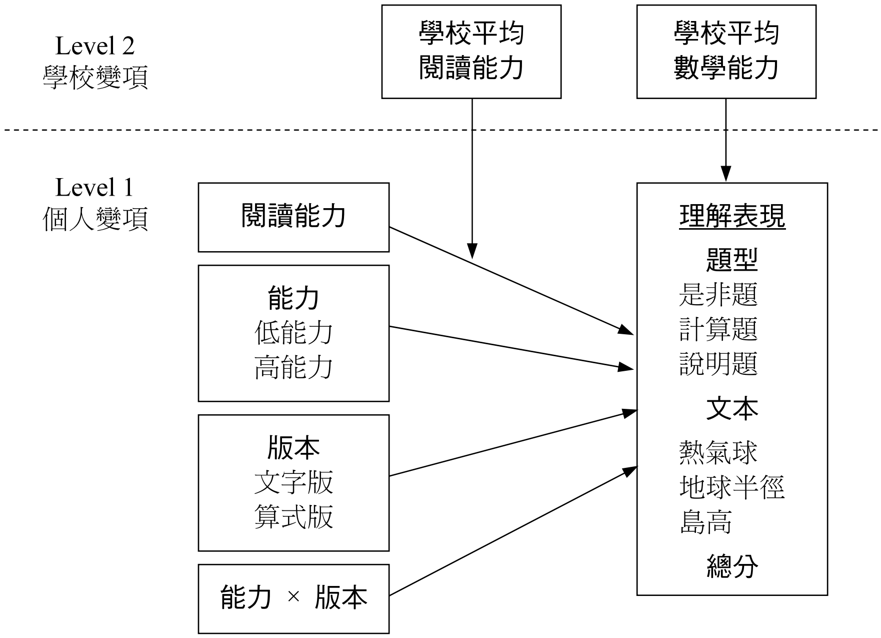
本研究控制兩種版本在非關鍵句與搭配的圖示都一模一樣，而文字版的關鍵句則控制字數以接近算式版的精簡，並試圖精準地傳達數量推理的根據與意涵（例如比例），參考同樣在子句層次上操弄算式版與文字版的文獻（Dee-Lucas & Larkin, 1991; Mayer & Jackson, 2005），預期精簡且精準的文字說明可降低理解語意的認知負荷，使其理解表現優於算式版。另外，依照Leung等人（1997）認知負荷觀點，推測文字版較佳的效果只發生在低能力學生，而高能力學生因具備解讀精簡算式的能力，預期在兩種版本應該都一樣好。

為使資料來自能力分布足夠廣泛且有數量夠多的受測者，本研究邀請了四所入學分數高低不同的高中職合作，算式版與文字版的分派採班級內隨機。以往類似研究處理這樣的資料，採版本與能力的二因子變異數是常見且廣被接受的統計方法。但本研究四所合作高中的學生能力不一，亦即能力變項是巢套於學校，近年來有愈來愈多學者採用HLM處理（Powell et al., 2017; Rau et al., 2017），透過多層次模式（multilevel modeling）將具巢套特性的資料分解為個體層次（如學生）與總體層次（如學校），並於總體層次設定組織單位的迴歸係數與隨機效果，從而解決使用單階的變異數分析常見的問題：將變異來源散計於學生個體，而忽略學校組織脈絡效果的偏誤。又因同一學校的學生之間有較多的相同經驗或共通性（如本研究四所高中的入學分數不同，但校內的學生能力相依程度高），一旦違反變異數同質性與誤差獨立性的假設，將導致低估標準誤、增加型I錯誤率，影響分析結果的解釋與推論（Hox et al., 2017; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012）。因此，本研究同時採用共變數分析與

HLM，瞭解兩種統計方法得出的結論是否有差異；亦即受試者來自不同學校或能力異質的群組時，有無控制學校層次變項對學生個人理解表現的影響。研究與分析架構如圖1所示。

圖1

研究與分析架構



本研究目的包括：

（一）數學科普文章以算式或文字來表達涉及推導數量關係的關鍵句，對高、低能力的高中生何者較佳？

（二）就上一問題分析來自多校的學生資料，共變數分析與HLM是否會得到不同的結論？

貳、研究方法

一、研究參與者

參與者來自臺北、新北、宜蘭、臺中四所高中職，共八個班級、306名高一學生。這四所學校入學登記分發最低錄取分數之PR值為69～89，學生學業程度在中上以上。刪除資料不全者七名，有效樣本為299名。本研究贈以文具作為學生的時間補償。

二、研究工具

（一）閱讀能力測驗

國內無適用於高中生的閱讀能力測驗，因受測者為高一上學期的學生，故採適用國一到國三學生的閱讀推理篩選測驗（柯華葳、詹益綾，2007）。該測驗旨在篩選語文閱讀有困難的學生，內部一致性係數為.81，再測信度係數為.78，效度方面可區辨不同年級與不同閱讀能力者的表現。本測驗施測簡便，需時僅20分鐘，滿分18分。

（二）數學先備知識測驗

此一自編測驗旨在瞭解學生的幾何基本知識，內容是三篇數學科普文章所涉及的「三角形」、「圓」、「平行」等數學知識，例如畢氏定理、內角和定理、外角定理、相似三角形的應用，圓心角、圓周長、弧長的計算，以及平行線截角性質等，滿分12分，內部一致性係數為.69，達可接受。

（三）文本

三篇數學科普文章改寫自科普書籍（曹亮吉，1996，2003），並設計了對應的理解測驗。文本的主題分別為熱氣球升空後的俯瞰距離、計算地球半徑的西洋數學史典故，以及求算離島高度的中國數學史典故，後文以「熱氣球」、「地球半徑」、「島高」稱之。文中描述與推導數量關係的關鍵句，若用算式呈現，為「算式版」；若以文字描述，則為「文字版」。為了讓兩個版本使用完全相同的圖示與理解測驗，文字版的內容也包含代數符號，以圖2熱氣球文本為例，兩版本非關鍵句介紹了各代數符號的意義，而文字版關鍵句見圖2黑框內文字，例如，「『最遠距離 s』的平方是『地球半徑r與熱氣球升空高度 h之和』的平方減去『地球半徑 r』的平方」；該關鍵句在算式版則以「 $s^2=(r+h)^2-r^2$ 」來表示。為確認文本內容與算式的正確性，以及文字表述的流暢性，除了由一位數學教育教授與一位科學教育教授多次修訂，並以41名大學生進行前導的行為實驗，以修改文本內容與理解測驗題目。

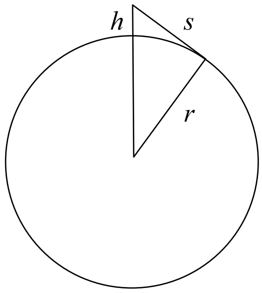
文本涉及的數學概念是國中幾何單元教過的內容，「熱氣球」主要涉及畢氏定理；「地球半徑」則包含角度公式、弧度公式及圓心角與弧長的關係；「島高」則與相似三角形對應邊成比例有關。三個文本的關鍵句分別有1、5、5句，文字版與算式版的對照請見附錄一。其餘非關鍵句的內容在兩版本都相同。每篇文本均附圖示，圖中都標有與內文敘述相對應的代數符號。排版上，每篇文本都以一頁橫式A4紙呈現，文字在左，圖示在右，採黑白印刷。

（四）理解測驗

搭配算式版與文字版的理解測驗是相同的，包括是非題、計算題與說明題三個分測驗，題目旁邊都會附上與文本內容相應的附圖，並標有代數符號（例如 r、h、s）。

圖2

「熱氣球」全文及文字版關鍵句（黑框）與算式版關鍵句（灰底）的對照



目前台東舉辦熱氣球嘉年華。想像隨熱氣球冉冉升空，俯瞰台東美麗的景致，真令人心嚮往之。到底搭乘升空的熱氣球可以俯瞰多遠距離外的地面景色呢？畫個圖來說明這個問題與解決的方法。

假設圓形代表地球，其半徑為 r，熱氣球升空距地面的高度為 h，所能看到的最遠距離為 s，此一視線為地表的切線。根據畢氏定理，

「最遠距離 s」的平方是「地球半徑 r 與熱氣球升空高度 h 之和」的平方減去「地球半徑 r」的平方。

$$s^2=(r+h)^2-r^2$$

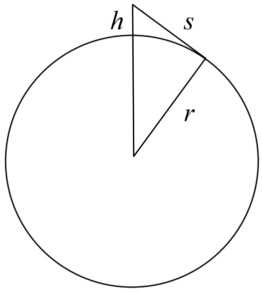
因此只要查書得知地球的半徑，就可以推算出，當熱氣球升空高度為多少時，可以看到多遠距離之外的地面景色。

是非題在於測驗學生能否正確判斷文本中提到的資訊，或根據文本進行推理，附圖使用原文本上的圖。內容有些與算式有關，有些則以文字表達。以「熱氣球」為例，如圖3所示。三篇文本是非題的題數分別為5、10及10題。依標準答案計分，三個文本得分最多為5、10及10分。

圖3

「熱氣球」的五題是非題

是非題：對的寫〇，錯的寫 X



1. (　) 圓的切線與過切線的半徑互相垂直。
2. (　) $s^2=2hr-h^2$
3. (　) 我們可以學到利用地球半徑和升空高度所形成的三角形求可視的最遠距離。
4. (　) 當熱氣球高度升高為原先的 1.5 倍時，可見的最遠距離亦會增為原先的 1.5 倍。
5. (　) 不同身高的人搭乘熱氣球時可看見的最遠距離會有所差異，例如甲比乙高 20 公分，則可見的最遠距離亦會比乙多 20 公分。

計算題就文中重要的定理或關係式，給定部分數字，設計成計算題型，檢驗學生能否運用題目的資料、注意到題目的單位換算，進而正確列式。附圖使用原文本上的圖。每篇文本各有一題計算題，以「熱氣球」為例，如圖4所示。計分依能否列出主要算式、正確代入數值，評為3、2、1、0分（參見附錄二）。

圖4

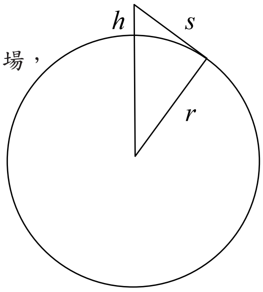
「熱氣球」的計算題

計算題：請盡可能寫出你的**計算過程**

台南舉辦熱氣球活動，已知熱氣球升空高度為 50 公尺時，可以俯瞰看到走馬瀨農場，

如果地球半徑是 r 公里時，

請問走馬瀨農場距離升空的熱氣球至少多遠？



說明題包含問答題或填充題兩類題型，附圖上的符號改以文字取代。「熱氣球」與「地球半徑」分別搭配一題問答題，設計成向家人或同學解釋的情況，請學生試著表達出文本的推導過程，如圖5所示。「島高」則因推導過程相當繁複，因此由四小題的填充題組成一個題組。計分依能否提到推導的關鍵公式而評為3、2、1、0等級（參見附錄三），而「地球半徑」牽涉多個原理（包含角度公式、弧度公式及圓心角與弧長的關係），故增加等級4以表示所有原理均能正確說明。

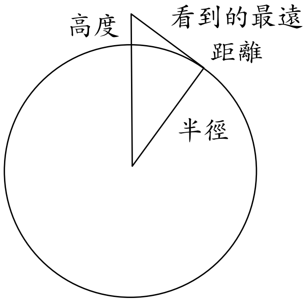
圖5

「熱氣球」的說明題

問答題

想像你要向家人解釋熱氣球高度會如何影響可俯瞰的最遠距離，

請利用**右圖的用語**來解釋說明**推導過程**。



因計算題與說明題的計分具主觀性，故從算式版與文字版各隨機抽取20份試卷（排除作

答有空白者），由兩名評分者建立評分者間信度，包括一致性百分比（percent agreement，總份數為分母，評分一致的份數為分子）（Supovitz et al., 1997）及Kappa係數（Sim & Wright, 2005）。結果三個文本計算題一致性百分比中位數為.85（全距為.80～.85）、Kappa係數中位數為.76（全距為.60～.80，*ps* < .001）；說明題的一致性百分比中位數為.88（全距為.70～1.00）、Kappa係數中位數為.81（全距為.58～.92，*ps* < .001），顯示具有良好的評分者間信度。是非題、計算題與說明題三部分的內部一致性係數，分別為.60、.67、.87，勉強可接受。

考量題型與文本的分析各有意義，分析時分成三個題型或者三個文本暨總分，共七項。另外，考量題數不一，故採正確率的概念轉換分數，也就是每項分數的滿分都是100。分數轉換方式詳見附錄四。

三、研究程序

本研究採兩階段團體施測，主試者是合作班級的數學教師，實施地點在原班教室。第一階段施測閱讀能力測驗與數學先備知識測驗，施測時間分別為20分鐘與10分鐘。第二階段為正式測驗，實施文本閱讀與理解測驗，時間約一節課（50分鐘）。

正式測驗的兩種版本採班內隨機分派，學生並不知道有不同版本。閱讀算式版的學生有149人、閱讀文字版的有150人。施測程序上，首先由主試者說明施測目的，並提醒每篇文本的閱讀與測驗有時間限制。隨後，學生依序閱讀與測驗熱氣球、地球半徑、島高三個文本，由主試者宣布開始與停止的指令。每當一篇文本的閱讀時間結束，主試者會請學生們將該文本的頁面黏合，接著進行該文本的理解測驗，以使學生受測時無法往前翻閱文本內容。閱讀與測驗時間是參考本團隊先前研究的數據，對絕大多數受試者而言是充足的。熱氣球、地球半徑、島高依序的閱讀時間分別是2、4、6分鐘；理解測驗的時間則是8、12、12分鐘。

參、結果與討論

一、能力分組

為區分本研究受測學生的能力高低，以閱讀能力分數與數學先備知識分數為預測變項，對理解表現總分進行線性逐步迴歸分析。結果顯示，閱讀能力與數學先備知識分別和理解表現之間的Pearson相關達顯著，$r(299) = .11$，$p = .03$；$r(299) = .50$，$p < .001$，而閱讀能力與數學先備知識之間的Pearson相關亦達顯著，$r(299) = .21$，$p < .001$。逐步迴歸的結果，閱讀能力分數未進入迴歸方程式，僅數學先備知識可以解釋理解表現總分24.5%的變異量，$F(1, 297) = 97.86$，$p < .001$，β係數達.50（$t = 9.89$, $p < .001$），表示數學先備知識分數愈高，理解表現總分愈高。因此，本研究以數學先備知識之得分將受測學生大略平分成高低能力兩群，10分以下的學生為低能力組（共126人，得分平均數8.58、標準差1.76）、11與12分的學生為高

能力組（共173人，得分平均數11.57、標準差0.50），兩組學生能力具顯著差異，$t(139.62) = 18.53$，$p < .001$。

為了確認算式版與文字版的受測學生能力是否相等，以高低能力（2）×版本（2）的卡方檢定結果為χ^2（1, $n = 299$）= 0.08，$p = .78$，顯示分派至兩版本的學生能力高低無顯著差異。若改以數學先備知識之原始得分進行兩版本學生的t檢定，$t(297) = 0.17$，$p = .86$，同樣顯示兩版本學生的先備知識相近。能力與兩種版本組成四組，數學低能力且閱讀算式版（以下稱低算組）64人、數學低能力且閱讀文字版（以下稱低文組）62人、數學高能力且閱讀算式版（以下稱高算組）85人、數學高能力且閱讀文字版（以下稱高文組）88人。

二、理解表現的共變數分析

理解表現涵蓋三種題型、三個文本及總分，以閱讀能力為共變項，並以高低能力、版本為自變項，對三個題型、三個文本進行MANCOVA，而對總分則進行ANCOVA。組內迴歸同質性考驗結果，七個理解表現僅於熱氣球文本顯著，$F(3, 298) = 3.56$，$p = .02$，表示其他六個理解表現的組內迴歸斜率皆具同質性（p = .31～.86），可以進行共變數分析。變異數同質性考驗結果，題型、文本的Box's M值皆達顯著（p = .04、.01），表示至少有一個理解表現未能符合變異數同質的假設。而在Levene's同質性檢定，計算題、熱氣球、地球半徑、總分等四個表現皆達顯著，$F(3, 295) = 6.23$，$p < .001$；$F(3, 295) = 5.37$，$p = .001$；$F(3, 295) = 4.49$，$p = .004$；$F(3, 295) = 2.92$，$p = .03$，表示變異數不同質。綜合以上，理解表現各細格的變異數或迴歸斜率具有異質性，顯示可能受到學校間變異或其他因素的影響。基於實驗研究與相關文獻慣常採用變異數或共變數分析，為能對照HLM分析控制學校變項後的差異，本研究仍提出共變數分析的結果。

共變項閱讀能力在題型、文本的MANCOVA與總分的ANCOVA都不顯著（p = .33、.28、.08），比較表2、表3細格的原始平均數與調整後平均數可發現，排除閱讀能力高低與否對平均數的影響非常小，可能因為閱讀推理篩選測驗有天花板效應（滿分18分，全體學生平均數為16.18分）所致。進一步檢定能力與版本在題型、文本、總分的交互作用效果，皆未達顯著（p = .53、.88、.81），表示數學能力的效果不會因為版本而不同。

在三種題型的主要效果（參考表2），能力主要效果皆是高能力學生顯著高於低能力學生，是非題$F(1, 294) = 59.83$，$p < .001$，$\eta^2 = .17$；計算題$F(1, 294) = 63.52$，$p < .001$，$\eta^2 = .18$；說明題$F(1, 294) = 38.53$，$p < .001$，$\eta^2 = .12$。版本主要效果在是非題未達顯著（$p = .20$），在計算題、說明題是算式版顯著高於文字版，$F(1, 294) = 9.56$，$p = .002$，$\eta^2 = .03$；$F(1, 294) = 7.63$，$p = .006$，$\eta^2 = .03$。

在三個文本及總分的主要效果（參考表3），能力主要效果皆是高能力學生顯著高於低能力學生，熱氣球$F(1, 294) = 26.49$，$p < .001$，$\eta^2 = .08$；地球半徑$F(1, 294) = 57.74$，$p < .001$，η^2

表 2

不同能力學生讀不同表徵版本於各題型的平均數及調整後平均數

題型	版本	低能力（n = 126）			高能力（n = 173）			平均		
		M	（*SD*）	*Adj. M*	*M*	（*SD*）	*Adj. M*	*M*	（*SD*）	*Adj. M*
是非題	算式版	71.09[a]	（12.19）	71.12	80.04[b]	（10.99）	80.10	76.20	（12.31）	75.61
	文字版	68.12[a]	（12.50）	68.23	79.73[b]	（ 9.72）	79.58	74.93	（12.33）	73.91
	平均	69.63	（12.38）	69.68	79.88	（10.34）	79.84	75.56	（12.32）	74.76
計算題	算式版	43.23[a]	（31.19）	43.28	68.89[b]	（22.28）	69.00	57.87	（29.29）	56.14
	文字版	34.77[a]	（25.73）	34.98	58.59[b]	（26.30）	58.29	48.74	（28.52）	46.64
	平均	39.07	（28.83）	39.13	63.65	（24.88）	63.65	53.29	（29.22）	51.39
說明題	算式版	33.77[a, b]	（28.41）	33.82	54.64[c]	（28.82）	54.76	45.67	（30.37）	44.29
	文字版	25.94[a]	（25.95）	26.18	45.14[b, c]	（25.60）	44.81	37.20	（27.35）	35.50
	平均	29.92	（27.40）	30.00	49.81	（27.56）	49.79	41.43	（29.16）	39.90

註：[a]、[b]、[c]相同者代表同質子集，亦即細格平均數的 95%信賴區間重疊。

表 3

不同能力學生讀不同表徵版本於各文本與總分的平均數及調整後平均數

文本	版本	低能力（$n = 126$）			高能力（$n = 173$）			平均		
		M	（*SD*）	*Adj. M*	*M*	（*SD*）	*Adj. M*	*M*	（*SD*）	*Adj. M*
熱氣球	算式版	$63.13^{a, b}$	（26.47）	63.15	76.21^{c}	（17.94）	76.26	70.59	（22.87）	69.70
	文字版	53.62^{a}	（24.51）	53.72	$68.31^{b, c}$	（22.94）	68.18	62.24	（24.62）	60.95
	平均	58.45	（25.87）	58.43	72.19	（20.96）	72.22	66.40	（24.09）	65.33
地球半徑	算式版	43.11^{a}	（22.88）	43.16	60.43^{b}	（20.22）	60.53	52.99	（23.00）	51.85
	文字版	35.74^{a}	（17.07）	35.95	53.39^{b}	（17.45）	53.10	46.09	（19.32）	44.52
	平均	39.48	（20.49）	39.55	56.85	（19.14）	56.82	49.53	（21.48）	48.19
島高	算式版	41.86^{a}	（25.88）	41.92	66.93^{b}	（26.36）	67.07	56.16	（28.89）	54.49
	文字版	39.47^{a}	（27.23）	39.73	61.76^{b}	（27.39）	61.41	52.55	（29.38）	50.57
	平均	40.68	（26.47）	40.83	64.30	（26.94）	64.24	54.35	（29.14）	52.53
總分	算式版	48.37^{a}	（21.59）	48.42	67.07^{b}	（17.98）	67.17	59.04	（21.63）	57.79
	文字版	41.92^{a}	（18.84）	42.12	60.09^{b}	（17.38）	59.82	52.58	（20.06）	50.97
	平均	45.20	（20.46）	45.27	63.52	（17.97）	63.49	55.80	（21.07）	54.38

註：^{a}、^{b}、^{c}相同者代表同質子集。

= .16；島高$F(1, 294) = 55.99$，$p < .001$，$\eta^2 = .16$；總分$F(1, 294) = 68.67$，$p < .001$，$\eta^2 = .19$。版本主要效果在熱氣球、地球半徑、總分皆為算式版顯著優於文字版，$F(1, 294) = 10.70$，$p = .001$，$\eta^2 = .04$；$F(1, 294) = 10.403$，$p = .001$，$\eta^2 = .03$；$F(1, 294) = 9.64$，$p = .002$，$\eta^2 = .03$，僅島高文本的兩版本沒有差異（$p = .21$）。

以上各理解表現均無能力與版本的交互作用，而主要效果若達顯著，均為數學高能力者優於低能力者，閱讀算式版優於文字版。以各細格平均數的95%信賴區間重疊者視為同質子集，四組在多數表現上是依能力分成兩個子集，也就是在是非題、計算題、地球半徑、島高、總分等五項表現上，都是「低文≒低算<高文≒高算」；說明題和熱氣球則分成三個子集：「低文≒低算<低算≒高文<高文≒高算」。顯示能力的影響力較大，而版本的效果略小。

三、理解表現的HLM分析

HLM採兩階層分析（如圖1），第一階為個人變項，包含三個題型、三個文本及總分等七項理解表現分數，低算、低文、高算和高文等四組，以及閱讀能力；第二階為學校變項，包含學校平均閱讀能力，以及學校平均數學能力（即學校平均數學先備知識分數）。本研究採用HLM 7.03版統計軟體，估計方法為限制最大概似估計法（restricted maximum likelihood），第二階迴歸係數為一般最小平方法（generalized least squares）。連續變數採總平減（grand mean centering），以利於迴歸係數的解釋。雖然本研究的第二階僅四所學校，但第一階學生數多達299位，且分析著重第一階的固定效果，HLM可提供強韌的估計值（McNeish & Stapleton, 2016）。

（一）零模型

以七項理解表現分數為依變項，零模型檢驗顯示組內相關係數（Intraclass Correlation Coefficient, ICC）介於.30～.63之間，表示學生理解表現有30%～63%的變異可被學校層次因素解釋；具高度組內相關，學校的組織效果不可忽視，宜採HLM分析。

（二）組別模型

為能比對共變數分析，區別數學高低能力學生閱讀不同版本的差異，進行組別模型分析。以低算組為基準，第一階加入其他三個組別變項，並納入閱讀能力為共變項，第二階為學校變項，以學校的平均數學能力與平均閱讀能力控制校間差異對學生理解表現的影響。模型如下：

Level 1:

$$ACH_{ij} = \beta_{0j} + \beta_{1j}\text{*低文組} + \beta_{2j}\text{*高文組} + \beta_{3j}\text{*高算組} + \beta_{4j}\text{*閱讀能力} + r_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{*學校平均數學能力} + u_{0j}$$

$\beta_{1j} = \gamma_{10}$
$\beta_{2j} = \gamma_{20}$
$\beta_{3j} = \gamma_{30}$
$\beta_{4j} = \gamma_{40} + \gamma_{41}$ *學校平均閱讀能力

ACH_{ij} 表示第j個學校之第i個學生的理解表現。r_{ij} 為學生 i 之隨機誤差項。γ_{00} 為控制所有其他變項之調節總平均數，即低算組之理解表現調節總平均，γ_{01} 為學校 j 平均數學能力對調節總平均數之影響，u_{0j} 為學校j的隨機效果。γ_{10} ～ γ_{30} 分別代表低文組、高文組、高算組，相對於低算組在理解表現之差異。γ_{40} 與 γ_{41} 為閱讀能力調節總平均數與學校 j 平均閱讀能力對理解表現之影響。

根據表4與表5，平均理解表現、學校平均數學能力皆達顯著（$ps < .05$），顯示控制其他變項後，低算組之理解表現調節總平均數於各題型、文本及總分為48.93～76.32分，當學校 j 平均數學能力較全體數學能力總平均數多1分，該校於各題型、文本及總分之平均理解表現增加7.75～22.52分。γ_{01} 顯示學校平均數學能力對學生平均理解表現的影響，這是在共變數分析中無法獲知的組織脈絡效果。

題型方面（參考表4）：在是非題，四組學生表現相當（$ps > .05$）。在計算題與說明題，相對於低算組，低文組學生的表現顯著減少9.39、8.44分（SE = 3.97、3.83；$ps < .05$）；高文組學生與低算組無顯著差異（$ps > .05$）；高算組學生的表現於計算題顯著增加11.45分（SE = 3.91, $p = .004$），而說明題則與低算組無顯著差異（$p > .05$）。

文本方面（參考表5）：相對於低算組，低文組學生於熱氣球、地球半徑及總分，理解表現顯著減少10.01、7.75、7.01分（SE = 3.67、2.87、2.54；$ps < .01$）；高文組學生與低算組無顯著差異（$ps > .05$）；高算組學生於地球半徑、島高及總分，理解表現顯著增加6.47、8.88、6.23分（SE = 2.83、3.67、2.50；$ps < .05$）。

閱讀能力對理解表現於各題型、文本及總分的影響，皆無顯著差異（$ps > .05$）；僅學校平均閱讀能力於熱氣球文本有負向影響（$\gamma_{41} = -10.03$, $SE = 3.69$, $p = .007$），主因是各校的熱氣球文本表現均偏高（皆有滿分），以致學校 j 的平均閱讀能力低於總平均1分，該校學生i的閱讀能力於熱氣球文本的理解表現仍佳〔(-10.03)*(-1)*閱讀能力$_{ij}$)〕。

隨機效果部分，除計算題與島高文本，其餘題型與文本的組間變異成分皆不顯著，表示各校間的理解表現已無顯著差異。相較於零模型，組別模型的解釋變異量，學校內為3%～9%，學生的個別差異尚待其他變項解釋；學校間達92%～99%，顯示透過學校平均數學能力與平均閱讀能力可有效控制學校間的變異，此即為HLM分層的效果。

綜合以上HLM的分析，在控制學校平均數學能力、學校平均閱讀能力及學生閱讀能力後，各組理解表現的關係為「低文 ＜ 低算≒高文 ＜ 高算」。相較於共變數分析結果，排除學校平均數學能力的顯著影響，使得四組學生表現的差距有所縮減，以計算題為例，高算與低算兩組的平均分數差異從25.72分降至11.45分，而高文與低算組的平均分數差異則從15.01分降

表 4

理解表現於各題型之組別模型分析結果摘要

固定效果	是非題		計算題		說明題	
	係數	標準誤	係數	標準誤	係數	標準誤
平均理解表現（γ_{00}）	76.32***	1.31	55.32**	3.70	48.93**	3.10
學校平均數學能力（γ_{01}）	7.75**	0.75	18.48*	2.87	22.52**	2.00
低算組[a]	0	0	0	0	0	0
低文組（γ_{10}）	-3.10	1.71	-9.39*	3.97	-8.44*	3.83
高文組（γ_{20}）	2.52	1.68	1.01	3.88	-6.10	3.75
高算組（γ_{30}）	3.06	1.69	11.45**	3.91	3.74	3.77
閱讀能力（γ_{40}）	0.20	0.38	0.23	0.89	-0.09	0.85
學校平均閱讀能力（γ_{41}）	-1.26	1.72	-2.86	4.11	-5.18	3.89
隨機效果	變異數	χ^2	變異數	χ^2	變異數	χ^2
理解表現差異（τ_{00}）	0.01	1.03	17.08*	7.82	3.81	3.00
組內變異（σ^2）	92.20		491.64		458.69	
偏差值（Deviance）	2182.37		2673.46		2651.74	

[a] 參數設為零，為比較基準。

*$p < .05$. **$p < .01$. ***$p < .001$.

表 5

理解表現於各文本及總分之組別模型分析結果摘要

	熱氣球		地球半徑		島高		總分	
固定效果	係數	標準誤	係數	標準誤	係數	標準誤	係數	標準誤
平均理解表現（γ_{00}）	72.52**	2.81	52.41**	2.39	56.38**	4.81	59.29**	2.29
學校平均數學能力（γ_{01}）	13.29*	1.61	14.24*	1.62	22.34*	4.48	16.43*	1.72
低算組[a]	0	0	0	0	0	0	0	0
低文組（γ_{10}）	-10.01**	3.67	-7.75**	2.87	-3.28	3.73	-7.01**	2.54
高文組（γ_{20}）	-4.90	3.60	-1.05	2.81	3.41	3.65	-1.01	2.48
高算組（γ_{30}）	2.77	3.62	6.47*	2.83	8.88*	3.67	6.23*	2.50
閱讀能力（γ_{40}）	-1.21	0.81	0.78	0.64	0.44	0.84	0.16	0.57
學校平均閱讀能力（γ_{41}）	-10.03**	3.69	0.59	2.93	-1.94	3.92	-2.88	2.62
隨機效果	變異數	χ^2	變異數	χ^2	變異數	χ^2	變異數	χ^2
理解表現差異（τ_{00}）	0.05	0.85	3.35	3.71	58.96***	21.78	5.60*	6.26
組內變異（σ^2）	423.56		257.82		434.21		201.54	
偏差值（Deviance）	2627.60		2483.87		2639.35		2412.77	

[a] 參數設為零，為比較基準。

*$p < .05$. **$p < .01$. ***$p < .001$.

至無顯著的1.01分，各組關係則從能力與版本沒有交互作用的「低文≒低算<高文≒高算」，轉變成算式版本縮減能力落差的「低文<低算≒高文<高算」。顯示當學生巢套於學校時，宜透過HLM辨別出學校與學生變項的變異對個人表現的影響，有助適當的資料估計、解釋與推論。

肆、結論與建議

首先比較兩種統計方法（研究問題二）的結果來總結研究發現（研究問題一)。依據預測理解表現較佳的數學先備知識，將學生分成高、低兩種能力，連同版本作為自變項，並以閱讀能力作為共變項，進行共變數分析。結果顯示，不論何種題型（是非題、計算題、說明題)、哪個文本（熱氣球、地球半徑、島高)，或者總分，共變項的效果都不顯著、兩個自變項都沒有交互作用，且高能力學生的表現都優於低能力學生，而版本效果除了是非題和島高文本之外，其餘五個表現都是算式版優於文字版。從四組的95%信賴區間可知，多數表現為「低文≒低算<高文≒高算」，僅說明題和熱氣球是「低文≒低算<低算≒高文<高文≒高算」。也就是只以學生層次的解釋變項進行分析時，能力的影響較大、版本的影響稍小，且兩種效果具可加性。由於資料來自入學成績不同的四所高中職，HLM零模型顯示理解表現具高度校內相關，故以二階層的線性模型控制學校變項，結果顯示學校平均數學能力顯著影響學生平均理解表現，排除學校變異後使得學生層次能力的變異減少，凸顯了版本效果，整體得到「低文 < 低算≒高文 < 高算」的結果。從兩種統計方法的比較可知，當受試者來自於能力異質的群組時，群體層次變項可能改變學生個體層次變項的解釋力與交互關係。放在本研究的脈絡下，可以想像同為低能力的學生但在不同平均數學能力的學校內學習，他們的數學學習經驗必然不一樣，教師使用的教學表徵以及作業要求的表徵方式不同，就使版本的表徵在他們身上產生不同的效果。因此，採用HLM或多層次模式是分析巢套或階層性資料較佳的統計方法。此外，三篇文本有表徵差異的句子不多，分別僅1、5、5句，但因內容涉及數學推導，即使兩版本大部分文字內容以及整個附圖都相同，以相同的測驗仍能發現表徵差異在理解上的效果，顯示算式與文字的表徵差異不容忽視。

本研究得到與預期不一致的結論，發現數學科普文章以算式表達數學推導優於文字。以往文獻多數得到文字優勢或者兩者相當的結果，Leung等人（1997）是少數發現算式優勢的研究。但Leung等人在實驗一得到和其他文獻一樣的文字優勢，尤其是低能力受試者。在後續實驗是透過加長文字版的文字長度與冗餘效果而獲得算式優勢，用以主張認知負荷在表徵效果中的角色。嚴格說來，Leung等人是指出，當文字冗長到認知負荷過重時，文字有助於理解的優勢就會消失。基本上，Leung等人比較的是變項用文字表示還是符號表示，亦即在詞的層次的差異。本研究的關鍵差異則在子句的層次，文字版關鍵句雖然較算式長，但編製時已盡量

精簡文字長度，因此本研究的發現並不能僅從文字版篇幅過長造成認知負荷來解釋。

本研究的算式優勢最核心的來源可能是關鍵句的語言特性。本研究為使文字版保有語文的特色，且達到控制文字版關鍵句之長度的目的，避免直譯算式，但此一作法可能造成讀者理解文字版關鍵句的困難。例如，地球半徑文本關鍵句5，對應到算式分母「$\theta_1 - \theta_2$」的文字版是「兩地陽光斜射夾角θ_1與θ_2的差」(見附錄一)，文字版除了語句篇幅較長之外，還把隱含在θ_1、θ_2這兩個符號背後的意義「A地陽光斜射夾角」、「B地陽光斜射夾角」的部分訊息呈現在語句中，同時，用「……的差」來名詞化「……減……」的子句，然而，讀者轉譯「θ_1與θ_2的差」到算式，可能需要從「$\theta_1 - \theta_2$」或「$\theta_2 - \theta_1$」中做一判斷。本研究原本預期攜帶了文本相關脈絡的文字版關鍵句，會比算式版精簡的符號來得容易形成整體的心智模型，也以為級轉移減少的文字量可降低長篇大論造成的負面效果。然而，從結果來事後解釋，文字版將相關脈絡壓縮在名詞化的語句，理解的難度可能過高。相對地，算式中的符號暫時脫離了脈絡，其精簡的符號便於數學推導，例如「$d = \dfrac{\theta}{360} \times 2\pi r$。而$360° = 2\pi$，所以$d = \theta \times r$……求得$\theta = \theta_1 - \theta_2$，因此可知$r = \dfrac{d}{\theta_1 - \theta_2}$」，亦即算式是在擱置符號背後的意義之下進行推導，待完成推導後再置回脈絡中解讀其意義；相對地，文字版是在攜帶著符號背後意義的狀態下進行數學推導，此一繁重的認知負荷，可能造成理解的困難。

另外，本研究發現的算式優勢也可從文本主題所屬領域、評量重點與測驗題型，以及受試者特性來與其他文獻相互比較。幾何是數學學門中探討圖形與空間的一支，閱讀幾何文本非常仰賴圖文整合(Epelboim & Suppes, 2001; Lee & Wu, 2018)。而多媒體研究指出，整合來自空間分離、有時間落差之不同來源的訊息(例如圖區與文區)會比單一來源的困難，稱作分散注意力效果(split attention effect)(Ayres & Sweller, 2014)，而此一現象的發現最早就是來自幾何教材(Tarmizi & Sweller, 1988)。文獻上其他研究的訊息在時空上是單一來源，而在本研究的幾何文本中，算式在文字的背景下相對明顯，或許比全為文字的版本容易在文區和圖區搜尋到互相關聯的訊息，而有助於圖文整合。其次，數學推導是本研究評量閱讀理解的重點，計算題當然側重解題，說明題的作業要求強調「向家人解釋……**推導過程**」，甚至是非題也有題目是和公式或數量推算有關，這樣的評量重點與測驗題型可能有利於算式版。而文獻上有些研究(Mayer & Jackson, 2005; Österholm, 2006)關注的是讀者形成的心智模型，也就是重視質性的理解甚於量化的推導。第三，臺灣學生認為數學就是計算或解決數學教科書上的題目的比例很高(李浩然、柳賢，2012)，數學教師教學時，亦鮮少強調透過閱讀學習數學(蘇慧珍等人，2017)。因此，本研究對象擅長閱讀算式，不擅長理解文字對數學概念的描述，可能也是造成算式優勢的理由之一。

本研究至少有下列幾個限制：第一，參與的學校數少，且為避免資料回收時空白率過高，並未邀請入學成績中下的高中職，使得本研究所討論的低能力學生實際上並非全體高中職學

生群體的低能力者。第二，學生層次的變項只有以數學先備知識能力高低與文本版本所組成的組別和閱讀能力，尚待納入其他可能影響學生理解表現的變項。第三，受試者對數學內容與算式表徵的熟悉程度可能是表徵效果重要的調節變項。數學是高一學生重要的學習科目，他們相較於其他年齡層的讀者更為熟稔數學語言與算式表徵，本研究的結論要推論到其他群體的受試者，尚須進一步探討。第四，可得知測量閱讀能力的工具過於簡單，無法有效區分學生的閱讀理解能力，致使閱讀能力無法預測數學科普文章閱讀理解的結論，必須再多加驗證，或發展適用之工具。

從上述討論可知，影響表徵效果的因素很多，至少包括特定算式與文字的語言結構、文本內容的主題領域、評量側重的面向與測驗題型，以及受試者的特性。綜合前兩段的討論與文獻回顧，以下提出對教育與未來研究的建議：選擇表徵與語言的組織方式應考量文章的目的；假如希望讀者能掌握數量變化的數學推導，採用算式會較文字說明來得好，尤其文章必須整合文字內容與圖或表格等空間上分離的資訊時，算式有利於不連續資訊的整合；但若希望讀者形成整體現象的心智模型、能質性地掌握文章的意義，文字可能比算式更適當；但是，合併文字說明與算式表徵似乎未能理想地讓讀者既形成整體的心智模型，又掌握數量推導；算式會吸引讀者大多數注意力，或者耗費大量認知資源在數量推導，進而削弱了心智模型的建構（Dee-Lucas & Larkin, 1991; Leung et al., 1997; Mayer & Jackson, 2005）。未來在算式與文字表徵之議題上，還需要更多的研究，包括擴大閱讀材料到各學科領域、對不同年齡階段與能力的學生、給不同閱讀目的的讀者，以及設計不同語言結構層次之表徵、操弄表徵之字數或句數，讓文字與算式互補（Kolloffel et al., 2009），方能得出更精準的結論。

誌謝

本研究承蒙教育部補助國立臺灣師範大學學習科學跨國頂尖研究中心之高等教育深耕計畫的經費支持，以及國科會（NSC 102-2511-S-003-020-MY3）、科技部（MOST 108-2511-H-003-014-MY3）的經費補助，並感謝國立臺灣師範大學數學系洪萬生教授、科學教育研究所楊文金教授在建構閱讀文本上的協助，以及國立中央大學學習與教學研究所趙子揚助理教授對資料分析的建議，特此致謝。

參考文獻

一、中文文獻

左台益、李健恆（2018）。素養導向之數學教材設計與發展。**教育科學研究期刊，63**（4），29-58。https://doi.org/10.6209/JORIES.201812_63(4).0002

【Tso, T.-Y., & Lei, K.-H. (2018). Design and development of mathematical literacy-oriented subject materials. *Journal of Research in Education Science, 63*(4), 29-58. https://doi.org/10.6209/JORIES.201812_63(4).0002】

吳昭容、曾建銘、鄭鈴華、陳柏熹、吳宜玲（2018）。領域特定詞彙知識的測量：三至八年級學生數學詞彙能力。**教育研究與發展期刊，14**（4），1-40。https://doi.org/10.3966/181665042018121404001

【Wu, C.-J., Cheng, C.-M., Cheng, C.-H., Chen, P.-H., & Wu, Y.-L. (2018). The measurement of domain-specific vocabulary knowledge: The mathematical vocabulary ability of third to eighth grade students. *Journal of Educational Research and Development, 14*(4), 1-40. https://doi.org/10.3966/181665042018121404001】

李浩然、柳賢（2012）。國三學生數學觀念之研究。**科學教育學刊，20**（3），267-294。https://doi.org/10.6173/CJSE.2012.2003.03

【Lee, H.-J., & Leou, S. (2012). Ninth grade students' conceptions of mathematics. *Chinese Journal of Science Education, 20*(3), 267-294. https://doi.org/10.6173/CJSE.2012.2003.03】

柯華葳、詹益綾（2007）。國民中學閱讀推理篩選測驗編製報告。**測驗學刊，54**（2），429-449。https://doi.org/10.7108/PT.200712.0429

【Ko, H.-W., & Chan, Y.-L. (2007). Reading comprehension screening test for junior high school students. *Psychological Testing, 54*(2), 429-449. https://doi.org/10.7108/PT.200712.0429】

曹亮吉（1996）。**阿草的葫蘆：文化活動中的數學**。遠哲科學教育基金會。

【Cao, L.-J. (1996). *A-Cao's gourd: Mathematics in cultural activities*. Yuan T. Lee Foundation Science Education for All.】

曹亮吉（2003）。**阿草的數學聖杯：探尋無所不在的胚騰**。天下遠見。

【Cao, L.-J. (2003). *A-Cao's holy grail of mathematics: Exploring the omnipresent patterns*. CommonWealth.】

陳世文、古志雄、楊文金（2018）。從系統功能語言觀點探討科學詞彙的歧義與解歧。**科學教育學刊，26**（3），241-259。https://doi.org/10.6173/CJSE.201809_26(3).0003

【Chen, S.-W., Ku, C.-H., & Yang, W.-G. (2018). Exploring lexical ambiguity and disambiguation of science terminologies from the lens of systemic functional linguistics. *Chinese Journal of Science Education, 26*(3), 241-259. https://doi.org/10.6173/CJSE.201809_26(3).0003】

陳世文、楊文金（2006）。以系統功能語言學探討學生對不同科學文本的閱讀理解。**師大學報：科學教育類，51**（1，2），107-124。https://doi.org/10.6300/JNTNU.2006.51.05

【Chen, S.-W., & Yang W.-G. (2006). The impact of a systemic functional linguistics-based science text and a conventional science text on students' reading comprehension. *Journal of Taiwan Normal University: Mathematics & Science Education, 51*(1, 2), 107-124. https://doi.org/10.6300/JNTNU.2006.51.05】

陳昭珍、宋曜廷、章瓊方、曾厚強（2020）。配合國小課程單元科普讀物人工分級推薦與系

統可讀性分析之差異研究。**圖書資訊學刊，18**（1），45-67。https://doi.org/10.6182/jlis.202006_18(1).045

【Chen, C.-C., Sung, Y.-T., Chang, C.-F., & Tseng, H.-C. (2020). Examining the differences of readability leveling of Chinese popular science books by experts and by CRIE system for elementary school children. *Journal of Library and Information Studies, 18*(1), 45-67. https://doi.org/10.6182/jlis.202006_18(1).045】

蘇慧珍、楊凱琳、陳佳陽（2017）。閱讀策略教學對高二學生數學學習表現的影響。**教育科學研究期刊，62**（1），33-58。https://doi.org/10.6209/JORIES.2017.62(1).02

【Su, H.-C., Yang, K.-L., & Chen, C.-Y. (2017). Effects of teaching reading strategies on senior high school student's mathematics performance. *Journal of Research in Education Sciences, 62*(1), 33-58. https://doi.org/10.6209/JORIES.2017.62(1).02】

二、外文文獻

Adams, T. L., & Lowery, R. M. (2007). An analysis of children's strategies for reading mathematics. *Reading & Writing Quarterly, 23*(2), 161-177. https://doi.org/10.1080/10573560601158479

Ainsworth, S. (2014). The multiple representation principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 464-486). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.024

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 206-226). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.011

Dee-Lucas, D., & Larkin, J. H. (1988). Novice rules for assessing importance in scientific texts. *Journal of Memory and Language, 27*(3), 288-308. https://doi.org/10.1016/0749-596X(88)90056-3

Dee-Lucas, D., & Larkin, J. H. (1991). Equations in scientific proofs: Effects on comprehension. *American Educational Research Journal, 28*(3), 661-682. https://doi.org/10.3102/00028312028003661

Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research, 41*(12), 1561-1574. https://doi.org/10.1016/S0042-6989(00)00256-X

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge. https://doi.org/10.4324/9781315650982

Jankvist, U. T. (2009). A categorization of the "whys" and "hows" of using history in mathematics education. *Educational Studies in Mathematics, 71*(3), 235-261. https://doi.org/10.1007/s10649-008-9174-9

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 247-262). Cambridge

University Press. https://doi.org/10.1017/CBO9781139547369.013

Kolloffel, B., Eysink, T. H. S., de Jong, T., & Wilhelm, P. (2009). The effects of representational format on learning combinatorics from an interactive computer simulation. *Instructional Science, 37*(6), 503-517. https://doi.org/10.1007/s11251-008-9056-7

Lee, W.-K., & Wu, C.-J. (2018). Eye movements in integrating geometric text and figure: Scanpaths and given-new effects. *International Journal of Science and Mathematics Education, 16*, 699-714. https://doi.org/10.1007/s10763-016-9790-2

Leung, M., Low, R., & Sweller, J. (1997). Learning from equations or words. *Instructional Science, 25*(1), 37-70. https://doi.org/10.1023/A:1002969618881

Lim, S. Y., & Chapman, E. (2015). Effects of using history as a tool to teach mathematics on students' attitudes, anxiety, motivation and achievement in grade 11 classrooms. *Educational Studies in Mathematics, 90*(2), 189-212. https://doi.org/10.1007/s10649-015-9620-4

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369

Mayer, R. E., & Jackson, J. (2005). The case for coherence in scientific explanations: Quantitative details can hurt qualitative understanding. *Journal of Experimental Psychology: Applied, 11*(1), 13-18. https://doi.org/10.1037/1076-898X.11.1.13

McNeish, D., & Stapleton, L. M. (2016). Modeling clustered data with very few clusters. *Multivariate Behavioral Research, 51*(4), 495-518. https://doi.org/10.1080/00273171.2016.1167008

Niss, M. A. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In A. Gagatsis & S. Papastavridis (Eds.), *3rd mediterranean conference on mathematical education - Athens, Hellas 3-4-5 January 2003* (pp. 116-124). Hellenic Mathematical Society.

O'Halloran, K. L. (1998). Classroom discourse in mathematics: A multisemiotic analysis. *Linguistics and Education, 10*(3), 359-388. https://doi.org/10.1016/S0898-5898(99)00013-3

O'Halloran, K. L. (2008). *Mathematical discourse: Language, symbolism and visual images*. A & C Black.

Österholm, M. (2006). Characterizing reading comprehension of mathematical texts. *Educational Studies in Mathematics, 63*(3), 325-346. https://doi.org/10.1007/s10649-005-9016-y

Powell, S. R., Driver, M. K., Roberts, G., & Fall, A. M. (2017). An analysis of the mathematics vocabulary knowledge of third- and fifth-grade students: Connections to general vocabulary and mathematics computation. *Learning and Individual Differences, 57*, 22-32. https://doi.org/10.1016/j.lindif.2017.05.011

Rau, M. A., Aleven, V., & Rummel, N. (2017). Supporting students in making sense of connections and in becoming perceptually fluent in making connections among multiple graphical representations. *Journal of Educational Psychology, 109*(3), 355-373. https://doi.org/10.1037/edu0000145

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.

Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading & Writing Quarterly, 23*(2), 139-159. https://doi.org/10.1080/10573560601158461

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy, 85*(3), 257-268. https://doi.org/10.1093/ptj/85.3.257

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage.

Supovitz, J. A., MacGowan III, A., & Slattery, J. (1997). Assessing agreement: An examination of the interrater reliability of portfolio assessment in Rochester, New York. *Educational Assessment, 4*(3), 237-259. https://doi.org/10.1207/s15326977ea0403_4

Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*(4), 424-436. https://doi.org/10.1037/0022-0663.80.4.424

Watkins, A. E. (1979). The symbols and grammatical structures of mathematical English and the reading comprehension of college students. *Journal for Research in Mathematics Education, 10*(3), 216-218. https://doi.org/10.2307/748810

Whitin, P., & Whitin, D. (2004). *New visions for linking literature and mathematics*. National Council of Teachers of Mathematics.

Young-Loveridge, J. M. (2004). Effects on early numeracy of a program using number books and games. *Early Childhood Research Quarterly, 19*(1), 82-98. https://doi.org/10.1016/j.ecresq.2004.01.001

附錄一 文字版與算式版關鍵句的對照

文本1：「熱氣球」

1. 「最遠距離 s」的平方是「地球半徑 r 與熱氣球升空高度 h 之和」的平方減去「地球半徑 r」的平方。

   $s^2 = (r+h)^2 - r^2$

文本2：「地球半徑」

1. 弧長可依圓心角佔360°的比率乘以圓周長求得。

   $$d = \frac{\theta}{360} \times 2\pi r$$

2. 角度360°以弧度表示就是2π。

   $360° = 2\pi$

3. 弧長等於所對應的圓心角弧度乘以半徑。

   $d = \theta \times r$

4. θ 為 θ_1 與 θ_2 的差。

   $\theta = \theta_1 - \theta_2$

5. 地球半徑就是「兩地距離除以這兩地陽光斜射夾角 θ_1 與 θ_2 的差」。

   $$r = \frac{d}{\theta_1 - \theta_2}$$

文本3：「島高」

1. $\overline{MF}$ 可當作是 $\overline{DF}$ 與 $\overline{DM}$ 之差。

   $\overline{MF} = \overline{DF} - \overline{DM}$

2. 三角形 PAC 與 CMF 相似，因此 d_2 與 d_1 的比值，等於 $\overline{PA}$ 與 $\overline{CM}$ 的比值。

   $\triangle PAC \sim \triangle CMF$，因此 $\dfrac{d_2}{d_1} = \dfrac{\overline{PA}}{\overline{CM}}$

3. 三角形 PRA 與 CDM 相似，因此 $\overline{PR}$ 與 $\overline{CD}$ 的比值，等於 $\overline{PA}$ 與 $\overline{CM}$ 的比值。

   $\triangle PRA \sim \triangle CDM$，因此 $\dfrac{\overline{PR}}{\overline{CD}} = \dfrac{\overline{PA}}{\overline{CM}}$

4. $\overline{PR}$ 就等於「$\overline{CD}$ 乘以 $\overline{PA}$ 與 $\overline{CM}$ 的比值」也等於「竿高乘以 d_2 與 d_1 的比值」。

$$\overline{PR} = \overline{CD} \times \frac{\overline{PA}}{\overline{CM}} = 竿高 \times \frac{d_2}{d_1}$$

5. 島高等於「竿高」加上「竿高乘以 d_2 與 d_1 的比值」。

$$島高 = 竿高 + 竿高 \times \frac{d_2}{d_1}$$

附錄二　計算題的評分

分數	計算題
3	能列出正確的算式，並正確地代入數據
2	能列出正確的算式，但錯誤地代入數據
1	算式不完整或錯誤
0	沒有計算過程，或僅出現無意義的文字

附錄三　說明題的評分

分數	說明題
4	能提出全部的關鍵公式，且說明正確（僅「地球半徑」有此等級）
3	能提出一個關鍵公式，且說明正確
2	能提出一個關鍵公式，但說明不完整或數據錯誤
1	提出之公式不完整或錯誤，且說明不完整或數據錯誤
0	沒有說明或推導過程，或僅出現無意義的文字

附錄四　理解表現七個分數的計算方式

理解表現		計算方式
題型	是非題	個別計算三個文本是非題的正確率，再平均後乘上100
	計算題	將各題的等級除以滿分轉換成正確率，再平均後乘上100
	說明題	1. 熱氣球和地球半徑文本：先將各題的等級除以滿分轉換成正確率 2. 島高文本：先分別計算四個小題正確率後再平均 3. 最後再將三個文本的正確率平均後乘上100
文本	熱氣球	個別計算是非題、計算題、說明題每個題型的正確率，再平均後乘上100
	地球半徑	同上
	島高文本	因說明題為四小題的題組，以各小題的正確率再平均作為說明題的正確率。其餘同上
	總分	1. 不分文本 2. 是非題25題，滿分為25分；計算題三題，滿分為9分；說明題三題，滿分為10分 3. 分別計算是非題、計算題、說明題每個題型的正確率，平均後乘上100

Journal of Research in Education Sciences
2021, 66(1), 107-139
https://doi.org/10.6209/JORIES.202103_66(1).0004

Reading Popular Mathematics from Equations or Words: Comparison of Analysis of Covariance and Hierarchical Linear Modeling

Chao-Jung Wu
Department of Educational Psychology and Counseling,
Institute for Research Excellence in Learning Sciences,
National Taiwan Normal University

Chien-Hua Cheng
Department of Educational Psychology and Counseling,
National Taiwan Normal University

Ling-Chia Chang
Department of Educational Psychology and Counseling,
National Taiwan Normal University

Abstract

Understanding mathematical reasoning is challenging. No conclusive evidence exists on which external representation is more beneficial to comprehension: equation or words. This study involved 299 high school students and examined the effects of external representations and participants' abilities on reading comprehension of popular mathematics. Because students were nested within schools, analysis of covariance (ANCOVA) and hierarchical linear modeling (HLM) were performed, and their results were compared. The materials included three popular mathematics and comprehension tests, all of which were in the domain of geometry. The main difference between the equation version and the verbal version was the method of representation used in key sentences (only one, five, and five sentences differed in each of the three passages, respectively). The other sentences, illustrations, and tests were the same in both versions. Students were randomly assigned into groups for each of the two versions and completed the reading comprehension tests, reading comprehension screening tests, and math prior knowledge tests. The ANCOVA results demonstrated that the equation readers outperformed the verbal readers and that the high-ability readers performed better than the low-ability ones, but the results did not indicate any interaction between version and ability. After the exclusion of the effects of school-average math ability, the HLM results found that

Corresponding Author: Ling-Chia Chang, E-mail: alcc@ntnu.edu.tw
Manuscript received: May 3, 2019; Revised: Jan. 13, 2020; Accepted: Mar. 2, 2020.

low-ability equation readers demonstrated a nonsignificant difference in performance compared with high-ability verbal readers. The benefits of using equations were discussed by comparing the linguistic features of the two external representations. Furthermore, the results were compared with previous research in terms of passage domain, cognitive load, focus of measurement, and participant characteristics.

Keywords: external representation, linguistics, popular mathematics, reading comprehension