
邱怡靜、曾建銘、吳昭容（2023）。
三至八年級學生數學文字題的表徵轉換與等價能力。
臺灣數學教育期刊，10（2），1-25。
doi: 10.6278/tjme.202310_10(2).001

三至八年級學生數學文字題的表徵轉換與等價能力

邱怡靜¹ 曾建銘² 吳昭容^{1,3}

¹ 國立臺灣師範大學教育心理與輔導學系

² 國家教育研究院測驗及評量研究中心

³ 國立臺灣師範大學學習科學跨國頂尖研究中心

代數思維是算術一般化的歷程，著重數量或變量間的關係，而掌握等價性為算術過渡到代數的關鍵之一。問題情境提供語意脈絡，有助於學生發展等價性。本研究為了瞭解橫跨國小到國中的等價性發展，針對 2677 位三至八年級學生設計數學文字題解題算式的多重是非題題組，採用認知診斷模型（cognitive diagnostic model [CDM]）界定出表徵轉換能力與等價能力兩種認知屬性，確認學生可分成困難者、單一表徵者、多元表徵者三種類型，並以試題反應理論（item response theory [IRT]）估計學生的能力值，以探究三至八年級學生在三種類型的百分比、能力值及其差異。研究結果顯示，隨著年級的增加，困難者百分比大致呈現下降，單一表徵者、多元表徵者百分比大致呈現上升的趨勢，而國小已有二至六成學生屬於多元表徵者，顯示國小階段即可培養初階代數能力，然而國中學生僅四成為多元表徵者。在能力值方面，困難者最低，單一表徵者居中，多元表徵者最高，且困難者文字題算式判斷能力成長遲滯，單一表徵者的能力隨年級增加明顯些，而多元表徵者的能力則隨年級有更顯著地增長，顯示等價能力會隨等式涉及的數學內涵而呈現多層次。此外，本研究示範了整合 CDM 與 IRT 以描繪長期發展之能力的研究方法，並檢討其未來優化的方向。

關鍵字：代數思維、表徵轉換、等價

通訊作者：吳昭容，e-mail：cjwu@ntnu.edu.tw
收稿：2023 年 7 月 20 日；
接受刊登：2023 年 10 月 3 日。

Chiu, Y. C., Cheng, C. M., & Wu, C. J. (2023).

Analyzing representation transformation and equivalence abilities across students in grades three to eight in mathematical word problems.

Taiwan Journal of Mathematics Education, 10(2), 1–25.

doi: 10.6278/tjme.202310_10(2).001

Analyzing Representation Transformation and Equivalence Abilities across Students in Grades Three to Eight in Mathematical Word Problems

Yi-Ching Chiu¹ Chien-Ming Cheng² Chao-Jung Wu^{1,3}

¹ Department of Educational Psychology and Counseling, National Taiwan Normal University

² Research Center for Testing and Assessment, National Academy for Educational Research

³ Institute for Research Excellence in Learning Science, National Taiwan Normal University

Algebraic thinking involves generalizing arithmetic principles and focusing on the relationships between quantities or variables. Mastering the concept of equivalence is a key step in transitioning from arithmetic to algebra. Mathematical word problems provide a semantic context and aid in developing equivalence. To investigate the development of mathematical equivalence from elementary to junior high school, the study devised sets of multiple true or false questions based on mathematical word problems and recruited 2,677 students spanning grades three to eight. Based on the cognitive abilities of representation transformation and equivalence ability, students were divided into three groups. The cognitive diagnostic model (CDM) was used to examine the percentages of three ability types within grades three to eight: difficulties, single representations, and multiple representations. The item response theory (IRT) was used to estimate the students' ability levels and discern variances across these three types of abilities for students in grades three to eight. The results showed that as students progressed through higher grade levels, the percentage of those difficult students decreased, while the percentages of students with single representation and multiple representation increased. It was observed that a considerable percentage of elementary students, ranging from 20% to 60%, exhibited proficiency in handling multiple representations, indicating that primary algebraic thinking could be nurtured at the primary school level. However, only about 40% of junior high school students demonstrated competency in multiple representations. In terms of ability levels, difficult students manifested the lowest values, followed by those with single representations; students with multiple representations displayed the highest values. Additionally, the growth of algebraic proficiency among difficult students lagged behind. Students with single representations showed noticeable improvement with increasing grade levels, while those with multiple representations exhibited more significant growth, indicating that equivalence abilities display multilevel characteristics depending on the mathematical content integrated into the equations. Furthermore, this study elucidated the research methodology that combines CDM and IRT to delineate long-term developmental abilities and discussed their potential optimization.

Keyword: algebraic thinking, representation transformation, equivalence

Corresponding author : Chao-Jung Wu , e-mail : cjwu@ntnu.edu.tw

Received : 20 July 2023;

Accepted : 3 October 2023.

壹、緒論

普及代數學習 (Algebra for all) 的呼籲彰顯了代數是一門關鍵課程 (gatekeeper course) (College Board, 2000; No Child Left Behind Act of 2001, 2002)。代數不只是中學與高等數學的基礎，也是許多職業要求的能力 (Domina et al., 2015)。早期數學課程採用先算術、再代數的階段式安排，近年來學者認為這種教學進程讓學生在算術學習階段過於重視數值運算而忽略關係的察覺，從而妨礙後續代數的學習，因此主張代數可以在小學低、中年級就及早開始 (Cai & Knuth, 2005; Kieran, 2004a, 2022)；實徵研究也支持小學低、中年級學生就能獲益於代數課程，而發展出核心的代數能力 (Blanton et al., 2015; Radford, 2018)。

討論代數思維 (algebraic thinking) 的內涵及其與算術思維 (arithmetic thinking) 之差異雖有多種不同的主張，但多數學者同意對等號意義的掌握，尤其理解等價性 (equivalence) 是關鍵能力之一 (Knuth et al., 2008; Knuth et al., 2005; Kieran, 2004a, 2004b)。探討學生發展初階代數能力的研究方式，常以如 $8 + 5 = \square + 9$ 的等式填空 (Kieran, 2004a) 或對 $13 + 11 = 12 + 12$ 的等式判斷對錯 (Molina & Ambrose, 2008) 來觀察學生是否理解等號不僅是指出左邊數字運算的結果 (operational meaning)，還掌握了等號左右兩邊數量之等價性的關係性意義 (relational meaning)，從而檢驗學生能否超越算術思維，進而掌握數量關係的代數思維 (Chesney & McNeil, 2014; Prediger, 2010)。

然而，放在脈絡意義更為豐富的文字題 (word problem) 情境下，一個問題的解決可採數種不同的列式呈現，同樣是一種等價性的關係理解，也是能在國小低、中年級就開始的學習活動，提供及早的代數活動經驗 (Carey, 1991)。例如「小明有 3 顆糖，小明和小華共有 8 顆糖，請問小華有幾顆糖？」學生若能理解列式能採 $8 - 3 = \square$ ，也可以是 $3 + \square = 8$ ，代表他能掌握這兩個算式的等價性，比僅能判斷其中一式正確的學生來得更具代數思維的能力，或至少已具備前代數思維 (pre-algebraic thinking)。

隨著學習加減、乘除、四則運算等課程內容，學生若不只能由已知數求算答案這類算術思維來看待各類文字題，還能掌握各類文字題解題存在不只一種可能的列式、從數量關係這類代數思維來理解問題情境，代表即使小學的早期亦能在文字題的脈絡下培養代數思維。反之，即使已經學習了方程式等代數課程，若學生無法判斷兩個等價的列式，那麼其代數思維仍有待加強。換言之，代數思維能力是橫跨國小到國中長期發展的能力。認知診斷模式 (cognitive diagnostic model [CDM]) 能透過作答反應分析學生的潛在能力、推估其具備或不具備哪些認知屬性 (如技能、概念等)；而試題反應理論 (item response theory [IRT]) 的垂直等化能將跨年級的學生能力定位在同一把連續量尺上，而水平等化則能把同年級多個題本的結果放在同一量尺，本研究採用跨年級、多複本的文字題解題列式之判斷作業，從 CDM 與 IRT 兩種角度探究學生的代數思維，並探討代數思維在不同年級的變化。

貳、文獻探討

一、算術思維與代數思維

學生在進行算術操作時，多以具體數值為出發點，透過對數值的運算求得結果 (Kieran, 2004a; Kilpatrick et al., 2001; Radford, 2018)；因此算術思維被視為以求得答案為目的的思維方式。例如前文「小明與小華」題的求解，不論學生採用 $8-3=\square$ 或是 $3+\square=8$ 的列式，如果他是針對具體數字 3 和 8 進行運算，例如「8 扣掉 3 是多少？」或者「3 湊上多少會是 8？」都屬於算術思維。因此，並非能解決含有未知數的算式就具備代數能力。面對 $3+\square=8$ 若能從等號左右等價的角度看到「因為 $3+\square$ 和 8 一樣多，所以 \square 等於 8 減 3」，這種思維方式就開始展現代數思維的特徵。因此，在某些含有未知數題目類型上，不是僅從能否正確解題就能區分算術思維與代數思維。

樣式一般化 (generalization) 與文字題解題 (problem solving) 的活動常被運用為算術思維過渡到代數思維的活動 (Radford, 1996; Sibgatullin et al., 2002)。在尋找樣式一般化的過程中，雖然所觀察到的是具象的樣式或具體的數列，但變化的數量是一種變量，活動的焦點並非確切的數值結果，而在於藉由數值間的變化建構抽象的等式。而文字題解題需要列出求算未知數的算式，未知數可視為一種變量，學生可透過等式尋找未知數。Radford (1996) 認為，變量和方程式的概念與樣式的一般化取向有著本質的聯繫，而未知數和等式的代數概念似乎與問題解決取向有著本質性的聯繫。

代數思維通常被視為包含了將算術操作一般化的認知歷程，這可從幾位學者對代數內涵之觀點加以了解。Kaput (1998, 2008) 主張代數有多個交織的面向，其中一般化和符號操作是兩個核心，前者通常是（但不限於）對算術樣式的一般化與一般化的表示方式，後者是由符號系統的語法引導對符號的操作。此二核心之上建構出學校代數課程的三個主題：一是從操作與關係抽離結構與系統，二是對函數、關係、聯變的探究，三是建模語言的應用。Kieran (1996, 2004b) 提出了另一種學校代數的理論模式，她描述了代數的三個相互關聯的主要活動：生成活動 (generational activities)、轉化活動 (transformational activities)、以及全域/元級活動 (global/meta-level activities)。生成活動涉及代數式和等式的產生，包括文字題、形體樣式或數列、數字關係的規則等；轉化活動涉及在語法指導下的形式化操作，如整併同類項、因式分解、去括弧、解方程等；全域/元級活動涉及以代數為工具的活動，包括：解題、建模和預測、探究結構和變化、分析關係、一般化和證明等。Hodgen 等人 (2018) 認為 Kieran 模式的生成活動對應到 Kaput 核心一（一般化），而 Kieran 模式的轉化活動則對應到 Kaput 的核心二（符號操作）。

二、等價性及其發展

對於代數思維的內涵雖有上述不同理論主張，但都指出等號的意義與等價性在發展代數思維上的重要性。例如，Kieran (2004a, p. 141) 認為算術思維過渡到代數思維時，重新

聚焦到等號的意義是五個重點之一，而 Kaput (1998, 2008) 也認同等號意義與等價性是算術一般化到代數的關鍵。學習數學初期，學生習慣將等號視為「運算結果是…」的符號；對等號的認知是連結算式與結果的符號，亦即等號左邊為運算程序，右邊是運算結果。此種狹隘的等號意義理解，使學生在面臨 $9+4+3=9+\square$ 的問題時，容易寫出 $9+4+3=9+16$ 或 $9+4+3=9+25$ 的錯誤形式，影響代數思維的形成 (McNeil, 2007)。因此等號意義的擴展被視為是代數學習的準備 (Knuth et al., 2008)，且實徵研究支持等號的理解預測兩年後的代數思維表現 (Matthews & Fuchs, 2020)，顯示發展代數思維必須轉換學生對等號意義的認識，建立等價性的概念。

等價性有兩個核心意義，一是相等，另一是可互換 (interchangeable) (Fyfe et al., 2022)。掌握等價性不僅止於將等號理解為一種關係符號，還包括正確地編碼等式、辨識等式的兩邊，並將等式兩邊的數或式理解為數學物件。也就是不論 $4+3$ 在等式的左邊還是右邊，都能不僅解釋為將 3 加到 4 的操作過程，還能視 $4+3$ 為一種可被操作的抽象物件，並知道數與式可以用多種可互換的方式表示 (McNeil et al., 2019)，包括數與數 (如 $9+4+3=9+\square$)、數與式 (如 $x^2=9$)、式與式 ($x^2+4=4x$) 的互換。加法等式填空問題雖是檢驗小學生掌握等價性常見的測驗題型 (Hornburg et al., 2022; McNeil et al., 2019)，然因數學存在著各式各樣的運算，掌握等價性涉及能否正確編碼等式兩邊的數學物件，故絕非僅用加法等式就能完整評估。有些研究運用了包括減法或乘除法的算式 (Matthews & Fuchs, 2020; Molina & Ambrose, 2008; Wardat et al., 2021)，豐富了不同層次之等價性的面貌。

算術思維雖是代數思維的基石，但運算的操作慣性與等號意義的侷限性會阻礙學生發展代數思維。根據「改變阻力理論」(change-resistance account) (Byrd et al., 2015; Hornburg et al., 2022; McNeil, 2014) 的觀點，學生受到早期運算經驗不斷強化的影響，認為等號右邊為左邊的運算結果，內化了此種算術思維的狹隘觀念；當學習代數課程時，固有的經驗阻礙對等號意義的重新認識、阻撓學生發展對數學等價性的理解，進而在代數學習上產生困難。McNeil (2007) 以七至十一歲學生完成如 $9+4+3=9+\square$ 的等式來驗證該理論。發現七至十一歲學生等價性的表現隨年齡增加呈現 U 型現象，也就是至少答對一題加法等式填空題的學生，在七歲和十、十一歲的百分比顯著高於八、九歲的學生。她解釋原因在於，國小一、二年級學生剛開始接觸算術，學習經驗不足以干擾等價性的表現，於三、四年級時因大量算術的練習，受到算術經驗的抑制，導致對加法等價算式之問題的表現下降；直至五年級後，由於解題的回饋與多元的運算經驗，降低算術經驗對等價表現的抑制，使得等價表現有所提升。因此，及早提供代數活動的經驗，有助於防範固著在算術經驗而造成改變的阻力。

三、文字題列式求解的算術思維與代數思維

文字題是用文字呈現問題情境的數學題，發展文字題解題能力是數學課程的重要目標之一 (教育部, 2018)。然而文字題解題對學生而言較計算題來得困難許多。Mayer (1992)

將文字題解題分為問題表徵與問題解決兩個階段，問題表徵再細分為問題轉譯、問題整合兩項子階段，問題解決則再細分為規劃和監控、執行兩項子階段。於問題表徵階段，學生首先必須對文字陳述有基本語意的理解，並將外部文字問題轉譯成可理解的內在心理形式，建構問題情境中的連貫而一致的表徵，如：符號、圖形、示意圖…等；接著運用基模區分訊息的重要性、辨認關係、以符合邏輯的方式整合訊息，為後續問題解決歷程做準備。而問題解決階段，學生先針對所面臨的問題進行策略評估與規劃，擬定問題解決方案，將資訊轉換為數學結構，執行列式、運算，並在規劃過程中自我監控及糾正，不斷修改解決方案；最後進行算式解題。許多文獻都指出文字題解題的難點在於問題表徵階段（Hegarty et al., 1995; Krawec, 2014; Passolunghi & Pazzaglia, 2005），解題者必須運用基模知識建構心理模型，掌握問題的已知數與未知數的數量關係，方能進行後續的解題規劃與執行監控。國際學生能力評量計畫（Programme for International Student Assessment [PISA]）數學素養被界定為真實情境脈絡下解決問題的數學推理能力，並透過形成（formulate）、應用（employ）、詮釋與評估（interpret and evaluate）三個歷程不斷循環來實現（Organisation for Economic Cooperation and Development [OECD], 2018）。其中，形成是指將問題的情境脈絡轉化為數學結構；應用則是以數學概念、事實、程序和推理得出具體的數學結果；詮釋與評估是將所獲得的數學結果與情境脈絡結合，針對數學結果進行情境脈絡的解釋、衡量方案的可行性。PISA 中的數學問題脈絡情境多以文字呈現，若將 PISA 的三個歷程對應 Mayer（1992）的文字題解題階段，形成與問題表徵階段關係較緊密，而應用、詮釋與評估則與問題解決階段相關聯。

為了說明文字題解題歷程與算術思維、代數思維的關係，以下以單步驟加減文字題的改變題（change problem）為例。依據廣被引用的單步驟文字題分類，Riley 等人（1983）將單步驟加減文字題依其語意分成合併題、改變題、比較題三類，且每類又因未知數在基模中的角色不同而產生難度各異的題型。以改變題為例，其基模為「**起始量**經增加/減少**改變量**後，成為**結果量**」，而文獻顯示，結果量未知（文獻上簡稱為改變 1、改變 2）最簡單、改變量未知（簡稱改變 3、改變 4）難度居中、起始量未知（簡稱改變 5、改變 6）最為困難（Cummins et al., 1988; Riley et al., 1983）。結果量未知-減少（改變 2）的題目如「美美有 8 顆巧克力，給了小莉 5 顆，美美現在有幾顆巧克力？」只需順著情境改變的時序和語意列式為 $8 - 5 = \square$ ，列式和求算答案都不算困難。起始量未知-減少（改變 6）的題目如「美美有一些巧克力，給了小莉 5 顆之後，美美現在有 8 顆巧克力。請問美美原本有幾顆巧克力？」若順著情境改變的時序和語意，應該列式為 $\square - 5 = 8$ ，但此時求算答案就較為困難。若學生尚未掌握等號的等價性，無法採取等號左右各加 5 來求算答案，就必須嘗試錯誤地湊出 \square 的數值。而一般課堂上出現改變 6 的問題，通常會要求學生列式為 $8 + 5 = \square$ 。這樣的列式雖然有利於求算答案，但學生必須以逆轉時序與顛倒語意（「給了」變成「拿回來」）的方式理解問題，語意理解與推理能力必須夠好、工作記憶容量的要求也較高。

其它類型的單步驟加減文字題如合併題或比較題，或單步驟乘除文字題如等組、倍數比較、面積、直積等題型（Greer, 1992）亦有上述特性。難度較低的題型通常依著語意列式

後，會是未知數單獨在等號右邊的算式，而難度較高的題型若依語意列式，會出現未知數在等號左邊，需要掌握等量公理方才便於求算答案；反之，若要列出未知數單獨在等號右邊、易求答案的算式，則需要有較佳的認知能力。兩步驟文字題因為問題整合的需求更高，更容易產生列式或求解的困難。例如「建國的身高比曉樂高 12 公分，小英的身高比曉樂高 9 公分。小英的身高是 154 公分，那建國的身高是幾公分？」不論列式為 $\square - (154 - 9) = 12$ ，或是 $154 - 9 + 12 = \square$ ，在語意理解與推論或運算求解上，都相當挑戰。解題者若能判斷其中一個算式正確，代表具備問題表徵能力甚至解題規畫階段（Mayer, 1992）或者能執行形成歷程（OECD, 2018），本研究稱為具備「表徵轉換能力」，以前述改變 6 的「美美」題為例，若能辨識出 $\square - 5 = 8$ 或者 $8 + 5 = \square$ 兩者之一為正確列式，就具備表徵轉換能力。若能判斷 $\square - 5 = 8$ 和 $8 + 5 = \square$ 兩個算式都可以合理的呼應解題規劃，本研究則稱為具備「等價能力」。

以往探討學生代數思維的研究多採等式填空的計算題作業，並未提供問題情境（McNeil, 2007; Knuth et al., 2008; Wardat et al., 2021; Hornburg et al., 2022）。從前文可知，文字題解題可以在具有語意脈絡的情境下探究學生對算式之間等價性的掌握狀況，而且可以隨著學生年級的增加，將加減的問題擴展到更多元的數學問題上。本研究為了探究學生在等價性之代數思維上的差異，採用每一道文字題均提供三小題的列式，讓學生就各小題進行對錯的判斷。本研究安排了兩類的題組：一類在三小題列式中僅有一小題為正確，可以偵測學生的表徵轉換能力；另一類文字題則有兩小題正確，足可偵測學生的等價能力與表徵轉換能力；此一安排一則避免題組若固定都有兩小題正確會增高學生的猜測率，另則如改變 2 之類較簡單的文字題，合理的列式只有一種，故需隨題目特性彈性調整正確小題數量為一小題或兩小題。

四、認知診斷模式與試題反應理論

認知診斷模式是以多維度的二元向量來表示個體是否掌握了這多個潛在的認知屬性，並提供認知屬性之組型的訊息（吳慧珉等人，2015），例如本研究可用認知診斷模式以了解受測者是否具備表徵轉換以及是否具備等價能力，也能知道學生屬於兩種能力組合下的何種組型。目前 CDM 已有多種認知診斷模式被開發，其中 DINA 模式（deterministic input, noisy “and” gate model）是常被應用在數學教育研究的模式（吳慧珉等人，2012；Xu et al., 2023）。DINA 模式假設學生能否答對試題，除了受到是否具備相關的認知屬性外，亦受到疏忽及猜測兩個參數的影響。當學生具備該試題的所有認知屬性卻答錯該試題的機率，即為疏忽參數；當學生未具備該試題的所有認知屬性但答對該試題的機率，即為猜測參數。在 DINA 模式的假設下，若一道試題含有多個認知屬性，學生須具備該試題的所有認知屬性才能答對該試題。由於 DINA 模式符合數學解題的假設，且其模式簡單易懂，故本研究採用 DINA 模式來評估學生的表徵轉換、等價能力，及二者的組型。

試題反應理論是描述受測者能力、試題參數（如難度、鑑別度等），與作答反應的數學模式，旨在估計受測者潛在能力，且可藉由共同題或安排重複考生的測驗等化技術，將多個測驗題本連結，使多份測驗可在同一個量尺上進行比較（郭伯臣、王暄博，2008）。IRT 包含了許多不同模式，其中 Rasch 模式適用二元計分的試題，僅採用難度做為試題的參數，將受測者能力與試題難度建立在同一個量尺上，單位為 logit；能力與試題難度的差距決定受測者答對該試題的機率，當能力值高於試題難度值的差異越大，答對該題的機率越高（王文中，2004）。

五、本研究論點

以往代數思維的研究未見運用認知診斷模式探討學生潛在的認知屬性及其組型。本研究設計兩類的數學文字題算式判斷題組，用來偵測表徵轉換與等價能力兩種認知屬性；依據兩種認知屬性的有無，可將學生分成四種類型（詳見後文表 5），但基於表徵能力與等價能力的發展具有順序性，等價能力發展奠基於表徵能力之上，理論上不存在具備等價能力卻無表徵能力者，因此本研究預期以 DINA 能將學生區分為三種類型：困難者（答題錯誤或形態雜亂）、單一表徵者（只能正確辨識一小題列式），以及多元表徵者（能正確辨識兩小題列式者）。

以往文字題解題或是代數之等價能力的研究也少有橫跨多個年級以觀察其能力的發展現象。本研究以三至八年級共六個年級學生為對象，採用相同的文字題解題算式判斷作業，且試題內容配合各年段的學習程度，並在各年段題本間設有共同題，以便能垂直等化。為了確保銜接的試題參數具足夠穩定性，在共同題的選擇上，選取前後兩年段學生皆可作答、試題參數較佳的試題，並挑選對前一年段難度稍高，但對後一年段難度稍低的試題作為共同題。此外，本研究同一個年段內有多個複本試題，前述垂直等化的共同題也同時是水平等化的共同題（詳見後文圖 1）。故可透過 Rasch 模式對橫跨三至八年級學生進行列式判斷能力的估計。

綜合上述，為了瞭解三至八年級學生文字題能力在代數思維的表現，本研究運用文字題解題算式的題組，界定表徵轉換能力與等價能力兩種認知屬性，以探討以下兩個問題。

- （一）三至八年級學生在困難者、單一表徵者、多元表徵者三種類型的百分比及其差異為何？
- （二）三至八年級學生在三種類型上的能力值及其差異為何？

參、研究方法

一、受試者與有效資料

正式施測的受試者參考國小、國中學校別資料（教育部統計處，2016a，2016b），採分層取樣方式邀請北一區 3 個縣市、北二區 3 個縣市、中區 4 個縣市、南區 4 個縣市、東區及離島 3 個縣不同規模的學校參與，最後有 19 所國小及 20 所國中參與。國小以每校三至

六年級各 1 班，國中則每校七和八年級各 1 班的學生為對象。三至八年級各年級參與人數在 407 至 511 人之間，總計 2677 位學生，如表 1。

表 1

本研究正式施測之各縣市學校規模、校數，與人數

區域	縣市	校數	國小學校規模 ^a				國中學校規模 ^b			
		小	中	大	總計(人數)	小	中	大	總計(人數)	
北一	臺北市	0	1	0	1 (203)	0	0	2	2 (112)	
	新北市	2	0	0	2 (79)	1	1	1	3 (155)	
	宜蘭縣	0	0	1	1 (85)	0	1	0	1 (48)	
北二	桃園市	—	—	—	—	1	0	0	1 (40)	
	新竹縣(市)	0	2	1	3 (295)	—	—	—	—	
	苗栗縣	0	0	1	1 (96)	—	—	—	—	
中	臺中市	1	0	0	1 (74)	1	1	1	3 (138)	
	彰化縣	0	1	0	1 (102)	0	0	1	1 (59)	
	南投縣	—	—	—	—	0	1	0	1 (55)	
	雲林縣	0	1	1	2 (201)	0	0	1	1 (58)	
南	嘉義縣(市)	0	0	1	1 (102)	0	0	1	1 (55)	
	臺南市	1	0	0	1 (70)	0	1	0	1 (54)	
	高雄市	1	0	0	1 (61)	2	0	0	2 (96)	
	屏東縣	0	1	0	1 (98)	—	—	—	—	
東＋離島	花蓮縣	0	1	0	1 (74)	1	0	0	1 (32)	
	臺東縣	1	0	0	1 (50)	1	0	1	2 (98)	
	澎湖縣	0	0	1	1 (87)	—	—	—	—	
	總計	6	7	6	19 (1677)	7	5	8	20 (1000)	

註：^a 國小學校規模 6 班以下為小校、7~24 班為中校、25 班以上為大校。

^b 國中學校規模 12 班以下為小校、13~36 班為中校、37 班以上為大校。

測驗的無效樣本比率約 1%，因此正式施測的有效樣本為 2,654 人。無效樣本的判定原則有三，包括：1.空白或作答不清楚的題數超過一半者。2.答對率在 1/4 以下，且答題具某種樣式（如：111222111222...等）者。3.學校提供的特教生名單。

二、研究工具

本研究使用自編的「數學文字題測驗」，數學文字題的組成包含文字描述的已知資訊（例如：小明有 438 元、甲買了一瓶沙拉油、…）、量的關係（小華給小明5 顆糖、平均分給 3 個人、…）、及問句（小華有幾顆糖？每人可以分得幾個？、…）。數字類型含括整數、分數、或小數。答題方式採多重是非題類型，詳見「命題」。

（一）題型分類

本研究根據解題時會運用到的算式將三至八年級的數學文字題分成四種類別：一步驟加減、一步驟乘除、二步驟計算、一元一次應用，並將各類別再細分成數類，共計 25 種的

文字題型。測驗區分為 A、B、C 年段（分別是三和四、五和六、七和八年級），各年段有四個複本，以每個複本最多 20 題估算，各年段抽選 20 個題型預備命題。確認題型架構後，其後進行分年段的題本設計、題數估算與數字（整數、分數、小數）分配。

（二）命題

四名數學教師參與命題，題型採多重是非題。每道文字題搭配三個算式，作答時只要判斷列式是否正確，不用計算答案。表 2 為數學文字題命題之示例題，各年段以一題組為例，呈現數學文字題的命題樣式。每個列式搭配「①對」與「②錯」兩個選項，若認為算式正確，則選①；若覺得算式錯誤，則選②。每道文字題與三個算式的判斷作業稱為一題組。三個列式中「至少有一個是正確算式，及至少一個是錯誤算式」，其中，一小題正確的題組有 63 題（如表 2「豆花店」），兩小題正確的題組有 71 題（如表 2「陳爺爺」）。

表 2
數學文字題命題之示例題

年段	類別	題型	數字	示例題
A	一步驟加減	比較	整數	<p>陳爺爺忘記自己今年幾歲，只知道陳爸爸比陳爺爺年輕 28 歲，陳爸爸今年 53 歲，陳爺爺到底幾歲？</p> <p>$53 + 28 = \square$ ①對 ②錯</p> <p>$53 - 28 = \square$ ①對 ②錯</p> <p>$\square - 28 = 53$ ①對 ②錯</p>
B	一步驟乘除	倍數	分數	<p>小樹豆花店製作豆花用掉 5 公升的糖水，美美豆花店製作豆花用掉 $4\frac{1}{3}$ 公升的糖水，小樹豆花店用掉的糖水是美美豆花店的幾倍？</p> <p>$5 \div 4\frac{1}{3} = \square$ ①對 ②錯</p> <p>$4\frac{1}{3} \div 5 = \square$ ①對 ②錯</p> <p>$5 \times 4\frac{1}{3} = \square$ ①對 ②錯</p>
C	二步驟計算	比例	小數	<p>一台掃地機器人掃地的速度固定不變，若掃 10.5 平方公尺，花掉 32.5 分鐘，則 25.5 分鐘可掃多少平方公尺？</p> <p>$10.5 \div 32.5 \times 25.5 = \square$ ①對 ②錯</p> <p>$25.5 \times 32.5 \div 10.5 = \square$ ①對 ②錯</p> <p>$25.5 \times (10.5 \div 32.5) = \square$ ①對 ②錯</p>

經由測驗專家建議，各年段每個題型含平行題需命 5 題，因此 A、B、C 年段各命題 100 題組。各年段在分數與小數的學習程度不同：於 A 年段，分數只做同分母分數的加減列式，小數只做加減列式，且三年級到小數第 1 位、四年級到小數第 2 位；於 B 年段，可做異分母分數加減、分數的整數倍、分數乘以分數列式，及多位數小數加減、小數的整數

倍、小數乘以小數、小數除以整數列式；於 C 年段，分數和小數都可進行四則混合列式。因此在命題時整數、小數、分數的數字類型分配上有所不同。

研究者先舉數個範例題供命題老師參考。命題卡上包括文字題所屬年段、類別、題型、數字類型，及題目內容、三個列式、正確答案、難度等。於命題初期，召開會議，設定命題原則：

1. 相似類別須區分

「一步驟加減」中的「比較」是對兩物之比較，即題目有「…比…」的敘述；「改變」則是對同一標的物（人）前後之改變。

2. 不過於侷限題目內容

「一步驟乘除」中的「倍數」非僅限於幾倍的敘述，亦包含倍、打折、比、比例、百分率等概念。

3. 避免無效資訊

情境描述儘量符合學生經驗，避免出現多餘或無用的敘述。

4. 各年段須有難度差異

同一文字題型，若於各年段都有命題，難度或程度應符合 $A < B < C$ 。

5. 列式以等式呈現，並注意數字表示

列式都呈現完整等式，故框框「□」可出現在等號的左邊或右邊；選項列式都以題目有出現的數字為主，避免呈現換算或化簡後的數字；列式雖不計算答案，但實際答案需合理。

6. 錯誤列式非任意組合數字

可以從學生易出現的迷思概念（例如「倍」字就表示用乘法）、或具誘答性的列式解答（例如題目有「某數比 A 少 3」的敘述，錯誤列式則以「 $\square - A = 3$ 」表示）來設計。

（三）審題與修題

300 道題組分成 6 個審查題本，分別交由 6 位數學專家審查。評分包括：**1.內容**須符合受測年段的學習程度，數字類型設定及題型分類應適當，避免出現罕見字而影響學生閱讀。**2.題幹**應該是學生熟悉的情境，語義要清楚明白、語順要流暢，字數應適中，句子不宜過長。**3.平行題**的敘述應與範例題一致，以使難度相當。**4.正確列式**須無爭議性，每題組至少要有有一個正確解答。**5.錯誤列式**應具備誘答力，每題組至少要有有一個錯誤解答。審查結果與命題老師於會議討論，並依共識修題。

（四）測驗與計分方式

組卷以 20 分鐘能答完試卷為原則，且 A 年段於題目與選項均加上注音，以避免識字量及認字關係而影響到題目閱讀與選答。作答方式考量三、四年級學生不熟悉畫卡，故 A

年段為直接作答於題本上；B、C 年段則採畫記答案卡。

1. 預試

各年段僅組一個題本進行預試，目的在估計各年段的題數以及挑選共同題。預試採便利取樣兩所國小和兩所國中（國中小為大校、小校各一所），三至八年級人數分別為 56、51、50、48、46、47 名，共 298 人。A 至 C 三卷題本均為 15 題，難度分配約為易 24%、中 68%、難 8%；類別分配 A 年段約為一步驟加減 47%、一步驟乘除 20%、二步驟計算 33%，B 年段約為一步驟加減 33%、一步驟乘除 13%、二步驟計算 54%，C 年段約為一步驟加減 13%、一步驟乘除 13%、二步驟計算 54%、一元一次應用 20%。數字分配約為整數 60%、分數 20%、小數 20%。結果發現 A 年段平均 8 分鐘即答完，而 B、C 年段平均 14 分鐘答完，但有極少數人在時間到時未能完成作答。

刪除過於簡單的試題後，根據題組的 CTT 難度、鑑別度，選出正式測驗的 8 個共同題組，以對低年段較難、對高年段較易，及包含整數、分數、小數的題目為原則，如表 3。

表 3
數學文字題正式測驗各年段間所採用之共同題組

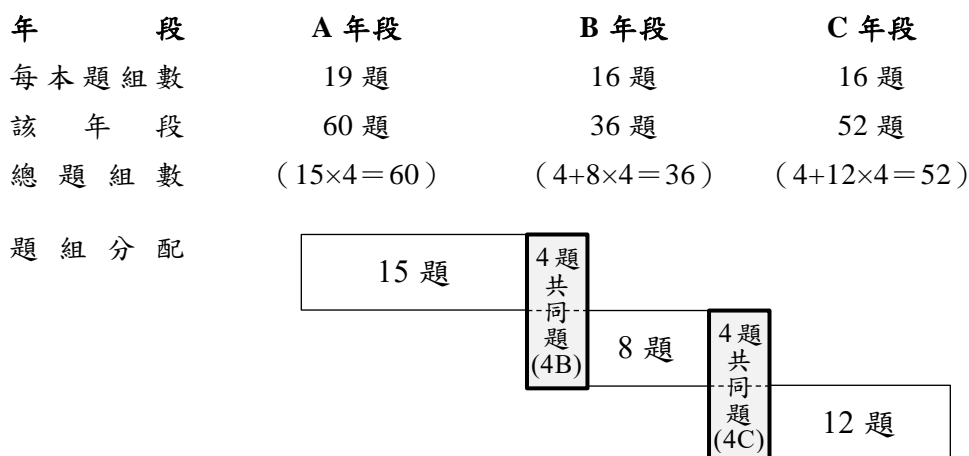
年段	類別	題型	來源 年段	數字 類型	題組 CTT 難度	題組 鑑別度
A、B	一步驟加減	比較（被比較量未知）	B	分數	0.72	0.33
	一步驟乘除	等量（連續量）	B	小數	0.73	0.35
	二步驟計算	加乘/乘加	B	整數	0.75	0.31
	二步驟計算	減除/除減	B	整數	0.70	0.41
B、C	一步驟加減	改變（起始量未知）	C	小數	0.66	0.61
	一步驟加減	比較（參照量未知）	C	分數	0.63	0.60
	二步驟計算	減減/減減	C	整數	0.66	0.58
	二步驟計算	減乘/乘減	C	整數	0.68	0.71

2. 正式測驗

A、B、C 三年段各四個複本，納入預試選出的共同題同時做垂直等化與水平等化。依預試的學生作答情形決定本次各複本題組數，A 年段為 19 個題組（57 小題）、B 年段為 16 個題組（48 小題）、C 年段為 16 個題組（48 小題）。由於複本間皆為平行題或相同題目（共同題），且相同題型的小題題序均相同，因此各複本的難度理應相當。其中 B 年段各題本和 A、C 年段各題本分別各有 4 個共同題組，如圖 1 所示，這些共同題不僅為跨年段共用，也是同年段各複本共用，例如 A、B 年段合計八個題本會共用 4 個相同的題組。共同題的比例則是參考過去學者的建議，以大於題本題數的五分之一為原則（Holland & Dorans, 2006; Livingston, 2004）。故 A、B、C 年段的總題組數各有 60、36、52 題，共 148 個題組。施測後刪除品質較差的題目，留用 134 個題組。

圖 1

數學文字題各年段題本的題組數及共同題組配置



3. 計分方法

首先，每一小題判斷正確即得 1 分，錯誤則為 0 分，以能檢視每一小題的鑑別度、難度與通過率；其次，以每一題組的三個小題全答對為 1 分，其中有一個小題判斷錯誤則為全錯，並以 0 分來計算，當學生能判斷文字轉譯成算式時，表示學生具備「表徵轉換能力」；當除了能列出算式之外，還能知道不同表達方式，表示學生具備「表徵轉換能力」及「等價能力」，藉此以檢測出學生文字題表徵及轉換能力。

(五) 試題參數估計與測驗等化

為了將不同題本的結果放在相同量尺上進行比較，本研究採用 IRT 的同時估計法 (concurrent estimation) 進行試題參數估計與測驗等化，利用共同題將所有題本的試題作答反應放在同一個檔案中進行估計，減少了測驗間連結時產生的誤差。國內外許多文獻指出，採用同時估計法能獲得較佳的精準度 (郭伯臣、王暄博, 2008; Hanson & Béguin, 2002)。

當使用同時估計法進行測驗等化時，會將受測群體的能力分佈視為單一群體並估計其平均數與標準差，這對於跨不同年級且明顯為不同能力群體的垂直等化設計 (vertical equating) 較不適合，會低估困難題本的試題參數，且高估簡單題本的試題參數 (DeMars, 2003)。因此，在垂直等化情境中為了降低上述偏誤，Wu 等人 (1998) 建議可將不同能力群體視為不同能力分布參數來進行同時估計，並以潛在迴歸 (latent regression) 模式來進行試題參數與能力估計。

在本研究中，為了精確估算各年級學生的代數能力值，因此除了正式測驗的作答反應資料外，納入系列研究中後續試題補充施測及常模建置施測的另外 68 個題組之作答反應，合併成一個 202 個題組的作答反應矩陣，進行 IRT 試題參數與能力估計，以便進行水平與垂直等化。矩陣內 1 與 0 的計分，以每一題組三個小題都答對為 1 分，其中有任一小題答錯則為 0 分。

(六) 編製 Q 矩陣

本研究以認知診斷模式中的 DINA 模式評估診斷兩個認知屬性的數學文字題測驗。進行 CDM 時，藉由 Q 矩陣表示第 j 個試題是否需具備認知屬性 k ，矩陣大小為 $J \times K$ ，以 $\alpha = (\alpha_1, \alpha_2)$ 表示受試者的二元精熟分數，若某個受試者的 $\alpha = (1, 0)$ ，表示他在第 1 個屬性是精熟的，而第 2 個屬性是不精熟的。因此本研究將 134 道題組分成具備表徵轉換能力 (K_1)、具備等價能力 (K_2) 2 個認知屬性，每一試題即為 Q 矩陣的一列，Q 矩陣如圖 2 所示。

圖 2

兩個認知屬性之 Q 矩陣

$$Q = \begin{matrix} & \begin{matrix} K_1 & K_2 \\ \downarrow & \downarrow \end{matrix} \\ \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} & \begin{matrix} \leftarrow \text{第 1 題} \\ \leftarrow \text{第 2 題} \\ \leftarrow \text{第 3 題} \\ \vdots \\ \vdots \\ \leftarrow \text{第 132 題} \\ \leftarrow \text{第 133 題} \\ \leftarrow \text{第 134 題} \end{matrix} \end{matrix}$$

為了檢視學生的表徵轉換能力、等價能力，本研究每個題組的 3 小題設計為「至少有一小題是正確算式，及至少一小題是錯誤算式」。當僅有一小題為正確答案時，學生必須知道如何將文字轉譯成數學算式，因此能判斷學生是否具備「表徵轉換能力」；而當題組的答案模式有兩個正確答案時，學生可能僅知道如何列出數學算式，或者還知道問題可有不同的算式表達方式，因此可以判斷學生是具備「表徵轉換能力」，或是具備「等價能力」。作答型態與兩種認知屬性的對應情況如表 4 所示。

表 4

數學文字題於具備表徵轉換與等價能力的答案型態

標準答案範例	正確題數	標準答案範例	正確題數
○ × ×	僅有一小題	○ ○ ×	兩小題
作答反應	能力判定	作答反應	能力判定
○ × ×	$K_1=1$	○ ○ ×	$K_1=1, K_2=1$
		○ × ×	$K_1=1, K_2=0$
		× ○ ×	
○ × ○		○ × ○	
○ ○ ×		○ ○ ○	
○ ○ ○	$K_1=0$	× ○ ○	$K_1=0, K_2=0$
× ○ ○		× ○ ○	
× ○ ×		× × ○	
× × ○		× × ×	
× × ×			

以表 2 中的「豆花店」及「陳爺爺」為例。「豆花店」題組中只有一個小題的算式是正確的，若學生能辨識出該小題正確，且判斷其它兩個小題錯誤，則顯示學生具有表徵轉換能力；而「陳爺爺」題組中有兩個小題的算式是正確的，學生能辨識出該兩小題的正確性並容許不同的表達方式，顯示學生具備表徵轉換及等價能力；若只辨認出其中一正確小題，且能拒絕錯誤小題，則顯示這位學生僅具備表徵轉換能力；其它組合的答案型態則代表學生既不具備等價能力、也不具備表徵轉換能力。

本研究運用 DINA 模式從學生各題的答題型態將學生區分成四種類型，如表 5 所示：(0, 0)代表「未具備表徵轉換與等價能力者」，(0, 1)代表「單具備高層次的等價能力，但無較低層次表徵轉換能力者」，(1, 0)代表「單具備表徵轉換能力，但無較高層次的等價能力者」，(1, 1)代表「兼具表徵轉換與等價能力者」。基於理論上不存在(0, 1)類型的學生（實際結果詳見「研究結果一」），因此將另外三類命名為困難者、單一表徵者、多元表徵者三種類型。

表 5
認知屬性與類型

認知屬性 (K_1, K_2)		表徵轉換能力 K_1	
		無	有
等價能力 K_2	無	(0, 0)困難者	(1, 0)單一表徵者
	有	(0, 1)併入(0, 0)分析	(1, 1)多元表徵者

三、程序

由各班導師或數學老師擔任主試者，施測程序依研究者提供的指導語及注意事項進行，實施地點在原班教室，各班施測時間為 20 分鐘。施測時，由主試者先發下題本及答案卡，四個複本均採每班隨機發放。施測完成後，本研究彙整資料，針對 2,654 位有效樣本進行試題參數與學生能力估計，以便進行水平與垂直等化，並在 DINA 模式下，使用試題對應認知屬性的 Q 矩陣分析，依學生代數能力進行分類，以及運用 Rasch 模式估計學生文字題解題算式的判斷能力。

肆、研究結果

一、認知診斷與試題品質

為檢核 DINA 模式的適配度，本研究將所有題本進行 DINA 模式與 G-DINA 模式（generalized deterministic input, noisy “and” gate model，其放寬 DINA 的一些前提限制、擴

展 DINA 模式的廣義性；de la Torre, 2011）的適配度檢驗，並以 BIC 指標作為檢驗指標；當 BIC 的值較小，則代表該模型的適配度較好。結果 11 個題本皆以 DINA 模式適配度較好，僅 B 年段 1 個題本呈現 G-DINA 模式適配度較佳。因此本研究的試題與 DINA 模式相適配。DINA 分析的結果顯示，試題猜測參數的中位數為 .27（四分差為 .17），疏忽參數中位數為 .16（四分差為 .07），數值皆偏高。推測原因在於本研究的作業採用多重是非題，而非是非題的缺點就是猜測度高，且學生不熟悉多重是非的題型，易導致試題有較高的疏忽參數值。DINA 認知診斷的組型顯示，表 5 四種類型中(0, 1)類型的學生人數如預期地極少，只佔全體樣本的 4%，可能因為猜測因素導致有極少數學生被歸為此類，因此後續將此類學生併入(0, 0)進行分析。

本研究三個年段所有題本之 Cronbach's α 在 0.82 至 0.87 之間，顯示具有良好的信度，並以 IRT 之 MNSQ 做為試題的建構效度之指標。根據 Wright 與 Linacre（1994）的建議，MNSQ 在 0.6 至 1.4 的範圍內，代表試題具良好的適配度，而本研究 89%題組的 MNSQ 介於 0.6 至 1.4 之間，代表大多數題組與 Rasch 單向度模式相適配，測驗試題測量到相同的潛在特質，具有建構效度。而 Rasch 難度值平均為 0.12 logit，代表整體試題對受測者而言屬中等難度，能適當地檢測出學生的能力。

二、不同類型的百分比及其差異

三至八年級學生在三類類型的百分比與總人次如表 6。卡方檢定分析結果顯著， $\chi^2(10, N=2654)=804.18, p<.001$ ，年級與類型有呈現顯著關聯性。進一步進行事後比較，以檢驗不同類型的百分比在各年級間是否具有差異，表中以不同字母的下標表示差異達顯著。結果於困難者中，三至六年級各年級均有差異（ $ps<.05$ ），而六、七、八年級的百分比無顯著差異，顯示於六年級開始，困難者百分比在 13%左右，呈現平緩的趨勢；在單一表徵者中，三至五年級逐年顯著增加，六年級的百分比卻降至與四年級相當，七、八年級的百分比則略增，超過四成百分比；至於多元表徵者，三年級的百分比佔兩成，顯著地低，四、五、七、八年級則佔三、四成百分比居次，而其彼此間無顯著差異，六年級的百分比則超過六成，顯著高於其它各年級。

表 6
三至八年級各類型的百分比及其差異

類型百分比 (%)	三年級	四年級	五年級	六年級	七年級	八年級
困難者	78 _a	46 _b	23 _c	11 _d	16 _{c, d}	12 _d
單一表徵者	2 _a	19 _b	36 _c	25 _b	43 _{c, d}	46 _d
多元表徵者	20 _a	35 _b	40 _b	64 _c	41 _b	42 _b
各年級人數 (人)	401	422	426	417	483	505

註：橫列各細格百分比的事後比較達 .05 顯著差異，即以不同字母的下標表示。

上述三類學生的百分比隨年級的變化大致如預期，亦即困難者的百分比隨年級增長而下降，單一表徵者與多元表徵者則上升，但有兩點值得關注。一是，六年級學生在整體趨勢中顯得特別地好，也就是困難者與單一表徵者特別少，而多元表徵者特別多。二是，若以多元表徵者作為代數思維在文字題列式判斷的一種體現，即使三年級也有兩成的學生展現這樣的能力，但已正式學習國中代數課程的七、八年級學生卻僅四成左右被歸為多元表徵者。

三、不同類型的能力值及其差異

Rasch 模式是依據每位學生的試題反應估計出單一的能力值，在本研究是能否判斷文字題解題算式的正確與否，因此估計的能力主要為表徵轉換；又因其中部分題組有等價的兩個正確列式，故該列式判斷能力也部分反映等價能力。三至八年級各類型的能力值見表 7 所示。類型（3） \times 年級（6）二因子變異數分析顯示，類型和年級的主要效果均達顯著， $F(2, 2636) = 2064.14$ ， $p < .001$ ， $\eta^2 = .61$ ； $F(5, 2636) = 75.69$ ， $p < .001$ ， $\eta^2 = .13$ ，且類型與年級交互作用達顯著， $F(10, 2636) = 17.96$ ， $p < .001$ ， $\eta^2 = .06$ ，因此接續進行單純主要效果與事後檢定。類型的單純主要效果在三至八年級全都顯著，分別為 $F(2, 2636) = 270.86, 465.84, 304.36, 296.06, 594.89, 479.35$ ， $ps < .001$ ，且三類學生的能力值在各年級的事後比較全都是困難者 $<$ 單一表徵者 $<$ 多元表徵者。年級的單純主要效果在三種類型也皆顯著，困難者、單一表徵、多元表徵的 F 值分別為 $F(5, 2636) = 20.86, 55.98, 77.71$ ， $ps < .001$ ，後續事後比較的結果在表 7 以不同字母的下標表示差異達顯著。其中困難者的三年級能力值顯著低於四至八年級；單一表徵者的三、四年級能力值顯著低於其它年級，其次為五至七年級，而八年級則顯著高於其它年級；而多元表徵者的三年級能力值最低，其次為四、五年級，再其次則為六年級，最高的七、八年級能力值顯著優於其它年級。

表 7
三至八年級各類型的能力值及其差異

平均能力值 (標準差)	三年級	四年級	五年級	六年級	七年級	八年級
困難者	-1.83 _a (0.70)	-1.52 _b (0.70)	-1.27 _b (0.55)	-1.37 _b (0.41)	-1.36 _b (0.58)	-1.22 _b (0.49)
單一表徵者	-1.31 _a (0.55)	-1.05 _a (0.58)	-0.30 _b (0.49)	-0.33 _b (0.46)	-0.20 _b (0.51)	0.02 _c (0.59)
多元表徵者	0.13 _a (0.51)	0.69 _b (0.69)	0.72 _b (0.81)	1.00 _c (0.82)	1.70 _d (0.96)	1.69 _d (0.92)
人數 (人)	401	422	426	417	483	505

註：能力值單位為 logit，且橫列各細格能力值的事後比較達 .05 顯著差異，即以不同字母的下標表示。

上述三類學生的能力值如預期地，困難者的能力最低、單一表徵者其次、多元表徵者最高；且隨年級的變化也大致如預期，三類學生都隨著年級而能力值提升，但有兩點值得關注。一是，困難者的年級差異較不明顯，僅有兩個能力層次，即三年級顯著較差，其餘四至八年級的困難者能力沒有差異；單一表徵者的年級差異稍微明顯些，分成了三個能力

層次；多元表徵者的年級差異則更為明顯，分成了四個能力層次。二是，雖然上一小節顯示六年級學生在三類的百分比顯得很獨特，但從本小節的能力值來看，三類型的六年級學生能力值在年級間都屬合理。

伍、結論與建議

本研究以兩類是非題組界定出表徵轉換能力與等價能力兩個認知屬性，CDM 結果支持這兩類能力的組合僅有三種：兩種能力都不具備的困難者、僅有表徵轉換能力的單一表徵者，以及具備兩種能力的多元表徵者。回應研究問題一，不同類型的百分比差異方面，隨著年級增加而大致呈現困難者百分比下降、單一表徵者與多元表徵者百分比上升的趨勢，其中六年級學生顯得特別好，有超過六成的多元表徵者。此外，即使是國小三年級的多元表徵者也有兩成，但國中七、八年級學生的多元表徵者僅四成左右。回應研究問題二，不同類型學生在文字題解題算式判斷能力值的差異方面，在三至八各個年級內皆如預期地以困難者的平均能力值最低，單一表徵者居中，多元表徵者最高；類型內也大致如預期地隨年級越高而能力值越高，這代表同一類型的學生能力值並非一致的。其中，困難者的能力值差異較不明顯，僅分成兩個能力階層；單一表徵者則分成三個能力階層；多元表徵者分成四個能力階層。此外，三類型的六年級學生能力值與前後年級的差異都屬合理，並未如百分比的資料般顯得獨特。本文從以下四點討論研究發現。

首先，CDM 結果如預期地得出三種類型的學生，且 IRT 估計所得的能力值也如預期地困難者最低、單一表徵者居次、多元表徵者最高，而且測驗的信效度為可接受，代表文字題列式判斷作業可作為代數思維的一種評估方式。文字題列式判斷提供學生連結問題情境與各種可互換之合理列式的經驗，這與文獻上多半採用沒有情境的等式填空或對錯判斷作業（Kieran, 2004a; McNeil et al., 2019; Molina & Ambrose, 2008）不同，它們是從數字運算的事實來展現對等價性的理解。不同型態題型可以提供不同的學習經驗，有助於學生對等價性意義的掌握。

第二，三至六年級多元表徵者占 20%~64%（表 6）的結果，顯示國小階段已有部分學生掌握一道單步驟或多步驟文字題可有多種不同的列式解法，展現了掌握等價性的相等與可互換性的特徵。也就是三、四年級部分學生已能正確判斷改變 6「美美」題的列式可採 $8+5=\square$ ，也可以用 $\square-5=8$ ，而五、六年級部分學生更能在已知面積與長、求算寬的文字題中，瞭解採 $310.8\div 21=\square$ 或 $310.8\div \square=21$ 均可，此結果符應 Blanton 等人（2015）、Radford（2018）、Carey（1991）的主張，國小即可培養學生代數能力，而判斷不同列式具備相等與可互換之功能的能力，也與我國數學課程在低年級已有加減互逆的學習內容、中年級則進一步學習乘除互逆的課程安排相呼應。

相對地，七、八年級單一表徵者和多元表徵者的比率都在 41%~46%（表 6），顯示雖然在國中大量代數課程中，在面對多步驟文字題和包含未知數的列式下，七、八年級學生

有四成只能判斷一種正確列式方式，另外四成則能判斷兩種正確的列式方式。由於本研究不同年級受試者接受的是經過等化的不同題本，題目內容包括晚近學習的內容，對學生而言，這類晚近學習的內容較為困難，尤其七、八年級的試題相較於其它年級多了代數符號的文字題與算式，其難度跨幅較大；而五、六年級的試題雖然也較三、四年級困難，但試題內容都同以數字或□呈現，使得三、四年級與五、六年級間的難度跨幅較小。因此七、八年級學生被歸為多元表徵者的比率未如預期，也屬合理現象。

第三，三至八年級困難者能力值差異小，僅分成三年級與四至八年級兩種能力層次，相較於分成三種與四種能力層次的單一表徵者和多元表徵者顯得成長遲滯，代表四至八年級有一群學生在文字題的表徵轉換能力上幾乎沒有學習；且困難者在六至八年級都維持大約一成的學生比率，並未進一步下降。這點就文字題解題教學而言，是一項警訊。

最後，七、八年級多元表徵者的比率顯著較六年級的少（表 6），但其能力值卻顯著較六年級的高（表 7），與第二點類似地，突顯了等價能力會隨等式涉及的數學內涵而呈多層次。如同 Fyfe 等人（2022）與 McNeil 等人（2019）所言，掌握等價性不僅有概念的成分，必須理解等號代表左右兩邊的數學物件相等，同時有技能的成分，必須要能編碼、辨識、操弄左右兩邊的數學物件。以第二點的兩道題目為例，面對改變 6 的文字題能表現出掌握等價性的能力較為容易，但在面積題要表現出能力則困難得多。由於六年級和七、八年級接受的是垂直等化後的不同題本，若單純從 CDM 分類後的類型百分比，會發現六年級在面對的文字題上能展現等價能力的比率較七、八年級的高，但六年級的試題難度相對於七、八年級的low，因此 IRT 估計出來的六年級多元表徵者的能力值就低於七、八年級。這顯現 CDM 與 IRT 可相輔相成，讓接受符應其學習內容之不同試題的不同年級學生，其認知發展能被適切地描述。

總結上述四點討論，顯示從 CDM 與 IRT 兩種資料面向可描繪長期發展之能力的類型與變化。CDM 能診斷學生特定能力是否精熟，但受限於不同年級學生能力與課程內容差異，難以用相同的試題進行跨年級測驗，因此研究大多以特定學科內容或年級為單位進行分析（Choi et al., 2015; Xu et al., 2023），鮮少跨年級甚至跨教育階段的調查，較難描繪一個長期發展的能力。IRT 則透過共同題或安排重複考生的方式進行等化，使不同測驗複本、不同年級學生數據可在同一個尺度中進行比較。雖然有部分實徵研究結合 IRT 與 CDM 的訊息，但多為單一年級的研究，很少有跨年級的調查（吳慧珉等人，2012; Abidin & Retnawati, 2019）。本研究採用具情境的文字題解題算式判斷作業探討橫跨三至八年級學生的表現差異，不僅可行，且提供了不同於無情境等式填空的作業。本研究三年級部分學生展現了等價能力，也支持國小中低年級就可及早提供代數學習機會的主張。

然而，本研究有以下兩點研究限制。一是多重是非題的缺點。多重是非題選項設計需考量誘答與多樣性，當題項出現明顯為互斥敘述時，例如在「陳爺爺」題組中，其中 $53 + 28 = \square$ 與 $53 - 28 = \square$ 兩個選項為互斥的敘述，可能會成為學生判斷對錯的線索。另外，多重是非題高猜測度與高疏忽參數亦為其缺點，低能力學生容易因猜對答案而被高估；題組內三個小題必須全部判斷正確才視為答對，高能力學生容易因疏忽而被低估。即使作

業有猜測度、疏忽參數偏高的情況，然而本研究的工具信效度及結果大致與預期相符，顯示仍具一定的可信度。二是未直接檢驗學生對同一文字題兩個正確的解題列式之等價性的判斷。本研究多元表徵者可能分別判斷三小題的列式能否符合題意與解題的需求，但他們並未外顯地表示兩個正確的解題列式是相等的。因此，本研究並不能確定，多元表徵者能順利從一個正確列式轉譯成另一個正確列式。

此外，本研究也衍生出兩個未來研究建議。一是運用整合性更適配的 CDM 與 IRT 模式探討學生的數學能力發展。本研究使用 Rasch 模型以一維度的連續潛在變量來估計學生整合表徵轉換與等價能力的解題算式判斷能力，而以 DINA 模型將學生在文字題解題判斷能力反應型態分為表徵轉換與等價能力兩個具階層關係的認知屬性，而非獨立的二維度的能力。目前 CDM 模式已有可處理認知屬性的非獨立性 (Köhn & Chiu, 2019; Leighton et al., 2004) 的模式，亦即有設定能力間存在階層關係之模型可以處理如同本文表徵轉換與等價能力之階層關係，可在處理資料時直接排除(0, 1)類型，亦即本研究採用這類處理階層關係的 CDM 模式會更能搭配一維的 IRT。由於本研究為結合 CDM 與 IRT 處理實徵資料的初探研究，未考量高階的 CDM 與 IRT 模式，例如以多向度 IRT (multidimension item response theory [MIRT]) 來搭配 CDM 的多個認知屬性，甚至採用題內多向度測驗 (within-item multidimensional test) (吳宜玲等人，2021) 估計出具階層性的多向度能力以搭配前述階層性 CDM (Köhn & Chiu, 2019; Leighton et al., 2004)，是未來可以嘗試的方向。近年來有越來越多測驗模型企圖整合潛在連續變項與潛在類別變項，例如整合 Q 矩陣與 IRT (Tseng & Wang, 2021)，或是整合 DINA 與 IRT (Hong et al., 2015; Hsu & Wang, 2015)，這些都將是探究學生數學認知能力發展的利器。二是代數思維的多面向不只能分為表徵轉換與等價能力，也可以有其它不同理論觀點的分法，例如前文提到國中階段代數符號相關的等價能力就和數字運算的等價能力有相當大的落差，而加減的等價性也不同于乘除的等價能力，後續研究者可針對不同的理論架構與研究目的劃分代數思維的面向，並配合評量架構採用如 MIRT、高階 Q 矩陣等更適切的統計模型，以了解代數思維不同能力的意涵與發展。

誌謝

本研究承蒙科技部專題研究計畫 (MOST 105-2511-S-003-039-MY3) 經費支持，並感謝研究助理鄭鈴華、研究生鄭俊彥在資料收集與分析上的協助，特此致謝。尤其感謝審查委員與編輯委員提供的意見，幫助本文更臻完善。

參考文獻

- 王文中(2004)。Rasch 測量理論與其在教育和心理之應用。《教育與心理研究》，27(4)，637–694。
[Wang, W. -C. (2004). Rasch measurement theory and application in education and psychology. *Journal of Education & Psychology*, 27(4), 637–694. (in Chinese)]
- 吳宜玲、楊心怡、吳昭容、陳柏熹(2021)。三至九年級學生數學運算能力等化測量與多向度分析。《清華教育學報》，38(2)，111–150。[Wu, Y.-L., Yung, H.-I., Wu, C.-J., & Chen, P.-H. (2021). Multidimensional test equating the arithmetical abilities of third to ninth grade students. *Tsing Hua Journal of Educational Research*, 38(2), 111–150. (in Chinese)]
[https://doi.org/10.6869/THJER.202112_38\(2\).0004](https://doi.org/10.6869/THJER.202112_38(2).0004)
- 吳慧珉、張育蓁、林宏昇(2012)。結合知識結構之 Q 矩陣設計於 DINA 模式之估計成效探究。《測驗統計年刊》，20(2)，1–30。[Wu, H.-M., Chang Y.-C., & Lin, H.-S. (2012). The research in DINA model with different Q matrix designs incorporating knowledge structures. *Journal of Research on Measurement and Statistics*, 20(2), 1–30. (in Chinese)]
<https://doi.org/10.6773/JRMS.201212.0001>
- 吳慧珉、鄭俊彥、施淑娟(2015)。認知診斷模式之理論與實務。《測驗學刊》，62(4)，303–328。
[Wu, H.-M., Cheng, C.-Y., & Shih, S.-C. (2015). Cognitive diagnostic model-theory and practice. *Psychological Testing*, 62(4), 303–328. (in Chinese)]
- 郭伯臣、王暄博(2008)。大型測驗中同時進行垂直與水平等化效果之探討。《教育研究與發展期刊》，4(4)，87–120。[Kuo, B.-C., & Wang, H.-P. (2008). A simultaneous vertical and horizontal equating of large-scale assessments. *Journal of Educational Research and Development*, 4(4), 87–120. (in Chinese)]
- 教育部(2018)。十二年國民基本教育課程綱要國民中小學暨普通型高級中等學校—數學領域。作者。[Ministry of Education. (2018). *Curriculum guidelines of 12-year basic education for elementary, junior high schools and general senior high schools – mathematics*. Author. (in Chinese)] <https://cirn.moe.edu.tw/Upload/file/27338/72246.pdf>
- 教育部統計處(2016a)。104 學年國民小學校別資料。[Department of Statistics, Ministry of Education. (2016a). *School-specific data for the 104th academic year in elementary schools*. (in Chinese)] https://stats.moe.gov.tw/files/detail/104/104_basec.xls
- 教育部統計處(2016b)。104 學年國民中學校別資料。[Department of Statistics, Ministry of Education. (2016b). *School-specific data for the 104th academic year in junior high schools*. (in Chinese)] https://stats.moe.gov.tw/files/detail/104/104_basej.xls
- Abidin, M., & Retnawati, H. (2019). A diagnosis of difficulties in answering questions of circle material on junior high school students. *Jurnal Penelitian dan Evaluasi Pendidikan*, 23(2), 144–155. <https://doi.org/10.21831/pep.v23i2.16454>
- Blanton, M., Stephens, A., Knuth, E., Gardiner, A. M., Isler, I., & Kim, J. S. (2015). The development of children's algebraic thinking: The impact of a comprehensive early algebra intervention in third grade. *Journal for Research in Mathematics Education*, 46(1), 39–87. <https://doi.org/10.5951/jresmetheduc.46.1.0039>
- Byrd, C. E., McNeil, N. M., Chesney, D. L., & Matthews, P. G. (2015). A specific misconception of the equal sign acts as a barrier to children's learning of early algebra. *Learning and Individual Differences*, 38, 61–67. <https://doi.org/10.1016/j.lindif.2015.01.001>
- Cai, J., & Knuth, E. J. (2005). Introduction: The development of students' algebraic thinking in earlier grades from curricular, instructional and learning perspectives. *ZDM—Mathematics Education*, 37(1), 1–4. <https://doi.org/10.1007/BF02655891>

- Carey, D. A. (1991). Number sentences: Linking addition and subtraction word problems and symbols. *Journal for Research in Mathematics Education*, 22(4), 266–280. <https://doi.org/10.5951/jresmetheduc.22.4.0266>
- Chesney, D. L., & McNeil, N. M. (2014). Activation of operational thinking during arithmetic practice hinders learning and transfer. *The Journal of Problem Solving*, 7(1), 24–35. <https://doi.org/10.7771/1932-6246.1165>
- Choi, K. M., Lee, Y. S., & Park, Y. S. (2015). What CDM can tell about what students have learned: An analysis of TIMSS eighth grade mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(6), 1563–1577. <https://doi.org/10.12973/eurasia.2015.1421a>
- College Board. (2000). *Equity 2000: A systematic education reform model*. Author.
- Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, 20(4), 405–438. [https://doi.org/10.1016/0010-0285\(88\)90011-4](https://doi.org/10.1016/0010-0285(88)90011-4)
- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76(2), 179–199. <https://doi.org/10.1007/s11336-011-9207-7>
- DeMars, C. E. (April 21-25, 2003). *Recovery of graded response and partial credit parameters in MULTILOG and PARSCALE* [Paper presentation]. American Educational Research Association Annual Meeting 2003, Chicago, IL. <https://www.learntechlib.org/p/94949/>
- Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming high and falling short: California's eighth-grade algebra-for-all effort. *Educational Evaluation and Policy Analysis*, 37(3), 275–295. <https://doi.org/10.3102/0162373714543685>
- Fyfe, E. R., Byers, C., & Nelson, L. J. (2022). The benefits of a metacognitive lesson on children's understanding of mathematical equivalence, arithmetic, and place value. *Journal of Educational Psychology*, 114(6), 1292–1306. <https://doi.org/10.1037/edu0000715>
- Greer, B. (1992). Multiplication and division as models of situations. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 276–295). Macmillan Publishing Co, Inc.
- Hanson, B. A., & Béguin, A. A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common item equating design. *Applied Psychological Measurement*, 26(1), 3–24. <https://doi.org/10.1177/0146621602026001001>
- Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, 87(1), 18–32. <https://doi.org/10.1037/0022-0663.87.1.18>
- Hodgen, J., Oldenburg, R., & Strømskag, H. (2018). Algebraic thinking. In T. Dreyfus, M. Artigue, D. Potari, S. Prediger, & K. Ruthven (Eds.), *Developing research in mathematics education: Twenty years of communication, cooperation and collaboration in Europe* (pp. 32–45). Routledge. <https://www.educ.cam.ac.uk/people/staff/ruthven/9781351625418preview.pdf>
- Holland, P.W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 187–220). Praeger.
- Hong, H., Wang, C., Lim, Y. S., & Douglas, J. (2015). Efficient models for cognitive diagnosis with continuous and mixed-type latent variables. *Applied Psychological Measurement*, 39(1), 31–43. <https://doi.org/10.1177/0146621614524981>

- Hornburg, C. B., Devlin, B. L., & McNeil, N. M. (2022). Earlier understanding of mathematical equivalence in elementary school predicts greater algebra readiness in middle school. *Journal of Educational Psychology*, 114(3), 540–559. <https://doi.org/10.1037/edu0000683>
- Hsu, C. L., & Wang, W. C. (2015). Variable-length computerized adaptive testing using the higher order DINA model. *Journal of Educational Measurement*, 52(2), 125–143. <https://doi.org/10.1111/jedm.12069>
- Kaput, J. J. (1998). Transforming algebra from an engine of inequity to an engine of mathematical power by “algebrafying” the K-12 curriculum. In National Research Council (Ed.), *The nature and role of algebra in the K-14 curriculum: Proceedings of a national symposium* (pp. 25–26). National Academy Press. <https://files.eric.ed.gov/fulltext/ED429801.pdf>
- Kaput, J. J. (2008). What is algebra? What is algebraic reasoning? In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 5–17). Lawrence Erlbaum Associates.
- Kieran, C. (1996). The changing face of school algebra. In B. Hodgson, C. Alsina Català, J. M. Álvarez Falcón, C. Laborde, & A. J. Pérez Jiménez (Eds.), *8th international congress on mathematical education: Selected lectures* (pp. 271–290). S. A. E. M. “Thales”.
- Kieran, C. (2004a). Algebraic thinking in the early grades: What is it. *The Mathematics Educator*, 8(1), 139–151.
- Kieran, C. (2004b). The core of algebra: Reflections on its main activities. In K. Stacey, H. Chick & M. Kendal (Eds.), *The future of the teaching and learning of algebra: The 12th ICMI study* (pp. 21–33). Kluwer Academic.
- Kieran, C. (2022). The multi-dimensionality of early algebraic thinking: Background, overarching dimensions, and new directions. *ZDM—Mathematics Education*, 54(6), 1131–1150. <https://doi.org/10.1007/s11858-022-01435-6>
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academy Press. <https://nap.nationalacademies.org/read/9822/chapter/1>
- Knuth, E. J., Alibali, M. W., Hattikudur, S., McNeil, N. M., & Stephens, A. C. (2008). The importance of equal sign understanding in the middle grades. *Mathematics Teaching in the Middle School*, 13(9), 514–519. <https://doi.org/10.5951/MTMS.13.9.0514>
- Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equivalence & variable. *ZDM—Mathematics Education*, 37(1), 68–76. <https://doi.org/10.1007/BF02655899>
- Köhn, H. F., & Chiu, C. Y. (2019). Attribute hierarchy models in cognitive diagnosis: Identifiability of the latent attribute space and conditions for completeness of the Q-matrix. *Journal of Classification*, 36(3), 541–565. <https://doi.org/10.1007/S00357-018-9278-6>
- Krawec, J. L. (2014). Problem representation and mathematical problem solving of students of varying math ability. *Journal of Learning Disabilities*, 47(2), 103–115. <https://doi.org/10.1177/0022219412436976>
- Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoaka's rule-space approach. *Journal of Educational Measurement*, 41(3), 205–237. <https://doi.org/10.1111/j.1745-3984.2004.tb01163.x>

- Livingston, S. A. (2004). *Equating test scores (without IRT)*. Educational Testing Service. <https://www.ets.org/Media/Research/pdf/LIVINGSTON.pdf>
- Matthews, P. G., & Fuchs, L. S. (2020). Keys to the gate? Equal sign knowledge at second grade predicts fourth-grade algebra competence. *Child Development*, 91(1), e14–e28. <https://doi.org/10.1111/cdev.13144>
- Mayer, R. E. (1992). *Thinking, problem solving, cognition*. W. H. Freeman.
- McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology*, 43(3), 687–695. <https://doi.org/10.1037/0012-1649.43.3.687>
- McNeil, N. M. (2014). A change-resistance account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives*, 8(1), 42–47. <https://doi.org/10.1111/cdep.12062>
- McNeil, N. M., Hornburg, C. B., Devlin, B. L., Carrazza, C., & McKeever, M. O. (2019). Consequences of individual differences in children's formal understanding of mathematical equivalence. *Child Development*, 90(3), 940–956. <https://doi.org/10.1111/cdev.12948>
- Molina, M., & Ambrose, R. (2008). From an operational to a relational conception of the equal sign. Thirds graders' developing algebraic thinking. *Focus on Learning Problems in Mathematics*, 30(1), 61–80. <http://hdl.handle.net/10481/4721>
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Organisation for Economic Cooperation and Development. (2018). *PISA 2022 mathematics framework (Draft)*. OECD Publishing. <https://pisa2022-maths.oecd.org/files/PISA%202022%20Mathematics%20Framework%20Draft.pdf>
- Passolunghi, M. C., & Pazzaglia, F. (2005). A comparison of updating processes in children good or poor in arithmetic word problem-solving. *Learning and individual differences*, 15(4), 257–269. <https://doi.org/10.1016/j.lindif.2005.03.001>
- Prediger, S. (2010). How to develop mathematics-for-teaching and for understanding: The case of meanings of the equal sign. *Journal of Mathematics Teacher Education*, 13(1), 73–93. <https://doi.org/10.1007/s10857-009-9119-y>
- Radford, L. (1996). Some reflections on teaching algebra through generalization. In N. Bernarz, C. Kieran, & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching* (pp. 107–111). Kluwer Academic Publishers.
- Radford, L. (2018). The emergence of symbolic algebraic thinking in primary school. In C. Kieran (Ed.), *Teaching and learning algebraic thinking with 5- to 12-year-olds: The global evolution of an emerging field of research and practice* (pp. 3–25). Springer.
- Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153–196). Academic Press.
- Sibgatullin, I. R., Korzhuev, A. V., Khairullina, E. R., Sadykova, A. R., Baturina, R. V., & Chazova, V. (2022). A systematic review on algebraic thinking in education. *EURASIA Journal of Mathematics, Science and Technology Education*, 18(1), 1–15. <https://doi.org/10.29333/ejmste/11486>
- Tseng, M. C., & Wang, W. C. (2021). The Q-matrix anchored mixture Rasch model. *Frontiers in Psychology*, 12, 564976. <https://doi.org/10.3389/fpsyg.2021.564976>

- Wardat, Y., Jarrah, A. M., & Stoica, G. (2021). Understanding the meaning of the equal sign: A case study of middle school students in the United Arab Emirates. *European Journal of Educational Research*, 10(3), 1505–1514. <https://doi.org/10.12973/eu-jer.10.3.1505>
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370–371. <https://www.rasch.org/rmt/rmt83b.htm>
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1998). *ACER ConQuest: Generalised item response modelling software*. ACER Press.
- Xu, T., Wu, X., Sun, S., & Kong, Q. (2023). Cognitive diagnostic analysis of students' mathematical competency based on DINA model. *Psychology in the Schools*, 60(9), 3135–3150. <https://doi.org/10.1002/pits.22916>