

國立臺灣師範大學教育心理與輔導學系
教育心理學報，民 98，40 卷，4 期，683—706 頁

國小加減法數學文字題歷程導向解題 診斷評量題組之編製發展與功能分析研究*

許 家 驊

國立嘉義大學
教育學系

本研究旨在參照數學解題歷程模式進行成份（認知、後設認知）分析，並搭配學習內容分析，編製歷程導向單步驟加減法文字題解題診斷評量題組。採合目標與叢集取樣抽取嘉義縣市一年級學生共 477 人，施測後進行統計分析。在項目分析方面，平均表現水準約 75%，單題、各成份及全題組的內部相關與極端組 t 考驗結果良好。在效度分析方面，主成份、試探性及驗證性因素分析均發現題組的二成份三因子結構（認知成份含基本解題及數量辨識技巧、後設認知成份含檢核技巧），且共同向度性、因素萃取、抽取變異比及模式適配度均屬良好，並具良好之效標關聯及預測效度。在信度分析方面，內部一致性係數及組合信度亦屬良好。而在個案剖面訊息方面，亦具診斷處方功能。

關鍵詞：後設認知、測驗信效度、解題評量、解題歷程、解題學習診斷處方

就認知心理學訊息處理（information processing）取向的觀點而言，若能針對個體學習所涉之心智運思歷程進行精細的成份分析（componential analysis）、測定及診斷，將有利於個體學習歷程及能力狀態之深入了解（Snow & Lohman, 1993）。準此，即使在數學解題特定學習領域（domain specific），若能依前述理念編製歷程導向（process-oriented）解題能力診斷評量工具，亦應有助於個體解題心智運思歷程成份之精細診斷。

而就解題歷程模式成份要素而言，除 Polya（1957）最早提出「了解、計劃、執行（execution）、回顧（reviewing）」之四階段基礎歷程模式外，其他研究者亦曾分別提出不同歷程成份內涵之數學文字題解題階段或策略使用模式，如 Krulik 與 Rudnick（1989）的「閱讀及了解問題、探索問題（exploration）、選擇策略、執行策略、回顧驗證（verifying）」五項策略技能模式。Garofalo 與 Lester（1985）的「導向（orientation）、組織、執行、驗證」四階段認知與後設認知整合解題模式。Montague、Warger 與 Morgan（2000）的「閱讀理解（read-comprehension）、轉述轉譯（paraphrase-translation）、具象化轉換（visualization-transformation）、假設（hypothesize）計劃、估

* 本論文係寫自行政院國家科學委員會補助之九十六年度專題研究中，有關先期工具發展之部份成果報告內容（編號：NSC 96-2413-H-415-001），謹此誌謝。本文文責由研究者自負。

計預測、運算計算、檢核評鑑 (check-evaluation) – 自我教導 (self-instruction)、自我質問 (self-questioning)、自我監控 (self-monitoring) 』認知與後設認知聯合解題策略模式。Schurter (2002) 的「了解、計劃、執行、回顧、自我質問」五階段結合認知與認知監控之解題教學模式。Desoete 與 Roeyers (2005) 的「語言理解、心智表徵 (representation)、脈絡 (context) 訊息、選擇適用訊息、數量訊息閱讀與產出、運算符號 (operation symbol) 閱讀與產出、程序性 (procedural) 計算、數字系統知識、數量感 (number sense) 』九項數學解題認知成份技能模式。

上述模式，除 Desoete 與 Roeyers (2005) 解題認知成份技能模式外，餘皆屬解題歷程模式，且均涉及 Hutchinson (1992) 所提認知及後設認知 (metacognitive) 兩大解題成份之運作，其中部份研究結果顯示，不論是認知或後設認知成份均對個體解題表現具有正向影響 (Desoete & Roeyers, 2005; Montague, Warger & Morgan, 2000; Schurter, 2002)。雖然後續學者所提歷程模式之成份要素分化較為精細，但其均源於 Polya (1957) 的基礎解題歷程模式，且可相容於此一模式中。因此若以 Hutchinson (1992) 所提之成份分析「精鍊」(parsimony) 原則觀之，此模式同時具有簡約及包含性，適合作為發展解題歷程模式之基礎架構。此外，就國民小學文字題解題教學實務而言，「讀題、理解、選擇方法、列算式、計算答案、檢查」乃為個體進行文字題解題學習所需之基本能力，而此些能力同樣可被包含在前述基礎架構內。

故若以 Polya (1957) 基礎解題歷程模式為架構，之後同時考量 Hutchinson (1992) 所提成份分析向度及「精鍊」原則、國小學生文字題解題學習所需能力來綜合歸納前述研究所提模式內容，在文字題解題認知成份方面，大致有閱讀理解、表徵轉譯、策略計劃、執行四項，其中表徵轉譯的表徵部份，在性質上與理解較有關聯，應可併入閱讀理解中，而轉譯部份在性質上則與計劃較有關聯，可併入策略計劃中，而執行又可分為列式 (enumeration) 及計算兩小項。在文字題解題後設認知成份方面，亦有回顧、驗證、覺察 (awareness) 監控三項，均與檢核運作有關，但其中回顧在性質上即可含括驗證及覺察監控兩項在內。如此分類歸納方式不僅符合前述解題歷程模式、成份分析向度之假定，且與國小文字題解題教學實務要求個體具備之各項解題能力相互對應，將有助於其解題學習歷程能力要素之診斷。

基於上述分析，本研究所謂解題心智運思歷程係指個體在從事數學文字題解題時所須歷經之閱讀理解、計劃、執行、回顧四大階段性要素運用過程，而成份分析則指分析此歷程中所須涉及之要素作為解題心智運思歷程成份指標。其中閱讀理解階段依 Gagne'、Yekovich 與 Yekovich (1993) 觀點，可再細分為字義 (literal) 理解及推論 (inferential) 理解兩項成份，前者係指字義觸接 (word accessing) 活動、後者則與脈絡意義理解有關，計劃階段包含策略選擇一項成份，執行階段包含列式及計算兩項成份，回顧階段包含檢核行動及檢核方式兩項成份，共四大類七小項成份，此將作為診斷評量題組編製的共同成份架構。

至於題組使用之題型內容，在現行一綱多本的實況下，即使課程綱要中同一年段的能力指標相同，但因各校使用版本不同，實際上各題型出現形式及教學時間點仍可能均不相同，為方便統一控制題目類型結構及難度並測試成份診斷功能，本研究將先參照 Fuson (1992)、Verschaffel、Greer 與 DeCorte (2007) 對單步驟 (single step) 加減法數學文字題型的語意基模 (semantic schema) 分類觀點，之後分析彙整現行不同主要版本 (康軒、南一、翰林及部編) 在同一教學期間 (一年級上下學期) 之數學加減法文字題學習內容及其所涉運算量數訊息類型後，再據之進行題組編製。經以前述方式分析市場佔有率較大之康軒 (楊瑞智等編，民95，民96)、南一 (陳冒海等編，民95；張英傑等編，民96)、翰林 (許瑛珍等編，民95，民96)、國立教育研究院籌備處 (鄭國順等編，民95，民96) 出版之一年級數學課本第一二冊相關單元內容，之後再針對所涉運算量數訊息進行解析後，可得到幾類型學習活動，如表1所列：

表1 各主要版本單步驟文字題型及運算量數類型交叉分析表

		版本 學期	康軒		南一		翰林		部編		備註
			上	下	上	下	上	下	上	下	
題型語意 基模	合併型	整體	+	*	*	*	*	*	*	*	*
		部份	-								*
	結果量未知	增加	+	*		*	*	*	*	*	*
		減少	-	*	*	*	*	*	*	*	*
	改變型	改變量未知	增加	+					*		
			減少	-					*		*
	添加	增加	+						*		
		取走	減少	-							*
	比較型	比多	-	*	*		*	*		*	*
		比少	-					*	*	*	*
	運算 類型	個位數對個位數	不進位加	+	*		*	*	*	*	*
			不借位減	-	*		*	*	*	*	*
二位數對個位數		進位加	+		*(>10)	*	*	*	*	*	*
		不進位加	+		*	*	*	*	*	*	*
		不借位減	-			*	*	*	*	*	*
		進位加	+			*	*	*	*	*	*
二位數對二位數		借位減	-		*	*	*	*	*	*	*
		不進位加	+			*	*	*	*	*	*
		不借位減	-	*		*	*	*	*	*	*
		進位加	+			*	*	*	*	*	*
		借位減	-			*	*	*	*	*	*
量數大小			<10	<20	<10	<50	<20	<70	<20	<50	
算式			橫	橫	橫	橫	橫	橫	橫	橫直	

註：*代表該版本已含該項內容

數字部份為版本所涉運算數量

等式或大小於符號部份為版本運算數量範圍界定

自表1中可發現，各版本在上下學期共同題型基模方面，分別為合併型（combine）—求整體量（whole），改變型（change）結果量未知（result unknown）—增加及減少，比較型（compare）—比多（more than）三型四類。在共同運算量數大小方面，上學期康軒及南一版為10以內，翰林及部編版為20以內，下學期除翰林版為70以內外，其餘版本皆為50以內。在上下學期共同運算量數類型方面，分別為個位數對個位數進位加、不進位加、不借位減法（一下時四版本使用者皆已學過），二位數對個位數不進位加、借位減法（一下時四版本使用者皆已學過），二位數對二位數不借位減法（一下時四版本使用者皆已學過，不進位加法除康軒版外、餘版使用者皆已學過）四類，此將作為診斷評量題組編製的共同內容基礎。

職此，本研究旨在參照個體數學解題歷程模式進行成份（認知、後設認知）分析，爾後搭配學習內容分析結果，希望據之發展良好適用之歷程導向解題診斷評量題組。依前述目的，本研究之具體問題可細列如下：

- （一）歷程導向解題診斷評量題組在解題成份測量上之效度為何？
- （二）歷程導向解題診斷評量題組在解題成份測量上之信度為何？
- （三）歷程導向解題診斷評量題組在解題成份測量上之鑑別力（全題組及各成份）為何？
- （四）歷程導向解題診斷評量題組在個案解題學習表現上之診斷處方功能為何？

研究方法

一、研究對象

由於目前九年一貫課程採一綱多本方式實施，為便於統一教學版本、協調單元教學進度及顧慮研究對象在校的數學學習經驗起見，故採合目標與叢集取樣 (purposeful and cluster sampling) 方式於使用適當版本之嘉義縣市國小普通班一年級學生中，以學校為叢集單位選取所需樣本。而依 Crocker 與 Algina (1986) 所提「每項目至少需五人」之標準計算，歷程導向解題診斷評量題組預定編製八大題，但每大題有九個單題，全題組共七十二個計分項目，每項目乘上五人，至少約需360人。再者本研究將運用結構方程模式 (structural equation model, SEM) 進行驗證性因素分析，綜合 Joreskog 與 Sorbom (1993)、Hair、Anderson、Tatham 與 Black (1998) 的建議，模式中每一個估計的自由參數至少要有五至十人以上的受試者，以降低抽樣誤差及適配度不佳的機率，且模式所得參數估計值亦將較為穩定。在本研究圖2模式自由估計參數數目25的狀況下，至少需有125人、但最好能在250人以上的受試者。

最後實際取樣人數，嘉義縣抽取四校八班 (康軒兩校四班116人、翰林及南一版各一校二班112人) 合計228人，嘉義市抽取四校八班 (康軒版兩校四班125人、部編及南一版各一校二班124人) 合計249人，全部總計477人，符合前述測驗項目編製及 SEM 統計分析之人數標準。此外，本研究取樣方式，在區域學校的整體代表性或許有所不足，但就使用同版本的學校而言，應仍具相當代表性。

二、研究架構

本研究分為題組發展及功能分析兩部份進行，前者將歷經整合解題理論及實徵研究之文獻探析、解題成份分析、學習內容分析及編製題組四個次階段，後者則包含項目 (難度及鑑別度) 分析、測驗 (信效度) 分析、個案診斷處方分析三個次階段，其架構及內容如圖1所示。

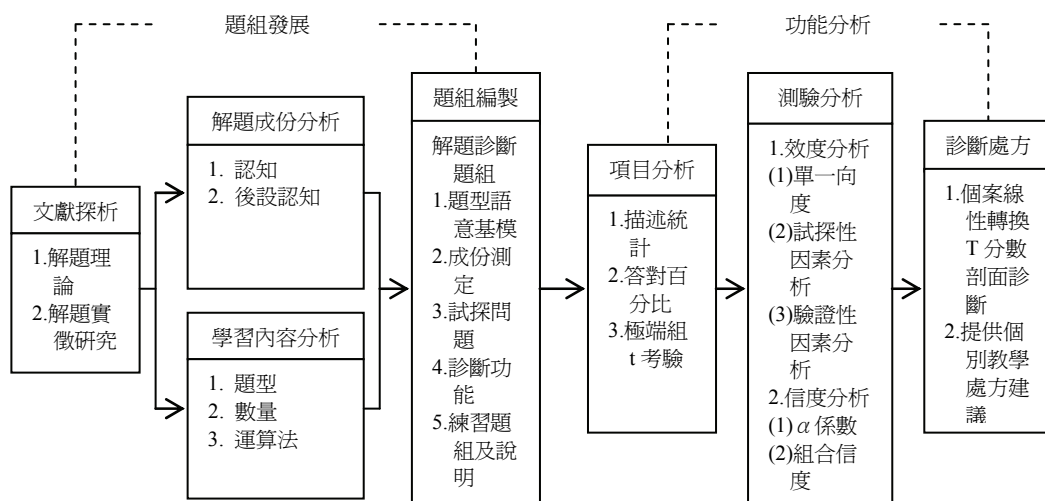


圖1 歷程導向解題診斷評量題組發展及功能分析研究設計架構

三、研究工具

(一) 單步驟加減法數學文字題歷程導向解題診斷評量題組

先針對市場佔有率較大之國小康軒版、南一版、翰林版及部編版一年級數學課本第一二冊相關單元進行內容分析(楊瑞智等編,民95,民96;陳冒海等編,民95;張英傑等編,民96;許瑛珍等編,民95,民96;鄭國順等編,民95,民96),歸納各版本共同題型後,再參考 Fuson(1992)、Verschaffel、Greer 與 DeCorte(2007)對單步驟加減法數學文字題型語意基模的分類,來分析其相關題型及計算難度(如前述表1)。之後搭配緒論所提成份分析結果,依序編製各題之成份試探單題。

題組內容包含合併型—求整體量、改變型—結果量未知—增加或減少、比較型—比多三型四類。其中合併型之語意結構包含兩個前提句及一個問句,前者為部份量(part)訊息,後者為求整體量之訊息。結果量未知改變型亦包含兩個前提句及一個問句,但其中一個前提句具有隱含改變關係之訊息。而比較型亦包含兩個前提句及一個隱含比較關係訊息之問句。計算難度分為十以內個位數對個位數進退位加減、二十以內二位數對個位數不進退位加減兩型,其編製細目向度如表2所示:

表2 歷程導向解題診斷評量題組編製細目表

計算難度	合併型	改變型(結果量未知)		比較型
	求整體量	增加	減少	比大
十以內個位數對個位數進位加	2	2	0	0
二十以內二位數對個位數借位減	0	0	2	2

註:依表1分析結果,以上兩種計算難度均在康軒、南一、翰林及部編版一年級下學期教學範圍內

任一題均細分為閱讀理解—含字義理解及推論理解,計劃—策略選擇,執行—含列式及計算,回顧—含檢核行動及檢核方法等四類七項成份,其中推論理解可再針對兩個前提句及一個問句進行試探。整合以上,每題所含之成份名稱、試探問題及診斷功能編製架構可細列如表3所示:

表3 歷程導向解題診斷階層設計型式及功能

向度	類別	成份	試探問題	診斷功能
認知	閱讀理解	字義理解	這一題的題目是?	是否有識題障礙?
		推論理解	根據兩個前提句分別提問? 題目要我們算什麼?	是否分別理解各前提句意義? 是否理解問句意義?
	計劃 執行	策略選擇	你覺得這一題要怎麼做才可以得到答案?	是否了解如何達成目標(運算法)?
		列式 計算	這一題的算式要怎麼寫? 這一題的答案是多少?	是否具有正確列式(橫式)能力? 是否具有正確計算能力?
後設 認知	回顧	檢核行動	寫完題目後你會再檢查嗎?	是否具有驗算檢核習慣?
		檢核方式	你是怎麼檢查的?	是否具有驗算檢核能力?

今以單步驟合併型—整體(實施時使用注音版)為例,依前述架構進行編製,其內容形式如表4所示:

經徵詢班級教師建議後,發現學生已有看注音讀題自選能力,為不擾亂其答題過程起見,故採自行選答方式實施。同時為降低猜測答對的可能性,每一單題正確選項答案均作隨機排列處理。

由於各成份單題無法獨立於完整解題歷程外被蒐集,故須搭配文字題型之解題歷程採一大題含九個單題的題組方式實施,但在進行分析時,則可再依成份屬性將各大題內之單題表現重新加以組合。依表2題型及數量觀之,全題組將含八個大題、七十二個單題,共分九個成份分題組,每個分題組內共有八個相同成份單題。

表4 歷程導向解題診斷評量題組及練習題組內容示例(含練習題組說明)

成份	試探問題	選項	診斷功能	練習題組說明內容
*題型：單步驟合併型—整體(實施時使用注音版)				
蛋糕店下午賣了6個蛋糕(前提一)，晚上又賣了8個蛋糕(前提二)，一共賣出幾個蛋糕(問句)?				
字義理解	這一題的題目是?	(1) 蛋糕店下午賣了6個蛋糕，晚上又賣了8個蛋糕，一共賣出幾個蛋糕? (2) 書店下午賣了6個書套，晚上又賣了8個書套，一共賣出幾個書套? (3) 蛋糕店上午賣了6個蛋糕，下午又賣了8個蛋糕，一共賣出幾個蛋糕?	是否有識題障礙?	小朋友!因為這一題的題目是蛋糕店下午賣了6個蛋糕，晚上又賣了8個蛋糕，一共賣出幾個蛋糕?所以要選1，在空格上寫1。
	蛋糕店下午賣了幾個蛋糕?	(1) 6 (2) 8	是否理解前提一意義?	小朋友!因為這一題的題目說蛋糕店下午賣了6個蛋糕，所以要選1，在空格上寫1。
	晚上又賣了幾個蛋糕?	(1) 6 (2) 8	是否理解前提二意義?	小朋友!因為這一題的題目說晚上又賣了8個蛋糕，所以要選2，在空格上寫2。
推論理解	題目要我們算什麼?	(1) 下午和晚上一共賣出幾個蛋糕? (2) 下午賣出幾個蛋糕? (3) 晚上賣出幾個蛋糕?	是否理解問句意義?	小朋友!因為這一題的題目問蛋糕店下午和晚上—共賣出幾個蛋糕?所以要選1，在空格上寫1。
	你覺得這一題要怎麼做才可以得到答案?	(1) 把下午賣的和晚上賣的蛋糕加一加 (2) 把晚上賣的和下午賣的蛋糕減一減	是否了解如何達成目標(運算法)?	小朋友!因為這一題的題目問蛋糕店下午和晚上—共賣出幾個蛋糕?所以要把下午和晚上賣出的蛋糕合起來算，要選1，在空格上寫1。
計劃	這一題的算式要怎麼寫?	(1) $6+8=$ (2) $8-6=$	是否具有正確列橫式能力?	小朋友!因為要把下午和晚上賣出的蛋糕合起來算，所以6和8要加起來，要選1，在空格上寫1。
計算	這一題的答案是多少?	(1) 14 (2) 2	是否具有正確計算能力?	小朋友!因為6和8要加起來的答案是14，所以要選1，在空格上寫1。
檢核行動	寫完題目後你會再檢查嗎?	(1) 有 (2) 沒有 (3) 一點點	是否具有驗算檢核習慣?	小朋友，寫完題目後你有沒再檢查呢?有就選1，沒有就選2，只有檢查一些就選3。
檢核方式	你是怎麼檢查的?	(1) 再唸唸題目、想想算法對不對?然後再算算答案看對不對? (2) 再唸唸題目、想想算法對不對? (3) 再唸唸題目、算算答案對不對? (4) 再想想算法對不對?然後算算答案看對不對? (5) 只有唸唸題目 (6) 只有想想算法對不對? (7) 只有算算答案對不對? (8) 其他	是否具有驗算檢核能力?	小朋友!如果你寫完題目後有檢查，那你是怎麼做檢查的呢?想想看再選一選、寫一寫?

註：以上為一大題、內含九個成份單題

至於題目的計分，除檢核行動及檢核方式分題組之單題為部份給分外，其餘七個成份分題組之單題均為二元計分，在認知成份題組方面，每大題共有七個認知成份計分點(單題項目)，當個體正

確回答一個試探問題，採計一分，累計加總後最高共7分。在後設認知成份題組方面，每大題共有兩個後設認知成份計分點（單題項目），第一個計分點採「有、一些、沒有」三類，以2、1、0部份計分，第二個計分點依檢核處之數量部份計分，題目、算式、答案此三項檢核處，涉及三者之組合得3分、涉及任二者之組合得2分、僅涉及一項得1分，累計加總後最高共5分，因此每一大題最高總分為12分。全題組共編製八大題（每個單一題型基模均編製兩大題，每大題各有九個不同成份單題），因此個人全題組整體表現最高總分為96分。此外若以成份而言，個體在接受全題組施測後，可得到九個分題組共九個成份分數，認知成份七個分題組每個最高八分共56分，後設認知成份題組第一個分題組最高16分、第二個分題組最高24分，故全題組九個分題組表現（成份分數）的最高總分亦與整體表現最高總分同為96分。分數的運用均以「分題組」或「全題組」總分為之。

因本工具係為目前數學課程單元學習內容，故在效度方面，除採課程本位、四位數學教育專家（兩位大學教授、兩位小學數學老師）審閱，同時兼顧題組題型結構基模同構性（isomorphic）之原則來支持工具的內容效度外，將再以個體在校數學學業考試成績為效標，以建立效標關聯（criterion-related）效度，並使用因素分析來提供構念本位（construct-related）或構念效度證據。在信度方面，將採內部一致性 α 係數及組合（composite）信度作為指標。在項目分析方面，將以答對百分比為難度指標、極端組 t 考驗為鑑別力指標。

（二）單步驟加減法數學文字題歷程導向解題診斷練習題組及說明

由於前項解題診斷題組係採歷程導向設計，其題組形式與一般自行解題題組不盡相同，因此將另行設計與其形式相同之練習題組及說明，提供個體在歷程導向解題診斷題組正式實施前作為練習之用。其形式及內容示例，如表4所示。

本練習題組將配合標準化文字及口述說明簡報實施（採 Microsoft powerpoint 製作，以單槍投影機播放），其內容包含共同作答說明、前七個單題注音文字視訊閃示及音訊口述說明（讀題引導及正確答案回饋錄音），而為免改變其原本檢核狀況或習慣，故有關後兩個檢核問題、不作練習教導，僅在前七個練習單題實施完畢時，於最後直接請學生按其狀況意見選答。另為降低猜測答對的可能性，練習題正確選項答案亦作隨機排列處理，其結果僅為了解個體答題狀況之用，不予計分。

四、實施程序

（一）題組編製及施測準備

1. 蒐集文字題解題相關文獻，進行內容歸納整理。
2. 自前所歸納之相關文獻內容進行文字題解題認知、後設認知成份分析選取。
3. 針對國小一年級學生的學習內容及難度進行題型、數量及運算法之分析（如表1內容）。
4. 以解題成份及學習內容分析結果為基礎，進行題組編製，其主要流程如下：
 - （1）編製目標：歷程診斷導向。
 - （2）編製細目表：依項目診斷功能、涉及成份、題數、運算法四大向度細列題組內容。
 - （3）依細目表進行初步項目及試探問題編擬。
 - （4）針對填答形式及程序，設計說明指導語。
 - （5）依細目表先由研究者及其學術同儕合力進行項目內容核對檢視後，再由四位數學教育專家（兩位大學教授、兩位小學數學老師）審閱。
 - （6）修正及編輯題組：

綜合上述審視結果，進行兩項調整，第一為使各單題選項方向一致起見，在第一及第二大題中，第一個單題第二個誘答選項之前提句用字及數量，本與正確選項不同，後來改為與正確選項相同，而僅修改問句或物項名稱，以便與後續其他單

題中選項的寫法方向一致。第二為符合各版本教學內容共同性，將計算類型由原先二位對二位不借位減法調整為二位對一位借位減法。

- 5.編製練習題組。
- 6.針對練習題組及施測說明製作標準化視訊與音訊錄音簡報。
- 7.針對練習及正式題組進行注音標示。

(二) 試作、正式施測及評分

1. 試作

於非取樣學校中，挑選三十位一年級學生進行試作，其程序說明如後。

(1) 填答練習

使用練習題組、單槍投影機配合標準化口述說明簡報實施，實施時先進行2分30秒鐘的共同作答說明（目的說明、填答事項及主題目讀題引導），之後按成份依序進行前七個單題，每一單題均先同時呈現注音文字視訊閃示及讀題引導音訊口述說明（約45~55秒鐘，依題目長短而定，46/46/45/52/55/51/45），之後自行選答待答20秒鐘，再呈現正確答案回饋說明之音訊口述說明（11~19秒鐘，依題目長短而定，19/13/12/14/24/19/11）。但為免改變其原本檢核狀況或習慣，故後兩個檢核問題、不作練習教導，僅在最後直接請小朋友按其狀況選答，全部耗費13分25秒。

(2) 自行作答

在填題練習結束後，再配合標準化施測簡報作正式答題前的共同說明33秒鐘（含注音文字視訊閃示及音訊口述說明），之後自行選答32分鐘，全部耗費32分33秒。結果發現無論在填答形式或程序上，實施狀況均甚良好。

2. 正式施測

依既定計劃及上述程序使用正式題組對所有研究對象實施。

3. 評分

研究者在助理協助下進行原始資料輸入建置，之後依前述正式題組所提原則，使用 SPSS 統計程式進行資料檔轉碼評分，各單題依其成份歸屬性質分別先加總為各分題組總分，之後再加總為全題組總分。

(三) 功能驗證

整合所得資料針對研究問題性質，進行項目、測驗信效度及個案診斷處方等分析。

五、資料處理與分析

首先在各項基本表現方面，採描述統計、平均數、標準差、答對百分比（如平均題組水準－題組所得分數除以題組最高得分）、皮爾遜積差相關（Pearson's product-moment correlation）、二系列（biserial）相關來說明受試的各項基本表現。次針對研究問題一，進行兩項分析，一為構念效度分析、二為效標關聯及預測效度分析，前者含共同向度之主成份分析（principal component analysis）、試探性及驗證性因素分析（exploratory and confirmatory factor analysis, EFA and CFA），後者含效標相關及迴歸預測分析。再針對研究問題二，進行兩項分析，一為傳統信度分析、二為組合信度分析。接著針對研究問題三，進行極端組 *t* 考驗分析。最後針對研究問題四，以個案之直線轉換（linear transformation）*T* 分數剖面訊息（profile）進行診斷及處方（diagnosis and prescription）分析。以上所涉統計分析，除驗證性因素分析係以 Amos 7.0 進行外，餘均採 SPSS 12.0 for Windows 系統為之。

結果與討論

一、各項描述統計分析

如表5至表9所示。

表5 全題組及分題組表現描述統計 (N=477)

統計數	總分	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
M	72.38	7.14	7.67	7.58	7.02	6.79	6.88	6.77	11.82	10.72
SD	13.23	1.41	0.87	1.08	1.50	1.63	1.58	1.69	5.27	6.20
P	.75	.89	.96	.95	.88	.85	.86	.85	.74	.45

註：P 為表現百分比

全題組最高總分 96 分

分題組最高總分前七個均為 8 分、第八個 16 分、第九個 24 分

若以標準參照評量的角度觀之，表5所示各項全題組及分題組表現數據，除第八、九分題組外，此次個體答題反應良好、難度偏易、大部份個體表現均達一定精熟水準以上（至少70%）。其中以推論理解一及二兩個分題組表現較佳且變異較小，但策略選擇、列式及計算三個分題組表現較弱且變異較大，字義理解及推論理解三兩個分題組表現居中，檢核行動及方式兩個分題組表現較差，代表個體檢核能力較弱，特別是在檢核方式上。

表6 各分題組總分與所屬單題表現之相關 (N=477)

單題	分題組總分								
一至八	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
相關	.47 至 .76	.32 至 .71	.40 至 .86	.39 至 .70	.36 至 .78	.40 至 .74	.33 至 .74	.954 至 .974	.978 至 .981

註：以上數值之概率值均為 .000

前七者數據為二系列相關，後二者數據為積差相關

自表6看來，各單題與各分題組相關良好，此亦顯示各單題與各分題組的測試能力間彼此具有良好關聯，各分題組與所屬單題間內部一致性良好。至於最後兩個分題組各項數據較大之原因可能為，此二分題組項目測量屬性為較穩定之解題檢核習慣及態度，而非較不穩定之解題認知表現，加上此二分題組均採部份給分而非二元計分，故各單題與分題組總分相關較高且一致。

表7 各分題組所屬單題表現與全題組總分之相關 (N=477)

單題	分題組									全題組
一至八	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式	總分
相關	.16 至 .48	.11* 至 .49	.19 至 .41	.15 至 .38	.31 至 .45	.28 至 .46	.29 至 .47	.65 至 .70	.62 至 .67	總分

註：以上數值之概率值，底線者為小於 .01、*者為小於 .05、餘均為 .000

前七者數據為二系列相關，後二者數據為積差相關

自表7看來，各分題組單題與全題組相關尚稱良好，此亦顯示各分題組單題與全題組測試能力間彼此具有良好關聯，全題組與各分題組單題間內部一致性良好。

表8 全題組總分與各分題組表現之相關 (N=477)

全題組	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
總分	.54	.55	.52	.55	.62	.61	.62	.69	.65

註：以上數值之概率值均為 .000

以上數據為積差相關

自表8看來，各分題組與全題組相關良好，此亦顯示各分題組單題與全題組測試能力間彼此具有良好關聯、內部一致性良好。

表9 各分題組表現間之相關 (N=477)

分題組	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
推論理解一	<u>.56</u>								
推論理解二	<u>.54</u>	<u>.86</u>							
推論理解三	<u>.64</u>	<u>.36</u>	<u>.35</u>						
策略選擇	<u>.61</u>	<u>.51</u>	<u>.49</u>	<u>.72</u>					
列式	<u>.66</u>	<u>.55</u>	<u>.53</u>	<u>.65</u>	<u>.83</u>				
計算	<u>.67</u>	<u>.57</u>	<u>.55</u>	<u>.63</u>	<u>.81</u>	<u>.92</u>			
檢核行動	<u>.15</u>	<u>.11*</u>	<u>.13</u>	<u>.14</u>	<u>.14</u>	<u>.12</u>	<u>.14</u>		
檢核方式	<u>-.04</u>	<u>.14</u>	<u>.08</u>	<u>.03</u>	<u>.04</u>	<u>.004</u>	<u>-.004</u>	<u>.40</u>	

註：以上數值之概率值，方框者為 .000、底線者為小於 .01、*者為小於 .05、餘大於 .05；以上數據為積差相關

自表9看來，前七個分題組表現間之相關界於.35至.92間，後兩個分題組表現間之相關為.4，且均達.000顯著水準，代表前七個分題組間及後兩個分題組間的各自內部相關良好，此亦顯示前七個分題組及後兩個分題組測試能力間彼此具有良好關聯、內部一致性良好。

其中發現字義理解分題組與第二至七分題組間均具.5以上的關聯，與推論理解三、策略選擇、列式、計算分題組間更具.6以上的關聯，代表接受問題字面訊息的能力（字義理解）是其他解題能力的基礎。此外兩項辨識問題數量訊息的能力（推論理解一及二）彼此間具.86的關聯、同質性相當高，再者辨識及理解問句訊息的能力（推論理解三）與策略選擇、列式、計算三項能力間均具.6以上的關聯，特別是與策略選擇能力間更具.72之相關，表示辨識及理解問句訊息的能力（推論理解三）與策略選擇非常有關。最後策略選擇能力與列式、計算兩項能力間均具.8以上的關聯，而列式與計算能力彼此間又具.92的相關，代表三者間具有密不可分之關聯。

但後兩個分題組與前七個分題組間因測試能力屬性不同，故僅具有低度相關。

二、構念效度分析

（一）共同向度分析

為了解各分題組內單題在所屬成份上之共同向度性，使用主成份分析法分別對各分題組內的各單題進行分析，如表10所示。

自表10看來，在前置分析部份，七個分題組主成份分析的 KMO 值界於 .61至 .98間，僅有第二分題組較低（約.61），其餘六個分題組均接近或大於 .7，依 Kaiser（1974）的規準，各次主成份分析之取樣適切性良好、居中等以上程度。另各個 Bartlett 近似卡方值大且達.000顯著水準，代表分析使用之相關矩陣並非單元矩陣，且各相關係數彼此不同並大於0，適合進行主成份分析。

表10 分題組所屬各單題之主成份分析成份負荷量 (N=477)

單題一至八	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
成份負荷量	.51 至.76	.29 至.77	.32 至.89	.32 至.66	.35 至.77	.45 至.72	.34 至.73	.97 至.974	.978 至.981
KMO	.73	.61	.72	.69	.70	.71	.73	.97	.98
Bartlett 近似卡方	414.41***	455.52***	706.60***	444.08***	581.91***	585.82***	559.51***	7821.67***	9020***

註：以上卡方檢定 df 均為 28；***表 $p < .001$
98.61%之成份負荷量均在 .3 以上

正式分析部份，各分題組主成份負荷量之分布，除第二分題組單題四與該分題組總分的主成份相關接近 .3 (.29) 外，其餘各分題組各單題與各分題組總分間的主成份相關均界於 .31 至 .981 間，此代表各分題組的共同向度性良好。至於最後兩個分題組可能因項目測量屬性為較穩定之解題檢核習慣及態度，而非較不穩定之解題認知表現，加上此二分題組均採部份給分而非二元計分，故各單題與成份相關較高且一致。

接下來，為了解全題組所屬各分題組在不同成份向度上之個別共同向度性，使用主成份分析法對全題組內的各分題組進行分析，如表11所示。

表11 全題組所屬各分題組之主成份分析不同成份負荷量 (N=477)

題組成份	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式	KMO	Bartlett 近似卡方
認知成份	.81	.73	.74	.76	.87	.90	.90	.21	.07	.82	3153.62***
後設成份	-.09	.14	.11	-.09	-.08	-.12	-.10	.77	.85		

註：以上卡方檢定 df 為 36；***表 $p < .001$

自表11看來，在前置分析部份，KMO 值為 .82、Bartlett 近似卡方值為3153.62 ($df=36$, $p < .001$)，KMO 值依 Kaiser (1974) 的規準，本次主成份分析之取樣適切性良好且接近極佳、居中上程度，而 Bartlett 近似卡方值相當大且達.000顯著水準，代表分析使用之相關矩陣並非單元矩陣，且各相關係數彼此不同並大於0，適合進行主成份分析。

正式分析部份，就大於 .3之標準而言，可發現前七個分題組的第一成份負荷量界於 .74 至 .90 間，後兩個分題組的第二成份負荷量界於 .77 至 .85 間，再參照本題組編製理論向度而言，亦即前七個分題組與全題組間的認知成份相關界於 .74 至 .9 間，而後兩個分題組與全題組間的後設認知成份相關界於 .77 至 .85 間，此代表全題組不同成份內的個別共同向度性相當良好。

由表10及表11看來，各分題組及全題組內的不同成份共同向度性均為良好，且符合本題組原初之編製理論架構，加上為避免產生 Byrne (2001) 所謂虛假 (bogus) 因素，以及陳密桃、李新民與黃秀霜 (民96) 所謂擬似 (spurious) 相關的問題，並降低因素分析參數估計所需樣本數之門檻，以提高分析的信效度起見，後續之試探性及驗證性因素分析均將在考量 Bandalos 與 Finney (2001) 的「題組內單一向度」原則下，審慎以同性質之分題組 (item parcel) 為單位進行。

(二) 試探性因素分析

由於假定本題組測量之潛在因素間應具有一定程度之相關存在，故使用主軸法 (principal axes method) 搭配斜交轉軸 (oblimin rotation) 進行試探性共同因素分析，其前置分析部份，KMO 值為 .82、Bartlett 近似卡方值為3153.62 ($df=36$, $p < .001$)，KMO 值依 Kaiser (1974) 的規準，本次因素分析之取樣適切性良好且接近極佳、居中上程度，而 Bartlett 近似卡方值相當大且達.000顯著水準，代表因

素分析使用之相關矩陣並非單元矩陣，且各相關係數彼此不同並大於0，適合進行因素分析。

正式分析部份，系統成功萃取三個因子（轉軸收斂於第六個疊代），在轉軸前三個因子的總和分別為4.55、0.94及0.76，其變異數百分比分別為50.53%、10.41%及8.42%，總累積變異解釋比為69.36%，第一與第二因子相關為 .58、第二與第三因子相關值為 .15、第一與第三因子相關值為 .12，符合使用斜交轉軸之假定。但由於三個因子具有低至中度正相關，因此系統無法加入平方和負荷量以取得總變異數，故無法估計轉軸後之總累積變異解釋比，僅能報告三個因子在轉軸後的平方總和，分別為4.25、3.13及0.95（接近1），此值相當於使用正交轉軸後所得之因子特徵值（eigen value），依 Kaiser（1960）建議大於1的特徵值可予保留之規準看來，前二者具相當影響，最後者雖僅近門檻，但所含二因素負荷量均在.3以上（如表12所示），為顧及理論結構完整性，仍宜予保留。

所得之組型（pattern）及結構（structure）矩陣如表12所示。

表12 分題組斜交轉軸因素分析結果

	組型矩陣			結構矩陣		
	第一因子	第二因子	第三因子	第一因子	第二因子	第三因子
字義理解	.62	.22	-.04	<u>.74</u>	.57	.07
推論理解一（數量句一）	.01	.96	.04	.56	<u>.97</u>	.18
推論理解二（數量句二）	.06	.89	.00	.55	<u>.89</u>	.14
推論理解三（問句）	.81	-.10	.04	<u>.76</u>	.37	.13
策略選擇	.89	-.01	.03	<u>.89</u>	.51	.14
列式	.90	.04	-.03	<u>.93</u>	.56	.09
計算	.87	.09	.03	<u>.91</u>	.58	.09
檢核行動	.11	-.03	.56	.16	.12	<u>.57</u>
檢核方式	-.12	.06	.72	.01	.10	<u>.72</u>

註：方框表分題組之因素歸屬

由於組型矩陣在性質上較接近轉軸後因素分數之加權係數（相對重要性），而結構矩陣在性質上則為因素與觀察變項間之相關，亦即因素負荷量（factor loading），較能反應因素與觀察變項間之關係，適合作為判定因素結構及命名之用（邱皓政，民90），故自結構矩陣看來，這些負荷量均在 .3 以上，代表每個分題組與兩個因素間均具相當程度之關聯，再自其在兩個因素的數值大小來判斷其因素歸屬，可發現字義理解、推論理解三（問句）、策略選擇、列式、計算五個分題組適合歸屬第一因子，推論理解一（數量句一）、推論理解二（數量句二）兩個分題組適合歸屬第二因子，檢核行動、檢核方式兩個分題組適合歸屬第三因子，歸納第一因子所屬五個分題組之性質，均屬數學解題之基本能力，故命名為「基本解題技巧」，歸納第二因子所屬兩個分題組之性質，均屬辨認數學文字題解題脈絡中相關數量之能力，故命名為「數量辨識技巧」，歸納第三因子所屬兩個分題組之性質，均屬數學文字題解題檢核相關能力，故命名為「檢核技巧」。若再配合表11各分題組不同成份之主成份分析結果及前述因子相關值，第一及第二因子應可歸為認知成份，而第三因子則可歸為後設認知成份。

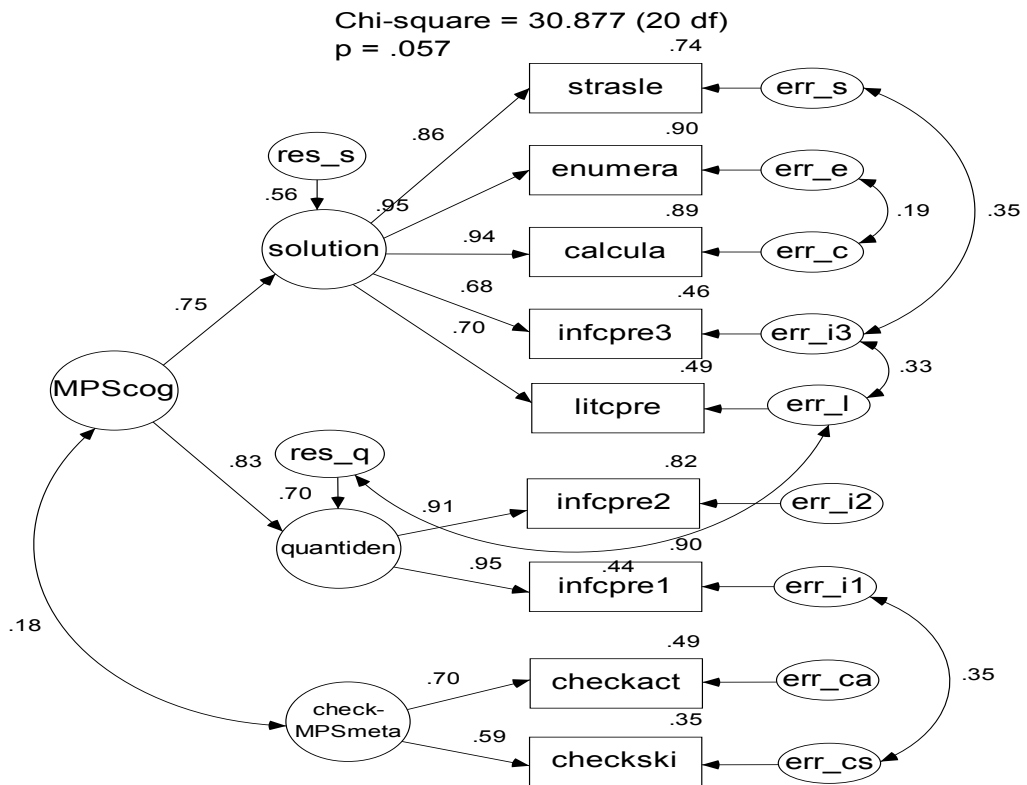
因此本題組所測試診斷之能力有三，一為基本解題技巧、二為數量辨識技巧、三為檢核技巧，前者包含接收題目字面訊息、找到問句、決定方法、列出式子及計算答案等五項能力，中者為辨認兩個前提句中之數量的能力，後者為從事解題檢核行動及方式之能力。前兩項為解題認知能力，第三項為解題後設認知能力。如此結果不僅與前述文獻所提之解題歷程成份元素及其向度歸屬相符，且進一步發現了其間隱含之彼此群聚隸屬的共同性質結構及向度。

本次分析之各題組再製共同性界於 .34至 .87間（其中八個大於 .53以上，佔89%）、題組間再製相關有二十五個大於 .30以上（64%）、且界於 .36至 .86間（餘為不同成份之相關故較低），而再製

相關與實際相關之非重複性殘差絕對值僅有三個大於 .05 (8%)，因素萃取狀況良好。

(三) 驗證性因素分析

前述分析結果雖已浮現成份、因子及觀察變項間之可能包含隸屬型態，但二成份三因子整體階層組織結構及徑路關係並不明朗，為深入了解其合理與可靠性，在樣本數(近500人)、遺漏值處理精確度及研究目的(僅針對樣本分析)之綜合考量下，將據此參照李茂能(民95)的建議，以分題組為單位使用 Amos 程式內定之「最大可能性」估計法(Maximum Likelihood, ML)進行二階驗證性因素分析，而不選用「不偏(unbiased)或「不受漸近分配限制」估計法(Asymptotically Distribution-Free, ADF)。初始估計模式雖整體指標適配比達70%，但因未顧及測量誤差共變路徑關係，致使 AGFI、RMSEA、CMIN/DF 三個重要指標未達接受基準，故再依修改指標及表9、21所示之關係組型，在未更動原因子結構路徑前提下，逐次調控測量誤差共變路徑反覆估計，獲致如圖2之修改模式，整體適配項目如表13所示。



變項圖註：

1. solution：基本解題技巧
 - (1) litcpre：字義理解 (2) infcpre3：推論理解三 (3) calcula：計算 (4) enumera：列式 (5) strasle：策略選擇
2. quantiden：數量辨識技巧
 - (1) infcpre1：推論理解一 (2) infcpre2：推論理解二
3. check-MPSmeta：檢核技巧(解題後設認知成份)
 - (1) checkact：檢核行動 (2) checkski：檢核方式
4. MPScog：解題認知成份
 - (1) solution：基本解題技巧 (2) quantiden：數量辨識技巧

圖2 歷程導向解題診斷評量題組驗證性因素分析結構

首先自圖2結構看來，此次被驗證之因素結構中，共分一個二階潛在因素、三個一階潛在因素、九個觀察指標變項，第一個一階潛在因素「基本解題技巧」影響「字義理解」、「推論理解三（問句）」、「策略選擇」、「列式」、「計算」五個觀察指標變項，第二個一階潛在因素「數量辨識技巧」影響「推論理解一（數量句一）」、「推論理解二（數量句二）」兩個觀察指標變項，第三個一階潛在因素「檢核技巧」影響「檢核行動」、「檢核方式」兩個觀察指標變項。二階潛在因素「解題認知成份」影響前兩個一階潛在因素，且與第三個一階潛在因素間具有關聯。

次從因素與觀察指標變項的相關（因素負荷量）看來，第一個一階潛在因素「基本解題技巧」與所屬五個觀察指標變項「字義理解」、「推論理解三（問句）」、「計算」、「列式」、「策略選擇」間之相關分別為 .70、.68、.94、.95、.86，第二個一階潛在因素「數量辨識技巧」與所屬兩個觀察指標變項「推論理解一（數量句一）」、「推論理解二（數量句二）」間之相關分別為 .95、.91，第三個一階潛在因素「檢核技巧」與所屬兩個觀察指標變項「檢核行動」、「檢核方式」間之相關分別為 .59、.70。二階潛在因素「解題認知成份」與前兩個一階潛在因素間之相關分別為 .75、.83，而此因素又與第三個一階潛在因素間之相關為 .18。

再就各觀察指標變項對潛在因素的測量變異解釋貢獻量觀之，第一組觀察指標變項（字義理解、推論理解三／問句、計算、列式、策略選擇）對第一個一階潛在因素「基本解題技巧」分具49%、46%、89%、90%、74%的測量變異解釋貢獻比，共同萃取測量變異解釋貢獻比為70%。第二組觀察指標變項（推論理解／數量句一、推論理解二／數量句二）對第二個一階潛在因素「數量辨識技巧」分具90%、82%的測量變異解釋貢獻比，共同萃取測量變異解釋貢獻比為87%。第三組觀察指標變項（檢核行動、檢核方式）對第三個一階潛在因素「檢核技巧」分具49%、35%的測量變異解釋貢獻比，共同萃取測量變異解釋貢獻比為42%，所有觀察指標變項對潛在因素的共同萃取測量變異解釋貢獻比為67%。前兩個一階潛在因素（基本解題及數量辨識技巧）對二階潛在因素「解題認知成份」分具56%、70%的測量變異解釋貢獻比，共同萃取測量變異解釋貢獻比為63%，若以李茂能（民95）的建議標準觀之，以上所提測量變異萃取解釋貢獻比大致都接近或在50%以上（除檢核方式外），表示大部份觀察指標變項對所測構念之代表性尚屬良好（變異萃取解釋比之計算公式請參見李茂能，民95）。

而在測量誤差共變關係方面，字義理解／讀題與推論理解三／問句、策略選擇測量誤差間分具.33及.35之相關，計算與列式測量誤差間具.19之相關，推論理解／數量句一與檢核方式測量誤差間具.35之相關，字義理解／讀題測量誤差與數量辨識技巧殘差間具.44之相關。此外第一與第二潛在因素相關為.63、第二與第三潛在因素相關為.15、第一與第三潛在因素相關值.13（另以調控測量誤差共變關係後之一階 CFA 分析所得）。

本模式估計之各迴歸加權係數顯著性界於.01至.000間，各共變數（與相關有關）顯著性界於.03至.000間，各變異數顯著性界於.047至.000間。

表13 模式整體適配度檢定項目分析

項目	絕對適配檢定				增值/相對適配檢定							精簡適配檢定		CMIN/DF
	$\chi^2(df=20)$	GFI	RMR	RMSEA	AGFI	NFI	TLI	CFI	RFI	IFI	PNFI	PCFI	AIC	
適配標準	儘可能小 $p > .05$	>.90	<.05	<.05 優 <.08 良	>.90	>.90	>.90	>.90	>.90	>.90	>.50	>.50	儘可能小	1~3
統計量	30.877 ($p=.057$)	.986	.292	.034 ($p=.874$)	.968	.990	.994	.997	.983	.997	.550	.554	80.877 (90~3197)	1.544
評估結果	+	+	-	+	+	+	+	+	+	+	+	+	+	+

註：以上項目及標準係參考自余民寧（民95）、李茂能（民95）、王保進（民93）、邱皓政（民92）之建議，+表符合，-表不符合

由於李茂能(民95)、王保進(民93)、邱皓政(民92)認為SEM整體模式適配度應綜合多重指標之結果衡量之,故依前述學者所建議之模式適配度檢定項目及標準看來,本次分析結果在表13十四個模式適配檢定項目中,有十三項達理想標準、僅RMR未符合,但依李茂能(民95)所言,此值易受測量單位大小之影響,解釋不易。準此評估,除RMR外,十三個適配檢定項目均達理想標準,適配比率為100%,再就余民寧(民95)所提分類挑選重要指標比對原則而言,AGFI、RMSEA、CMIN/DF等亦達標準,故本模式整體適配程度應在可接受範圍內,亦即綜合主成份及試探性因素分析所得之二成份三因子因素結構假定目前應得以成立、未被推翻。

整體而言,本次驗證性因素分析所得結果,無論在因素與觀察指標變項的相關(因素負荷量),抑或是各觀察指標變項對潛在因素的個別及共同平均測量變異解釋萃取量上大致良好,且因素結構及因子相關均符合前述試探性因素分析結果之假定,加上因素模式結構適配程度尚屬適當。因此目前本題組所測試診斷之成份及能力組成結構(如圖2所示)應可暫時得到確認。其中雖然「檢核技巧」之測量共同萃取變異(42%)低於一般門檻(50%),但仍保有二分之一弱之共同萃取變異且差距並非相當的大,加上凡有測量必有誤差,SEM的分析亦非樣本獨立(sample independent),所以未來有必要再蒐集樣本加以確認。

三、效標關聯及預測效度

(一) 效標關聯效度

為了解本題組與外部效標間之關聯,以個體一年級在校的數學學業考試成績(含單步驟加減法文字題單元學習內容之紙筆評量分數)為效標,採皮爾遜積差相關分析解題診斷全題組、認知成份題組與效標間之關係,如表14所示。

表14 全題組及因素歸屬後分題組(各成份)效標關聯效度分析(N=477)

效標	全題組	認知成份題組	基本解題技巧	數量辨識技巧	檢核技巧	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
數學學業	.38	.57	.53	.53	.04	.45	.50	.52	.43	.51	.49	.46	.07	-.01

註：以上數值之概率值,方框者均為.000、餘大於.05

自表14可看出,無論全題組或認知成份題組,二者均與效標數學學業考試成績間具有顯著之中度正相關($r = .38, .57, p < .001$),除檢核技巧分題組(含行動及方式)因測試能力屬性與數學學業考試成績較不相同,導致相關甚低外,其餘分題組與數學學業考試成績間亦具有界於.43至.53之顯著中度正相關($p < .001$)。自上述說明看來,解題診斷全題組及認知成份題組均具良好之效標關聯效度。

(二) 預測效度

為檢視本題組對外部效標間之預測力,以個體一年級在校的數學學業考試成績為被預測變項、解題診斷全題組及認知成份題組總分為預測變項,分別進行直線迴歸預測,如表15所示。

表15 題組預測效度迴歸分析摘要表(N=477)

模式	投入變項	R	R ²	調整後R ²	F值	原始分數迴歸係數(B)	標準化迴歸係數(β)	t值
1	全題組總分	.38	.14	.14	78.94***	.17	.38	8.89***
2	認知成份題組總分	.57	.33	.33	231.91***	.41	.57	15.23***

註：全題組及認知成份題組總分分別為預測變項,數學學業考試成績為被預測變項;***表 $p < .001$

表15係分別以個體的全題組及認知成份題組總分來預測其數學學業考試表現，結果發現前二者對後者均具相當預測力，其相關為.38及.57、調整後相關平方為.14及.33、迴歸考驗均達.000顯著水準（ $F=78.94, 231.91, p<.001$ ），原始分數及標準化迴歸係數亦均達.000顯著水準（ $t=8.89, 15.23, p<.001$ ）。這代表個體的全題組、認知成份題組表現與其數學學業考試成績間分具約14%及33%的共同變異，且可由全題組、認知成份題組表現這兩個變項來作預測解釋，亦即二者對個體數學學業考試成績分具約14%及33%的預測解釋量，且無論是原始分數或標準化迴歸係數均具相當之預測加權重要性，顯示二者分別對其數學學業考試成績具有良好預測力。至於全題組總分可能因內含與學業考試表現較無關之檢核技巧題組，故預測力較低。

四、信度分析

為了解本題組之測量可靠性，故進行以下分析。

（一）傳統信度

如表16所示。

表16 全題組及因素歸屬後分題組內部一致性信度分析（ $N=477$ ）

項目	全題組	基本解題技巧	數量辨識技巧	檢核技巧	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
係數	.94	.93	.85	.97	.71	.64	.75	.70	.72	.73	.75	.99	.99
刪題建議	無	無	無	無	無	無	無	無	無	無	無	無	無
題數	72	40	16	16	8	8	8	8	8	8	8	8	8

註：上述「無」表即使任刪一題，均無助於所得係數小數點以下第一位之增加。方框為標準化 α 、餘為庫李。

自表16看來，因受題數影響，除部份分題組數值較低（如第二分題組）外，其餘分題組均接近或大於.7，尚屬可接受範圍，而因素歸屬後之三個分題組係數分別為.93、.85及.97，全題組係數.94，代表因素歸屬後之分題組與全題組內部一致性相當良好。至於最後兩個分題組可能因項目測量屬性為較穩定之解題檢核習慣及態度，而非較不穩定之解題認知表現，加上此二分題組均採部份給分而非二元計分，故數據較高。

（二）組合信度

依李茂能（民95）所提使用標準化因素負荷量之計算公式（係轉化自 Reuterberg 與 Gustafsson 於1992所提使用非標準化迴歸係數之計算公式）對前述 Amos 進行 CFA 所得資料進行組合信度分析，如表17所示。

表17 全題組及因素歸屬後分題組組合信度分析（ $N=477$ ）

項目	全題組	基本解題技巧	數量辨識技巧	檢核技巧
組合信度	.95	.92	.93	.60

若以李茂能（民95）的建議標準觀之，表17前三項數據都在.70以上（第四項檢核技巧因潛在變項萃取測量變異比率較低導致數據較低），表示觀察指標變項對所測構念之測量可靠性尚屬良好。

五、鑑別力分析

為了解全題組、各分題組及各單題的鑑別力，採傳統測驗理論作法，將個體總分由高至低排序後，選取居前27%及後27%之個體為高低分組，後以此二組在全題組、各分題組及各單題上的表現進

行獨立樣本 t 考驗。

(一) 全題組及分題組 (各成份) 極端組鑑別力分析

如表18所示。

表18 全題組及分題組 (各成份) 極端組鑑別力分析

項目	題 組 別													
	全題組	認知成份題組	基本解題	量數辨識	檢核	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
t 值	26.30	13.96	14.63	6.97	18.15	10.31	7.62	6.03	11.88	12.97	12.15	12.64	16.10	12.59

註：由於系統採人數累積百分比 (PR) 選取兩組個體，而同分數的個體必須同時選取，故造成兩組人數及變異數不等之狀況 (高分組 165 人、低分組 144 人)，因此上述 t 及概率值均採「不假設變異數相等」之數據。以上數值之概率值均為 .000。

自表18看來，全題組及各分題組極端組表現的 t 考驗結果，均達.000顯著水準，代表全題組及各分題組對不同能力個體的鑑別力良好。

(二) 分題組 (成份) 單題極端組鑑別力分析

如表19所示。

表19 各分題組各單題極端組鑑別力分析

單題 一至八	分題組								
	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
t 值	2.53 (.012)	1.95 (.053)	1.74 (.083)	4.08	3.32 (.001)	3.32 (.001)	4.23	15.26	11.96
	至 9.56	至 6.13	至 5.35	至 8.06	至 9.43	至 10.49	至 8.96	至 15.56	至 12.44

註：上述 t 及概率值依表 20 註解所提，亦採「不假設變異數相等」之數據。以上未有括號標示之數值，其概率值均為 .000

自表19看來，各分題組內各單題極端組表現的 t 考驗結果，除第二及三分題組各有一題數據較低外，餘均達 .05以上顯著水準，代表各分題組絕大部份單題對不同能力個體的鑑別力良好。

六、因素歸屬後各項組合分數統計

九個分題組在經前述主成份分析、試探性及驗證性因素分析後，因素歸屬後分題組表現之組合分數訊息如表20及表21所示。

表20 因素歸屬後組合分題組表現之描述統計 ($N=477$)

組合分題組	M	SD	P	與全題組總分之相關
認知成份	49.84	8.11	.89	.69***
基本解題技巧	34.59	6.87	.87	.67***
數量辨識技巧	15.25	1.88	.95	.55***
後設認知成份 (檢核技巧)	22.55	9.62	.56	.79***

註：P 表表現百分比；***表 $p < .001$

表21 因素歸屬後組合分題組總分與各分題組表現之相關 (N=477)

組合分題組	字義理解	推論理解一	推論理解二	推論理解三	策略選擇	列式	計算	檢核行動	檢核方式
認知成份	<u>.81</u>	<u>.71</u>	<u>.70</u>	<u>.78</u>	<u>.89</u>	<u>.91</u>	<u>.91</u>	.16	.03
基本解題技巧	<u>.80</u>	.58	.56	<u>.82</u>	<u>.91</u>	<u>.93</u>	<u>.92</u>	.16	.01
數量辨識技巧	.57	<u>.96</u>	<u>.97</u>	.37	.52	.56	.58	.12	.11
後設認知成份 (檢核技巧)	.06	.15	.19	.09	.10	.07	.08	<u>.81</u>	<u>.87</u>

註：以上數值之概率值均為 .000

方框代表該組合分題組總分與各內含分題組表現之相關

底線代表該組合分題組總分與同一歸屬成份組合內的非內含分題組表現之相關

未註記代表該組合分題組總分與不同歸屬成份組合內的非內含分題組表現之相關

自表20、21看來，不同成份內之內含或非內含分題組間相關均在 .3以上且達 .000顯著水準，代表即使在因素歸屬後，三個分題組在所屬成份之內部一致性仍然良好，且與不同成份歸屬之分題組間具有低度相關。

基於上述，各項因素歸屬後之組合分數訊息，將可提供分析及解讀個體表現之新分數組合向度，而此架構乃源於文獻所提理論假定、主成份分析、試探性因素分析、驗證性因素分析之結果而衍生，故兼具理論與實徵意義。準此，本評量題組共可使用九個獨立分題組分數、基本解題技巧題組分數、數量辨識技巧題組分數、檢核技巧題組分數、認知成份題組分數、後設認知成份題組分數、全題組總分等各項組合分數，搭配剖面圖 (profile) 來對個體不同面向的解題能力進行評估診斷及解讀，但其最高總分不同，若需涉及個體內與個體間之差異比較，則須以共同量尺分數轉換後為之。

七、題組診斷處方功能分析

為了解本解題診斷題組對低能力個體的診斷力起見，以下針對兩個個案於線性轉換後之各解題認知成份及一般解題平均表現 T 分數進行剖面分析及對照 (因二者計分量尺不同，加上一般自行解題題組無法提供檢核訊息且題組性質不同，故將檢核行動及方式除外)，以說明其診斷及處方功能。其中一般解題平均表現係指個體在參照現行一年級課本內容自編之單步驟加減法文字題解題題組上的平均表現 (解題題組之題型結構與解題診斷題組所用題型及題數完全相同、涉及運算數量值範圍相近，僅有題組形式不同，亦即採一般自行解題而非題組選答方式進行)。另外此二位個案在本診斷題組的各題均有檢核行動，但均僅針對答案進行檢核而已。

(一) S110之解題認知成份與平均表現剖面及診斷處方分析

本個案於線性轉換後之解題認知成份平均表現及一般解題平均表現 T 分數分別為34.19及31.93，其能力層次約為低於共同量尺平均數以下一個半至兩個標準差之間，其各解題成份 (34.78, 42.32, 44.64, 49.87, 32.9, 31.79, 27.7) 與一般解題平均表現之剖面對照如圖3所示。

自圖3可知，S110之字義理解、策略選擇與列式三個成份表現與一般解題平均表現相差不多，推論理解三個成份表現明顯優於其一般解題平均表現，僅有計算成份表現低於一般解題平均表現，以上七個成份表現僅有推論理解三此成份表現達共同量尺平均數，其餘都在平均數以下。

綜合上述訊息，雖然 S110的成份平均表現及一般解題平均表現都明顯低於共同量尺平均數，但經剖面分析後，仍有推論理解三分題組的辨認問句能力達共同量尺平均數。其餘成份表現雖在共同量尺平均數以下，但推論理解一及二分題組的數量辨認能力明顯高於一般解題平均表現，字義理解、選擇策略及列式與一般解題平均表現相差無幾，僅有計算能力較其一般解題平均表現為低。故依 S110目前的成份平均表現及一般解題平均表現層次與相符程度看來，個體現前解題能力中，以問句辨識能力最佳，量數辨識能力次之，字義理解、選擇策略及列式能力尚可，計算能力最弱。

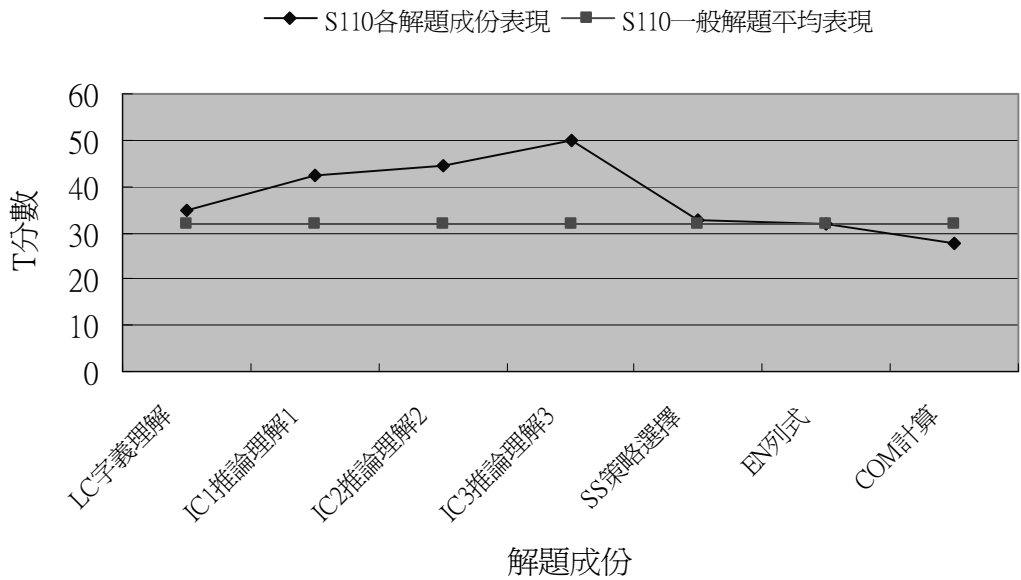


圖3 S110之解題認知成份與平均表現剖面

就處方角度視之，教學者未來可再針對 S110的數量辨識、字義理解、選擇策略、列式及計算能力設計相對應學習內容進行補救提昇，特別是計算能力部份。

(二) S41之解題認知成份與平均表現剖面及診斷處方分析

本個案於線性轉換後之解題認知成份平均表現及一般解題平均表現 T 分數分別為19.41及19.72，其能力層次約為低於共同量尺平均數以下三個標準差，其各解題成份 (34.78, 19.27, -1.67, 36.51, 26.76, 31.79, 15.86) 與一般解題平均表現之剖面對照如圖4所示。

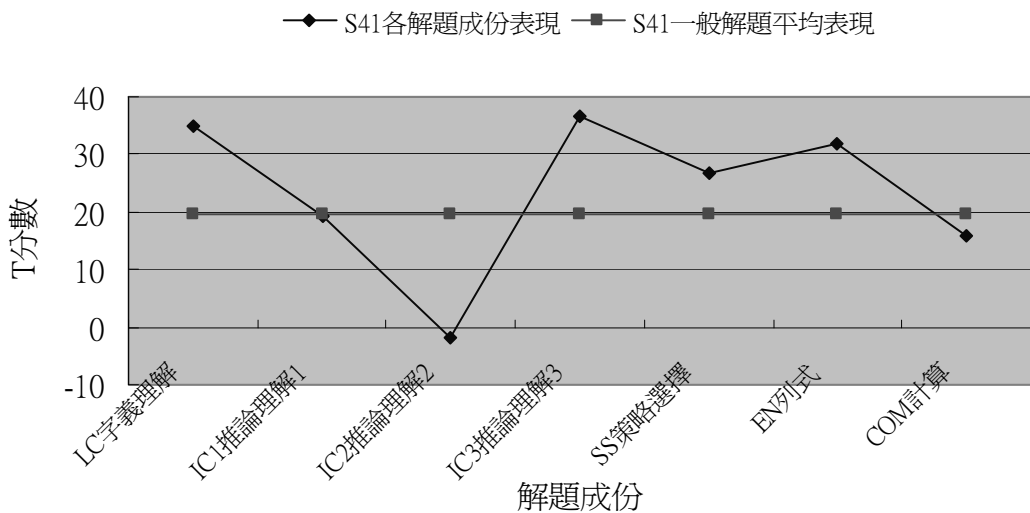


圖4 S41之解題認知成份與平均表現剖面

自圖4可知，S41之推論理解一成份表現與一般解題平均表現相差不多，字義理解、推論理解三、策略選擇與列式四個成份表現明顯優於其一般解題平均表現，推論理解二及計算成份表現低於一般解題平均表現，其中尤以推論理解二為甚，以上七個成份表現均未達共同量尺平均數，所有成份表現都在平均數以下。

綜合上述訊息，S41的成份平均表現及一般解題平均表現都明顯低於共同量尺平均數，經剖面分析後，所有成份表現雖在共同量尺平均數以下，但字義理解、問句辨識、策略選擇與列式能力明顯高於一般解題平均表現，推論理解一分題組之第一量數辨識能力與一般解題平均表現相差無幾，僅有第二量數辨識及計算能力較其一般解題平均表現為低，特別是第二量數的辨識能力。故依 S41目前的成份平均表現及一般解題平均表現層次與相符程度看來，個體現前解題能力中，以字義理解、問句辨識、策略選擇與列式能力較佳，第一量數辨識能力尚可，計算能力較弱，第二量數辨識能力最弱。

就處方角度視之，教學者未來可再針對 S41的字義理解、量數及問句辨識、選擇策略、列式及計算能力設計相對應學習內容進行補救提昇，特別是計算能力及第二量數辨識部份。

八、綜合歸納與討論

觀諸題組各項性能指標之分析結果，無論在題組的內部相關與決斷值（鑑別力），或是構念效度（試探性及驗證性因素分析，含單一向度性、因素萃取、抽取變異比及模式適配度），抑或是內部一致性 α 係數及組合信度各方面所得測試數據均屬良好，而個案剖面分析結果亦具診斷處方參考價值，故本題組在測驗工具編製上具有相當程度之良好功能。

若再就所發現之題組二成份三因子結構（認知成份含基本解題及數量辨識技巧、後設認知成份含檢核技巧）而言，此證據清楚指出本題組之內容架構不僅達成原初編製研究目標，且與 Desoete 與 Roeyers (2005)、Krulik 與 Rudnick (1989)、Garofalo 與 Lester (1985)、Hutchinson (1992)、Montague、Warger 與 Morgan (2000)、Polya (1957)、Schurter (2002) 的模式相容，亦即本題組在理論基礎及目標達成上均具良好適配度。

結論與建議

一、結論

基於前述結果發現，各項研究問題暫已獲致相當程度之釐清，分述如下：

（一）歷程導向解題診斷評量題組具有良好之解題成份測量效度

自前述結果可知本題組具良好之效標關聯及構念效度。在效標關聯效度方面，本題組不僅與外在效標數學學業成就間具中度正相關，並具良好預測性。在構念效度方面，本題組之各能力測量共同向度性良好，並隱含認知成份「基本解題」及「量數辨識」技巧、後設認知成份「檢核」技巧二成份三因子結構。

（二）歷程導向解題診斷評量題組具有良好之解題成份測量可信賴度

自前述結果可知本題組具良好之傳統內部一致性及組合信度，前者意指本題組各單題、各分題組及全題組的內部一致性良好，後者則意指本題組在三因子結構下之各成份及因子構念測量上的內部一致性良好。

（三）歷程導向解題診斷評量題組具有良好之解題成份測量鑑別力（全題組及成份）

自前述結果可知本題組具良好之單題及成份測量鑑別力，意指本題組各單題、各分題組在成份

測量上的區別性良好，能有效區分不同能力個體在各單題及各成份上的表現。

(四) 歷程導向解題診斷評量題組具有解題教學上之診斷處方功能

自前述結果可知本題組具良好之成份診斷及教學處方功能，意指本題組能有效診斷個體之解題學習（成份）弱處，並可由此轉化成未來教學時之重點建議，具解題教學實務應用價值。

二、建議

(一) 教學及評量方面

由於本題組在各指標方面均達良好標準，故可在單步驟加減法文字題教學歷程的中後階段，運用本題組進行不同目的之評量。可於教學中，配合教學題型使用以進行形成性及診斷性評量，亦可於教學後，配合教學內容使用以進行總結性及診斷性評量。不論何種使用方式均有助於個體各解題歷程成份學習弱處之診斷處方。

另外在二步驟加減法文字題教學前，則可作為解題歷程學習先備成份能力測試及診斷處方之用。教學者亦可在教學歷程中適時加入解題檢核方法之教導，以培養個體檢核習慣及其效率。

(二) 未來研究方面

由於本題組在各指標方面均達良好標準，故基於本次題組發展之經驗，後續研究可再針對不同年級、不同題型（加減法、乘除法、四則）、不同步驟（二步驟、三步驟、多步驟）進行各類解題成份診斷題組之編製發展。

在理論研究方面，可擴大解題歷程成份診斷工具之發展及驗證範疇，並作為探討個體解題學習弱處之研究工具。而在實務應用方面，則可配合不同教學目的，增加更多不同性質之診斷工具，此將有助於實務工作者在教學現場中，針對學習者的解題學習歷程進行診斷及處方。

三、研究限制

雖然本研究目前所得結果均為良好，但仍有兩項潛在限制，其一為顧慮教學版本及進度的統一起見，樣本數及取樣區域受限，故需考慮研究結果是否受樣本依賴因素的影響，特別是預測效度可能受脈絡背景或人口統計變項影響而隱含「階層」效應，但因本研究目的並不在於此，故未再進行階層線性模式（Hierarchical Linear Model, HLM）之分析。此外因班級教師建議題組設計之選項數應考量一年級學生的閱讀能力，不宜過多，故另需考量研究結果是否受個體作答猜測率的影響。

參 考 文 獻

- 王保進（民93）：多變量分析—套裝程式與資料分析。台北：五南。
- 余民寧（民95）：潛在變項模式：SIMPLIS 的應用。台北：高等教育。
- 李茂能（民95）：結構方程式模式軟體 Amos 之簡介及其在測驗編製上之應用。台北：心理。
- 邱皓政（民90）：社會與行為科學的量化研究與統計分析：SPSS 中文視窗版資料分析範例解析。台北：五南。
- 邱皓政（民92）：結構方程模式：LISREL 的理論、技術與應用。台北：雙葉。
- 陳冠海等編（民95）：國小數學課本（第一冊）。台南：南一。
- 張英傑等編（民96）：國小數學課本（第二冊）。台南：南一。
- 陳蜜桃、李新民、黃秀霜（民96）：台灣地區國民小學教師實用智能、情緒智能與專業表現之相關研究。教育心理學報，39卷，2期，295-316頁。

- 許瑛珍等編 (民95)：國小數學課本 (第一冊)。台南：翰林。
- 許瑛珍等編 (民96)：國小數學課本 (第二冊)。台南：翰林。
- 楊瑞智等編 (民95)：國小數學課本 (第一冊)。台北：康軒。
- 楊瑞智等編 (民96)：國小數學課本 (第二冊)。台北：康軒。
- 鄭國順等編 (民95)：國小數學課本 (第一冊)。台北：國立教育研究院籌備處 (翰林印行經銷)。
- 鄭國順等編 (民96)：國小數學課本 (第二冊)。台北：國立教育研究院籌備處 (翰林印行經銷)。
- Bandalos, D. L., & Finney, S. J. (2001) Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp.269-296). NJ: Lawrence Erlbaum Associates.
- Byrne, B. M. (2001). *Structural equation modeling with Amos: basic concepts, applications and programming*. NJ: Lawrence Erlbaum Associates.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando, Florida: Holt, Rinehart and Winston, Inc.
- Desoete, A., & Roeyers, H. (2005). Cognitive skills in mathematical problem solving in grade 3. *British Journal of Educational Psychology, 75*, 119-138.
- Fuson, K. C. (1992). Research on whole number addition and subtraction. In D. A. Grouws (Eds.), *Handbook of research on mathematics teaching and learning: A project of the national council of teachers of mathematics* (pp. 39-48). NY: Macmillan.
- Gagne', E. D., Yekovich, C. K., & Yekovich, F. R. (1993). *The cognitive psychology of school learning* (2nd ed.). NY: HarperCollins College Publishers.
- Garofalo, J., & Lester, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education, 16*(3), 163-176.
- Joreskog, K. G., & Sorbom, D. (1993). *Structural equation modeling with SIMPLIS command language*. Chicago: Scientific Software International.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black (1998). *Multivariate data analysis*. NJ: Prentice-Hall International.
- Hutchinson, N. L. (1992). The challenges of componential analysis : Cognitive and metacognitive instruction in mathematical problem solving. *Journal of Learning Disabilities, 25*(4), 249-252 & 257.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Psychometrika Educational and Psychological Measurement, 20*, 141-151.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-36.
- Krulik, S. K., & Rudnick, J. A. (1989). *Problem solving: A handbook for senior high school teachers*. Boston, MA: Allyn & Bacon.
- Montague, M., Warger, C., & Morgan, T. H. (2000). Solve it! strategy instruction to improve mathematical problem solving. *Learning Disabilities Research and Practice, 15*(2), 110-116.
- Polya, G. (1957). *How to solve it*. Princeton NJ: Princeton University Press.
- Schurter, W. A. (2002). Comprehension monitoring: An aid to mathematical problem solving. *Journal of Developmental Education, 26*(2), 22-33.
- Snow, R. E., & Lohman, D. F. (1993). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed, pp. 263-332). Phoenix, Indiana: Oryx Press.

Verschaffel, L., Greer, B., & DeCorte, E. (2007). Whole number concepts and operations. In F. K. Lester (Eds.), *Second Handbook of research on mathematics teaching and learning: A project of the national council of teachers of mathematics* (pp. 557-628). Charlotte, NC: Information Age Publishing Inc.

收稿日期：2008年03月14日

一稿修訂日期：2008年07月29日

二稿修訂日期：2008年09月11日

接受刊登日期：2008年09月12日

Bulletin of Educational Psychology, 2009, 40 (4), 683-706
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

The Development and Function of Process-Oriented Diagnostic Assessment Battery for First Graders' Learning on Word Problem Solving

Chia-Hua Hsu

Department of Education
National Chiayi University

The purpose of this study was to develop the Process-Oriented Diagnostic Assessment Battery (PODAB) for single-step word problem of addition and subtraction through componential and content analyses. Participants were composed of 477 first graders from Chiayi by convenient and cluster sampling. Participants were administered the PODAB. Item analysis suggested good passing percentage, internal correlation and critical ratio, regardless whether in single items, sub-battery or the whole battery. As for construct validity, three factors were suggested by exploratory factor analysis and confirmatory factor analysis (basic solving, quantity identification, and checking skills), and there were appropriate dimensionality, variance proportions of factor extraction, model fit, as well as criterion-related and predictive validity. For reliability, results showed appropriate internal consistency coefficient and composite reliability. Finally, because of individualized information, the battery could provide diagnostic profile and subsequent prescriptions.

KEY WORDS: diagnosis and prescription in mathematical problem solving (MPS), metacognition, reliability and validity of test, the process of MPS, the assessment of MPS