

國小四年級學童數常識三階段診斷測驗工具的發展與應用之研究

余明杰¹ 李茂能² 楊德清^{3,*}

¹臺南市永康區大灣國民小學

²國立嘉義大學 教育學系

³國立嘉義大學 數理教育研究所

摘要

本研究之主要目的在編製「國小四年級數常識三階段診斷測驗工具」，並以此測驗工具檢驗小四學童數常識的發展情形與信心指數。166位南部某公立小學四年級學生參與本研究。研究結果顯示：一、「四年級三階段數常識診斷測驗工具」具有良好的信、效度；二、四年級學童在數常識各因素之發展具有顯著差異，且學童於數常識之表現有待加強；不同數常識程度之樣本學童在數常識各成分間的發展上具有顯著差異以及數常識能力與學童之數學學習成就具有高度正相關性；樣本學童在使用數常識方法解題之比率偏低；三、第三階段之信心指數能夠進一步探索學生回答問題之真正想法。最後，依研究發現提出建議。

關鍵詞：三階段診斷測驗、四年級、數常識

壹、研究動機與目的

一、研究背景與動機

許多研究指出，對數字的感覺是導引數學學習的重要因素，並且是發展數學學習的主要來源(Berch, 2005; Chrysostomou, Pitta-Pantazi, Tsingi, Cleanthous, & Christou, 2013; Dehaene, 2011; Yang & Li, 2013)；以及強調數常識在數學學習上扮演一個重要的角色(Dyson, Jordan, & Glutting, 2013; McIntosh, Reys, & Reys, 1992; Yang & Li)。美國數學

教師協會(National Council of Teachers of Mathematics [NCTM], 2000)所發表的《學校數學課程之原則與標準》(*The Principles and Standards for School Mathematics*)中數與計算的標準更主張「本標準的中心乃在發展學生數常識的能力」(頁32)。而我國九年一貫數學領域課程綱要在第二階段數與量的主題中，強調教師應培養學童流暢的數字感(教育部, 2008)。且許多研究指出數常識的發展將影響學生數學成就的表現(Jordan, Glutting, & Ramineni, 2010; Yang, Li, & Lin, 2008)；而缺

*通訊作者：楊德清，dcyang@mail.ncyu.edu.tw

(投稿日期：民國104年8月27日，修訂日期：民國105年2月20日，接受日期：民國105年2月22日)

乏數常識將影響數學學習(Dyson et al.; Jordan, Glutting, et al., 2010)。由上述探討可見數常識的重要性。因此，協助兒童發展數常識已被許多國際研究與報告所強調(Berch; Dunphy, 2007; NCTM, 2000; Jordan, Raminent, & Watkin, 2010; Sood & Jitendra, 2007; Verschaffel, Greer, & De Corte, 2007; Yang & Li)。

由於數常識已被視為21世紀數學教學的重心(NCTM, 2000; Yang & Li, 2013)，因此，近年來國內外有相當多數常識相關的研究(例如：楊德清，2002；劉曼麗，2006；Dunphy, 2007; Dyson et al., 2013; Jordan, Glutting, et al., 2010; Markovits & Sowder, 1994; McIntosh, Reys, Reys, Bana, & Farrel, 1997; Yang & Li)。上述研究所使用之工具包括紙筆測驗、訪談或線上測驗；其中楊德清與李茂能(2007)所發展之「國小二階段數常識電腦化診斷測驗系統」(第一階段答案選項，第二階段乃原因選項)，不僅能以此模式診斷出學童之數常識表現與思維方式，亦可藉此幫助教師瞭解學童數常識發展情形，以作為未來教學調整之依據。研究亦發現此二階段測驗模式不但具備量化特點，能夠透過線上測驗減少人力、物力與時間的浪費以大量地蒐集資料與檢驗大樣本學生之數常識表現，同時能夠進一步探究樣本學生解決問題時所使用之方法與可能之迷思概念(Li & Yang, 2010)。然而上述二階段測驗之缺點是無法瞭解答題學生之信心指數為何？Peşman與Eryilmaz (2010)依據科學教育研究之需要開發三階段診斷測驗模式，研究發現除了可以達到二階段的效果外，並能發展具備良好信、效度之診斷測驗工具以瞭解受測學生答題之信心指數，進而確認學生之學習情況。

因此，本研究除了吸取二階段診斷測驗的優點之外，並參照Peşman與Eryilmaz (2010)

所開發之三階段診斷測驗模式進行研究，期盼此測驗系統不僅能對答案來源(數常識方法、迷思概念、傳統算則、猜測)做到進一步之判斷，於第二、三階段收雙層測試之效，以提高診斷測驗之信效度。更重要的是，它能幫助教學者判斷學生數常識發展的分類層次、探索學生對此概念之真正想法與信心，並以此為根據協助學生進行學習。此外，四年級乃是中年級學生進入高年級的一個重要學習里程，故本研究認為四年級具有高度的研究價值。因此，若能透過四年級三階段數常識診斷評量工具的編製，來瞭解學生數常識的發展情形，教師便能根據所呈現的結果來改善教學，此乃本研究之主要動機。

二、研究目的與研究問題

本研究旨在編製「國小數常識三階段診斷測驗工具」，並探究小四學童數常識的發展情形與信心指數。因此，研究問題如下：

- (一)檢驗「四年級數常識三階段診斷測驗工具」是否具有良好之信、效度？
- (二)藉由「四年級數常識三階段診斷測驗工具」，檢驗完成國小四年級數學課程之學童
 - 1.數常識的表現情形？
 - 2.分析不同數常識程度學童在數常識各因素之表現，並探討四個因素與學童數學學習成就之關聯性？
 - 3.分析樣本學童在在回答數常識問題所使用之解題方法？
- (三)藉由「四年級數常識三階段診斷測驗工具」，分析四年級學童數常識中信心指數於各因素間之表現，並探討其

與猜答、迷思、數常識、紙筆等作答原因之答題關聯性？

貳、文獻探討

一、數常識的意義、重要性與組成架構

(一)數常識的意義

數常識可以視為個人對數字、運算、以及數字與運算所交織成情境的理解，此種理解包含可以使用靈活與彈性的方法產生解題策略來解決數字與運算問題，以及生活中涵蓋數字與運算的相關問題，並能判斷答案合理性(Markovits & Sowder, 1994; McIntosh et al., 1992; Yang & Li, 2013)。

(二)數常識的重要性

許多學生認為將一連串定理、公式與算則進行背誦與熟練就是把數學學好(Markovits & Sowder, 1994; Verschaffel et al., 2007)，但是學生在解決數字情境問題時，他們並不理解算則隱含的意義，而如此重視算則的傳統教學，不單會侷限了學生分析、理解及解題能力的發展(McIntosh et al., 1992; Yang & Li, 2013)，也會因為缺乏合理答案的認知及對解題欠缺思考，容易造成學生數學學習挫敗的經驗(Sowder & Kellin, 1993; Reys & Yang, 1998)。然而數常識強調概念的理解，有意義的學習，並能彈性靈活地發展不同的策略以解決問題，此更凸顯數常識的重要。

Lock與Gurganus (2004)指出數常識能將數學課程中的數字、規則及運算的知識做連結，並可以直覺地運用彈性策略來處理問題。楊德清(2002)主張「數常識是一種概念，亦是多種能力的組合，所以概念即是對數字與運算符號之意義的理解，這種理解儲存於長期記憶區，當有需要時，便可自動自長期

記憶區取出使用」(頁234)。從上述的觀點來看，數常識是連結數字、規則與運算的良好概念網路，它能使一個人直覺地發現數字與數字、數字與運算間的關係，然後採行彈性且適切的策略解決問題(Sowder, 1992)。在Yang等(2008)的研究中指出，數常識的不足會影響學童數學表現，因此建議教師應多提供學童數常識的學習機會。上述論點進一步指出數常識的重要性。

(三)數常識的組成架構

根據前述之數常識的意義與重要性，以及本研究的目的主要在探討國小四年級學童在數常識的相關表現，研究者歸納統整國內、外學者(楊德清, 2002; McIntosh et al., 1992, 1997; NCTM, 2000; Sowder, 1992; Yang et al., 2008)所提之數常識組成架構，並參照楊德清與李茂能(2007)所發展之「國小四年級數常識電腦化診斷測驗系統」為基礎，進一步加入第三階段信心量表，以發展成小四數常識三階段診斷測驗。本研究以楊德清與李茂能之研究為基礎，數常識組成架構包括下列4大成分：1.具備瞭解數字的基本意義的能力；2.具備比較數字大小的能力；3.具備數字分解與合成的能力；以及4.具備判斷答案合理性的能力。

二、數常識的相關研究

過往之中小學數學教育由於過於重視良好計算能力與技巧的結果，以致教學太過機械化，學生可能學會了各種不同的解題技巧和數學概念，卻忽略了數學的彈性應用與發展性(林文生、鄔瑞香, 1999; 楊德清, 2002; NCTM, 2000)。因此，許多國家皆強調數常識在小學數學教與學中應扮演重要的角色以及幫助兒童發展數常識的重要性，進

而激發許多數常識相關的研究(Dunphy, 2007; Dyson et al., 2013; Jordan, Glutting, et al., 2010; Markovits & Sowder, 1994; McIntosh et al., 1997; Menon, 2004; NCTM; Yang & Li, 2013; Yang & Wu, 2010)。例如：Markovits與Sowder以及Menon的研究發現美國中小學學生在回答數常識相關問題時，缺乏彈性靈活的發展估算或心算等策略解題的能力，亦缺乏發展運用參考點與判斷答案合理性的能力，此顯示美國中小學學生數常識能力普遍不足。McIntosh等(1997)比較澳洲、瑞典、美國與臺灣的國小學生之數常識表現，亦發現類似情形，兒童普遍缺乏發展估、心算能力、缺乏發展運用參考點與判斷答案合理性的能力。劉曼麗(2006)針對臺灣國小四年級學生整數數感能力的研究，一發現小四學生普遍在整數數感能力的不足。Reys與Yang (1998)的研究結果顯示學生具有良好的筆算能力並不代表他們亦能發展良好的數常識，同時臺灣國小兒童亦普遍缺乏數常識能力。

三、三階段診斷測驗的相關研究

(一)三階段診斷測驗的意義、背景及相關研究

三階段診斷測驗之發展主要源自於科學教育領域，目前在科學教育領域內則有數篇與三階段診斷測驗相關之研究(e.g., Caleon & Subramaniam, 2010; Cetin-Dindar & Geban, 2011; Peşman & Eryilmaz, 2010)。所謂三階段診斷測驗之意義乃包括第一階段為答案選項(檢驗內容知識)，第二階段為相對答案之原因選項(檢驗探究知識)，第三階段為信心選項(檢驗回答者的信心強度)。Peşman與Eryilmaz以及Cetin-Dindar與Geban這兩篇研究使用的信心指數是1.有信心及2.沒有信心；Caleon與Subramaniam所使用的信心指數是1.猜測、

2.非常沒信心、3.沒信心、4.有信心以及5.非常有信心；同時學者指出Confidence rating可以被視為是個體對其認同答案之內心預估的信仰(Renner & Renner, 2001)，以及個體被要求去評估它們在回答認知任務之表現的正確性(Stankov & Crawford, 1997)。科學教育學者發展三階段診斷測驗之主要背景原因乃是因為二階段診斷測驗的優點是可以檢驗學生的答案選項與相對應答案選項之原因選項，但是卻無法真正瞭解學生回答上述兩階段問題與信心指數。基於上述原因，相關研究學者乃開發三階段診斷測驗以補足二階段診斷測驗之缺點(e.g., Caleon & Subramaniam; Cetin-Dindar & Geban; Peşman & Eryilmaz)。

Caleon與Subramaniam (2010)發展三階段診斷測驗以檢驗10年級學生波的性質與傳播概念的理解情形、回答問題之信心強弱度、以及相關迷思概念的現況。他們的研究發現三階段診斷測驗具有良好的信、效度，同時在探討學生的科學概念理解情形是很有用。此外，Caleon與Subramaniam的研究進一步發現許多受測學生的波概念不足，且透過信心強弱度可以確認學生的另有概念(alternative conceptions)是真的還是偽的另有概念。Cetin-Dindar與Geban (2011)發展三階段診斷測驗以檢驗高中學生酸鹼性概念的理解情形。他們的研究發現此三階段診斷測驗具有良好的信、效度，比二階段測驗更具信度，並且可以用來檢驗學生的另有概念、學生的先備知識等，以作為教學之參考依據。Peşman與Eryilmaz (2010)亦發展三階段診斷測驗以檢驗高中生簡易電子電路概念的理解情形，其研究發現亦類似於Peşman與Eryilmaz以及Caleon與Subramaniam的研究結果。

(二)數學教育中三階段診斷測驗的相關研究

目前國內外文獻中有關數常識診斷測驗編製的研究，皆以二階段診斷測驗方式並透過電腦進行線上施測(許清陽，2006；楊德清、李茂能，2007；Yang & Tsai, 2010)，目前此數常識二階段診斷測驗在英文期刊發表除了作者外並無其他相關研究論文發表，國內研究亦以作者所發表或指導之碩士論文為主。上述二階段診斷測驗的第一階段檢驗受測者回答數常識問題之答案，第二階段則進一步依據受測者的答案選擇相對應的原因。研究指出相較一般傳統紙筆測驗而言，二階段診斷測驗不但具備量化特點可以同時大量地蒐集資料與瞭解學生的數常識表現，同時能夠質性地進一步探究學生回答問題的解題方法與診斷學生可能產生之迷思概念(Yang et al., 2008; Yang & Tsai)。上述二階段診斷測驗系統雖然具備前述優點，然而其缺點是無法瞭解學生回答二階段問題的信心度，以進一步確認其概念的穩固程度。此外，目前國內外數學教育領域並無三階段診斷測驗之相關研究，作者乃參考科學教育領域三階段診斷測驗相關研究(e.g., Caleon & Subramaniam, 2010; Cetin-Dindar & Geban, 2011; Peşman & Eryilmaz, 2010)，發展數常識三階段診斷測驗。

基於上述論點，本研究依據上述相關研

究發展三階段診斷測驗系統，此三階段診斷測驗之第一、二階段內容要求與二階段診斷測驗相同，目的在探討學生數常識之表現與解題方法；而第三階段則為信心指數，目的在反應學生對第一、二階段回答數常識問題時信心展現的程度。此模式不僅能判斷學生數常識發展現況、探索學生對此概念之真正想法，更能進一步判斷出一個錯誤答案是否源自於迷思或是知識貧乏，於第二、三階段可收雙層測試之效，以提高診斷測驗之信效度。本研究數常識三階段診斷系統結構圖如圖1所示。

參、研究方法

一、研究設計及架構

根據研究動機與目的及相關文獻分析所得之結果，提出以下之研究架構圖(圖2)。

圖2右框框內之不同數常識程度學童，乃本研究將受測學生在數常識之表現分成4群，分別是高數常識程度組學生之成績為數常識測驗成績在前25%，中上程度組為數常識測驗成績在26%~50%，中下程度組為數常識測驗成績在51%~75%，低程度組為數常識測驗成績在75%之後。

二、研究對象

本研究選取南部某公立小四學生為樣

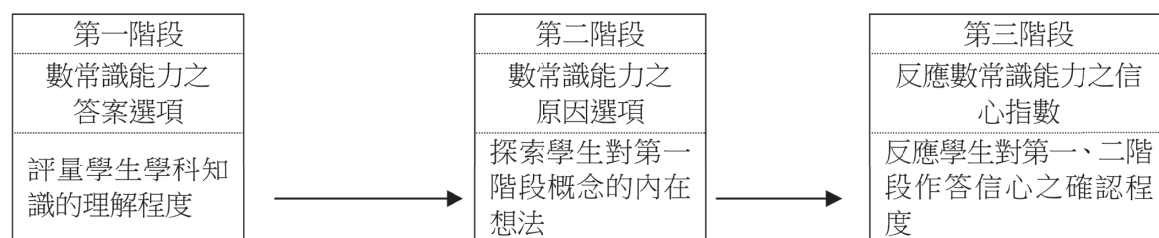


圖1：數常識診斷測驗三階段因素結構圖

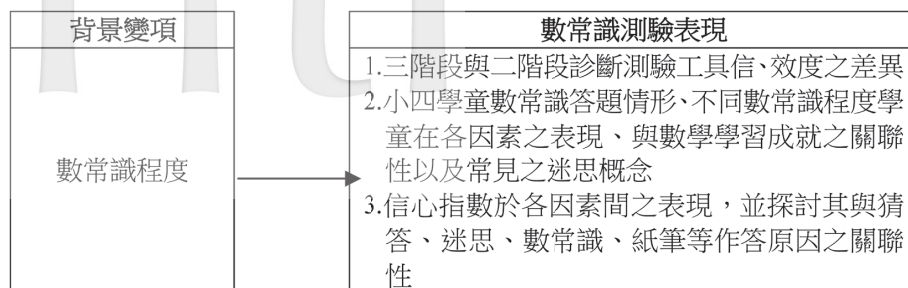


圖2：本研究之架構圖

本，班級數共五班，總計166位學生(男生有89位，女生有77位)，年齡約10歲。研究樣本之學生家庭背景大多以工、商與教為主，多數的學生除了在校求學之外，亦有夜間加強或是學習才藝的普遍現象。由於數學課程內並未正式編排數常識相關單元，由於學生的學習機會較少，因此樣學生在數常識之基本概念則較缺乏。

三、研究工具

工具內容：本研究以先前相關研究國小四年級數常識二階段診斷測驗工具作為基礎(楊德清、李茂能，2007)，加入第三階段之信心指數選項編製而成。上述數常識二階段診斷測驗工具包含四個數常識因素：(因素一)具備瞭解數字基本意義的能力、(因素二)具備比較數字大小的能力、(因素三)具備數字分解與合成的能力以及(因素四)具備判斷答案合理性的能力。每個數常識因素各有8題，共32題。數常識三階段診斷測驗工具範例說明如下(如圖3)。

上述三階段測驗之第一階段為數常識問題所對應之答案選項，接著第二階段為個答案選項所對應之原因選項(包括：數常識方法、算則方法、迷思概念與猜測)，最後是第三階段則是信心選項(包括：靠心中的直覺來判斷、紙筆計算、刪除不可能的答案、或者

用猜測的)，此部分之信心選項主要是再次用於確認學生回答完原因選項後，其內心回答問題之真正的想法為何。

評分方式：本研究因採三階段選擇的模式，第一、二階段之數常識計分乃依據先前之相關研究採配對給分的方式(Yang et al., 2008; Yang & Tsai, 2010)，而第三階段反應數常識能力之信心指數部分屬於情意範疇，但為便於進行分析之用，乃針對數常識、紙筆、迷思、猜答等四種答題原因採取對應計分標準，茲說明如表1。

在第二階段和第三階段各選項的得分分別設定為4、2、1、0，主要考量是因為本研究之主題是探討小四學生數常識的相關研究，「數常識」是本研究的主軸，因此，我們認為測驗分數應該加權，以彰顯其重要性。

四、資料收集與分析

(一)收集與整理

國小四年級數常識三階段診斷測驗共包含32題。考量測驗時間因素，將題目分成兩部分，每一部分16題，約需一節課的時間，因此，本測驗共需2節課的時間完成。本研究共選取五個班級參與測驗，考量各班彈性時間不同，經與各班導師討論後，乃於該學年度期末考後以2節課時間，各班選取各自彈性

答案選項	原因選項	確認選項
1 元	因為哥哥有 1000 元，弟弟有 999 元，所以相差 1 元 因為哥哥有 400 元，弟弟有 399 元，所以相差 1 元 因為 $4-3=1$ ，所以相差 1 元 因為是用猜的	我是用 <u>猜測</u> 的 我是靠心中的 <u>直覺</u> 來判斷答案 我是把 <u>不可能的選項刪去</u> 來判斷答案的 我是用 <u>紙筆計算</u> 的方式
2 元	因為哥哥有 1001 元，弟弟有 999 元，所以相差 2 元 因為哥哥有 41 元，弟弟有 39 元，所以相差 2 元 因為哥哥有 401 元，弟弟有 399 元，所以相差 2 元 因為是用猜的	我是用 <u>猜測</u> 的 我是靠心中的 <u>直覺</u> 來判斷答案 我是把 <u>不可能的選項刪去</u> 來判斷答案的 我是用 <u>紙筆計算</u> 的方式
10 元	因為哥哥有 40 元，而弟弟有 30 元，所以相差 10 元 因為四位數比三位數多一位數，所以相差 10 元。 因為是用猜的	我是用 <u>猜測</u> 的 我是靠心中的 <u>直覺</u> 來判斷答案 我是把 <u>不可能的選項刪去</u> 來判斷答案的 我是用 <u>紙筆計算</u> 的方式
100 元	因為哥哥有 1000 元，弟弟有 900 元，所以相差 100 元 因為哥哥有 400 元，弟弟有 300 元，所以相差 100 元 因為四位數中最小的是 1111，三位數中最大的是 999，相減大約是 100 因為是用猜的	我是用 <u>猜測</u> 的 我是靠心中的 <u>直覺</u> 來判斷答案 我是把 <u>不可能的選項刪去</u> 來判斷答案的 我是用 <u>紙筆計算</u> 的方式

圖3：數常識三階段診斷測驗工具範例

註：例題為哥哥的存款是「四位數中最小」的數字，而弟弟的存款是「三位數中最大」的數字，請問兩人相差多少錢？(左邊圈圈代表第一階段選項，中間圈圈代表第二階段選項，右邊圈圈代表第三階段選項)。

表1：計分摘要表

答題狀況					
第一、二階段 (數常識能力)					
答案選項	正確答案			錯誤答案	
	4分			0分	
原因選項	數常識	算則	迷思	猜測	不予計分
	4分	2分	1分	0分	0分
給分	8分	6分	5分	4分	0分
第三階段 (反應數常識能力之信心指數)(答案對時才給分)					
信心指數	我是靠心中的直覺來判斷答案	我是用紙筆計算的方 式	我是把不可能的選項 刪去而得到答案的	我是用猜測的	
給分	4	2	1	0	

的時間分別完成第一與第二部分的測驗，並由研究者親自指導測驗的進行。共170位小四學生參與本測驗，但有4位學生的測驗資料不完整(缺第一或第二部分之資料)，因此，實際有效樣本數為166位。本研究透過班級測驗以收集完整之測驗結果，施測完成後，先將所有受試樣本的原始資料以逐題的方式檢閱，刪除基本資料及作答有缺誤的部分，再將完整之原始資料轉入SPSS17.0、AMOS6.0統計套裝軟體與EXCEL2007文書軟體進行資料分析。

數學學習成就指的是小四樣本學生完成該學年度完整課程(上、下學期)所獲得之數學平均成績。此項成績由學校提供。

(二)資料分析

本研究將所蒐集得到的資料，採用統計分析方法進行分析，說明如下：

本研究之研究問題(一)為檢驗「四年級三階段數常識診斷測驗工具」是否具有良好之信、效度？因此以SPSS17.0進行信度分析，求得內部一致性信度，用以檢驗 α 係數能否達到.75以上，並以AMOS6.0進行各因素題組間的驗證性因素分析，求得標準化徑路係

數鍵入建構信度的計算增益集sem_stats.xls軟體中(李茂能，2006)，檢驗建構信度能否達到.70以上。

研究問題(二)-1檢驗完成國小四年級數學課程之學童數常識的表現情形？乃運用SPSS17.0統計軟體列出觀察值之平均值、標準差等描述性統計量進行探討，以便於對答題分布情形做分析描述。

研究問題(二)-2分析完成國小四年級數學課程之學童在數常識各因素之表現，及四個因素與學童數學學習成就之關聯性為何？為了檢驗數常識各因素間的差異及關聯性，除了以SPSS17.0進行單因子重複量數變異數分析，並進行事後比較外，還應用AMOS6.0結構方程模式進行多組分析比較。

研究問題(二)-3分析分析樣本學童在在回答數常識問題所使用之解題方法？此部分之資料乃透過學生回答之原因選項(猜答、迷思、數常識、紙筆等作答原因)進行次數統計。

研究問題(三)分析四年級學童數常識中信心指數於各因素間之表現，並探討其與猜答、迷思、數常識、紙筆等作答原因之關聯

性為何？乃運用SPSS17.0統計軟體列出觀察值之描述性統計量，並配合EXCEL2007文書軟體，檢驗信心指數與各因素之間的作答情形，並探討其與猜答、迷思、數常識、紙筆等作答原因之答題關聯性。

肆、研究結果與討論

一、三階段數常識診斷測驗之信、效度分析

三階段數常識診斷測驗分成四個因素，每個因素八題，共32題。為避免多分相關因素分析所可能導致非正定矩陣的出現，於是將各因素中的題目依其測驗面向採隨機的方式將2題分為1組，數常識能力(第一、二階段，以二N表示)與反應數常識能力之信心指數(第三階段，以三C表示)各分為16個題組(Byrne, 2001; Holt, 2004)。而在題組的編碼上，則是將因素一的題組1之配對編號為二N_11、三C_11，題組2之配對編號為二N_12、三C_12……，因素二的題組1之配對編號為二N_21、三C_21，題組2之配對編號為二N_22、三C_22……，以下依此規則類推，題組分配如表2所示，信、效度之分析結果說明如下：

(一)信度分析

三階段數常識診斷測驗將針對內部一致性係數、建構信度二方面來進行考驗。

1. 內部一致性係數

本診斷測驗的內部一致性係數以Cronbach's α 係數進行考驗，因此，除了以SPSS17.0視窗版統計套裝軟體考驗四個分測驗及整份測驗之外，亦針對第一、二階段(數常識能力)與第三階段(反應數常識能力之信心指數)加以分析，比較結果如表3。

由表3可以看出，三階段各分測驗的 α 值分別是 .703、.777、.785、.734，總測驗的 α 值為 .907。從結果中可發現三階段之「具備比較數字大小的能力」以及三階段之「具備數字分解與合成的能力」之 α 值接近於 .800的水準。整體而言，三階段診斷測驗模式之總測驗 α 值皆符合 .750以上的水準外。以上結果說明了本研究整份測驗具有良好之信度。

2. 建構信度

建構信度亦是聚斂效度的指標，代表測量指標能測到潛在構念的程度，最好大於0.7 (Fornell & Larker, 1981)。建構信度的估計需透過SEM或CFA的途徑，主要是利用標準化因素負荷量進行計算。本研究的建構信

表2：三階段數常識診斷測驗各分測驗題組分配表

數常識因素	數常識能力	題組			
	信心指數	1	2	3	4
因素一	二N_1	Q01、Q04	Q02、Q07	Q03、Q05	Q06、Q08
	三C_1				
因素二	二N_2	Q09、Q10	Q11、Q14	Q12、Q15	Q13、Q16
	三C_2				
因素三	二N_3	Q17、Q20	Q18、Q24	Q19、Q23	Q21、Q22
	三C_3				
因素四	二N_4	Q25、Q28	Q26、Q30	Q27、Q31	Q29、Q32
	三C_4				

表3：三階段類型數常識診斷測驗在各分測驗及整份測驗內部一致性係數摘要

測驗別	數常識因素				總測驗
	因素一	因素二	因素三	因素四	
題數	8	8	8	8	32
三階段之 α 值	.703	.777	.785	.734	.907

度，係利用AMOS 6.0統計分析軟體，進行驗證性因素分析，獲得以下數常識之四大因素(具備瞭解數字的基本意義的能力、具備比較數字大小的能力、具備數字分解與合成的能力、與具備判斷答案合理性的能力)上各題標準化因素負荷量之後，由於SPSS與AMOS程式無法計算建構信度，所以用sem_stats.xls軟體，因此將此標準化因素負荷量輸入建構信度的計算增益集表單中，計算結果如表4所示。

由表4得出三階段數常識診斷測驗之各因素建構信度值均大於 .600，表示模式內在品質佳，此外，其結果亦符合李茂能(2003)認為建構信度最好大於 .700之判別標準，可見三階段測驗具有良好的建構信度。

(二)效度分析

在效度考驗方面，以內容效度、效標關聯效度與建構效度等方面進行分析。

1. 內容效度

王文科(1999)認為如果所編製之測驗題目內容能達到教材內容所要求之目標，即兩者內容程度相符，便可稱此份測驗具有內容效度；亦即，若能透過三階段數常識診斷測驗測得個別學生之數常識表現，即可認定此測驗具內容效度。因此，透過文獻探討，針對四年級數學課程內容與數學領域能力指標

進行深入分析，並參酌楊德清與李茂能(2007)所設計之國小四年級數常識電腦化診斷測驗系統，在研究群共同討論下，為了讓數常識診斷測驗能涵蓋更多診斷訊息，在先前二階段研究的基礎下，再加上第三階段之信心指數選項來蒐集學生回答題目之信心，以協助教師依據診斷訊息調整教學，幫助學生補強不足之處。本工具為了使測驗具有診斷功能，在測驗的設計上雖採選擇題型，但將選擇題從兩個部分修正為三個部分，分別為答案選項、原因選項、信心指數選項以蒐集資料。研究者為確保本研究三階段診斷工具的可行性與適用性，在編製的過程中除了與研究同儕及國小教師討論外，為力求嚴謹並多次與數學教育學者及統計專家進行討論，確認本研究之測驗內容能夠對數常識的因素做有效的反映，並能達到診斷學生迷思概念的目的。至此，乃完成標準化「國小四年級三階段數常識診斷測驗工具」。透過以上一連串的編撰過程，本研究之三階段數常識診斷測驗工具應具有良好的內容效度。

2. 效標關聯效度

本研究採用SPSS17.0進行數常識與數學學習成就之積差相關考驗，結果顯示數學學習成就與三階相關係數為 .661，呈顯著的中度正相關，且顯著值較嚴謹。此結果可呼應文獻探討中發現兒童的數常識對於數學學習具有某種程度的影響力之論點。因此，本三階段測驗具有良好的效標關聯效度。

表4：三階段類型數常識診斷測驗之各因素建構信度

測驗別	數常識因素			
	因素一	因素二	因素三	因素四
三階段	.826	.872	.869	.845

3. 建構效度

本研究之數常識診斷工具乃根據楊德清與李茂能(2007)研究中已通過模式適配度指標評估、四因素模式適配度考驗之二階段數常識診斷測驗試題加以增修，即加上第三階段之信心指數選項來蒐集學生回答題目之真正原因。因此，研究者乃以先前研究為基礎，於AMOS模式中加入第三階段信心指數，透過驗證性因素分析，確認四因素模式在三階段診斷測驗中的適配性依然具有可行性及探討價值，足以顯示此份診斷測驗與過去研究之因素模式得以相互呼應。從因素分析的結果，可以看出三階段診斷測驗具有良好的建構效度。

此外，Kline (1998)主張所蒐集之資料在進行相關分析前，須先確保資料是否符合常態性。因此，在進行驗證性因素分析前，先透過AMOS6.0對資料進行檢驗，結果發現：四因素模式的大部分偏態skew值介於-1.047 ~ 1.158之間，皆未達偏態指數關切值3。因此，整體而論，所得之偏態值仍屬合理範圍。另就峰度值來看，全部數值皆介於-1.299 ~ 0.317之間，符合峰度值不可大於8的門檻。因此，本研究無論從偏態值或峰度值皆可說明此份診斷測驗資料之常態性。

為了對所蒐集資料檢驗其與診斷測驗中的四因素模式是否相適配，乃運用AMOS6.0針對受試資料($N = 166$)進行此四因素模式之驗證。首先研究者應用AMOS6.0繪製本研究之因素架構圖，以MI指標(< 10)作為修正依據，從不同指標中探討整份測驗中四因素模式的適配性，並採用而吳明隆(2007)建議的絕對適配度指數檢定(χ^2 , RMR, RMSEA, GFI, AGFI)、增值適配度指數檢定(NFI, RFI, IFI, TLI, CFI)、簡約適配度指數檢定(PGFI, PNFI, PCFI, CN, χ^2 自由度比, AIC, CAIC)等三方面來

探討因果模式之適配度，以下就檢定結果呈現如表5。

由表5可看出，自由度等於427，卡方值為780.081，顯著性機率值 $p < .001$ ，表示三階段數常識診斷測驗的四因素模式與實際資料並不契合。若從整體模式適配度之其他統計量則可發現，除了RMSEA值= .071 ($< .080$)、PGFI值= .588 ($> .500$)、PNFI值= .664 ($> .500$)、PCFI值= .756 ($> .500$)、 χ^2 自由度比= 1.827 (< 2.000)及AIC值、CAIC值，符合適配標準之外，其他如RMR值= 1.040 ($> .050$)、GFI值= .727 ($< .900$)、AGFI值= .663 ($< .900$)、NFI值= .771 ($< .900$)、RFI值= .734 ($< .900$)、IFI值= .881 ($< .900$)、TLI值= .859 ($< .900$)、CFI值= .879 ($< .900$)、CN值= 101 (< 200)，則未達到模式適配的標準。針對以上結果，研究者進一步探討造成此結果之內在因素，發現在自由度較大，以及樣本數相對偏少的情況下，除了RMR值、CN值相較超過適配的臨界值許多外，其餘相關統計量雖也未達適配標準，但卻已距臨界值相對接近。可見在初次加入以人內在感知為選項之信心指數後，雖適配狀況未盡完美，但應可適用。

綜合以上結果本研究所編製之三階段數常識診斷測驗工具，在信度方面達到 .750以上之內部一致性係數及 .700以上之建構信度水準外，在四種效度指標：內容效度、效標關聯效度、建構效度，亦都有不錯的表現。因此，可看出本診斷測驗工具具備良好的信、效度。

二、國小四年級學童數常識之發展情形

(一)學童在數常識各因素間答題之分布情形

表6呈現學童在數常識診斷測驗各因素之描述統計資料。

表5：三階段數常識診斷測驗之驗證性因素分析模式適配度之評估結果($N = 166$)

統計檢定量	適配的標準或臨界值	四因素檢定結果數據	模式適配度判斷
	χ^2 $p > .05$ (未達顯著水準)	780.081 ($p < .001$)	否
絕對適配度指數	RMR	< .05 1.040	否
	RMSEA	< .08 (若< .05優良；< .08良好)	是
	GFI	> .90以上 .727	否
	AGFI	> .90以上 .663	否
	NFI	> .90以上 .771	否
增值適配度指數	RFI	> .90以上 .734	否
	IFI	> .90以上 .881	否
	TLI(NNFI)	> .90以上 .859	否
	CFI	> .90以上 .879	否
	PGFI	> .50以上 588	是
簡約適配度指數	PNFI	> .50以上 .664	是
	PCFI	> .50以上 .756	是
	CN	> 200 101	否
	χ^2 自由度比	< 2.00 1.827	是
	AIC	理論模式值小於獨立模式值，且同時小於飽和模式值 982.081 < 1,056.000 982.081 < 3,470.461	是
	CAIC	理論模式值小於獨立模式值，且同時小於飽和模式值 1,397.392 < 3,227.130 1,397.392 < 3,602.044	是

表6：國小四年級學童在數常識各因素之描述統計 ($N = 166$)

因素順序	題數	平均分數	標準差	總平均分數	標準差
因素一	8	5.732	2.467	45.860	19.733
因素二	8	6.390	2.446	51.120	19.572
因素三	8	5.381	2.659	43.050	21.270
因素四	8	5.342	2.476	42.730	19.812

註：各試題之滿分均為12分。

表6數據顯示四年級學童在單題平均分數介於5.300 ~ 6.400分之間，其中又以「具備比較數字大小的能力」最高，平均得分6.390分，而「具備判斷答案合理性的能力」則最低，平均分數為5.340分。表6資料亦顯示學

童在數常識各因素測驗總分之平均數與標準差，在各因素總分最高96分的情況下，四年級學童的平均得分除了「具備比較數字大小的能力」稍高於50%水準外，其他三個因素的平均得分皆低於50%水準(45% ~ 48%)，

可見表現並不理想，與過去之研究發現相似(Yang et al., 2008; Yang & Tsai, 2010)，顯見學童的數常識能力有待加強。

上述結果顯示四年級學童在三階段數常識診斷測驗中的得分普遍不高，表現最好的成分(具備比較數字大小的能力)也只有51.120分(滿分96分，共8題)，平均單題得分6.390分(最高12分)，而在表現最差的因素(具備判斷答案合理性的能力)中，單題平均則僅有5.340分，顯示學童數常識能力不足。進一步將本研究與先前研究做比較發現：本研究在「具備判斷答案合理性的能力」的表現上最弱，正好與先前研究不謀而合(Yang, Hsu, & Huang, 2004; Yang et al., 2008; Yang & Tsai, 2010)。深入探討原因，可能和學校課程在此部分未特別規劃教授，以致於學童並不熟悉此類題目。此外，也可能與此項因素較其他成分來得靈活，其所涵蓋的數常識概念屬於較高層次的範疇；例如：「A水果3斤50元，B水果7斤100元，不用紙筆計算，請問哪一種水果比較便宜？」在缺乏紙筆計算的情況下，學童必須充分瞭解3斤50元與7斤100元所代表的意義與差別。然後，要能夠彈性靈活發展解題方法，如將3斤50元可先轉換成6斤100元，進而在總價(100元)相同的條件下，比較可以買到較多斤數者較便宜。然而多數學童解此問題時，習慣使用算則方式比較(求50除以3與100除以7)，由於無法整除，以致陷入計算的迷思。因此，該因素中的試題對必須仰賴傳統算則才能求出正確答案或對於相關概念薄弱之學童而言是較為困難的。

而在「具備比較數字大小的能力」的表現上，亦與先前研究相呼應(Yang et al., 2004, 2008; Yang & Tsai, 2010)，顯示此一成分之相似題因為是數學教科書常見的題型，故學童在此概念有一定的認知，所以能獲得較高的分數。

(二)不同數常識程度之學童在各因素問題之表現，及其與數學學習成就之關聯性

表7呈現不同數常識能力之學童在各因素上發展的差異情形。本研究將學生在數常識之表現分成4群，分別是高程度組(A)、中上程度組(B)、中下程度組(C)、低程度組(D)

表7資料顯示1.低數常識程度之學童的表現為：因素二>因素一>因素四>因素三。2.中下數常識程度之學童的表現為：因素二>因素一>因素三>因素四。3.中上數常識程度之學童的表現為：因素二>因素一>因素四>因素三。4.高數常識程度之學童的表現為：因素二>因素三>因素一>因素四。由以上結果可以得知，不同數常識程度學童普遍在因素二表現較佳。另外值得一提的是：中下、中上數常識學童在各因素間之表現，相較於低、高數常識程度學童而言，則顯得落差更大。因此，當大眾關注在資源教育與資優教育的同時，我們也更應多多瞭解中程度學童的需求。

接著探討此四種數常識程度之學童其數常識能力之表現與其數學學習成就表現之差異情形，因此，將成績轉換為 t 分數以利比較，以進一步瞭解數常識與數學學習之間相互的關聯性。表8報告不同程度學生在數學學習成就的描述統計結果。

依據變異數分析結果，將相關之推論摘要如表9所示。

上述結果顯示不同數常識程度之學童在各因素的表現優劣皆呈現，高程度>中上程度、中上程度>中下程度、中下程度>低程度，此顯示數常識能力具有預測性，也就是：整體數常識能力高(低)者即表示其個別因素能力亦高(低)。此外，不同數常識程度之學童，所呈現之數學學習成就已出現明顯的差

表7：不同數常識程度學童於數常識各組成因素間之作答表現比較表

因素別	程度別	平均數	標準差	個數
具備瞭解數字的基本意義的能力	低程度(A)	25.571	11.876	42
	中下程度(B)	39.714	12.943	42
	中上程度(C)	51.585	13.223	41
	高程度(D)	67.195	11.959	41
具備比較數字大小的能力	低程度(A)	28.738	13.204	42
	中下程度(B)	45.928	10.644	42
	中上程度(C)	57.658	10.341	41
	高程度(D)	72.829	9.861	41
具備數字分解與合成的能力	低程度(A)	22.309	10.077	42
	中下程度(B)	35.523	13.596	42
	中上程度(C)	46.439	13.312	41
	高程度(D)	68.609	14.242	41
具備判斷答案合理性的能力	低程度(A)	23.857	10.415	42
	中下程度(B)	33.571	11.820	42
	中上程度(C)	47.902	11.787	41
	高程度(D)	66.292	12.929	41

表8：不同數常識程度之學童在數學學習成就之描述統計

程度別	平均數	標準差	個數
低程度(A)	41.260	9.805	42
中下程度(B)	47.520	8.698	42
中上程度(C)	53.530	6.100	41
高程度(D)	57.970	4.969	41

註：平均數為數學學習成就滿分為100分轉化後之t分數。

異，藉由對照平均分數：高程度數常識學童之數學學習成就($M = 57.970$) > 中上程度數常識學童之數學學習成就($M = 53.530$) > 中下程度數常識學童之數學學習成就($M = 47.520$) > 低程度數常識學童之數學學習成就($M = 41.260$)，更能清楚知道擁有高數常識能力之學童，其數學學習成就相對表現較佳，反之則否。由此可知，數常識能力與數學學習成就具有密切之正相關，此與先前之研究發現一致(Jordan, Glutting, et al., 2010; Yang et al., 2008)。因此，若能幫助學童培養更豐富之數常識似乎對數學學習成就之提升會有所幫助。

(三)學童在數常識各組成因素所採用之解題方法比較

以下將各因素中學童所採用之解題方法(原因選項)百分比，以圖4並輔與直條圖說明如下。

以上結果顯示由於因素一(具備瞭解數字的基本意義的能力)為最基本之知識，因此學童較能在此因素中發揮較佳的表現(46.7%)，但其比例與「猜測」作答(43.9%)相去不遠，顯示學童在學習數學知識時，容易忽略數字基本意義的重要性，在未透徹瞭解概念之下

表9：不同數常識程度之學童在數學學習成就之單因子變異數分析摘要表

	變項名稱	平方和 (SS)	自由度 (df)	平均平方和 (MS)	F值	p	η^2	事後比較
因素一	組間	38,881.283	3	12,960.428	82.774	< .001	.605	B > A
	組內	25,365.247	162	156.576				C > B、A
	總和	64,246.530	165					D > C、B、A
因素二	組間	43,247.661	3	14,415.887	117.003	< .001	.684	B > A
	組內	19,959.929	162	123.209				C > B、A
	總和	63,207.590	165					D > C、B、A
因素三	組間	47,702.308	3	15,900.769	95.605	< .001	.637	B > A
	組內	26,943.306	162	166.317				C > B、A
	總和	74,645.614	165					D > C、B、A
因素四	組間	42,342.811	3	14,114.270	101.988	< .001	.654	B > A
	組內	22,419.526	162	138.392				C > B、A
	總和	64,762.337	165					D > C、B、A

註：A：低數常識程度、B：中下數常識程度、C：中上數常識程度、D：高數常識程度。

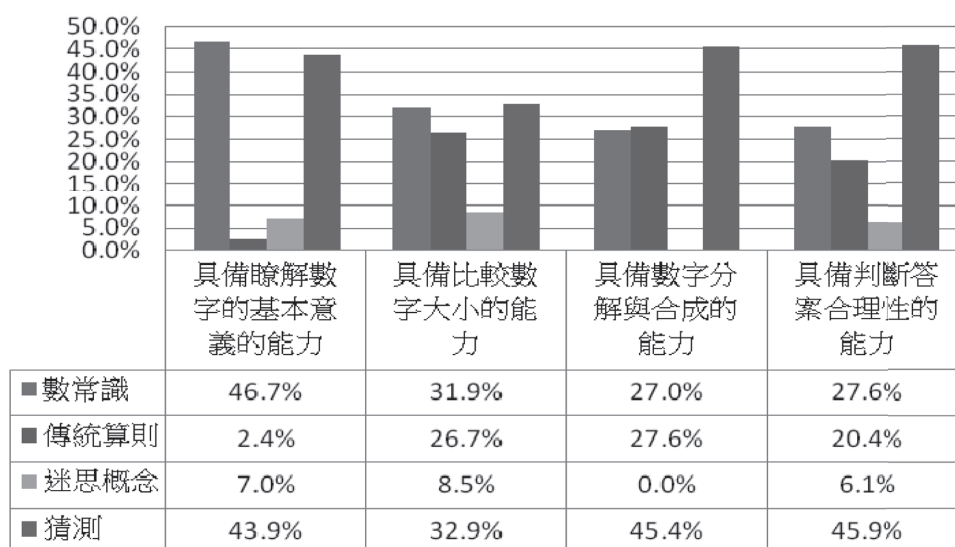


圖4：學童在數常識各組成因素所採用之解題方法比較圖表(N = 166)

註：各解題方法之% = 學生在數常識測驗四因素之各原因選項之人數和 ÷ (166 × 8) × 100%。

進行運算練習，無法全面且通盤地對數字的基本意義有正確的認知。因此，在教導「數字的基本意義」方面仍有很大的進步空間，這是第一線教師需要瞭解的部分。此外，就因素二(具備比較數字大小的能力)而言，「猜

測」作答占了最高的比例(32.9%)，但若將採用「數常識」方法(31.9%)與「傳統算則」解題(26.7%)加總，則可發現此類型之題目由於是平常較為常見，故對部分學童來說，反而會容易選擇「傳統算則」解題而忽略了採用

「數常識」方法的契機；而在因素三(具備數字分解與合成的能力)部分，可以發現採用「數常識」方法(27.0%)與「傳統算則」解題(27.6%)的比例相當接近，顯然部分學童面對此種須靈活思考之題型時，寧可使用較為有把握的運算技巧取得答案。最後，在因素四(具備判斷答案合理性的能力)部分，若將採用「數常識」方法(27.6%)、「傳統算則」解題(20.4%)，與具有「迷思概念」(6.1%)、以「猜測」作答(45.9%)個別加總進行比較，則可發現前者之和(48.0%)顯然小於後者之和(52%)，差距多達4.0%，顯示具有較高層次數常識概念的題目對學童是較為困難的，以致在此因素之表現相對較為薄弱。

三、國小四年級學童數常識信心指數之高低表現

本節主要目的在瞭解三階段診斷工具是否具有判斷學生數常識發展分類層次的優點。基於以上原因，搭配四種方法(數常識、傳統算則、迷思概念、猜測)所設計之第三階段信心指數選項(我是靠心中的直覺來判斷答案、我是把不可能的選項刪去而得到答案的、我是用紙筆計算的方式、我是用猜測的)進行深入的分析探討。

表10說明各因素中學童所採用之解題方法百分比及信心指數高低百分比。

透過學生在原因選項(數常識、算則、迷思概念、猜測)對照信心選項(直覺、紙筆計

表10：學童在數常識各組成因素所採用之解題方法及信心指數比較表(N = 166)

組成成分	解題方法	原因選項百分比	信心指數百分比	(信心 - 原因)百分比
因素一	數常識	46.7%	26.1%	-20.6%
	傳統算則	2.4%	5.2%	+2.8%
	迷思概念	7.0%	21.2%	+14.2%
	猜測	43.9%	47.5%	+3.6%
因素二	數常識	31.9%	30.9%	-1.0%
	傳統算則	26.7%	8.1%	-18.6%
	迷思概念	8.5%	23.6%	+15.1%
	猜測	32.9%	37.5%	+4.6%
因素三	數常識	27.0%	25.5%	-1.5%
	傳統算則	27.6%	10.8%	-16.8%
	迷思概念	0.0%	16.7%	+16.7%
	猜測	45.4%	47.0%	+1.6%
因素四	數常識	27.6%	28.3%	+0.7%
	傳統算則	20.4%	2.2%	-18.2%
	迷思概念	6.1%	18.8%	+12.7%
	猜測	45.9%	50.8%	+4.9%
總平均	數常識	33.3%	27.7%	-5.6%
	傳統算則	19.3%	6.6%	-12.7%
	迷思概念	5.4%	20.0%	+14.6%
	猜測	42.0%	45.7%	+3.7%

算、刪除不可能答案、猜測)之百分比比較，可以清楚瞭解兩者之差異。以下將就「單一因素解題策略vs.信心指數」及「跨因素解題策略vs.信心指數」等兩個層面進行比較分析。「單一因素解題策略vs.信心指數」目的乃在針對學童於單一因素中所採用解題策略之原因選項百分比及信心指數高低百分比進行比較；而「跨因素解題策略vs.信心指數」則是在分析學童於不同因素間以同一種解題策略進行解題之原因選項百分比及信心指數高低百分比使用情形。探討如下：

(一)在「比較單一因素解題策略vs.信心指數」方面

首先，就採用「數常識」策略的情形而言，除了在因素四(具備判斷答案合理性的能力)信心指數比例(28.3%)略高於原因選項(27.6%)之外，我們可以發現，「數常識」策略在其他三個組成因素中信心指數比例(26.1%、30.9%、25.5%)皆低於原因選項(46.7%、31.9%、27.0%)，且在因素一比例差距更高達-20.6%，因素二、因素三則落差不到2%，顯然學童在因素一(具備瞭解數字的基本意義的能力)信心指數表現遠不如原因選項；其次是使用「傳統算則」解題的學童，除了在因素一(具備瞭解數字的基本意義的能力)信心指數比例(5.2%)略高於原因選項(2.4%)之外，在其他三個組成因素的比較情形則都是信心指數(8.1%、10.8%、2.2%)遠遠落後原因選項(26.7%、27.6%、20.4%)。接著探討的是具有「迷思概念」的使用情形，關於此項答題策略的表現，學童在四個因素中皆呈現出信心指數(21.2%、23.6%、16.7%、18.8%)高於原因選項(7.0%、8.5%、0.0%、6.1%)的趨勢，且其差距都達10.0%以上；最後要探討則是以「猜測」作答解題現象，結果呈現與「迷思概念」的使用情形雷同，

也都出現在四個因素中皆呈現出信心指數(47.5%、37.5%、47.0%、50.8%)高於原因選項(43.9%、32.9%、45.4%、45.9%)的趨勢，只是兩者間的差距都不到5.0%。

表10資料顯示學童在各因素之迷思概念百分比皆相當高(因素一：43.9%，因素二：32.9%，因素三：45.4%，因素四：45.9%)，此顯示學生數常識概念薄弱，且回答數常識問題時所產生之迷思概念情形比例高，此代表未來學校數學教學應將數常識概念融入。此外，因素四之迷思概念最高，達45.9%，同時信心指數亦最高，達50.8%，此結果顯示學生在回答因四四的問題時有高度的信心，但卻不知本身概念的迷思情形；此結果與過去的研究相似，臺灣國小學生在因素四(判斷答案的合理性)的表現最差，原因是國小數學教科書內缺少相關問題的探討，由於學生接觸此類問題的機會較少，因此，表現就相對較弱。

(二)在「比較跨因素解題策略vs.信心指數」方面

根據四個因素所呈現的結果，進一步計算出學童在四個因素間所採用解題方法之原因選項及信心指數的整體平均值，分別為「數常識」策略：原因選項33.3%，信心指數27.7%，比例下降幅度為5.6%；「傳統算則」解題：原因選項19.3%，信心指數6.6%，比例下降幅度為12.7%；具有「迷思概念」：原因選項5.4%，信心指數20.0%，比例上升幅度為14.6%；以「猜測」作答：原因選項42.0%，信心指數45.7%，比例上升幅度為3.7%。基於以上結果，進一步針對具備數常識能力之答案選項正確學童，進行原因選項與信心指數作答情形的深入分析探討，並將結果整理如表11。

卡方考驗的檢定可以發現：不同數常

表11：答案選項正確之不同解題方法學童在信心指數選項作答時的表現及卡方考驗摘要表

		原因選項				χ^2 (p)
		猜測	迷思概念	傳統算則	數常識	
信心指數	我是用猜測的	個數	209	3	1	27
		百分比	6.3%	0.1%	0.0%	0.8%
	1,350.095 ($< .001$)					
	我是把不可能的 選項刪去而得到 答案的	個數	42	88	16	141
		百分比	1.3%	2.7%	0.5%	4.2%
	我是用紙筆計算 的方式	個數	68	315	190	451
		百分比	2.0%	9.5%	5.7%	13.6%
	我是靠心中的直 覺來判斷答案	個數	116	659	142	852
		百分比	3.5%	19.8%	4.3%	25.7%

識能力之學童對於信心指數作答百分比的整體差異，所呈現之 χ^2 值等於1350.095， $p < .001$ ，達到 .05顯著水準。表示不同數常識能力之學童在信心指數的作答表現具有顯著的不同。

上述結果顯示雖然因素一(具備瞭解數字的基本意義的能力)是學童在原因選項中採用「數常識」策略最高者，且高達近一半的比例(46.7%)，但在信心指數的表現卻只有四分之一的比例(26.1%)，信心指數足足比原因選項少了20.6%。如此的落差表示學童對於數字的基本意義所涵蓋的內涵並無法確實掌握，學習過程中僅著重於計算而忽略了數字概念的理解。這樣的結果也在其他三種策略皆呈現出信心指數(5.2%、21.2%、47.5%)高於原因選項(2.4%、7.0%、43.9%)的趨勢中獲得確認。透過以上的探討可提供課程設計者與執行者一個反思的方向。此外，就因素二(具備比較數字大小的能力)、因素三(具備數字分解與合成的能力)、因素四(具備判斷答案合理性的能力)而言，可以發現三者採用「數常識」策略上，信心指數與原因選項的比例相差(-1.5% ~ 0.7%)並不大，且在具有「迷思概念」、以「猜測」作答的部分也都和因素一(具備瞭解數字的基本意義的能力)有著異

曲同工之妙，即信心指數與原因選項的比例相差在這兩個部分都呈現同樣的趨勢(迷思概念：12.7% ~ 16.7%；猜測：1.6% ~ 4.9%)。此外，不同於因素一(具備瞭解數字的基本意義的能力)的部分，「傳統算則」解題在這三個因素中，都呈現信心指數(8.1%、10.8%、2.2%)低於原因選項(26.7%、27.6%、20.4%)。從以上得知，學童在面對這三類問題時，在原因選項能夠透過篩選選項敘述中所呈現的列式、算則中找到答案，但到了選答信心指數時則只能真實呈現其作答時的內在歷程。因此，確認出學童其實並不全然運用紙筆獲得答案，而只是選擇了一個在其學習過程中較具印象的選項。由此可見，學童在解數常識層次漸高的問題之能力仍是欠缺的。

若就跨因素解題策略vs.信心指數的使用情形來看，以及表11結果，可得出以下現象：1.以「猜測」作答及採用「數常識」策略兩方面，表現較為一致，無論是原因選項或是信心指數皆是比例中之前兩名，顯見學童在面對數常識題型時，相對於其他數學問題是較為陌生的。2.「傳統算則」解題在原因選項與信心指數的選答趨勢，除了在因素一(具備瞭解數字的基本意義的能力)較為類似外，其餘三個因素皆出現信心指數遠低於原

因選項的現象，差距分別為18.6%、16.8%、18.2%，可見學童可以憑印象選擇出正確的答案卻無法準確進行解題步驟，表示學童知識性的理解並不完整。3.具有「迷思概念」學童在原因選項與信心指數的選答上，恰與「傳統算則」解題的情況相反，四個因素都出現該項信心指數比例遠高於原因選項的現象，差距分別為14.2%、15.1%、16.7%、12.7%，顯然在信心指數的二次確認後，受試者較能清楚地知道本身在回答試題時的真正想法為何，結果也正代表學童之迷思概念實際比例依然不少，需要教學者進行補救教學予以改善。若就整體表現而言，可發現採用「數常識」策略及「傳統算則」解題兩方面，在信心指數確認後學童的表現並未如原因選項般，而是較為低落，可見學童在原因選項選答時，往往能以消去法或猜測獲得正確的答案，但到了選擇信心指數時，則以更貼近事實的答案呈現真實的想法，因此有如此之現象。而當討論到具有「迷思概念」及以「猜測」作答兩個部分時，所看到的結果則和採用「數常識」策略及「傳統算則」解題恰好背道而馳，也就是信心指數在確認後學童的表現都高過原因選項。由此可知，學童在數常識能力的表現並不如原因選項選答所呈現的，而是更弱，這也是研究者與教學者所該正視的，如何降低這兩類的比例是未來努力的方向。

再從深入探討答案選項正確而採用不同解題方法之學童，在信心指數選項作答時的表現來看，除了採用「迷思概念」解題的學童無法正確在信心指數選項中呈現出一致的想法之外，其餘採用三種策略的學童皆能以第一順位使用與其原因指數相互搭配之信心指數選項，並且都鮮少用猜測的，比例皆小於不到1%，可見在具備數常識能力學童的

身上，除了在原因選項與信心指數的作答上發揮了雙層測試的實際效果，也展現了這些學童數學認知上的穩固。此外，以「迷思概念」策略解題的學童在面對數常識問題時，並無法確實掌握答題的關鍵，常常是採取並不完備的觀點對試題進行回應。因此，對於具有「迷思概念」的學童，教學者有必要提供更多的關注與協助，以幫助他們穩固數學學習的基礎。

透過以上的討論可以發現，雖然學童之數常識能力的表現從原因選項的作答結果已經可以獲得初步的診斷訊息。但研究結果告訴我們，若能再以信心指數的回饋資訊進一步確認學童內心的想法，則所得結果就會更貼近真實的回答狀況。透過此研究全體受試樣本所得到的分析結果，研究者發現，採用「數常識」策略及「傳統算則」解題的比例在原因選項中分別是33.3%、19.3%，皆高於信心指數中之27.7%、6.6%，若以合併比較而言，前者為52.6%，後者為34.3%，差距多達18.3%，顯見學童之數常識能力並非如原因選項中所呈現的那麼好。正確來看，僅有三分之一的學童數常識能力良好，這樣的訊息值得我們加以關注；從另一個層面進行探討發現，具有「迷思概念」及以「猜測」作答的比例在原因選項中分別是5.4%、42.0%，皆低於信心指數中之20.0%、45.7%，若以合併比較而言，前者為47.4%，後者為65.7%，差距同樣是多達18.3%。由此可見學童之數常識能力遠比原因選項中所呈現的來得差，也就是說，約有三分之二的學童在數常識能力的表現上有待加強。藉由此發現，確實提供了我們一個反思的契機。若再從答案選項正確而採用不同解題方法之學童進行分析，則可發現此類型學童在原因選項與信心指數的表現上呈現一致的結果。因此，此份研究所編

製之三階段數常識診斷工具便是希望可以發揮它應有的功效，即是不僅能對答案來源(數常識方法、迷思概念、傳統算則、猜測)做到進一步之判斷，於第二、三階段收雙層測試之效，以提高診斷測驗之信效度。更重要的是，它能幫助教學者判斷學生數常識發展的分類層次、探索學生對此概念之真正想法，並以此為根據協助學生進行學習。

伍、結論、貢獻與建議

一、結論

根據前述研究所得之結果與討論，歸納出以下幾點結論：

(一)「四年級三階段數常識診斷測驗工具」具有良好的信、效度

結果顯示本研究所編製之三階段數常識診斷測驗工具，在信度方面內部一致性係數達到 .90 以上之及建構信度在 .80 以上之水準外，在四種效度指標：內容效度、效標關聯效度、建構效度，亦都有良好的表現。因此，本診斷測驗工具具備良好的信、效度。

(二) 四年級學童於數常識之表現

1. 四年級學童於數常識各組成因素間之發展具有顯著差異且表現有待加強

研究者發現學童在數常識各組成因素間的發展具有顯著差異。其中在「具備比較數字大小的能力」因素中表現最好；而在「具備判斷答案合理性的能力」因素中表現則最差。此外，四年級學生之數常識平均答對率約 50%，此結果顯示樣本學生之數常識有待加強。

2. 不同數常識程度之四年級學童在數常識各因素間的發展上具有顯著差異

研究結果顯示不同數常識程度(低、中下、中上、高)之學童在數常識各組成因素之能力上有顯著差異。整體而言，不同數常識程度學童普遍在因素二表現較佳。至於中下、中上數常識學童在各因素間之表現，則相較於低、高數常識程度學童來說，顯得落差更大。

3. 數常識能力與學童之數學學習成就具有高度正相關性

擁有高數常識能力之學童，其數學學習成就相對表現較佳，反之則否。由此可知，樣本學生之數常識能力與數學學習成就具有密切之正相關。

4. 樣本學童在使用數常識方法解題之比率偏低

研究發現樣本學童在使用數常識方法回答問題之原因選項上偏低，此結果顯示學生之數常識能力有待加強。

(三) 第三階段之信心指數能夠進一步探索學生回答問題之真正想法

研究發現學童在數常識能力的表現並不如原因選項選答所呈現的，顯然在信心指數的二次確認後，受試者較能清楚地知道本身在回答試題時的真正想法為何，由此可見，信心指數於數常識三階段診斷測驗具有一定的參考意義存在。整體表現而言，採用「數常識」方法及「傳統算則」解題的比例在原因選項中分別是 33.3%、19.3%，皆高於信心指數中之 27.7%、6.6%，顯見學童之數常識能力並非如原因選項中所呈現的那麼好；而具有「迷思概念」及以「猜測」作答的比例在原因選項中分別是 5.4%、42.0%，皆低於信心指數中之 20.0%、45.7%，由此可見學童之數常識能力遠比原因選項中所呈現的來得弱。

綜合以上結論，研究者發現數常識三階段診斷測驗工具在診斷學童數常識能力之表現時，不但能透過原因選項發現學童的錯誤想法，更能將學童主要的內在思路呈現出來，可見加入了第三階段「信心指數」在此研究中發揮了它應有的雙層測試效用。除此之外，研究者經由與先前研究(楊德清、李茂能，2007；Yang et al., 2004, 2008; Yang & Tsai, 2010)之結果做分析比照，發現除了在診斷工具上皆具有良好的信、效度，檢驗學生的數常識表現以及回答問題所使用之方法外，同時能夠進一步瞭解學生回答問題之內心的真正想法。

研究者認為，「信心指數」不僅可運用於診斷學童數常識之能力外，更重要地，它能廣泛地運用於數學教育中的各個學習領域中，例如：資優學童診斷、低成就學童診斷等，它的作用猶如驗算，可以對所得結果做二次確認的動作，避免學童因為猜答或抄寫等方式取得正確答案，卻未被教師發現而無法獲得補救教學。因此，本研究之最大啟示就是：當我們探究學童某些數學能力的同時，是否也該著重在瞭解學童內心學習狀態，避免為了配合學習進度的推展，而忽略了學童原本脆弱的數學學習心態。倘若能藉由「信心指數」的分析瞭解，以進一步瞭解學生真正的學習狀況。

二、本研究對數學教育之主要貢獻

本研究對數學教育之主要貢獻有三點：

(一)創新性：目前在數學教育領域並未有三階段診斷測驗之相關研究，因此，本研究發展之三階段診斷測驗可視為數學教育領域內相關研究之先驅，除了可運用於學童數常識之研究外，亦可做為其他數學知識探究之參考依據。

(二)實務性：本三階段數常識診斷測驗工具不但具有二階診斷測驗之功效，透過信心指數的檢驗，可以進一步的瞭解學生回答問題時的內心想法，以發揮雙層測試的功效。

(三)推廣性：透過電腦化網路測驗，國內任何地方只要具備電腦與網路連線，學校老師即可對學生進行測驗，並透過後臺資料的輸出，以進一步瞭解學生的數常識能力與信心，以做為未來教學或補救教學之參考依據。

三、建議

依研究發現提出下列建議：

(一)數常識應融入數學課程與教學中，以協助學生發展數常識能力

本研究結果發現參與測驗學生數常識答對率偏低，且普遍具有高比率的迷思概念。此結果顯示國內學生數常識能力有待加強，特別是日前國際間許多研究皆強調幫助學生發展數常識的重要性與必要性(Berch, 2005; Dehaene, 2011; Dyson et al., 2013; Jordan, Glutting, et al., 2010; McIntosh et al., 1997; Yang & Li, 2013; Yang & Tsai, 2010)。然而目前國內國小數學教科書在數常識相關的教學活動並不多，因此，建議教科書設計者應將數常識相關活動融入國小數學教科書，以協助學生發展數常識能力。

(二)協助國小教師善用本系統，以增強教師瞭解其學生之數常識與迷思概念

本研究所發展之三階段診斷測驗系統，不僅能有效診斷學生之數常識表現，並對答案來源(數常識方法、迷思概念、傳統算則、猜測)做到進一步之判斷，於第二、三階段收雙層測試之效，以提高診斷測驗之信效度。

更重要的是，它能幫助教學者判斷學生數常識發展的分類層次、探索學生對此概念之真正想法與信心，並以此為根據協助學生進行學習。因此，建議舉辦相關之研習活動以協助教師能夠善用本系統，以協助教師更能掌握學生之學習成效。

(三)未來加強相關之研究

未來三階段診斷測驗之相關研究可增加研究樣本數目，以及選取更具全國代表性之樣本，以進一步驗證數常識三階段診斷測驗之可行性。

誌謝

本研究蒙科技部(原國科會)專題計畫補助，計畫編號MOST 104-2511-S-415-002-，特誌申謝；文中所提論點純屬作者個人之意見，並不代表國科會立場。作者衷心感謝總編輯暨審查委員對本文所提供之寶貴意見，由於您的協助，方能使本文能夠以最佳的方式呈現。

參考文獻

1. 王文科(1999)。教育研究法(五版)。臺北市：五南。
2. 吳明隆(2007)。結構方程模式AMOS的操作與應用。臺北市：五南。
3. 李茂能(2003)。圖解式結構方程模式軟體AMOS之簡介與應用。國民教育研究學報，11，1-39。
4. 李茂能(2006)。結構方程模式軟體Amos之簡介及其在測驗編製上之應用：Graphics & Basic。新北市：心理。
5. 教育部(2008)。國民中小學九年一貫課程綱要：數學學習領域課程綱要。臺北市：作者。
6. 許清陽(2006)。國小學童數感理論模式建構與電腦化數感診斷測驗系統之研究。未出版之博士論文，國立高雄師範大學教育學系，高雄市。
7. 林文生、鄔瑞香(1999)。數學教育的藝術與實務——另類教與學。新北市：心理。
8. 楊德清(2002)。從教學活動中幫助國小六年級學生發展數字常識能力之研究，科學教育學刊，10(3)，233-260。
9. 楊德清、李茂能(2007)。九年一貫國小階段數常識電腦化診斷測驗系統之開發與應用(NSC96-2521-S-415-001)。臺北市：行政院國家科學委員會。
10. 劉曼麗(2006)。整數數感融入國小四年級數學科教學之研究。科學教育學刊，14(2)，121-147。
11. Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities*, 38(4), 333-339.
12. Byrne, B. M. (2001). *Structural equation modeling with Amos: Basic concepts, application, and programming*. Mahwah, NJ: Lawrence Erlbaum.

13. Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*, 32(7), 939-961.
14. Cetin-Dindar, A., & Geban, O. (2011). Development of a three-tier test to assess high school students' understanding of acids and bases. *Procedia Social and Behavioral Sciences*, 15, 600-604.
15. Chrysostomou, M., Pitta-Pantazi, D., Tsingi, C., Cleanthous, E., & Christou, C. (2013). Examining number sense and algebraic reasoning through cognitive styles. *Educational Studies in Mathematics*, 83(2), 205-223.
16. Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed.). New York: Oxford University Press.
17. Dunphy, E. (2007). The primary mathematics curriculum: Enhancing its potential for developing young children's number sense in the early years at school. *Irish Educational Studies*, 26(1), 5-25.
18. Dyson, N. I., Jordan, N. C., & Glutting, J. (2013). A number sense intervention for low-income kindergartners at risk for mathematics difficulties. *Journal of Learning Disabilities*, 46(2), 166-181.
19. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
20. Holt, J. K. (2004). *Item parceling in structural equation models for optimal solutions*. Paper presented at the 2004 Annual Meeting of the Mid-Western Educational Research Association. Columbus, OH.
21. Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*, 20(2), 82-88.
22. Jordan, N. C., Ramineni, C., & Watkin, M. W. (2010). Validating a number sense screening tool for use in kindergarten and first grade: Prediction of mathematics proficiency in third grade. *School Psychology Review*, 39(2), 181-195.
23. Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford.
24. Li, M.-N. F., & Yang, D.-C. (2010). Development and validation of a computer-administered number sense scale for 5th-grade children in Taiwan. *School Science and Mathematics*, 110(4), 220-230.
25. Lock, R. H., & Gurganus, S. (2004). Promote number sense. *Intervention in school and clinic*, 40(1), 55-58.
26. Markovits, Z., & Sowder, J. T. (1994). Developing number sense: An intervention study in grade 7. *Journal for Research in Mathematics Education*, 25(1), 4-29.

27. McIntosh, A., Reys, B. J., & Reys, R. E. (1992). A proposed framework for examining basic number sense. *For the Learning of Mathematics*, 12, 2-8.
28. McIntosh, A., Reys, B. J., Reys, R. E., Bana, J., & Farrel, B. (1997). *Number sense in school mathematics: Student performance in four countries*. Perth, WA: Mathematics, Science & Technology Education Centre, Edith Cowan University.
29. Menon, R. (2004). Elementary school children's number sense. *International Journal for Mathematics Teaching and Learning*. Retrieved January 10, 2008, from <http://www.cimt.plymouth.ac.uk/journal/default.htm>
30. National Council of Teachers of Mathematics. (2000). *The principles and standards for school mathematics*. Reston, VA: Author.
31. Peşman, H., & Eryilmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *Journal of Educational Research*, 103(3), 208-222.
32. Renner, C. H., & Renner, M. J. (2001). But I thought I knew that: Using confidence estimation as a debiasing technique to improve classroom performance. *Applied Cognitive Psychology*, 15(1), 23-32.
33. Reys, R. E., & Yang, D.-C. (1998). Relationship between computational performance and number sense among sixth- and eighth-grade students in Taiwan. *Journal for Research in Mathematics Education*, 29(2), 225-237.
34. Sood, S., & Jitendra, A. K. (2007). A comparative analysis of number sense: Instruction in reformed-based and traditional mathematics textbooks. *Journal of Special Education*, 41(3), 145-157.
35. Sowder, J. T. (1992). Estimation and number sense. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 371-389). New York: Macmillan.
36. Sowder, J. T., & Kelin, J. (1993). Number sense and related topics. In D. T. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp. 41-57). New York: Macmillan.
37. Stankov, L., & Crawford, J. D. (1997). Self-confidence and performance on test of cognitive abilities. *Intelligence*, 25(2), 93-109.
38. Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 557-628). Charlotte, NC: Information Age.
39. Yang, D.-C., Hsu, C. J., & Huang, M. C. (2004). A study of teaching and learning number sense for sixth grade students in Taiwan. *International Journal of Science and Mathematics Education*, 2(3), 407-430.
40. Yang, D.-C., & Li, M. N. (2013). Assessment of animated self-directed learning activities mod-

ules for children's number sense development. *Journal of Educational Technology and Society*, 16(3), 44-58.

41. Yang, D.-C., Li, M. N., & Lin, C. I. (2008). A study of the performance of 5th graders in number sense and its relationship to achievement in mathematics. *International Journal of Science and Mathematics Education*, 6(4), 789-807.
42. Yang, D.-C., & Tsai, Y. F. (2010). Promoting sixth graders' number sense and learning attitudes via technology-based environment. *Educational Technology and Society*, 13(4), 112-125.
43. Yang, D.-C., & Wu, W. R. (2010). The study of number sense realistic activities integrated into third-grade math classes in Taiwan. *Journal of Educational Research*, 103(6), 379-392.

The Development and Application of Number Sense Three-Tier Test for Fourth Graders

Ming-Chieh Yu¹, Fred Mao-Neng Li² and Der-Ching Yang^{3,*}

¹Tainan Municipal Yongkang District Dawan Elementary School

²Department of Education, National Chiayi University

³Graduate Institute of Mathematics and Science Education, National Chiayi University

Abstract

The purposes of this study were to develop and apply the number sense three-tier test for fourth graders. One hundred and sixty six fourth graders from a public elementary school in southern Taiwan were selected to participate in this study. The findings were as follows. First, the Number Sense Three-tier Test for fourth graders has a good reliability and validity. Second, there is a significant difference among the four factors of number sense and sample students performed poorly on number sense. There is a significant difference on number sense performance among different levels of student achievement and there is a high positive relationship between mathematics achievement and number sense performance. The methods used by sample students were different for each number sense component. Third, the confidence index used in this study has important implications. Finally, the suggestions were discussed.

Key words: Three-Tier Diagnostic Test, Grade Four, Number Sense

* Corresponding author: Der-Ching Yang, dcyang@mail.ncyu.edu.tw