



數學素養導向—— 應用真實評量於國小學生數感概念

謝如山^{*}

摘 要

服膺108課綱數學素養國家政策，本研究目的為依數學素養導向，即以真實評量探究國小學生在數感概念的表現。真實評量的優勢在於與生活結合的真實性。參與者為桃園縣某國小四至六年級學生，四年級有93位，五年級有95位，六年級有95位，共283位學生參與。統計方法採用獨立樣本 t 考驗與順序性邏輯迴歸進行分析；質化方法依評分標準的四等分制進行學生解題策略分析。依 t 考驗檢視四至六年級學生，三題的題目難度均有顯著差異。依順序性邏輯迴歸分析結果，四至六年級學生於題1並未出現差異，但於題2與題3，四至六年級學生則出現不同的答對機率，即學生的表現依年級愈高而有較好的表現。於質化分析結果，學生的解題策略於題3出現較大的變化，四至六年級學生均出現較多猜測性的解題策略，其原因可能為學生對於生活情境的概念不理解所致。本研究建議有三：首先，教師應將真實的生活情境帶入數學課程，以落實數學素養的學習趨勢；其次，應考量資料特性採用適當的統計模式，如順序性邏輯迴歸分析，能推算出四至六年級學生於題目分數的比率；最後，從多重計分的標準可釐清學生的認知學習概念。

關鍵詞：真實評量、順序性邏輯迴歸、數感概念、數學素養

* 謝如山，國立臺灣藝術大學師資培育中心教授

E-mail: t0336@ntua.edu.tw

投稿日期：2022年5月4日；採用日期：2022年6月28日

The Orientation of Mathematics Literacy: Utilizing Authentic Assessment on the Concept of Number Sense for Elementary Students

Ju-Shan Hsieh^{*}

Abstract

To meet the trend of 108 mathematics curriculum standards, the aim of the study is to fulfill the authentic assessment to explore primary students' performance on the concept of number sense. The advantage of authentic assessment is to connect with life context beyond school. Total 283 participants were from 93 fourth graders, 95 fifth graders, and 95 sixth graders. Students were sampled from one elementary school in Taoyuan city. For the quantitative methodology, the *t*-test for independent sample methods was taken to explore whether students have different performances. Additionally, the ordinal logistic regression was applied to analyze the response rate among the three questions for different grade levels. The qualitative methodology was adopted to analyze solving strategies of students in different questions by the four levels of coding criteria. The results of the ordinal logistic regression reveal that fourth to sixth grade levels are indifferent in item 1, but response rates are diverse in item 2 and item 3. The results indicate that students' performances are better while their grade levels are higher. In qualitative analysis, the results expose the variant

^{*} Ju-Shan Hsieh, Professor, Teacher Education Center, National Taiwan University of Arts
E-mail: t0336@ntua.edu.tw
Manuscript received: Mar. 4, 2022; Accepted: Jun. 28, 2022

guessing-problem solving strategies among different graders in item 3. The possible reason can be ascribed to their incomprehension. Three suggestions are recommended in education fields. Firstly, to meet the trend of mathematics literacy, primary teachers should design the real-life situation into mathematics lessons. As a second, based upon the nature of data, it is crucial to take advantage of the proper statistical model. Finally, the problem-solving strategies of students can be corresponded to their cognitive learning levels via the authentic assessment.

Keywords: authentic assessment, ordinal logistic regression, number sense, mathematics literacy

壹、緒論

教育部（2014）於《十二年國民基本教育課程綱要總綱》（以下簡稱108課綱）提出未來教育的方向——培養國民具備全人教育的精神，以「自發」、「互動」及「共好」為理念，以「成就每一個孩子——適性揚才、終身學習」為願景。而數學領域課程的理念與願景，從數學是一種語言、也是一種人文素養出發，課程設計應提供每位學生有感的學習機會，培養學生正確使用工具的素養。為使學生具備數學素養，數學學習應依下列五項理念進行教學與評量：

一、數學是一種語言，宜由自然語言的題材導入學習，而自然的語言在於易使學生理解數學學習，透過實例的操作與解說，瞭解概念與算則後，再進入抽象理論的學習。

二、數學是一種實用的規律科學，教學宜重視跨領域的統整，使數學廣泛地應用在日常生活的需求，如統計、會計、經濟、金融、天文、科技等理論與產業領域。數學是一門善於處理規律的科學，故數學應與生活結合。

三、數學是一種人文素養，宜培養學生的文化美感，數學能成為一種與自然界對話的語言，是經過人類數千年來一連串探究、歸納、臆測與論證的成果。人類各種族文明造就出不同的思維文化，認識數學的文化面向，不僅有助於讓數學學習從工具性層次延伸到智識性層次，也更彰顯數學知識的人文價值，達到「適性揚才」與「終身學習」的教育目標。

四、數學應提供每位學生有感的學習機會，數學與其他領域的差異在於其結構層層累積，其發展既依賴直覺，又需要推理。同齡學生的數學認知發展又有個別差異，學習者若未能充分理解前一階段的概念，必然影響後續階段的學習。

五、數學教學應培養學生正確使用工具的素養，工具對於數學教學助益極大。除了傳統教具如圓規、三角板、方格紙等，資訊時代的計算機（calculator）、電腦（computer）、網路、多媒體、行動工具等都是有用的學習工具。實施時機以國民中學及高級中等學校教育階段為宜，教師並可在適當時機使用電腦輔助教學。

依數學素養學習的五項理念，數學教學與評量的趨勢應以學生生活為主軸，應用生活的自然語言，培養數學學習的思維，經由循序漸進的有感經驗，協助學生使用科技的工具，達到終身學習的目標。

其實教育可分為三個關鍵的向度：一為課程，從規範教師教學與學生學習的基礎，達到教育的宗旨；二是從課程與評量的設計方式，經由教師教與學的過程以達成教育目的；三要能檢核學生在課程的學習成效，才能評量在教學與學習的成效（Merta et al., 2015）。

評量在教育領域之所以扮演重要的角色，主要在於評量能反映教師教學與改善學生的學習（Darling-Hammond et al., 2020）。再者，評量是教室學習的重點之一（Mundia, 2020），如教師所設計的活動在學習上便相當的重要（Ahmad, 2020）。而活動的設計就是要應用真實評量的概念，這一種評量方式就是教師在學生學習的過程中，要能協助學生的學習（Barrio et al., 2015），且能檢視教師教學與學生學習（Azim & Khan, 2012）。

真實（authentic）的定義即為實際的、真正的，就如同字面上的意義，而真實評量就是在生活中的評量（Palm, 2008）。Palm（2008）比較真實評量與實作評量的不同，在於學校就是生活，課堂的實際操作及學生的學習就是真實的生活情境，所以真實評量不僅是評量，而是教學、課程與生活結合的範疇。因此，真實評量的特性在於設計學生生活的情境，是否能理解生活的語言，使學生有感地經驗學習，與未來數學素養的評量目標一致。相對地，實作評量則注重學生在實作的表現，實作評量的目標在於以明確步驟觀察學生在複雜問題的表現，如學生於數學課的測量過程是否精準無誤等。

游自達（2019）提出數學素養導向是將知識、能力和態度整合運用在情境化、脈絡化的學習過程，注重學習歷程、方法與策略，從實踐的表現來評量學習的成效。因此，素養導向教學有四個基本原則（林永豐，2018），分別為：一、整合知識、能力與態度；二、情境脈絡的學習；三、強調學習歷程、學習方法及策略（學會學習）；四、在生活及情境中整合活用。依此可看出數學素養導向趨勢與真實評量的目標一致。

臺灣學童普遍缺乏數感概念，主要原因為大部分的學校教學仍以計算為導向，即教師在教數學時，還是要求學生反覆不斷的練習，而反覆不斷的練習則會影響學生在數感概念的發展（Yang & Wu, 2010）。

其實數學的學習重點應放在學生的理解與應用，而評量的目的僅在確定學生不會的觀念，但是學生的觀念並未透過課堂間的教學進行概念的瞭解。再者，許多學生並無法靠著自行理解來提升數學成就。因此，如何經由真實的活動來設計教學的情境，因應數學素養導向的教學，會是未來的趨勢。

為能結合數學與生活，設計數學素養的評量內涵為未來教育的發展趨勢，因國內對於真實評量在數感概念的應用文獻不多，故本研究目的為應用真實評量的情境設計以探討國小學生不同年級於數感概念的表現是否有所差異？

貳、文獻探討

以下依真實評量與數感概念的相關文獻分別論述。

一、真實評量的相關文獻

Archbald與Newmann（1988）將真實的用語引入教學與評量，主要原因是他們發現傳統的教室忽略了很多可應用於學校之外情境的需求，他們認為評量不僅是用於檢測各科的學習成就，更重要的是要瞭解受測者對於意義化使用的熟練度。而對生活情境的熟練度，可提升智力的優質性，且對許多具貢獻性的人類成就更顯重要。

Newmann（1997, p. 361）對真實的解釋是智力品質的關鍵面向，它的範圍是一個課程、一個評量活動或是學生表現的一個範例，可以展現出所規範的探究活動，經由建構知識的過程，他們的學習遠超過於學校生活中所有的意義與價值。

真實評量是一種應用實際生活情境的評量方法（Hodgman, 2014），有不同的評量面向，如撰寫紀錄、檔案評量、觀察方法或是活動方式（Mundia, 2020; Sasmaz-Oren & Ormanci, 2011）。重要的是，真實評量是真實發生在生活的情境，情境具有挑戰性，也會更有難度，每個活動也互有關聯，可依此發展出測驗與活動（Martin et al., 2016; Olfos & Zulantay, 2007）。

Wiggins（1998）對真實評量提供了一些指標：一、真實評量需有真實性的特色；二、需要學生來「做」這個主題；三、其他人可以複製或模擬來進行測驗；四、對於一個複雜的問題，能有效地評量學生多方面的技巧；五、學生能在教學或課程中有更好的練習或學習機會。

Stenmark（1991, p. 16）定義真實評量是學校數學課程中一種特別的評量方式，其包括了在數學上的特殊意義，還特別指出數學的規範使真實評量在數學有不同的定義。數學注重「過程與結果」，並且將校外的生活情境引入，Stenmark指出，數學的規範需用真實評量的過程來體現，且學生可從真實評量看到結果的價值。Stenmark也對於數學在真實評量的應用做出釐清，重點在於數學活動的探究，

需發現規律後歸納、進行演繹，產生數學模式，經由辯證、簡化計算以及數學活動的重組，之後應用於生活（Stenmark, 1991, p. 3）。

Gulikers等人（2004）說明學習與評量是一體之兩面，兩者相互影響。為了能使數學與生活情境結合，在教學與評量均應設計真實的情境，以利於學生學習。Gulikers等人亦特別指出在真實評量中，讓學生自己從真實情境中學習，想出解題的策略以解決問題特別重要，而這些情境要能具備代表性、與學習的相關性與意義化的特性等。Vu與Dall’Alba（2014）則認為真實評量在於給予學生一個能投入解決的問題，去思考在我們周遭世界所發生的事件，而真實性就是能使學生成為社會人的主要途徑。

真實評量有三個主要的特性（Palm, 2008）：一、過程與結果。真實評量可以看到認知的歷程，學生如何建構知識、歷程的展現與成果的產出；有些教師認為結果很重要，有些教師則認為過程與結果一樣重要；而真實評量重視的是歷程與結果，學生所專注的解題歷程可將生活與學校結合，而此目標與國家的課程目標一致。二、情境。在真實的情境下，學生可以進行教學活動，這些活動會在生活中發生；當課堂給予時間限制時，意味著學生在真實生活中可能面臨到同樣的情況，會在侷限的時間內做出決定，所以情境就是生活；在學習與教學的過程，真實評量可以促進學生有效的學習。三、圖表的說明。學生可以將題目中的圖表說明應用在學校之外的生活情境，也就是說，學生可將生活的資訊表格化，以更有效地解決問題。

例如，經濟合作暨發展組織（Organization for Economic Co-operation and Development [OECD], 1999, p. 51）所進行的「國際學生能力評量計畫」（Programme for International Student Assessment, PISA），對數學素養（mathematical literacy）的定義便著重於圖表的情境設計，而不是採用一般的數學學習題。而這些圖表的情境需來自於生活，也就是真實性的概念。McFeetors等人（2021）探討四至五年級學生應用真實評量於幾何活動的情境設計，經由分組合作學習與創意的空間活動，學生在幾何活動的評量表現有顯著的成長。

真實評量發展至今已三十多年，其目的在於應用實際發生的生活情境進行設計，具備真實的主題情境，以使學生能將學校所學應用至生活，然而，國內於真實評量的應用較少，雖有從科技面來探討真實評量的應用（陳立庭，2017）與幼兒園的教學評量（張靜文，2012），但僅有一篇探討社會科的研究（洪德健，2001）。故本研究方向為國內應用真實評量概念的實證研究，將能為數學領域提供

具體的研究成果。

二、數感概念的相關文獻

美國數學教師協會（National Council of Teachers of Mathematics [NCTM], 2000）定義數感為學生能拆解與組合數字，能使用數字如1/2或100作為參考，來解決相關的數學問題。學生要能瞭解十進位數字系統來進行估測，以理解數字的相關大小與絕對性。數感概念包括數數、瞬間計數、理解數字規律、比較數字大小、估測與關鍵數字的拆解與轉換等（Berch, 2005; Jordan et al., 2006; Lago & DiPerna, 2010）。很多學生從學齡前即已具備數感概念（Gersten & Chard, 1999），有些學者發現數感概念甚至從嬰幼兒期就已經開始（Antell & Keating, 1983; Wynn, 1992）。

有關數感的重要，Yang與Wu（2010）提出四個理由：（一）數感可以表現出彈性、創造、合理與效率的思考模式（Dunphy, 2007）。例如，算式 $12 + 9 = ?$ ，一年級學童在做加法運算時，可以運用湊十的特性來解決，應用 $12 + 10 - 1$ 的方式來得到答案。（二）對於數量、數字、運算的主要概念，能有效率及有彈性地應用至生活中。例如，當一台玩具汽車需要599元，一個音樂盒需要399元時，小明要買兩個，需花多少張100元？學生要能理解599元需要幾張100元的概念。（三）成年人的數學思考發展也與數感有直接的關係。（四）過度的計算不僅會限制孩子在數學的思考與理解，也會阻礙孩子在數感上的發展（Burns, 1994; Kilpatrick et al., 2001; Reys & Yang, 1998; Yang & Li, 2008）。例如，學生對於 $25 \times 0.45 = ?$ ，答案應該是25的一半左右，學生的答案如果是1.125，則他們缺乏數感的能力。Yang與Wu亦指出傳統的數學教學若過於重視紙筆運算，會使學生缺乏數感能力。

在小學階段的相關研究顯示，學生在數感概念的表現於不同的數感內涵（王郁芳，2018；朱麗雯，2005；Can & Özdemir, 2020; Yang & Li, 2008; Yang et al., 2008; Yang et al., 2008）。朱麗雯（2005）探討三至六年級的數感概念結構，發現三年級學生答對率為44%，四年級為51%，五年級為62%，六年級為67%，呈現年級愈高，數感概念結果愈佳的趨勢。Can與Özdemir（2020）對國小四年級學生設計情境題與非情境題，以比較學生於數感概念表現的差異，結果顯示四年級學生於非情境題的表現較佳，且學生使用參考點（reference point）以判斷數字大小比未使用參考點的表現較佳，由此也發現數學課程應結合生活情境以促進學生數感概念的必要性。Yang與Li（2008）探討五年級學生在判斷數字大小的表現有所差異，

研究結果發現程度最差的學生不知要如何估算計算題的結果，而這些估算能力，有些三年級學生即已具備。王郁芳（2018）探討國小六年級學生於數感概念的解題錯誤類型分析，研究結果發現學童數感錯誤形成的原因，如數感策略的誤用、猜錯答案、習慣列式的精算、消極猜測、缺乏數學推理、缺乏檢核錯誤的能力等。

Yang與Sianturi（2019）進行六年級的數感概念，將數感概念區分為五個標準向度：（一）要能理解基本的數字運算，如在四年級進行萬位數字的運算，學生要能對於加減法計算有基本的概念；（二）要能理解數字的大小，例如千位、萬位的大小，哪些物品的價值約幾萬元，哪些物品的價值會是幾千元；（三）要能使用多步驟的運算符號，如加減混和運算；（四）要能瞭解如何應用運算符號解決問題，要能理解加法與減法的含意；（五）要能判斷運算結果的合理性。由於本研究題目設計的限制，本研究參照前四項的設計標準。

研究也發現學生在不同層面的數感表現，課室內的教學與教材設計對他們的表現有顯著的影響（周玉婷，2013；陳慶林，2004；劉穆華，2006；Kayhan Altay, 2010; Li & Yang, 2010; Yang & Li, 2008; Yang et al., 2008; Yang & Wu, 2010）。周玉婷（2013）應用實驗教學檢視學生的數感概念，設計建構教學於國小一年級學生數感概念的研究，經過教學後能增強學童數與運算間的關係。劉穆華（2006）對國小二年級學生進行數感概念的實驗，研究結果發現實驗組學生數感概念在位值與數的運算構面有顯著的提升。陳慶林（2004）對國小三年級的學生進行建構式教學後，數感概念中以瞭解整數基本意義及大小之能力的表現最好，經建構教學的學生數感概念之整體表現出現顯著差異。

目前國內應用真實情境以探討不同年級學生的數感概念，尚未有相關的實徵研究，故本研究目的在於設計真實評量以探討國小四至六年級學生於真實情境下，不同年級學生在數感概念的表現是否有顯著的不同。

參、研究方法

以下依研究對象、研究工具、評分標準、統計分析與質化分析分別論述。

一、研究對象

本研究從桃園縣某國小取樣四至六年級，每一年級有三個班級參與施測，四年級有93位，五年級有95位，六年級有95位，共283位學生參與。

二、研究工具

依真實評量的特性，工具編製以結合生活的理念進行設計。數感測驗題為檢視學生數與計算概念，要求學生能正確地進行加減法計算。對應之數學指標為「N-2-2延伸加、減、乘、除與情境的意義，使能適用來解決更多的生活情境問題，並能用計算器械處理大數的計算」（教育部，2014）。本研究的數感概念題型如下：

小丸子準備了 70,000 元要購買液晶電視、音響、錄放影機，下列表格為各項商品的價錢。

品項	液晶電視	音響	錄放影機
廠牌 A	39,900 元	13,000 元	3,900 元
廠牌 B	38,500 元	15,700 元	4,100 元

本題組的設計有三個子題：

題 1：液晶電視、音響、錄放影機要各買一台，請根據小丸子的預算，幫她選好廠牌，並計算價錢是多少？

題 2：如果小丸子接受你的建議，她還會剩下多少錢？

題 3：假如小丸子第二題剩下的錢花完了，她又另外準備 8,000 元的存款購買影片，她先提出 2,500 元買洋芋片，接著又存了 1,880 元。請寫出小丸子提款與存款的算式，並算出所餘的存款。

研究者自行設計真實評量試題，試題審核過程經兩位具教學實務經驗之數學輔導團教師進行審查，數學輔導團教師與研究者經過三次討論，修正試題詞彙，以符合國小四至六年級學生的閱讀能力與能力指標。以下對兩位試題審核者背景分別敘述：

審核者一：現為桃園市國民小學校長，具數學教育碩士學位，曾任數學領域專任輔導員，具教學經驗30年，擔任國小數學領域輔導員20年，曾參與桃園市數學基測出題教師、示範數學教學教師，曾參與教育部數學深耕計畫。

審核者二：現為桃園市退休教師，具數學教育碩士學位，曾任數學領域專任輔導員，具教學經驗35年，擔任國小數學領域輔導員25年，曾參與桃園市數學基測出題教師、示範數學教學教師，曾參與教育部數學深耕計畫。

以下依三個題目的評量層次敘述如下：

題1：依據Yang與Sianturi（2019）的基礎，本題以基本的加法運算為主，並需學生進行估測，故本題的設計同時具備兩個設計向度，依Bloom的分類（Anderson & Krathwohl, 2001），需具備分析與應用的能力。

題2：本題需具備減法的運算能力，當學生進行估算後，於本題需進行基本的減法計算，即驗證其估算是否符合題意的要求，本題依Bloom的分類為分析的向度。

題3：本題與平常課堂所出的題目不同，較符合生活的情境，依據Yang與Sianturi（2019）的設計，具備第三與第四的設計向度，學生需瞭解提款與存款即為加法與減法的運算，當學生列出算式後，要能計算出正確答案。此題依Bloom的分類屬於較高層次的評鑑向度，學生要能判斷本題與前兩題完全不同。

依據Yang與Sianturi（2019）及數學指標進行題組設計，本題組依真實評量設計與具備數感概念理論的內容效度。由於僅有三個題目，無法進行量化的信度分析，故本題組的信度將經由質化分析與量化分析進行一致性的驗證。

三、評分標準

本題組設計三個子題，每一題組以4分制進行評分，最高為4分，最低為1分。評分標準如表1，得4分為完全正確，得3分為少部分錯誤，得2分為部分正確，1分為完全錯誤。

表 1
評分標準

分數	評分指標
4	完全正確、清楚 能使用適當的策略與步驟來完成解題 能展現出學生有清楚的概念與解題的邏輯 學生的解釋很清楚 學生能使用圖像或其他方式來表達解題的過程 能正確與合理地回答三個子題問題
3	部分正確、清楚 某些問題能清楚地回答，但並不完整 學生在某些概念是很清楚的，但並不是完全清楚 有些學生的解釋很清楚，但需再多些解釋才會更加完整 學生的圖像或是其他的表現方式，對解題並未有明顯的幫助 有些答案是正確的，但有些答案是錯誤的

（續下頁）

分數	評分指標
2	少數正確、清楚 有些解題的過程省略或是不正確，僅有少部分的答案是正確的 少量的證據顯示學生瞭解這些概念和步驟 學生的解釋和表現很難理解，有些解題的過程需再多做說明 學生的圖像或是其他的表現方式，對解題可能無關 大部分的答案都是不正確的
1	不正確，不完整、不清楚 所有的問題都是錯誤的 學生並未瞭解關鍵的數學概念 學生的解題策略與方法令人無法理解 學生的圖像或是其他的表現方式，對解題無關 所有的答案都是不正確的

為評量學生多元計分結果，有兩位評分者進行評閱，評分者經研究者訓練，對於評分內容與學生填答表現進行試閱，經過檢核過程，才進行試卷批閱。兩位評分者的Sperman相關係數為0.98 ($p < .001$)。

四、統計分析

研究者使用IBM SPSS Statistic 22.0統計軟體進行分析，採用兩種統計分析，一為 t 考驗，另一為順序性邏輯迴歸 (ordinal logistic regression)，以下分別敘述。

(一) t 考驗 (t test)

為瞭解學生對於不同題目的表現， t 考驗獨立樣本 (independent sample)。例如，四年級學生對於題1與題3或是題2與題3是否出現顯著差異等。

(二) 順序性邏輯迴歸

為進一步瞭解不同年級在不同題目的表現，本研究採用順序性邏輯迴歸分析，由於依變項為多重計分，計分方式為1到4，愈高分者表現愈好，為順序性的資料型態。獨立變項為年級，為四到六年級，亦為順序性的資料，因採用 t 考驗無法顯現年級愈高在不同題目的表現，故本研究採用順序性邏輯迴歸分析。

依Liao (1994, p. 40) 順序性邏輯迴歸的模式，與二元邏輯迴歸的差異在於會有不同的 μ_s ，於公式的左側亦可發現順序性迴歸分析是一累進式 (cumulative) 的

邏輯分析 (Agresti, 1990)，如公式 (1)：

$$\ln \left[\frac{P(y \leq j|x)}{1 - P(y \leq j|x)} \right] = \mu_j - \sum_{k=1}^j \beta_k x_k \quad j=1, 2, 3, \dots, J-1 \dots\dots\dots (1)$$

為能計算出不同階段的機率值，其公式 (2) 如下：

$$\begin{aligned} \text{Prob}(y=1) &= \ln[-\sum_{k=1}^1 \beta_k x_k] \\ \text{Prob}(y=2) &= \ln[\mu_2 - \sum_{k=1}^2 \beta_k x_k] - \ln[-\sum_{k=1}^1 \beta_k x_k] \\ \text{Prob}(y=3) &= \ln[\mu_3 - \sum_{k=1}^3 \beta_k x_k] - \ln[\mu_2 - \sum_{k=1}^2 \beta_k x_k] \\ \text{Prob}(y=j) &= 1 - \ln[\mu_{j-1} - \sum_{k=1}^{j-1} \beta_k x_k] \dots\dots\dots (2) \end{aligned}$$

為能計算出順序性邏輯迴歸的機率，如公式 (3) (Liao, 1994, p. 40)：

$$\text{Prob}(y \leq j) = \text{prob}(y^* \leq \mu_j) = \frac{e^{\mu_j - \sum_{k=1}^j \beta_k x_k}}{1 + e^{\mu_j - \sum_{k=1}^j \beta_k x_k}} \dots\dots\dots (3)$$

本研究中， μ_j 代表於題1至題3得到不同分數的機率， $j=1, 2, 3, 4$ 。 β_k 即代表不同年級的斜率， x_k 代表1, 2, 3, 1為4年級，2為5年級，3為6年級。

五、質化分析

依表1的評分標準，學生的解題策略將經由評分者進行計分，再對學生的算式歷程進行個別分析，分析過程會將所有可能出現的解題方式列出，進行策略統計，會分析全對、部分對與全錯的算式，從解題過程可以對學生的學習認知進行判斷。

肆、研究結果

以下分量化結果與質化結果兩部分進行分析。

一、量化結果

量化分析可從兩部分來探討。一為使用 t 考驗來檢視題目間難易度的顯著差異。由於 t 考驗依其統計常態分配的限制，僅能檢視平均數的顯著差異，並無法進一步對於順序性的資料進行分析。二為檢視不同年級在不同題目的得分比率，研究者採用順序性邏輯迴歸以發現不同年級在不同題目的表現。

(一) t 考驗分析結果

t 考驗於不同年級學生對不同題目相對比較之表現如表 2。四年級學生於題 1 與題 2、題 1 與題 3 及題 2 與題 3 均有顯著差異。從三個題目的平均數來看，可發現題 3 難度最高，題 1 的難度最低。五年級學生在題 1 與題 2 並未出現顯著差異，但題 2 與題 3、題 1 與題 3，均出現顯著差異，對五年級學生而言，題 3 的難度明顯高於題 1 與題 2。六年級學生於題 1 與題 2、題 1 與題 3 及題 2 與題 3 均有顯著差異。從三個子題的平均數來看，可發現題 3 難度最高，題 1 的難度最低，本結果與四年級學生表現一致。

表 2

t 考驗於不同年級學生對不同子題相對比較之表現

子題比較	子題	<i>n</i>	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
四年級							
題 1 對題 2	1	93	3.85	.44	92	2.06	.04*
	2	93	3.70	.72			
題 1 對題 3	1	93	3.85	.44	92	4.55	.00**
	3	93	3.43	.87			
題 2 對題 3	2	93	3.70	.72	92	3.05	.00**
	3	93	3.43	.87			
五年級							
題 1 對題 2	1	95	3.65	.81	94	.42	.67
	2	95	3.61	.85			
題 1 對題 3	1	95	3.65	.81	94	4.73	.00**
	3	95	3.17	1.05			
題 2 對題 3	2	95	3.61	.85	94	3.73	.00**
	3	95	3.17	1.05			

(續下頁)

子題比較	子題	<i>n</i>	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
六年級							
題 1 對題 2	1	95	3.78	.53	94	3.93	.00**
	2	95	3.58	.86			
題 1 對題 3	1	95	3.78	.53	92	5.3	.00**
	3	95	3.24	1.09			
題 2 對題 3	2	95	3.58	.86	94	3.45	.00**
	3	95	3.24	1.09			

$p^{**} < .01$.

(二) 順序性邏輯迴歸分析結果

為推估不同年級學生在多重計分題型的表現，題1至題3於不同年級變項的順序性邏輯迴歸分析如表3，不同年級間並未出現顯著差異，然在不同計分的機率卻有顯著差異，如題1在分數1與分數2、分數2與分數3、分數3與分數4等均有顯著差異（ $p < .01$ ）。再者，應用順序性邏輯迴歸分析於不同年級各子題得分的機率估測如表4，可算出四至六年級學生在分數1至4的機率分布，如四至六年級答對4的機率為60%，以四年級得分1的機率如公式（4）：

$$\text{Prob}(1) = (\text{EXP}(-3.5 + .39 \cdot 1) / (1 + \text{EXP}(-3.5 + .39 \cdot 1))) \dots\dots\dots (4)$$

Wald χ^2 為 $\frac{\hat{\beta}}{\text{SE}}$ 的值，即SE為 H_0 虛無假設，即題目內的得分與年級表現並無差異，而 $\hat{\beta}$ 則為 H_1 對立假設的估測值，即題目內不同的分數表現存在顯著差異。從順序性邏輯迴歸的機率分析，可以檢視不同年級於不同子題間的表現結果。應用順序性邏輯迴歸分析於不同年級各子題得分的機率估測如表4，於題1的機率分布，四至六年級學生得分機率相近。

於題2，題1至題3於不同年級變項的順序性邏輯迴歸分析如表3，不同年級間未出現顯著差異，然在不同計分的機率卻有顯著差異，如在分數1與分數2、分數2與分數3、分數3與分數4等，均有顯著差異（ $p < .01$ ）。應用順序性邏輯迴歸分析於不同年級各子題得分的機率估測如表4，四至六年級在答對4的機率降為50%以下，有約27%的學生的分數為3。於題2的機率分布，四至六年級學生的表現相近。

表 3

題 1 至題 3 於不同年級變項的順序性邏輯迴歸分析

Model		β	SE	Wald χ^2	df	p
題 1						
Threshold	(item1=1)	-3.50	.42	69.83	1	.00**
	(item1=2)	-3.08	.36	70.78	1	.00**
	(item1=3)	-1.54	.27	32.96	1	.00**
Location	(grade=1)	.39	.41	.90	1	.34
	(grade=2)	-.20	.37	.30	1	.59
	(grade=3)					
題 2						
Threshold	(item2=1)	-3.29	.39	70.47	1	.00**
	(item2=2)	-1.61	.26	38.77	1	.00**
	(item2=3)	-1.31	.25	27.55	1	.00**
Location	(grade=1)	.29	.37	.61	1	.43
	(grade=2)	-.12	.37	.30	1	.74
	(grade=3)					
題 3						
Threshold	(item3=1)	-2.58	.29	81.50	1	.00**
	(item3=2)	-1.01	.21	22.35	1	.00**
	(item3=3)	-.46	.25	4.85	1	.00**
Location	(grade=1)	.24	.30	.65	1	.42
	(grade=2)	-.12	.28	.48	1	.48
	(grade=3)					

註：grade = 1 為四年級；grade = 2 為五年級；grade = 3 為六年級。

$p^{**} < .01$.

於題3，題1至題3於不同年級變項的順序性邏輯迴歸分析如表3，不同年級間未出現顯著差異。在不同計分的比較，如分數1與分數2、分數2與分數3、分數3與分數4等，均出現顯著差異（ $p < .01$ ）。應用順序性邏輯迴歸分析於不同年級各子題得分的機率估測如表4，四年級在答對4的機率降為15%，五年級為41%，六年級為53%。大部分四年級學生得分為3，約占45%，五、六年級於分數3的機率較為接近，分別為26%與30%。於此結果可看出，六年級全對的機率高於五年級，五年級

高於四年級，結果顯示年級愈高，得分機會愈高。

表 4

應用順序性邏輯迴歸分析於不同年級各子題得分的機率估測

子題與年級	得分			
	1	2	3	4
題 1				
四年級	4%	6%	24%	65%
五年級	4%	6%	24%	65%
六年級	5%	8%	28%	59%
題 2				
四年級	5%	21%	27%	48%
五年級	4%	20%	26%	50%
六年級	5%	22%	28%	45%
題 3				
四年級	9%	32%	45%	15%
五年級	9%	20%	30%	41%
六年級	4%	17%	26%	53%

註：得分 1 代表全錯，2、3 代表部分對，4 代表全對。

二、質化分析結果

以下依三個子題的結果分別論述。

（一）第一子題為液晶電視、音響、錄放影機要各買一台，請根據小丸子的預算，幫她選好廠牌，並計算價錢是多少？

學生在填答本問題時，只要答案合理，學生可以選擇同個廠牌或是不同廠牌的答案，各班級於題1的解題策略分析如表5，於四至六年級可以看出全班有20人以上均可回答正確，答對率除六年級有一班在70%以下，其他都在80%以上。從平均答對率來看，四至六年級的表現都相當接近，並未出現年級愈高、答對率愈高的趨勢。

表 5
各班級於題 1 的解題策略分析

答題分布	401	402	403	501	502	503	601	602	603
答對人數	25	31	25	24	24	28	29	27	22
選同一廠牌 (4)	12	23	16	20	20	24	26	21	8
選不同廠牌 (4)	13	8	9	4	4	4	3	6	14
選同一廠牌但計算錯誤 (3)	2	1	4	6	1	0	1	1	4
選不同廠牌但計算錯誤 (3)	3	0	1	1	1	1	0	2	6
選同一廠牌用 70,000 減 (2)	0	0	0	0	2	0	0	0	0
廠牌 a 的互減 (1)	0	0	0	0	0	0	0	1	0
a、b 廠牌全部加起來 (1)	0	0	0	0	0	0	1	0	0
未作答 (空白) (1)	0	0	1	0	2	2	1	0	0
合計人數	30	32	31	31	32	32	32	31	32
答對 %	83%	97%	80%	77%	81%	91%	91%	87%	69%

註：於答題分布，(4) 代表 4 分，(3) 代表 3 分，(2) 代表 2 分，(1) 代表 1 分。

學生選擇同一廠牌或是不同廠牌，列式正確但是計算錯誤，會得3分；若將第二題的答案寫在第一題，則獲得2分，因為未瞭解第一題的題意。於答對的解題策略面向，僅401與603兩個班級選擇不同廠牌的策略高於同一廠牌，而其他的班級選擇同一廠牌的人數高於選擇不同廠牌。可以發現401與603兩個班級的同學在選擇的策略上較為靈活，會考慮不同廠牌的價格進行計算。

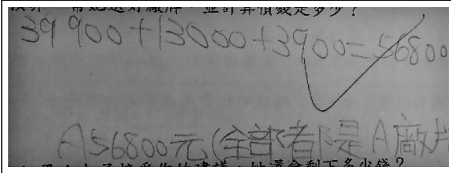
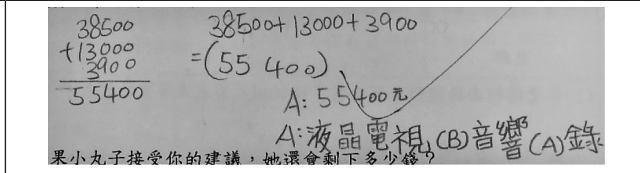
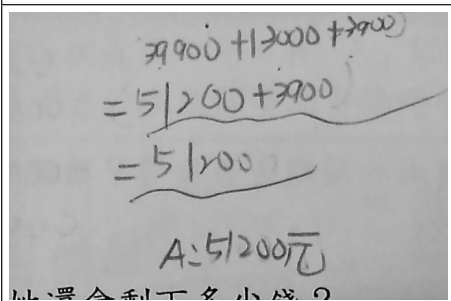
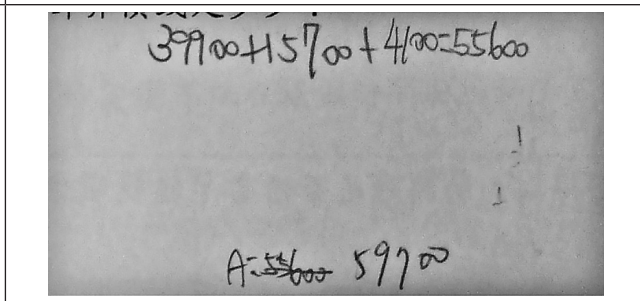
若學生隨意猜測及未作答，則獲得1分，如將廠牌a的答案互減，或是將a和b廠牌答案全部相加等。獲得3分的學生比率較高，都在10人以下，其他2分與1分的比率都較低，在2人以下。題1的部分同學解題算式說明如圖1，列出學生答對與計算錯誤的範例。需注意的是，603班有六位同學出現計算錯誤的問題，計算錯誤的比率最高，甚至高於四年級與五年級的班級，可以班級的教學是否需針對學生的錯誤計算進行調整。

(二) 如果小丸子接受你的建議，她還會剩下多少錢？

於本子題學生需要能使用減法策略進行解題，也就是用題1的結果，進行減法計算，各班級於題2的解題策略分析如表6。於本子題，每一班均超過20位同學能正確解題，答對率除401、501與603三個班級外，其他班級的答對率皆在80%以上，

圖 1

題 1 的部分同學解題算式說明

	
選同一廠牌	選不同廠牌
	
選同一廠牌但計算錯誤	選不同廠牌但計算錯誤

如402班幾乎全班都能答對，相對的，603班有10人算錯。所以學生並未因高年級而減低錯誤的機率。於401、501與603三個班級，因題1錯而題2錯的人數高於同年級的其他班，可以發現這三個班級在計算的錯誤需要在教學時進行要求。於題2的表現，也並未出現年級愈高、答對率愈高的趨勢。

表 6

各班級於題 2 的解題策略分析

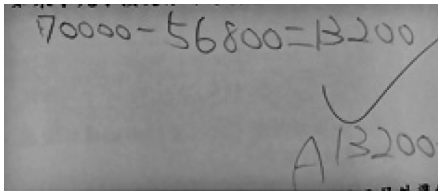
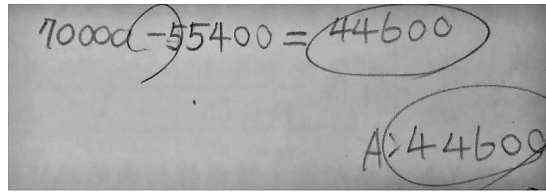
答題類型	401	402	403	501	502	503	601	602	603
全對 (4)	21	31	25	24	25	28	28	25	22
列式對算錯 (3)	3	0	3	0	0	2	1	2	0
題 1 錯故題 2 錯 (3)	5	1	1	7	2	0	1	3	10
題 1 答案寫到題 2 (2)	0	0	0	0	2	0	0	0	0
3,900-4,100 (1)	0	0	1	0	0	0	0	0	0
列式錯計算錯 (1)	0	0	0	0	1	0	0	0	0
空白 (1)	1	0	1	0	2	2	2	1	0
合計人數	30	32	31	31	32	32	32	31	32
答對 %	70%	97%	81%	77%	78%	88%	88%	81%	69%

註：於答題分布，(4) 代表 4 分，(3) 代表 3 分，(2) 代表 2 分，(1) 代表 1 分。

當列式與答案都正確，則為4分，但學生列式正確、計算錯誤，則會得3分。若題1算錯，則題2也錯，若是錯的答案正確，則獲得3分。將題1答案寫到題2，獲得2分，因為未瞭解題意。若學生隨意猜測及未作答，則獲得1分。題2部分學生解題算式說明於圖2，列出學生答對與計算錯誤的範例。

圖 2

題 2 部分學生解題算式說明

	
全對	錯誤類型：列式對，但計算錯誤

(三) 假如小丸子第 2 題剩下的錢花完了，她又另外準備 8,000 元的存款購買影片，她先提出 2,500 元買洋芋片，接著又存了 1,880 元。請寫出小丸子提款與存款的算式，並算出所餘的存款。

由於本題為生活中的情境題，於口語敘述的方式有別於課本的布題方式，故平均答對率低於前兩題。於本子題，學生需能理解加減法題意後，進行解題，各班級於題3的解題策略分析如表7，四至六年級學生的答對率較前兩個子題為低，402班答對率為53%，501班與502班分別為55%與50%，603班為59%。於此可發現，402班同學在前兩題的答對率高達97%，而在第三題的答對率為全年級最低，學生在基本數學運算較佳，然而在語文理解的變化上，若出現不同於課本的生活題型時，答對率反而較其他班級為低。類似情況亦出現在501、502與603等班級。

因正確的策略是先減後加，採錯誤的連減策略於402、501、502、603等班至少都有五位同學，所以使用錯誤的策略並未因年級較高而減少，而是分布在三個年級的班級中。再者，使用將第二題的答案，再用先減後加的錯誤策略，也平均分布在各年級，尤其以501班的人數最多。從以上這兩個錯誤策略來看，501班已超過10人，占該班人數的33%，所以501班的學生使用猜測的錯誤策略情況也較其他班級高。

表 7

各班級於題 3 的解題策略分析

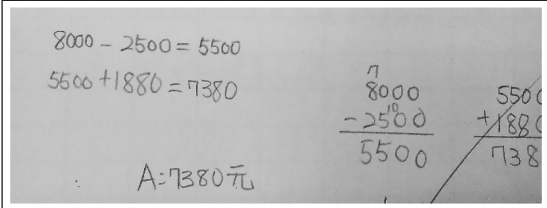
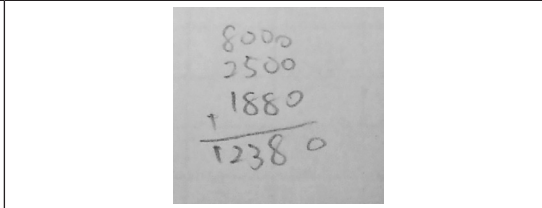
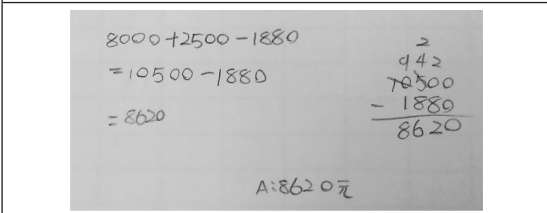
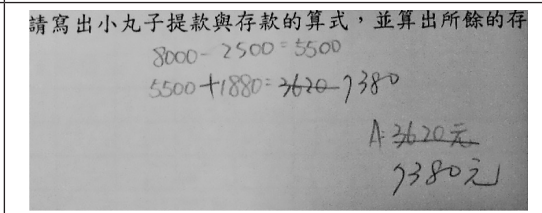
答題分布	401	402	403	501	502	503	601	602	603
先減後加 (4)	21	17	22	17	16	21	22	19	19
列式對但計算錯誤 (3)	6	7	3	2	5	3	1	4	3
先加後減 (2)	0	0	0	0	1	0	0	0	0
連加 (1)	0	0	2	0	0	0	0	0	0
連減 (1)	0	6	1	6	5	3	4	2	5
未作答 (空白) (1)	1	1	1	1	4	2	2	3	0
其他錯誤：2,500+8,000 (1)	0	0	0	0	0	0	0	0	0
把第二題剩下的算進去，先減後加 (2)	2	1	2	5	1	2	0	2	4
第二題剩下的錢去減第三題答案 (1)	0	0	0	0	0	1	0	0	0
第二題算進去連減 (1)	0	0	0	0	0	0	1	1	0
第二題算進去先加後減 (1)	0	0	0	0	0	0	0	0	1
2,500-1,800 (1)	0	0	0	0	0	0	1	0	0
8,000-2,500 (1)	0	0	0	0	0	0	1	0	0
合計人數	30	32	31	31	32	32	32	31	32
答對 %	70%	53%	71%	55%	50%	66%	69%	61%	59%

註：於答題分布，(4) 代表 4 分，(3) 代表 3 分，(2) 代表 2 分，(1) 代表 1 分。

當列式與答案都正確，則為 4 分，但學生列式正確、計算錯誤，則會得 3 分，若先加後減，表示學生知道要使用加法與減法，只是順序錯誤，得 2 分，其他隨意猜測及未作答，則獲得 1 分。題 3 部分，學生解題算式說明如圖 3，列出學生答對與計算錯誤的範例。於本題可看出，學生看不懂題意，而隨意猜測的比例高於前面兩題。學生未看清楚題意，直接用連減的方式，於四至六年級都很常見。再者，用第二題所剩下的答案，再先減後加，於四至六年級都常見。於本題可看出錯誤的解題策略明顯較多。對於生活常見的題型，如提款與存款等概念，多數學生已產生困難，在不瞭解題意的情況下，產生了多種猜測性錯誤的解題策略。於圖片中可看出學生使用先減後加、先加後減、連加與連減的解題策略，其中只有先減後加的策略是正確的。

圖 3

題 3 部分學生解題算式說明

	
先減後加	連加（列式錯誤）
	<p>請寫出小丸子提款與存款的算式，並算出所餘的存</p> 
先加後減（列式錯誤）	連減（列式錯誤）

伍、結論與建議

一、結論

可從量化結果與質化結果進行討論。

（一）量化結果

從 t 考驗結果顯示，除五年級學生對題1與題2未出現顯著差異外，其他年級學生於同一年級對三種題目均出現顯著差異，每一題組於不同年級均出現顯著差異。結果顯示，題1較簡單，題2較難，題3最困難。為進一步檢視不同年級於題1至3的表現，依順序性邏輯迴歸分析，四至六年級在題1與題2表現一致，但在題3的表現，六年級學生的表現優於五年級，五年級學生優於四年級。六年級學生表現較佳的原因，可能在於學生生活經驗相對於四年級學生較為多元所致。

（二）質化結果

從質性的解題分析，於題1，僅401與603兩班選擇不同廠牌的策略高於同一廠牌，可發現教師的教學靈活度對學生於真實評量的解題策略會有影響，此結果與文獻所提課室內的教學設計影響學習表現的結果一致（周玉婷，2013；陳慶林，2004；劉穆華，2006）。相對的，401、501與603班因題1錯而題2錯的人數高於同年級的其他班，由此亦可發現，課室教學對學生平常學習的要求差異亦會影響學生的學習表現。

於題1與題2的綜合表現來看，四至六年級錯誤率都較低，約20%~30%，於題3錯誤率則為30%~50%。四至六年級學生均對生活情境題的表現較差，當題目敘述較長或是與課本題目不同時，不論年級，學生會採取很多猜測的策略，如出現連減的策略。而會出現這些錯誤，有可能是誤解題目，不瞭解題目意義，亦有可能對於題目的出題方式不熟悉，或是對生活上的情況不瞭解所致，所以當使用真實情境布題後，在題目理解的難度增高，以致年級愈高，表現愈佳。

由於題3的問題較接近生活的語述，所以四至六年級學生於題3的整體表現，較題1與題2為低。可見，學生對生活的語述於四至六年級都有困難。從整體表現來看，四至六年級的平均答對率並未出現差異性，反而在個別班級的表現出現差異，由此可看出真實評量的結果反映教師的教學差異，也就是當課室教學與生活的情境愈接近時，學生的表現愈佳，如401、403、503與601等班級。從題3的表現，教師對於課堂的教學設計應有更多的教學活動結合生活。

從量化與質化研究結果發現，四至六年級學生的整體表現於題3較低，題目的設計與生活上的用語接近，如題3：「假如小丸子第2題剩下的錢花完了，她又另外準備8,000元的存款購買影片，她先提出2,500元買洋芋片，接著又存了1,880元。請寫出小丸子提款與存款的算式，並算出所餘的存款」。本題的自然語述在於剩下的錢花完了，她先提出……，接著又存……，寫出提款與存款……，算出所餘的存款。這樣的語述方式與一般學校的數學題目不同，例如一般教科書的題目設計為「小丸子有8,000元，先買了2,500元的洋芋片，曉華又給了小丸子1,880元，請問小丸子現在有多少元？」而教科書的題目設計與真實評量的情境完全不同。

從評量的結果來看，不論學生在哪一年級，都對於生活的語述有理解的困難，本研究發現數學素養課程的實施相當迫切，學校教師將生活的數學語言落實於課堂，才能使學生理解生活的語言，才能達到使用自然語言於生活，讓他們有感學

習數學的全人教育目標（教育部，2014）。

二、建議

本研究採用真實評量的情境布題，以多重計分分析學生於數感概念的解題策略，從研究結果提出以下的建議：

（一）從數學素養的學習趨勢，應將真實的生活帶入數學課程

從量化的分析結果，於題1與題2的結果發現四至六年級學生於數學基本運算的表現並無顯著差異，但在題3的表現，有年級愈高、表現愈佳的趨勢，如高年級答對的表現為53%，而四年級卻只有15%，看到學生於數學情境題的文字理解出現困難，反映了學校內的數學課程需要與生活結合的情境設計。從高年級的學生表現來看，通常學生對於生活經驗的敏感度會較四年級學生高，但從六年級學生53%的答對率，約一半的學生答對，可發現他們在生活情境題的答對率較前兩題為低。透過質化分析的結果，可比較每班學生使用解題策略的差異性，與教師教學相關，故於數學課堂設計生活化的課程，協助學生將數學與生活結合，為未來教育政策的發展方向。

（二）從順序性邏輯迴歸分析，發現資料的真實性

若從 t 考驗的檢視，僅能檢視學生在不同題目的表現是否有差異，如四年級學生在題1、題2與題3均出現顯著差異，僅能看出題目的難易度。但若由順序性邏輯迴歸分析，可看出不同年級在這三個題目間的表現機率，如題3可鑑別四至六年級學生表現的機率分布。結果發現順序性資料的結果，有別於進行變異數分析，僅比較平均數是否具備顯著差異的特性，由機率的分布可更清楚呈現差異的實際情況。

（三）從解題策略的分析，可看到學生的認知學習階層

本研究分析學生繁複的解題策略，從策略中發現學生的想法，不論學生是部分瞭解，還是不瞭解就進行錯誤的猜測，都可以從其算式分析看到學生的想法。本研究提供部分計分的評分標準與分析方式，對於學生錯誤解題的策略與計算錯誤等，給予部分計分的依據，而不是將學生的答題僅依對與錯的方式給予全錯或全對的結果。

由上述三點建議來看，數學的學習即為有感經驗的學習，學生於學校所學的數學概念，需應用於生活。本研究參與研究的四至六年級學生，雖在108課綱實施前受測，但從研究結果發現，學生因不瞭解生活中數學的自然語言，而產生錯誤的解題策略。本文提供了實施數學素養課程政策的重要性，也對於設計數學生活的課程與評量提供具體可行的建議。

參考文獻

- 王郁芳（2018）。國小六年級學童數感解題之錯誤類型分析（未出版之碩士論文）。國立臺南大學。
- 【Wang, Y.-F. (2018). *Analysis of the error patterns made by sixth graders using number sense to solve math problems* [Unpublished master's thesis]. National Tai-Nan University.】
- 朱麗雯（2005）。國小學童數感概念結構分析研究（未出版之碩士論文）。國立臺中教育大學。
- 【Chu, L.-W. (2005). *The structure analysis of number sense for elementary students* [Unpublished master's thesis]. National Taichung University of Education.】
- 周玉婷（2013）。建構式教學對國小一年級學童數常識能力之影響（未出版之碩士論文）。國立臺北教育大學。
- 【Chou, Y.-T. (2013). A study of building first graders' ability of number sense through constructive instruction [Unpublished master's thesis]. National Taipei University of Education.】
- 林永豐（2018）。素養導向教學設計的要領。載於周淑卿、吳璧純、林永豐、張景媛、陳美如（編），*素養導向教學設計參考手冊*（頁1-4）。教育部。
- 【Lin, Y.-F. (2018). The essentials of instructional design for literacy based. In S.-C. Chou, P.-C. Wu, Y.-F. Lin, C.-Y. Chang, & M.-J. Chen (Eds.), *The referenced book of instructional design for literacy based* (pp. 1-4). Ministry of Education.】
- 洪德健（2001）。國小社會科真實評量實施之研究（未出版之碩士論文）。國立嘉義大學。
- 【Hung, D.-C. (2001). *The implementation of authentic assessment in the teaching and learning of elementary social studies* [Unpublished master's thesis]. National Chiayi University.】
- 陳立庭（2017）。科技教育的真實評量——以篩選裝置為例。*科技與人力教育季刊*，3（3），20-39。https://doi.org/10.6587/JTHRE.2017.3(3).2
- 【Chen, L.-T. (2017). Authentic assessment in technology education: Screening machine. *Journal of Technology and Human Resource Education*, 3(3), 20-39. https://doi.

org/10.6587/JTHRE.2017.3(3).2】

張靜文（2012）。從真實評量觀點看幼兒園之教學評量。臺灣教育評論月刊，1（11），12-14。

【Chang, C.-W. (2012). The instruction and assessment in kindergarten from the perspectives of authentic assessment. *Taiwan Educational Review Monthly*, 1(11), 6-12.】

教育部（2014）。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學：數學領域。

【Ministry of Education Institute (2014). *Twelve Years Basic Education Curriculum Standards for Elementary School to High School: Mathematics Field*.】

陳慶林（2004）。透過建構式教學探討國小三年級學童數感表現之研究（未出版之碩士論文）。國立臺南大學。

【Chen, C.-L. (2004). *The study of the third grade primary pupils' performance on number sense through the constructing teaching method* [Unpublished master's thesis]. National Tainan University.】

游自達（2019）。素養導向教學的實踐：深化學習的開展。臺灣教育評論月刊，8（10），6-12。

【Yu, T.-T. (2019). Literacy based to the practice of instruction: The development of incentive learning. *Taiwan Educational Review Monthly*, 8(10), 6-12.】

劉穆華（2006）。國小二年級學童融入數感教學活動之實驗研究（未出版之碩士論文）。國立臺中教育大學。

【Liu, M.-H. (2006). *The experiment study of number sense instruction for second grade students* [Unpublished master's thesis]. National Taichung University of Education.】

Agresti, A. (1990). *Categorical data analysis*. John Wiley & Sons.

Ahmad, Z. (2020). Summative assessment, test scores and text quality: A study of cohesion as an unspecified descriptor in the assessment scale. *European Journal of Educational Research*, 9(2), 523-535. <https://doi.org/10.12973/eujer.9.2.523>

Anderson, W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's educational objectives*. Longman.

Antell, S. E., & Keating, D. P. (1983). Perceptions of numerical invariance in neonates.

- Child Development*, 54, 695-701. <https://doi.org/10.2307/1130057>
- Archbald, D. A., & Newmann, F.M. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school*. National Association of Secondary School Principals.
- Azim, S., & Khan, M. (2012). Authentic assessment: An instructional tool to enhance students learning. *Academic Research International*, 2(3), 314-320.
- Barrio, M. I. P., Escamilla, A. C., García, M. N. G., Fernández, E. M., & García, P. de la R. (2015). Influence of assessment in the teaching-learning process in the higher education. *Procedia-Social and Behavioral Sciences*, 176, 458-465. <https://doi.org/10.1016/j.sbspro.2015.01.497>
- Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities*, 38(4), 333-339. <https://doi.org/10.1177/00222194050380040901>
- Burns, M. (1994). Arithmetic: The last holdout. *Phi Delta Kappan*, 75, 471-476.
- Can, D., & Özdemir, I. E. Y. (2020). An examination of fourth-grade elementary school students' number sense in context-based and non-context-based problems. *International Journal of Science and Mathematics Education*, 18(7), 1333-1354. <https://doi.org/10.1007/s10763-019-10022-3>
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, 24(2), 97-140. <https://doi.org/10.1080/10888691.2018.1537791>
- 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. <https://doi.org/10.1080/03323310601125088>
- Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education*, 33(1), 18-28. <https://doi.org/10.1177/00224669903300102>
- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research*, 52, 67-86. <https://doi.org/10.1007/BF02504676>

- Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, 77, 153-175. <https://doi.org/10.1111/j.1467-8624.2006.00862.x>
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.
- Hodgman, M. R. (2014). Using authentic assessments to better facilitate teaching and learning: The case for student portfolios. *Journal of Studies in Education*, 4(3), 59-65. <https://doi.org/10.5296/jse.v4i3.6149>
- Kayhan Altay, M. (2010). *İlköğretim ikinci kademe öğrencilerinin sayı duyularının; sınıf düzeyine, cinsiyete ve sayı duyusu bileşenlerine göre incelenmesi*, [An investigation of middle grade students' number sense in terms of grade level, gender, and components of number sense] [Unpublished doctoral dissertation]. Eğitim Bilimleri Enstitüsü, Hacettepe Üniversitesi, Ankara.
- Lago, R., & DiPerna, J. (2010). Number sense in kindergarten: A factor-analytic study of the construct. *School Psychology Review*, 39(2), 164-180. <https://doi.org/10.1080/02796015.2010.12087771>
- Li, M. F., & Yang, D. C. (2010). Development and validation of a computer-administered number sense scale for fifth-grade children in Taiwan. *School Science and Mathematics*, 110(4), 220-230. <https://doi.org/10.1111/j.1949-8594.2010.00024.x>
- Liao, T. F. (1994). *Interpreting probability models logit, probit, and other generalized linear models*. Sage.
- Martin, A., Arrambide, M., & Holt, C. (2016). The impact of flipped instruction on middle school mathematics achievement. *Journal of Education and Human Development*, 5(3), 84-96. <https://doi.org/10.15640/jehd.v5n4a10>
- McFeetors, P. J., Marynowski, R., & Candler, A. (2021). Generative unit assessment: Authenticity in mathematics classroom assessment practices. *Educational Sciences*, 11(7), 1-22. <https://doi.org/10.3390/educsci11070366>
- Merta, I. M. E. D., Suarjana, I. M., & Mahadewi, L. P. P. (2015). Analysis of authentic assessment. *E-Journal PGSD Ganesha University of Education*, 3(1), 1-10. <https://doi.org/10.23887/jjpgsd.v3i1.5818>

- Mundia, L. (2020). Assessment of skills development in brunei trainee teachers: Intervention implications. *European Journal of Educational Research*, 9(2), 685-698.
- National Council of Teachers of Mathematics. (2000). *The principles and standards for school mathematics*.
- Newmann, F. M. (1997). Authentic assessment in social studies: Standards and examples. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, adjustment and achievement* (pp. 359-380). Academic Press.
- Organization for Economic Co-operation and Development. (1999). *Measuring student knowledge and skills. A new framework for assessment*.
- Olfos, R., & Zulantay, H. (2007). Reliability and validity of authentic assessment in a web based course. *Educational Technology and Society*, 10(4), 156-173.
- Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research, and Evaluation*, 13, 1-12. <https://doi.org/10.7275/0qpc-ws45>
- 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, 225-237.
- Sasmaz-Oren, F., & Ormanci, U. (2011). Teacher candidate levels of familiarity with the methods, techniques and tools composing the alternative assessment approaches. *Social and Behavioral Sciences*, 15, 3476-3483. <https://doi.org/10.1016/j.sbspro.2011.04.321>
- Stenmark, J. K. (1991). *Mathematics assessment: Myths, models, good questions, and practical suggestions*. NCTM.
- Vu, T. T., & Dall'Alba, G. (2014). Authentic assessment for student learning: An ontological conceptualization. *Educational Philosophy Theory*, 46(7), 778-791. <https://doi.org/10.1080/00131857.2013.795110>
- Wiggins, G. (1998). *Education assessment: Designing assessments to inform and improve student performance*. Joey-Bass.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358(27), 749-750. <https://doi.org/10.1038/358749a0>

- Yang, D. C., & Li, M. F. (2008). An investigation of 3rd-grade Taiwanese students' performance in number sense. *Educational Studies*, 34(5), 443-455. <https://doi.org/10.1080/03055690802288494>
- Yang, D. C., Li, M. F., & Li, W. J. (2008). Development of a computerized number sense scale for 3-rd graders: Reliability and validity analysis. *International Electronic Journal of Mathematics Education*, 3(2), 110-124. <https://doi.org/10.29333/iejme/222>
- 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, 789-807. <https://doi.org/10.1007/s10763-007-9100-0>
- Yang, D. C., & Sianturi, I. A. J. (2019). The earliest teaching and learning of probability in Singapore, the US, and Indonesia: From the perspectives of textbooks analysis. *Irish Education Studies*, 38(4), 535-559. <https://www.tandfonline.com/doi/full/10.1080/03323315.2019.1664313>
- Yang, D. C., & Wu, W. R (2010). The study of number sense: Realistic activities integrated into third-grade math class in Taiwan. *The Journal of Research*, 103(6), 379-392. <https://doi.org/10.1080/00220670903383010>

