

國小六年級學童在數常識異常作答之探究

葉俊谷¹ 楊德清² 黃財尉³

^{1,2} 國立嘉義大學數學教育研究所

³ 國立嘉義大學輔導與諮商研究所

(投稿日期：95 年5 月2 日；修正日期：95 年6 月7 日；接受日期：95 年7 月4 日)

摘要

本研究旨在探討國小六年級學童在數常識(Number Sense)測驗中的異常作答情形，進而比較學童在數常識不同組成份中的表現、不同性別以及不同數學學業成就下，學生發生異常作答的差異情形。本研究以「內注意係數(Winthin-Ability-Concern Index, W*)」與「外注意係數(Beyond-Ability-Surprise Index, B*)」(黃財尉，2005)偵測學生異常作答情形，研究工具為「國小第二階段數常識線上評量測驗」(林姿飴，2005)，施測對象為全國各地區部份公立國民小學之六年級學童，人數總計1212名。所得資料透過SPSS 統計套裝軟體，以描述性統計、Pearson 積差相關、逐步多元迴歸分析法、t 考驗及單因子變異數分析等統計方法加以分析，結果顯示：(一)學童在「比較數字相對大小」成份中的表現，對於W*及B*有較佳的預測力；(二)男童在數常識異常作答情形顯著高於女童；以及(三)數學學業低成就的學生在數常識異常作答情形顯著高於中、高成就的學生。

本研究針對以上研究結果加以分析與討論，並提出教學上及未來研究上的建議，以供其他教師、家長或教育相關單位作為參考。

關鍵詞：數常識、異常作答、內注意係數、外注意係數

壹、研究動機及目的

數常識(Number Sense)之教與學已受許多先進國家教育相關單位的重視(Japanese Ministry Education, 1989; National Council of Teachers of Mathematics [NCTM], 1989, 2000; Australian Education Council [AEC], 1991), 同時, 許多國家之專家學者亦積極從事數常識的相關研究(Sowder, 1992a, 1992b; Emaanelsson & Johansson, 1996; Yang, 2003; 楊德清, 2000, 2002; 許清陽、楊德清、李茂能, 2001; 林素微, 2002; 李威進, 2004; 林姿飴, 2005), 致使數常識的發展如雨後春筍般相繼應運而生, 國際間亦已將數常識之教學視為數學教育中不可忽視的課題。

評量應該支持重要數學主題的學習以及提供教師與學生有價值的資訊(NCTM, 2000); 而數常識既已成為數學教育中重要的主題之一, 發展有效的評量工具更突顯其重要性。雖然數常識評量工具具備良好的信、效度是必要的條件, 然而不論在評量測驗或診斷系統的學生作答結果中, 經常透露出學生不合常理的異常作答(Aberrant Response), 亦即受試者在回答相對低於其能力水準的題目時應該答對卻沒有答對, 而回答相對高於其能力水準的題目時應該答錯卻沒有答錯, 違反了Guttman(1944)的理想作答組型, 因此直接產生的負面影響便是關係到整個評量測驗或診斷系統的準確性。藉由分析這些異常作答的情況所產生之異常作答組型(aberrant response pattern), 進而將受試者的某些特質顯現出來, 所獲得之資訊將可以用來幫助解決研究者與教學者對於受試者學習上的問題, 同時藉此改良評量測驗或診斷系統之品質, 最後達到增進整個教學過程的學習效果(Kogut, 1986)。

承上所述, 由於國內外數常識評量工具相當缺乏, 再加上與異常作答相關理論的結合探討仍然乏人問津, 因此, 在國內欲極力開發出更完善之數常識評量工具的同時, 研究者藉此探討國小六年級學童在數常識評量工具的診斷後, 在數常識不同組成成份的解題表現中, 何者與異常作答指標的關係最為密切, 進而去探討學童與該組成成份的關聯性, 並從中分析學童在數常識不同組成成份的表現與異常作答之間的相關, 讓數常識評量測驗的結果可以更具意義, 可以幫助研究者與教學者更清楚掌握學童的數常識發展概況。

此外, 許多研究指出不同性別之學童在數學學習上具有顯著不同的表現, 例如: 數學學習的積極態度, 男生優於女生(曹宗萍、周文宗, 1998); 對於數字的記憶, 女生優於男生(陳淑絹, 1999); 在類似紙筆測驗的開放式評量測驗上, 尤其是觀念及程序部份, 女科學教育研究與發展季刊第四十四期

60

生優於男生(李佳香、黃佩晶、黃千玲、張麗麗, 2000); 關於乘、除法的運算概念, 女生優於男生(吳德邦、鄭嫻嫻, 2000); 在數學推理方面, 男生優於女生(劉慧如 & Garfield, 2002)。然而, 性別差異對於學童發生異常作答情形是否也會造成顯著的不同結果, 令研究者好奇並欲進一步探討。另一方面, 研究者亦想了解學童數學學業成就的高低與其發生異常作答的情形是否具有顯著的相關, 並探究其原因。

基於上述所言, 本研究之研究目的分述如下:

- 一、探究小六學童在數常識不同組成成份中的表現對異常作答指標之影響。
- 二、比較小六之男、女學童在數常識異常作答的差異情形。
- 三、比較小六之低、中、高數學學業成就學童在數常識異常作答的差異情形。

貳、文獻探討

一、數常識的意義與理論架構

數常識(Number Sense)簡單來說是在面對數字時的某種思考方式, 可以解釋為對數字

的理解並有直觀上的處理方法，在數字、運算及數字與運算之間不同的情境可以發展出有效的解題策略，並且在其判斷上具有變通性與靈活性，用以解決各種相關聯的問題(楊德清，2000，2002)。數常識亦是一種數學知識及經驗不斷累積與成熟的過程(Reys, 1994)，至今仍然不易給予出固定與明確的定義(Resnick, 1989；Reys et al., 1991；Silver, 1989；Sowder, 1992b)，因為它並非有限的知識本體，而是某種思考方法(Sowder, 1992b)。然而透過許多國內外學者長時間不斷地分析與探討，我們仍然可以從廣泛的數常識意義中，找尋出較清晰的輪廓以及概括性的瞭解。

相關之研究報告(NCTM, 1989, 2000；Resnick, 1989；Reys, 1994；Sowder, 1989)指出擁有良好數常識的孩童應該具備的特徵陳述如下：

一、瞭解數字的基本意義與具備數字間之多重表徵的能力(NCTM, 1989, 2000；Resnick, 1989；Reys, 1994；Sowder, 1989)

例如：知道10個0.01是0.1而不是0.010；明白0與1之間有許多小數存在。

再如：能順利轉換

1

3所代表的圖形表徵；知道0.5與

1

2是相等的。

二、對於數字之間的相對大小有充份的認知(NCTM, 1989, 2000；Resnick, 1989；Sowder, 1989)

國小六年級學童在數常識異常作答之探究

61

例如：可以分辨0.99與0.100的大小；能比較

11

7與

11

8的大小。

三、瞭解運算對數字的影響(NCTM, 1989, 2000；Resnick, 1989；Reys, 1994；Sowder, 1989)

例如：知道「40多」與「50多」的和可能二位數也可能三位數。

四、能夠使用合適的參考點(referents)來解決問題(NCTM, 1989, 2000；Resnick, 1989；Reys, 1994)

例如：在計算 99×2.5 時，會使用 $(100 - 1) \times 2.5$ 來運算。

五、可以找尋出最有利的解題策略以及有效判斷計算結果的合理性(Resnick, 1989；Reys, 1994；Sowder, 1989)

例如：當計算

3

1

4

3

3

21++時，能將

3

1

3

21+先行運算，再計算

4

3

的部分。

再如：知道 314×0.512 比314 的一半還大一點點。

綜上所述，數常識雖然不容易被定義，但它卻是包含許多能力的一種知識，這些能力涵蓋了對數字與運算的了解、比較數字的大小、使用參考點進行測量以及對計算結果合理性的判斷等等。數常識亦是一種對數字的感覺或是一種直覺，這是需要長時間的培養並且不容易被教導的特質。進言之，具備數常識的孩童在數字、運算或數字與運算之間所產生的不同情境，都可以發展出最有利的解題策略並且可以靈活彈性地運用以解決許多的問題。

本研究根據林姿銓(2005)對於國小六年級學童數常識發展，經由試題的內容分析、項目分析、探索式因素分析以及驗證性因素分析，最終採用數常識四種組成份為其組織架構，即「比較數字的相對大小」、「數與運算的多重表徵」、「運算結果之合理性的判斷」以及「數與運算的關係」。此乃數常識測驗工具評定國小六年級學童所應具備的數常識解題能力。

二、異常作答的理論研究

異常作答(Aberrant Response)是指受試者在評量測驗或診斷系統時，發生不合常理或是非預期中的回答反應，本研究所言之異常作答，係指違反了Guttman(1944)的理想作答組型，亦即受試者在回答相對低於其能力水準的題目時，應該答對卻答錯；而回答相對高於其能力水準的題目時，應該答錯卻答對的情況。

Huang(2002)認為在考慮只有受試者與試題的一般測驗上，誤差的因素可能概略地來科學教育研究與發展季刊第四十四期

62

自受試者、試題及此兩者間的交互作用。受試者可能因為本身的因素例如粗心、猜測、疲倦等等，造成測量時的干擾而產生不準確或無意義的結果；同樣的，試題可能也會因為不同的受試者而有不同的反應；至於兩者間的交互作用產生的影響可能又更加複雜。因此當我們特別針對受試者因素所產生的作答誤差型態所進行之探討，皆屬異常作答的研究範圍。

更進一步的，當我們歸納出許多不同情況的異常作答進行探究，即成所謂的異常作答組型(aberrant response pattern)。Guttman(1944)曾提出理想的作答組型應該是受試者要答對相對低於其能力水準的題目並且答錯相對高於其能力水準的題目。據此，若是受試者在一個測驗中的作答情形有出現不符合Guttman 所提的理想作答組型，即是屬於異常作答組型。

三、內、外注意係數的探究

內、外注意係數的來源有「試題」及「受試者」兩個層面，然而本研究的重心在於學童的異常作答情形，所以在此只考慮「受試者」層面的內、外注意係數，定義如下：內注意係數(Winthin-Ability-Concern Index, W*)是指受試者應該答對其能力水準之內的題目卻答錯的一種程度偵測指標；外注意係數(Beyond-Ability-Surprise Index, B*)是指受試者應該答錯其能力水準之外的題目卻答對的一種程度偵測指標。此由Huang(2002)建立在D'Costa(1993a, 1993b)與Cheng(1997)的研究基礎上所設計出來的。

Huang(2002)針對異常作答的鑑別提出「差異距離(discrepancy distance)」的概念，定義為介於受試者的「修正能力」與其所未預期作答之試題難度間的距離總合，當以「受試者」的作答為主體時，差異距離的概念便是以試題的「難度」平均值來估計。藉此，Huang(2002)以理論的「最大差異距離(maximum discrepancy distance)」控制計算的差異距離，避免試題長度不同的影響，提出新修定的「內、外注意係數」指標公式，而黃財尉(2004)再次修正此公式中的差異距離，使其控制在單倍的理論「最大差異距離」之內，最後得到以「受試者」作答為主體的「內、外注意係數」指標公式如下：

(1) 內注意係數

*

$iW) =$
 受試者能力「內」差異距離總合
 最大理論差異距離 $\times 100$

$$= \left[\frac{1}{2} \right] 100$$

$$2 /) 1 ($$

$$) () 1 ($$

1

*

×

-

$$- \times - \sum =$$

K

$$q q u$$

iT

j

$$j iT ij$$

國小六年級學童在數常識異常作答之探究

63

(2) 外注意係數

*

$$iB) =$$

受試者能力「外」差異距離總合
 最大理論差異距離 $\times 100$

$$= \left[\frac{1}{2} \right] 100$$

$$2 /) 1 ($$

$$) ($$

1

*

.

×

-

$$- \times \sum$$

+ =

K

$$q q u$$

K

Tj

$$iTj ij$$

i

在此公式中， $i j u$

表示作答結果，答對為1，答錯為0； $s q'$ 是由易至難排列之試題難度值，而

*

iTq 則是當受試者總分為T 分時，所對應依難度順序排列之第T 道與第(T+1)道

試題的難度平均值；K 表示試題總數，而 $[\]$ 表示「下高斯整數」符號(lower Gauss integer)。

針對上述內、外注意係數指標公式，進而判斷「受試者」是否屬於異常作答的決斷

值，Huang(2002)利用排序技巧對不同受試者的能力比(ability ratio)與作答的錯誤比(error ratio)進行組合，依照不同能力與難度水準所建立的常模參照(norm referenced)標準為主，

透過每一個能力比X 錯誤比的細格內，找出第90、95 及99 百分等級時所對照的內、外注意係數值，超過所對應的決斷值，判斷作答的異常顯著性。經由統計的程序找尋出異常作答的判斷標準，使得內、外注意係數指標更具準確性與客觀性。

另外值得一提的是，同屬異常作答偵測指標的SCI 指標(Sato, 1975)及MCI 指標(Harnisch & Linn, 1981)，因其具有以下三種困難：首先，對於異常作答的「程度」無法有效地辨識，即不同的異常作答組型卻可能得到同樣的指標值；其次，對於兩種完全不

同的心理作答特徵無法有效地區別，例如粗心的失誤(carelessness)與好運的猜測(guessing)；最後，該指標缺乏統計上的判斷標準，僅透過某些經驗樣本或是最佳經驗法則而定，缺乏統計上的依據(黃財尉，2005)。然而，內、外注意係數W*及B*指標可以順利克服以上的困難，藉此本研究遂採用此較佳的異常作答偵測指標以探討國小六年級學童在數常識測驗之異常作答情形。

科學教育研究與發展季刊第四十四期

64

參 考 文 獻

一、研究架構與研究樣本

配合本研究的動機及目的，分別以「在數常識各組成成份的表現」、「性別」及「數學學業成就」為自變項，W*(內注意係數)及B*(外注意係數)為依變項作為本研究的架構，如圖3-1 所示，藉此了解內、外注意係數在不同變項間各自的影響力或差異情形為何？本研究之研究樣本為全國北、中、南及東區部份公立國民小學六年級學童，共1212 名，其中男生有623 名，女生有589 名；另外，為了區分樣本學童的數學學業成就，茲向施測學校收集學生在校數學學習成績，部分學校因考量學生成績的保密性，而無法提供以致於沒有全數收集之。需要注意的是，當研究樣本來自不同的團體時，必須將其變項上的分數轉換為百分等級量數或標準分數才能進一步比較或統計分析(王保進，2002)，於是研究者依序將各校所收集的學生數學分數轉為t 分數，而後作為其數學學業成就的判斷標準，區分為低成就(後33%)、中成就(中34%)及高成就(前33%)的學生，分別有242 人、248 人、241 人，共731 人。

自變項

在數常識各組成成份

的表現

性別

數學學業成就

依變項

W* (內注意係數)

B* (外注意係數)

圖3-1. 本研究之架構圖

國小六年級學童在數常識異常作答之探究

65

二、研究工具

本研究採用林姿銓(2005)所編製的「國小第二階段數常識線上評量測驗」為研究工具，共有23 道試題。此研究工具之試題經由內容分析、項目分析、探索式因素分析以及驗證性因素分析，決定出數常識的四種組成成份以及試題的題數(比較數字的相對大小6 題、數與運算的多重表徵6 題、運算結果之合理性的判斷6 題以及數與運算的關係5 題)，各組成成份之試題僅以下述作為代表：

(一)比較數字的相對大小：

(1-1)如果不用直接計算，你可以知道

5

3

與

9

4

何者大？

5

3

9

4

一樣大 無法比較

(二)數與運算的多重表徵：

(2-1)不用紙筆計算，請估算

100

98 98.0 + 的答案大約是多少？

98.98 10 2 1

(三)運算結果之合理性的判斷：

(3-1)下列何者的答案最大？

999.9×99.99 999.9×999.9 9999×9.999 99.99×99.99

(四)數與運算的關係：

(4-1)不用紙筆計算，請估算下列哪一種水果最便宜？

西瓜3 斤50 元，柳丁7 斤100 元，鳳梨1 斤35 元

西瓜 柳丁 鳳梨 無法估算

科學教育研究與發展季刊第四十四期

66

另一方面，該測驗試題之數常識四種組成成份的內部一致性係數依次

為.7521、.7788、.6558、.5771，整份測驗的 α 值為.8696，四種組成成份的建構信度依次

為.847、.861、.707、.634，整份測驗的建構信度為.916，平均變異抽取量為.734，重測信

度達.832，由此可知，該測驗工具具有良好之信效度。

三、 分析工方法

本研究採用黃財財(2005)所建立之內、外注意係數W*與B*指標來診斷孩童的異常作答情形。根據Huang(2002)的研究指出，內、外注意係數是一種以原始觀察值為分析元素並以個別化解釋為主之指標，透過Huang 在不同的條件下檢驗W*與B*之分配特性，發現其偏態(Skewness)係數與峰度(Kurtosis)係數都與0 相差不遠，近似常態分配(normal distribution)，且W*與B*指標範圍皆是大於等於0 之連續數值，可作為分析時之依變項。

由於W*(內注意係數)與B*(外注意係數)之分配特性符合t 考驗、變異數分析、相關分析及迴歸分析等統計方法之基本假設，研究者遂針對本研究目的進行合適之統計分析。針對目的，以Pearson 積差相關之統計法來探討學童在各組成成份答對總數與W*及B*的相關情形，並以逐步多元迴歸分析法來考驗學生於各組成成份答對總數在W*及B*的預測力；針對目的二，以t 考驗之統計法比較國小六年級不同性別學童在數常識異常作答的差異情形；針對目的三，以單因子變異數分析之統計法比較國小六年級不同數學學業成就學童在數常識異常作答的差異情形。另外，透過Levene 法進行變異數同質性檢定，若考驗結果發現其差異情形F 值達統計上之顯著水準時，即樣本變異數不具同質性時，將於t 考驗中改用校正公式進行假設檢定(林清山，1993)；於單因子變異數分析中選用適合變異數不具同質性之Dunnett T3 檢定方法(王保進，2002)。

肆、研究結果與討論

一、小六學童在數常識不同組成成份中的表現對異常作答指標之影響

響

表4-1 顯示學童在數常識各成份題目的答對總數之平均數與標準差，以及與W*(內注意係數)、B*(外注意係數)之相關。需要說明的是，由於W*與B*指標採用受試者能力差異距離的概念，因此將試題作答全對者與全錯者予以刪除不列入考量(黃財尉，2005)，共刪除10人全對及1人全錯，有效研究樣本總計1201人。其中除了成份四(數與運算的關係)之測驗題數為5題外，其餘皆為6題。結果顯示，成份一(比較數字的相對大小)的平均答對率約為76.8%，為全部最高，而成份三(運算結果之合理性的判斷)的平均答對率

67

約40.8%則為最低。另外，全體W*及B*之平均數分別為3.26與2.78，標準差為2.47與1.92(並未列於表中)。

表4-1：學童於數常識各成份答對平均數、標準差與W*及B*相關表(人數=1201)

相關

自變項 平均數(答對率) 標準差

W* B*

比較大小答對總數 4.61(76.8%) 1.52 -.637*** -.639***

多重表徵答對總數 2.69(44.8%) 1.87 .011 -.420***

合理判斷答對總數 2.45(40.8%) 1.59 .070** -.264***

數與運算答對總數 2.23(44.6%) 1.33 .027 -.288***

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

註：各成份答對總數係指所有研究樣本在數常識各組成成份試題中所答對的題目總數，在此用以作為預測W*及B*的自變項。

在與W*(內注意係數)的相關方面，學童在各成份答對總數以成份一(比較數字的相對大小)與成份三(運算結果之合理性的判斷)達到顯著，尤以成份一呈現極顯著的負相關，而成份三則為正相關。這裡所代表的意義是，學童在成份一答對總數愈高，其W*愈低，亦即應該答對其能力水準之內的題目卻答錯的情況比較不嚴重；另一方面，學童在成份三答對總數愈高，其W*愈高，亦即應該答對其能力水準之內的題目卻答錯的情況比較嚴重。

在與B*(外注意係數)的相關方面，學童在各成份答對總數皆與B*呈現極顯著的負相關。這裡所代表的意義是，學童在各成份答對總數愈高，其B*愈低，亦即應該答錯其能力水準之外的題目卻答對的情況比較不嚴重。表4-2與表4-3將針對此結果作更進一步的探討與說明。

表4-2 顯示學童在各成份答對總數之變項對於依變項W*(內注意係數)的逐步迴歸預測分析結果。由表4-2可知，由於各成份答對總數對於W*的預測皆達統計上的顯著($p < .001$)，因此各成份答對總數對於W*皆具有良好的預測力，可聯合解釋W*共53.6%的變異量，亦即對於W*的預測，各成份答對總數的四個變項可以提供53.6%的解釋能力。其中「比較大小答對總數」的解釋量為40.6%，為主要預測變項。就 β 值而言，僅「比較大小答對總數」的 β 值為負值，其餘皆為正值，其標準化迴歸方程式為「W*(內注

68

意係數) = $-0.810 \times$ 比較大小答對總數 + $0.210 \times$ 合理判斷答對總數 + $0.189 \times$ 多重表徵答對總數 + $0.114 \times$ 數與運算答對總數」。

表4-2：各變項對於W*(內注意係數)的預測分析結果摘要表(人數=1201)

選入變數 多元相關係

數(R)

決定係數

(R²)

R²

改變量F 值 標準化迴歸

係數β t 值

比較大小答對總數.637 .406 .406 818.936*** -.810 -36.958***

合理判斷答對總數.700 .489 .084 196.227*** .210 9.279***

多重表徵答對總數.725 .525 .036 90.350*** .189 7.873***

數與運算答對總數.732 .536 .010 26.965*** .114 5.193***

*** $p < .001$

由表4-2 可知，學童在「比較大小答對總數」對於W*(內注意係數)具有最高的解釋量，並與W*呈現負相關，此意謂著學童在成份一(比較數字的相對大小)的答對總數愈高，其W*可能相對愈低，亦即應該答對其能力水準之內的題目卻答錯的情形比較不嚴重；至於學童在其他成份答對總數對於W*雖也具有預測力，但是解釋量皆不大。

另外需要說明的是，在表4-1 中顯示學童於成份二(數與運算的多重表徵)及成份四(數與運算的關係)答對總數與W*並不具有顯著的相關，然而在表4-2 中卻呈現其迴歸係數是顯著的。發生此現象的原因係由於數常識乃是一種涵蓋眾多技能之綜合能力，各組成成份間難免涉及彼此概念以致於無法完完全全獨立出來(李威進，2004；林姿飴，2005)，而導致學童於各成份答對總數之自變項間，可能產生輕微的線性重合問題，使其與依變項原本相關不顯著之自變項，在迴歸分析時，卻造成迴歸係數反而顯著之矛盾現象，因而在此並不能具以作為解釋自變項與依變項間是否相關之依據(王保進，2002)。

表4-3 顯示學童在各成份答對總數之變項對於依變項B*(外注意係數)的逐步迴歸預測分析結果。由表4-3 可知，僅「合理判斷答對總數」外(因其對於B*的解釋量太小，對於B*的預測力太低所以未被選入)，其餘變項對於B*的預測皆達統計上的顯著($p < .01$)，因此對於B*皆具有良好的預測力，可聯合解釋B* 共44.6%的變異量，亦即對於B*的預測，此三個變項可以提供44.6%的解釋能力。其中「比較大小答對總數」的解釋量為40.9%，為主要預測變項。就β值而言，三個變項的β值皆為負值，其標準化迴歸方程式為「B*(外注意係數) = $-.552 \times$ 比較大小答對總數 + $-.179 \times$ 多重表徵答對總數 + $-.063$ 國小六年級學童在數常識異常作答之探究

69

× 數與運算答對總數」。

表4-3：各變項對於B*(外注意係數)的預測分析結果摘要表(人數=1201)

選入變數 多元相關

係數(R)

決定係數

(R²)

R²

改變量F 值 標準化迴歸

係數β

t 值

比較大小答對總數.639 .409 .409 828.704*** -.552 -23.346***

多重表徵答對總數.665 .443 .034 73.046*** -.179 -7.216***

數與運算答對總數.668 .446 .003 7.163** -.063 -2.676**

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

由表4-3 可知，學童在「比較大小答對總數」對於B*(外注意係數)具有最高的解釋

量，並與B*呈現負相關，此意謂著學童在成份一(比較數字的相對大小)的答對總數愈高，其B*可能相對愈低，亦即應該答錯其能力水準之外的題目卻答對的情形比較不嚴重；至於學童在其他成份答對總數對於B*雖也具有預測力並呈現負相關，但是解釋量皆不大。綜合上述研究結果，成份一(比較數字的相對大小)的答題表現對於內、外注意係數有較密切的關係，皆高達約41%的解釋量並且呈現負相關，亦即學童在成份一的答題狀況愈好，其發生異常作答的情況愈低。根據林姿飴(2005)與許清陽等人(2001)的研究結果，皆發現學童在「比較數字的相對大小」組成份中有較良好的表現，並且對學童而言，此類型的題目是屬於比較容易的。這裡反映了一個現象，在一般學童們皆認為比較容易的題目中，如果學童可以掌握得愈好，其內、外注意係數會愈低，異常作答情形愈不嚴重，亦即應該答對其能力水準之內的題目時比較不會錯過，而且也比較不至於會答對其能力水準之外的題目。這是一個我們在初步判斷學童異常作答情形時，值得注意的參考標準。至於學童在其他成份的答題表現，由於對內、外注意係數變異的解釋量並不大，所以在此並不作過多的推論。

二、小六之男女學童在數常識異常作答的差異情形

不同性別之國小六年級學童在數常識異常作答的差異情形，以t 考驗來加以分析，其結果如表4-4 所示(扣除10 人全對，1 人全錯，總計1201 人)：

科學教育研究與發展季刊第四十四期

70

表4-4：不同性別之小六學童在數常識之異常作答差異情形(人數=1201)

注意係數 性別個數 平均數 標準差 自由度 t值 差異情形

W* 男 614 3.46 2.61

女 587 3.04 2.31

1192.26 2.969** 男>女

B* 男 614 2.87 1.94

女 587 2.68 1.89

1199 1.695

** $p < .01$

由表4-4 可知，在W*(內注意係數)方面，因性別之不同則具有顯著的差異存在($t = 2.969$, $p < .01$)，就其平均數而言，男生平均數(3.46)顯著高於女生(3.04)，表示在數常識測驗中，發生能力內應該答對卻答錯的異常作答情形，男生比女生嚴重。發生此現象的原因，可能來自受試者本身的特質，例如依賴紙筆計算、粗心大意或閱讀能力不足等；亦可能來自試題內容的影響，例如測驗內容或方式與傳統迥異等。實際情形仍需進一步證實，然而，男生發生能力內應該答對卻答錯的異常作答情形比女生嚴重，乃是我們需要關切的問題。

另外，在B*(外注意係數)方面，以顯著水準 $\alpha = .01$ 來考驗時，沒有因性別之不同而具有顯著的差異($t = 1.695$, $p = .09 > .01$)，即發生能力外應該答錯卻答對之異常作答情形，男、女生沒有顯著的不同，因此我們不予贅述。但值得注意的是，男生平均數(2.87)仍然大於女生(2.68)。

三、小六之低、中、高數學學業成就學童在數常識異常作答的差異情形

不同數學學業成就之國小六年級學童在W*(內注意係數)及B*(外注意係數)之平均數、標準差及各成就樣本數如圖4-1 所示，經單因子變異數分析後，其差異情形如表4-5

所示：

國小六年級學童在數常識異常作答之探究

71

圖4-1. 不同學業成就學童之W*與B*之平均數(標準差)比較圖(人數=731)

表4-5：不同學業成就之學童在數常識異常作答之變異數分析摘要表(人數=731)

注意係數 變異來源 平方和 自由度 平均平方和 F 值 事後比較

W* 組間 210.492 2 105.246 17.558*** 低>中

低>高

組內 4363.762 728 5.994

總和 4574.255 730

B* 組間 442.485 2 221.242 70.788*** 低>中>高

組內 2275.303 728 3.125

總和 2717.788 730

*** $p < .001$

2.67(2.09)

3.04(2.32)

3.95(2.87)

1.82(1.36)

2.74(1.77)

3.73(2.10)

0

0.5

1

1.5

2

2.5

3

3.5

4

4.5

低(242) 中(248) 高(241)

學業成就(人數)

平均數(標準差)

W*

B*

科學教育研究與發展季刊第四十四期

72

由表4-5 可知，在W*(內注意係數)方面，因數學學習成就之不同則具有顯著的差異

存在($F = 17.558, p < .001$)，就其平均數而言，低成就學生平均數(3.95)顯著高於中成就學生(3.04)以及高成就學生(2.67)，而中成就學生平均數與高成就學生則無顯著的差異。這裡表示在數常識測驗中，發生能力內應該答對卻答錯的異常作答情形，低成就學生會比中、高成就學生來得嚴重。

另外，在B*(外注意係數)方面，因數學學習成就之不同則具有顯著的差異存在($F = 70.788, p < .001$)，就其平均數而言，低成就學生平均數(3.73)顯著高於中成就學生(2.74)，而中成就學生平均數又顯著高於高成就學生(1.82)。這裡表示在數常識測驗中，發生能力外應該答錯卻答對的異常作答情形，其嚴重程度依序為低成就 > 中成就 > 高成就學生。綜合上述研究結果，樣本中不論是W*(內注意係數)或B*(外注意係數)的平均數皆與數學學業成就之高低成反比，即成就愈低，W*與B*的平均數愈高，反之則愈低。此即表示，數學學業成就愈高的學生，相較於成就低的學生可能較不易有異常作答的情況發生，此情況包含能力內應該答對卻答錯以及能力外應該答錯卻答對的題目。發生此現象的原因，可能來自先前所言之受試者本身的特質與試題內容的影響，實際情形仍需進一步證實，然而，低成就學生發生異常作答情形比高成就學生來得嚴重，乃是我們需要善加留意的問題。

伍、結論與建議

一、結論

(一)數常識組成成份中「比較數字的相對大小」與異常作答指標呈顯著負相關
「比較數字的相對大小」組成份同時是學童們認為比較容易的題型，而學童們異常作答情形的徵兆亦在此透露些蛛絲馬跡，即學童在「比較數字的相對大小」組成份的答題表現愈好，內、外注意係數會愈低。雖然不能因此而武斷地判定學童若能在他們認為比較容易的題目中表現愈好，發生異常作答的情形就必定不嚴重，但對於我們在測驗評量中初步探究學童們異常作答的情形時，仍提供了一項值得注意的參考標準，關鍵則在於學童們面對他們認為比較容易的題型時，若能掌握得愈好，其異常作答情形較不嚴重，反之，異常作答情形會較嚴重。

(二)小六男童相較於女童更容易發生異常作答的情形

在數常識測驗中，男童發生能力水準內應該答對卻答錯之異常作答的情形比女生嚴重，而能力水準外應該答錯卻答對之異常作答情形，雖沒有顯著的不同，國小六年級學童在數常識異常作答之探究

73

但男童仍是略比女生嚴重。雖然造成此現象之潛在因素可能涉及許多複雜的層面，需要進一步地探究，但此項結果對於往後各式的測驗評量，乃是一個值得注意的問題。

(三)小六數學低成就的學生相較於中、高成就學生更容易發生異常作答的情形
數學學習低成就固然包含了許多的原因，然而異常作答的情況依舊反映在此種現象上，不論是能力水準內應該答對卻答錯的題目或是能力水準外應該答錯卻答對的題目，數學低成就的學生發生異常作答的情形仍然比中、高成就學生嚴重，雖然其背後的因素可能涵蓋各種可能的原因，需要更深入的探討才能得知，然而，這仍是一項不容忽視的指標。

二、建議

(一)在極力發展各式評量測驗的同時，了解學童異常作答情形有其必要性
任何評量工具的測驗結果都希望可以更具意義，可以幫助研究者與教學者更清楚掌握學童的概念發展情形，相信這是每位教師的出發點與共同的目標。而本研究提供教師在診斷學童異常作答情形的一個範例，讓教師們了解評量測驗的內容、不同的性別以及不同學業成就的學生與測驗時異常作答情形的關聯性，讓測

驗的結果有更進一步的參考標準以及更具意義的詮釋。

(二)教師應重視學童異常作答的情形，適當給予協助與叮嚀，盡量避免犯錯。透過異常作答指標的偵測，發現學童在測驗評量時具有異常作答的情形，教師不應該忽略此問題，應該進一步探究其造成異常作答的原因，尤其是學生能力水準內應該答對卻答錯的情況，教師應給予學童適當的協助，叮囑他們需抱持著更加謹慎小心的態度來處理問題，如此異常作答的情形應可較為改善。以本研究結果而論，對於發生異常作答情形相較嚴重之男生及學業低成就的學生，教師更應當特別注意。

(三)從各式不同的測驗中探究學童異常作答的情形，找尋其可能的常模標準。透過本研究可得知：學童在他們認為比較容易的題型中表現愈好，則異常作答的情形較不嚴重；男生及學業低成就的學生其異常作答情形較女生及學業高成就的學生嚴重。由於本研究只侷限於單一的評量工具，若欲以此研究進行廣泛之推論，仍須謹慎小心。更甚者，若欲建立其具有參考指標的常模標準，有賴後續研究作更進一步的分析及更多元化的探討，相信不論對於教師或學生，在教學成效上將會是一大貢獻。

科學教育研究與發展季刊第四十四期

74

參考文獻

- 王保進(2002)：視窗版SPSS 與行為科學研究。台北：心理。
- 吳德邦、鄭媿嫻(2000)：臺灣中部地區國小二至六年級學生乘除法概念之研究。國立台中師院學報，14，341-364。
- 李佳香、黃佩晶、黃千玲、張麗麗(2000)：接受不同教學法之國小三年級學童在傳統評量與實作評量上表現之比較研究。學生學術論文年賽優勝作品專輯，87 學年度(上)(頁173-233)，屏東：國立屏東師範學院。
- 李威進(2004)：資訊融入九年一貫數學領域第一階段數學測驗之研究—以數字常識為例。嘉義：國立嘉義大學國民教育研究所碩士論文。
- 林清山(1993)：心理與教育統計學。台北：東華。
- 林素微(2002)：國小高年級學童數感特徵暨數感動態評量發展之探討。台北：國立台灣師範大學教育心理與輔導研究所博士論文。
- 林姿飴(2005)：電腦化數常識量表之編製及其發展之研究—以九年一貫數學領域第二階段學童為例。嘉義：國立嘉義大學數學教育研究所碩士論文。
- 陳淑絹(1999)：不同性別與學習潛能之國小學童在學習適應、注意力與記憶力之比較研究。國立台中師院學報，13，73-99。
- 許清陽(2001)：國小高年級學童數常識發展之研究。嘉義：國立嘉義大學國民教育研究所碩士論文。
- 許清陽、楊德清、李茂能(2001)：國小高年級學童數字常識評定測驗編製之研究。科學教育學刊，9(4)，351-374。
- 曹宗萍、周文宗(1998)：國小數學態度量表編製之研究。教育學術研討會論文集：數理教育組，87 學年度，1211-1246。
- 黃財尉(2004)：內外注意係數指標發展之研究(I)。國科會研究成果報告，NSC92-2511-S-415-006。
- 黃財尉(2005)：內外注意係數指標決斷值的建立與檢視。嘉義：濤石文化。
- 楊德清(2000)：國小六年級學生回答數字常識問題所使用之方法。科學教育學刊，8(4)，379-394。
- 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之研究。科學教育學刊，10(3)，233-260。

劉慧如 & Garfield, J. B. (2002)：統計推理的性別差異。教育心理學報，**34(1)**，123-137。

Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Melbourne: Curriculum Corporation.

Cheng, W.D. (1997). *An exploratory study examining the relationships among total score and different caution and fit indices*. Unpublished thesis. The Ohio State University.

D'Costa, A. (1993a). *Extending the Sato caution index to define the within and beyond ability caution indexes*. Paper presented at convention of National Council for Measurement in Education, Atlanta, GA.

D'Costa, A. (1993b). *The validity of the W, B and Sato Caution indexes*. Paper presented at the Seventh International Objective Measurement Conference, Atlanta, GA.

Emanuelsson, G., & Johansson, B. (1996). *Kommentarer till kursplan I matematik, Lpo 94*. (Nn-statutory guidance to the new Swedish curriculum of mathematics).

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, 9, 139-150.

Harnisch, D. L., & Linn, R. L. (1981). Analysis of item response patterns: Questionable test data and dissimilar curriculum practices. *Journal of Educational Measurement*, 18(3), 133-146

Huang, T. W. (2002). The power of the Beyond-Ability-Surprise Index and the Within-Ability-Concern Index for detecting person/item aberrances: *A Monte Carlo study*. Unpublished doctorate dissertation, The Ohio State University.

Japanese Ministry of Education. (1989). *Curriculum of mathematics for the elementary school*. Tokyo: Printing Bureau.

Kogut, J. (1986). *Review of IRT-based indices for detecting and diagnosing aberrant response patterns* (Research Report No. 86-4). Enschede, Netherlands: University of Twente, Department of Education.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *The principles and standards for school mathematics*. Reston, VA: NCTM.

Resnick, L. B. (1989). Defining, assessing, and teaching number sense. In J. T. Sowder, & B. P. 科學教育研究與發展季刊第四十四期

Schappelle(Eds.), *Establishing foundations for research on number sense and related topics: Report of a conference* (pp.35-39). San Diego: San Diego University, Center for Research in Mathematics and Science Education.

Reys, B. J., Barger, R., Dougherty, B., Hope, J., Lembke, L., Markovits, Z., Parnas, A. et al. (1991). *Developing number sense in the middle grades*. Reston, VA: National Council of Teachers of Mathematics.

Reys, B. J. (1994). Promoting number sense in the middle grades. *Mathematics Teaching in the Middle School*, 1(2), 114-120.

Sato, T. (1975). *The construction and interpretation of S-P tables*. Tokyo: Meiji Tosho.

Silver, E. A. (1989). On making sense of number sense. In J. T. Sowder, & B. P. Schappelle (Eds.), *Establishing foundations for research on number sense and related topics: Report of a conference* (pp.92-96). San Diego: San Diego University, Center for Research in Mathematics and Science Education.

Sowder, J. T. (1989). Introduction. In J. T. Sowder, & B. P. Schappelle (Eds.), *Establishing foundations for research on number sense and related topics: Report of a conference* (pp.1-5). San Diego: San Diego University, Center for Research in Mathematics and Science Education.

Sowder, J. (1992a). Estimation and number sense. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 371-389). New York: Macmillan.

Sowder, J. (1992b). Making sense of numbers in school mathematics. In G. Leinhardt, & R. Hatterp (Eds.), *Analysis of arithmetic for mathematics teaching* (pp.1-51). Hillsdale, NJ : Erlbaum.

Yang, D. C. (2003). Teaching and learning number sense--An intervention study of fifth grade students in Taiwan. *International Journal of Science and Mathematics Education. 1*(1), 115-134.

國小六年級學童在數常識異常作答之探究

77

A Study of Aberrant Responses on Number Sense for 6th Grade Elementary School Students

Chun-Ku Yeh¹ Der-Ching Yang² Tsai-Wei Huang³

^{1,2}The Graduate Institute of Mathematics Education, National Chiayi University

³The Department of Counseling, National Chiayi University

Abstract

The purpose of this study was to investigate aberrant responses on the number sense test (Chih-I Lin, 2005) for elementary school students who finished 5th grade courses. The Within-Ability-Concern Index (W*) and the Beyond-Ability-Surprise Index (B*) were used to detect aberrant responses of students on three different conditions of components of number sense, gender, and mathematics achievements. There were 1212 sixth grade elementary school students in the test. Through many analyses, results showed that (1) Component of “comparing number magnitude” contributed most percent of aberrances in both within-ability-concern and surprise-ability-surprise types of misfit, (2) male students responded on number sense remarkably aberrantly higher than female students, and (3) students with lower mathematics achievements performed aberrant responses remarkably more severe than the medium and the high achievements students.

According these findings, several recommendations for teachers, parents, education units were also provided in this study.

Key words : number sense, aberrant response, W*(within-abiliby-concern index), B*(beyond-abiliby-surprise index)

—