

國立台東大學教育系（所）教學科技碩士班

碩士論文

指導教授：李偉俊 博士

國小高年級成語電腦適性測驗之發展研究

The logo for NTU Library is a large, semi-transparent watermark in the background. It features a stylized green eye shape with a blue globe as the pupil. A green hand is positioned above the globe, and an orange pen nib is positioned below it. The text 'NTU Library' is written in a light blue, sans-serif font across the top of the eye shape.

研究生：郭依婷 撰

中華民國九十七年八月

國立台東大學教育系（所）教學科技碩士班

碩士論文

國小高年級成語電腦適性測驗之發展研究



研究生：郭依婷 撰

指導教授：李偉俊 博士

中華民國九十七年八月

國立台東大學

學位論文考試委員審定書

系所別：教育學系（所）教學科技碩士在職專班

本班 郭依婷 君

所提之論文 國小高年級成語電腦適性測驗之發展研究

業經本委員會通過合於  碩士學位論文 條件  
 博士學位論文

論文學位考試委員會：

吳振孝

(學位考試委員會主席)

郭達淳

李偉俊

(指導教授)

論文學位考試日期：97 年 8 月 9 日

國立台東大學

- 附註：1. 本表一式二份經學位考試委員會簽後，送交系所辦公室及註冊組或進修部存查。  
2. 本表為日夜學制通用，請依個人學制分送教務處或進修部辦理。



# 誌謝辭

首先，誠摯地感謝指導教授李偉俊博士，悉心的教導使我得以一窺電腦測驗這塊領域之奧妙，在其中不時的討論指引我正確的方向，使我在撰寫論文的這段期間獲益匪淺。再來，也很感謝郭達源博士及黃振榮博士對我論文嚴謹度提昇的建議，讓此份論文的研究更加地完善。

這份論文的完成另外亦得感謝，文山國小的潘老師、中山國小的麥老師、後庄國小的宋老師、五甲國小的林老師及金潭國小的江老師，有你們的協助幫忙，也才能讓此研究的進行得以順利圓滿。

三個暑期來的研究所生涯，教室裡生活的共同點滴，師長們的諄諄教誨、同學們互相扶持的陪伴，讓這段研究生活變得充實而精彩。

感謝教授不厭其煩地指出我研究中的缺失，感謝朋友總能在我迷惘時給我解惑與鼓勵，也感謝同學們的幫忙，恭喜我們順利地走過這三個暑期，一起邁向新的里程碑。

家人們在背後的默默支持更是我前進的動力，若沒有他們的體諒與包容，相信這段研究的日子，將是另一番不同的光景。

最後，謹以此文獻給我所有該感謝的人，謝謝你們！

郭依婷 謹誌

中華民國九十七年八月

# 國小高年級成語電腦適性測驗之發展研究

作者：郭依婷

國立台東大學 教育系所

## 摘要

本研究是結合編擬題庫、工具發展、效能及滿意度評估的研究。旨在發展「國小高年級成語電腦適性測驗」，透過 Flash 的施測介面提供師生在學習成語的過程中，擁有一個立即回饋資訊的環境。

研究目的有三：

- 一、編擬與建立國小高年級學童成語電腦適性測驗。
- 二、探討國小高年級學童成語電腦適性測驗之發展模式。
- 三、探討成語電腦適性測驗之效能及滿意度情形。

本研究的進行，首先對目前國小教科書之成語部分作內容分析後，再以所整理出的成語及收集坊間題庫為參考，設計成語題庫試卷，並以國小一至六年級的學生，每個年級各抽一班來做試卷預試。所得的預試樣本，再以 Tester 2.0 軟體，完成題目難度、鑑別度的計算後，依每個題目的難度及鑑別度整理排列成適性測驗金字塔階層的出題順序。完成適性測驗題庫的建立後，研究者研擬使用 Flash 多媒體的呈現方式讓學生使用電腦來完成此一電腦適性測驗。

研究結果方面，本研究得到如下結論：（一）利用 Tester 2.0 軟體進行試題分析具便利性（二）利用金字塔式適性測驗便於達成學生能力的估計（三）Flash 施測介面與金字塔適性測驗題庫內容可整合為一（四）成語電腦適性測驗具有多項功能（五）成語電腦適性測驗能力估算效能良好（六）使用者對成語電腦適性測驗感到滿意。

**關鍵詞：**國小高年級、成語、電腦適性測驗

# **A Research In The Development of the Idiom Computer Adaptive Testing for High Grade of Elementary School.**

Yi-Ting Kuo

## **Abstract**

This research is to develop the idiom computer adaptive testing, which is comprised of compiled and drafted the exam pool, teaching tool development, efficiency evaluation and approval. By using this system, teachers and students can get the feedback information immediately from the flash interface during learning the idiom.

There are three main purposes for this research, as followings,

- A. To design and to build the idiom computer adaptive testing of high grade of elementary school.
- B. To study the development method for the idiom computer adaptive testing of high grade of elementary school.
- C. To study the efficiency evaluation and approval for the idiom computer adaptive testing of high grade of elementary school.

For this research, first we analyzed the content of the idiom part from the current textbook of the elementary school; then, we designed the exam pool for the idiom test, which referred and collected from the current textbook of the elementary school and the existed exam pool; third, we selected one of classes from each grade in the elementary school to do the idiom test. From the test result, we used the Tester 2.0 to calculate the difficult and distinguish range; then, we built the idiom computer adaptive testing triangle level by following the difficult and distinguish range from each test item. After finishing getting the adaptive testing after the foundation of exam pool, researcher will grind, plan, and have students have access to computers for Flash multimedia programs to finish this computer adaptive testing.

For result of study, this research intends to get the following conclusion: (A) Utilize Tester 2.0 software to carry on the examination question for analyzing conveniences. (B) It is easy to reach the estimation of student's ability to utilize the adaptive testing of pyramid. (C) Flash multimedia programs can construct and examine interfaces and can be combined exactly with the adaptive testing exam pool content of the Pyramid into one. (D) The idiom computer adaptive testing has multiple functions. (E) The idiom computer adaptive testing, systematic ability estimate efficiency to be good. (F) The user feels satisfied with the idiom computer adaptive testing.

**Keyword : High Grade of Elementary School, Idiom, Computer Adaptive Testing**

# 目次

<b>第一章 緒論</b> .....	<b>8</b>
第一節 問題背景與研究動機.....	8
第二節 研究目的.....	10
第三節 名詞釋義.....	10
第四節 研究範圍.....	11
<b>第二章 文獻探討</b> .....	<b>12</b>
第一節 成語的探究.....	12
第二節 目前語文教學的相關研究.....	18
第三節 國小成語教學的相關研究.....	22
第四節 電腦測驗的相關研究.....	27
<b>第三章 研究方法</b> .....	<b>38</b>
第一節 研究架構.....	38
第二節 研究對象.....	39
第三節 研究設計.....	40
第四節 研究工具.....	40
第五節 資料分析.....	41
第六節 研究流程.....	42
<b>第四章 結果與討論</b> .....	<b>44</b>
第一節 成語電腦適性測驗建置過程與結果.....	44
第二節 成語電腦適性測驗之發展模式.....	51
第三節 成語電腦適性測驗之滿意度評估.....	70
<b>第五章 結論與建議</b> .....	<b>80</b>
第一節 主要發現.....	80
第二節 結論.....	83
第三節 建議.....	86
第四節 研究限制.....	89
<b>參考文獻</b> .....	<b>90</b>

附錄一.....	98
附錄二.....	102
附錄三.....	106
附錄四.....	110
附錄五.....	112



# 表次

表 2-1：對試題反應理論發展有實際貢獻的代表性作者和著作.....	35
表 3-1：研究時間進度表.....	43
表 4-1：國小成語測驗題庫分析一覽表.....	46
表 4-2：五、六年級預試之內部一致性係數和差異係數整理表.....	47
表 4-3：正式施測題庫按難易度排列及編號表.....	48
表 4-4：正式施測之抽樣學校及班級數、人數一覽表.....	70
表 4-5：正式施測之十題問卷題目.....	71
表 4-6：正式施測時間規劃流程表.....	72
表 4-7：正式施測利用 TESTER 2.0 分析結果摘要表.....	73
表 4-8：「成語電腦適性測驗」與「傳統紙筆測驗」之相關分析.....	74
表 4-9：使用者評估結果次數分配表.....	76



# 圖次

圖 2-1：金字塔式適性測驗示意圖.....	30
圖 3-1：研究架構圖.....	39
圖 3-2：研究流程圖.....	42
圖 4-1：七層金字塔式適性排列的線上 FLASH 題庫.....	50
圖 4-2：「成語電腦適性測驗」之資料流程圖.....	55
圖 4-3：「成語電腦適性測驗」之開頭歡迎畫面.....	57
圖 4-4：「成語電腦適性測驗」之注意事項說明頁.....	58
圖 4-5：「成語電腦適性測驗」之初始題.....	59
圖 4-6：「成語電腦適性測驗」其餘 27 題之作答畫面.....	64
圖 4-7：「成語電腦適性測驗」測驗結果之所得級分畫面.....	65



# 第一章 緒論

本研究是結合編擬題庫、工具發展、效能及滿意度評估的研究。旨在發展「國小高年級成語電腦適性測驗」，透過 Flash 的施測介面提供師生在學習成語的過程中，擁有一個立即回饋資訊的環境。

## 第一節 問題背景與研究動機

本節共分二部份，依次說明：「問題背景」及「研究動機」。

### 壹、問題背景

中國文化博大精深，擁有五千年悠久長遠的歷史，其中成語是世界上最具創意的一種語文結構，充滿著高度的智慧和趣味（敖幼祥，2006）。成語是先民語言的智慧，經過一再的使用和提煉所形成一種長期慣用、形式簡潔、意思精闢的定型片語或短句，所以成語是中國文化中一個非常特殊且優美的語言形式（史遷，2006）。成語乃出自於歷史典故而流傳於後世，是最富意義及智慧的文化結晶，千百年來，中國一直是個文學的民族，成語在其中更是扮演著不可或缺的角色（吳霞雲，1999）！

成語是語文教學與學習的主要核心主題，能活用成語，便是精通中文的第一步。而中文是小朋友最根本的能力，如果中文沒有辦法學得好，其他學科的理解能力就會降低，所以家長們才會如此重視中文能力（李書瑩，2007）。但是，成語有的來自歷史典故、有的來自古代文獻，除了典故本身的涵意外，還有後來用法的引申義，且有其特定的用法，小朋友們若不太容易理解這些成語的意思，不僅不容易記憶，而且還會經常誤用而鬧笑話。

董國峰（1990）指出將適當的成語穿插於文句之中，會使文章顯得簡潔有力；正確地使用成語，更能避免文章拖泥帶水。因此在學習成語時必須通盤了解、徹底讀通，才能將之靈活運用，使語文更為精練生動，有助於高層次的語文認知學習。

隨著網路時代的降臨，網路語言也隨之快速發展，時下年輕人溝通的文字已漸由火星文等網路符號所取代，甚至出現於國中、大學的作文基測中，顯示出學生國文程度的低落。有感於此，為挽救國粹，成語教學與評量的方式實有需要隨著時代的腳步而多元化。

沈中偉（2005）指出評量是指教師用來測驗學生學習表現和態度的方法，其目的在於診斷學生學習困難、評估其學習成效與學習歷程，同時也評量教師的教學設計、教學方法、教學媒體是否有其成效。而目前所使用的傳統紙筆測驗早已缺失百出、不敷使用，所以沈中偉（2005）同時也提出，近年因應資訊科技時代的來臨，教學評量與資訊科技的結合，以電腦化適性測驗（Computerized Adaptive Testing）與線上網路題庫進行測驗是未來教學評量的趨勢。

綜上所述，成語是中國文化的智慧結晶，但隨著時代的改變，學生成語能力普遍低落，國人實有需反思，是否成語的教學與評量方式已不符合時代的趨勢。若能結合資訊科技發展出一套成語電腦適性測驗，不僅能診斷學生學習的困難也能評估教師的教學是否有其成效，再經檢討精進，才能讓學生的成語學習達到更適性的發展。

## 貳、研究動機

基於以上的研究背景，研究者的研究動機有三，分述如下：

國語翰林版教科書（2006）第十一冊中曾提到，說話、作文能恰當的使用成語，就能很清楚的表達意思；但是成語用錯，就會鬧笑話了。因此在說話及作文方面若能將成語運用其中是可以提昇表達能力的。但在現今學生的基測作文中常有網路火星文的現象發生，也常有誤用及鬧笑話的情形。試想，會有這樣的情形產生，由研究者所整理的國小教科書成語教材之內容分析—以翰林版國語課本為例（2007）的文章結論中，可以發現國小教科書目前尚無針對成語教材作有規劃的課程與教學，學生的成語能力又從何培育而來呢？而這也是研究者會想針對國小成語這個主題部份做相關研究，此為動機之一。

岳修平（2003）等學者指出學校教育在教授成語時，教師多半要求學生背誦成語辭典，評量時也往往流於線性的回憶或再認反應，學習者無法有效學習，尤其不會實際運用，學習效果不佳。一個好的評量於學生方面要能評估出學生的學習成效及學習困難處，再針對其錯誤之處進行精熟學習；於教師方面也要能評估出教學是否有達到成效，若未達標準，教師應反思自己的教學方式是否有需再行修正的地方，此為動機之二。

柯光霖（2006）的研究指出實施適性化教學是解決學生個別差異的良策，正因為如此，在教學評量上已不宜再以傳統的評量方式來進行，而是需要一套合理的多元化評量方式，才能讓學生達到適性的學習與發展。因此，研究者欲研擬建立一套成語電腦適性測驗，提供目前國小教師一個多元化的評量方式，以期讓學生對於成語有更適性的學習，此為動機之三。

## 第二節 研究目的

基於以上的研究背景與動機，本研究之研究目的有三，臚列如下：

- 壹、 編擬與建立國小高年級學童成語電腦適性測驗。
- 貳、 探討國小高年級學童成語電腦適性測驗之發展模式。
- 參、 探討成語電腦適性測驗之效能與滿意度情形。

## 第三節 名詞釋義

茲將本研究所涉及的重要名詞，分別界定如下：

### 壹、 國小高年級學童

國小高年級學童，是指就讀於國小五、六年級的學生。

### 貳、 成語能力

成語簡單的界說是：它或者有出處，或者是俗諺，或者是習慣用語。有出處的成語，或者出於一則很具涵義的古事，或者出於某一典籍中的精言妙語。而且既謂之成語，則必然具備約定俗成、固定通用的語言型式，非一般人可以隨興杜撰。即使在運用之際，偶會更動一二字眼，或倒裝句式，也不得改易其義（顏崑陽，1992）。故成語能力為學生懂得成語的解釋及用法，並能將之運用於說話、作文中。

## 參、電腦適性測驗

電腦適性測驗不僅可以節省施測時間，更可以達到精確估計考生能力或某種潛在特質的目的。因為當測驗的難度能夠適合考生的能力程度時，這時測驗所測量到的考生能力最為精確。所以，最理想的施測狀況是：能夠針對每位考生不同的能力程度，來提供適合個別情況的測驗方式，這也就是電腦適性測驗所欲探討的內容。

### 第四節 研究範圍

本研究的研究範圍，就以下三個部份，說明如下：

#### 壹、就取樣範圍而言

本研究以高雄縣的國小為取樣範圍，共立意取樣五校，無法與高雄縣以外地區學校做進一步的比較。

#### 貳、就研究對象而言

本研究之研究對象，為高雄縣國小高年級的學生，無法推論至其它學習階段的學生。

#### 參、就研究主題而言

本研究所發展的電腦適性測驗是針對國小教材的成語內容部份，無法推論至其它書籍或教材的成語範圍上。

## 第二章 文獻探討

本章文獻探討共分四節，第一節是成語的探究，探討成語的界定、特性及使用性。第二節則是目前語文教學的相關研究，有語文的重要性、相關的理論與研究、教學現況整理，而語文教學中與本研究最有相關的作文與成語，也放至在本節探討中。第三節是研究者整理目前國小一至六年級的語文教科書中對於成語這方面所得的發現與結論，另外還有一些學者對於成語的原則、方法與教學部份的相關整理。最後第四節是電腦適性測驗的相關研究，共分為意義、發展類型及相關理論，分別說明整理於此節中。

### 第一節 成語的探究

成語是中國文化智慧特有的結晶，也是現代漢語的固定語中表現最生動、最為廣大民眾所喜用的部分之一，是現代漢語詞彙中的重要成員（余桂林，2001）。以下就成語的部份，分析如下：

#### 壹、成語的界定

江惜美（1994）提到成語通常以四個字為基本的結構，但事實上，並不是四個字的就是成語，成語只是以四個字居多。以下就成語的定義作界定，才不易造成成語觀念混淆，而有誤用、濫用的情形發生：

關於「成語」的界說，各家說法如下引述：

牛頓出版社 1992 年出版的《故鄉實用成語辭典》對成語的釋義是這樣的：「既謂之成語，則必然具備約定俗成、固定通用的語言型式，非一般人可以隨興杜撰。即使在運用之際，偶會更動一二字眼，或倒裝句式，也不得改易其義。」

吉林文史出版社 1997 年出版的《語文基礎知識詞典》中對成語的釋義是這樣的：「漢語成語以四個字組成的為基本格式。有來源於神話寓言的；有來源於歷史故事的；有來源於詩文語句的；有口頭流傳的。」史遷（2006）也提到成語有的來自歷史典故、有的是慣用語，有些成語的涵義較為深入、且有其特殊的用法。

國語日報印行《新編國語日報辭典》(2000):「社會上習用的古語以及流行的詞語。成語結構多樣,來源的途徑很多,有可以從字面看出的,有本來有典故的,有很典雅的,有很粗俗的。」

何謂成語?劉月華(2005)認為成語應該是形式和意義都很固定的、習用的,以四字形式為主的短語。陳桂成(2005)認為成語必須是短語,不能是詞;有固定格式—四言短語是其主要形式;是書面語,不能是口語;必須有字面外的深層涵義、高度凝練的豐富內涵。

成語指的是人們長期以來慣用的、簡潔精闢定型的詞組或短語,用成語表達不僅言簡意賅,並且能增強修辭效果。成語也是一種結構固定而凝練、內涵豐富而深刻的語言現象,更是人類的歷史和文化在人們所使用語言中的體現和反映(徐康,2005)。

由賀非雨(2006)主編的《成語300首》最後附錄的成語常識中指出:

成語是表示一般概念的固定詞組或句子,絕大部份由四個字組成。少於四個字的和多於四個字的,在成語中都占絕對少數。

成語和專名不同:專名和成語雖然都是由幾個字組成的固定詞組,但是專名是表示特定概念的,例如「李庄小學」只是一個學校獨有的名稱(當然,可能有重名)。成語則是表示一般概念的,例如「七零八落」,可以用來形容凋零的花朵,也可以用來形容被打敗的軍隊,還可以用來形容稀疏零散的人群。

成語跟諺語不同:諺語絕大部份是句子而不是詞組。諺語往往具有沈厚的口語色彩,不像成語那樣具有文言色彩。諺語的形式,不像成語那樣整齊,例如「坐山觀虎鬥」。當然,成語和諺語的區別並不是非常分明的。

成語跟歇後語不同:歇後語是一種固定的句子,一定要由兩部分構成,前一部分是以事物作比喻,後一部分從正面作說明。如「泥菩薩過河,自身難保」是歇後語,而不是成語。

成語和由四個字組成的普通詞組也不一樣。凡是一個詞組裡的詞可以抽換的,就應該把它看做普通詞組。成語是一個整體,不能隨意抽換其中的成分。

成語具有意義的整體性、結構的凝固性等兩個基本特徵。成語雖然是一種特殊詞彙,可以進行結構分析,可視為短語,但它在語言中的性質和作用,仍然和普通詞彙一樣,它的兩個基本特徵及其在語言中使用中的可重複性都決定了它可以作為句法分析的最小單位—詞的出現(朱麗芳,2007)。

綜上所述，成語有固定的組成型式，雖偶有倒裝，但也不改其義。成語的來源有的來自歷史典故、神話寓言等，有其特定的用法。成語多以四字組成，與專名、諺語、歇後語不同。最後，從中國文化研究院成語部分的網頁中可以知道唯有了解成語的來源與典故，才能知道成語真正的意涵。

## 貳、 成語的特性

成語是人們長期以來習用的、形式簡潔而意思精闢定型的詞組或短句。語言是文化的載體，漢語成語更是一種濃縮的文化。濃縮的結晶需要稀釋，成語的魅力也需要發掘才能展現出來（呂杰，2007）。雖然成語的來源各有不同，但其本身卻有共同的語文特點：

### 一、 字數：

成語絕大部份均是由四個字組成而來的，常被用來表示一般概念的固定詞組或句子，其中，少於四個字的成語和多於四個字的成語，在成語中都是占絕對的少數。所以成語一般來說都是四個字的，間或有不同，但都比較少見。

### 二、 基本結構：

成語的結構基本固定，不過結構的類型是多種多樣的。一般的四字成語可分為兩個部分，前兩個字為一部分，後兩個是另一部分。兩個部份之間有某種關係，有的是限制與被限制的關係，如「一心一意」、「千變萬化」、「搖頭晃腦」、「半斤八兩」；有的是修飾與被修飾關係，如「驚弓之鳥」、「鼎鼎大名」、「忐忑不安」、「嫣然一笑」；有的是陳述與被陳述關係，如「夜郎自大」、「水火不容」、「唇齒相依」、「能者為師」；有的是支配與被支配關係，如「別開生面」、「不咎既往」；還有行為與動機的關係，如「刻舟求劍」、「揠苗助長」。四字成語除了有「二二」組合外，還有「三一」組合，如「目不識丁」、「名副其實」、「一衣帶水」。此外，還有一些不止由四個音節組成的成語，其組成部分之間的關係就複雜一些。如「醉翁之意不在酒」，當中已包含了形容詞、名詞、助語及副語，想要找出其組成成分的關係，便十分困難。不過這些成語多半帶有慣用語、諺語、俗語性質，和四字成語有所不同。

### 三、 修辭傾向：

許多成語本身就是借助修辭手法造成的，而大部份的成語都傾向於採用比喻和誇張的修辭技巧。比喻：「金玉之言」－以金玉比喻一些意見的寶貴。「如飢似渴」－以飢渴的狀態比喻要求十分迫切。「狼心狗肺」－以狼狗的心思比喻壞人的心胸險惡。誇張：「口若懸河」－誇張地形容人講話時，像急流的水傾瀉下來一樣，滔滔不絕。「彈丸之地」－誇張形容地方像一顆彈丸般細小。有些成語則採用了以下的修辭手法：疊字：心心相印－「心」一字被重疊使用。「十年樹木，百年樹人」－兩個語句對稱，「十年」對「百年」，「樹」字在兩句中意思相同。雙聲：不倫不類－「倫」和「類」的聲母相同。疊韻：天真爛漫－「爛」和「漫」的韻母相同。

### 四、 語法運用：

成語的意思各有不同，若把它們放在不同的句子中，便會擔當不同的功能。

以上的特性，是依照古人求學過程、當時的現象、環境因素等不同原因，並經過長期以來的修潤、充實、選擇，一直流傳下來。因此成語也具有文化的傳承性，為約定俗成的習慣用語，其組織結構不可以任意變易，亦不得隨興杜撰，每句成語均蘊含著古人智慧的光芒，更可視為先民生活的寫照。

## 參、 成語的使用

成語的文字非常簡煉，又有固定的形式，本身具有易學易記的特點。所以無論學習、或者使用成語，都不很困難。但是如果想要很好地掌握它，也有應該注意的幾點，在維基百科的網頁上 (<http://zh.wikipedia.org/w/index.php>)，關於成語的使用性說明整理如下：

### 一、 了解成語的涵義

徹底理解學習和使用成語，跟學習和使用字、詞一樣，首先要徹底理解它的涵義，不能「不求甚解」。如果對某一個成語的意思還沒有理解透徹，或者對成語中某個字的音義還說不準確，就隨便拿來使用，那就難免用得不恰當，甚至鬧出笑話。成語的來源不一，有的容易懂，有的比較難懂，可是其中以容易懂的佔多數。有些成語，懂得了字面的意思，還要特別注意它的習慣上的用法。

## 二、 視場合使用成語

在不同的場合恰當使用成語，應當力求當下場合的恰當，不應當濫用。此外，還有一種情況，使用成語時也應該特別注意，即它的歷來的用法。對於在某一場合看起來好像都可以使用的幾個成語，更要仔細地辨別它們的不同之處，體會文章所需要表達的意思是什麼，選用哪個成語才最為恰當。

## 三、 分清成語引申的義涵

分清褒義貶義成語裡面，除了表示本身的意義之外，還表示人們的喜愛或者憎惡的感情。表示喜愛的，通常說它含有褒義；表示憎惡的，通常說它含有貶義。這兩類成語千萬不可誤用。

## 四、 注意成語的規範化

成語是固定的詞組或句子，一般不能隨便換字或改變它的結構。不過有少數成語，也有從早就有兩三種不同的說法的。如「一髮千鈞」也作「千鈞一髮」，「含沙射影」也作「含沙射人」，「鸛蚌相持，漁人得利」，「持」也作「爭」，「人」也作「翁」，等等。這是多年來形成的差異，我們只能根據「約定俗成」的原則，承認幾種說法都可以。我們自己使用成語，除非為了取得修辭的效果，對某一成語確有改變的必要時，都應該注意成語的規範化，也就是盡量地照原來的文字使用，保持語文的純潔和健康。因為這樣改變起來，就容易造成語文中的混亂現象，甚至使人誤解。

## 五、 靈活運用成語

成語的靈活運用一般說來，成語是「固定詞組」，有一定的固定性。可是，我們不能就認為它是一成不變的。在實際運用方面，是有它極大的靈活性的。何永清（2005）指出，要將成語靈活運用得宜，必須要了解成語的修辭技巧，說明整理如下：

### (一) 摹況、譬喻、轉化：

摹況修辭的成語：摹況也稱為摹寫，對人事物作描述時是運用感官的直覺。譬如呱呱墜地：「呱呱」是形容嬰兒的哭聲，這是利用聽覺來修辭的成語；萬紫千紅：形容群花盛開的美麗景色，這是運用視覺來修辭的成語；臭氣沖天：形容奇臭無比，這是利用味覺來修辭的成語。譬喻修辭的成語：譬喻的修辭是一種打比方的修辭方法，可分為「明喻」、「略喻」及「借喻」。明喻：深藏若虛、安之若素、趨之若鶩、受寵若驚、浮生若夢、虛懷若谷等，以上使用喻詞「若」。略喻：省略喻辭，如：人心鼎沸。借喻：井底之蛙，用來比喻見識狹窄的人。轉化修辭的成語：轉化分為擬人和擬物，擬人的如：草木皆兵；擬物的如：笑容可掬。

### (二) 映襯、借代、對偶、排比：

映襯修辭的成語：映襯是運用對比的修辭法，如：紅男綠女、貪生怕死、天南地北、南船北馬、唇槍舌劍等。借代修辭的成語：用別名來代替本名的修辭，如：目不識丁，是用「目」來借代「文字」。對偶修辭的成語：詞性及結構相互對應的修辭法，如：平心靜氣、山窮水盡、得心應手、四平八穩、四捨五入、千頭萬緒等。排比修辭的成語：將三個以上的語句並列的鋪敘修辭法，如：直諒多聞。

### (三) 類疊、頂真、回文、鑲嵌、倒裝：

類疊修辭的成語：同一字詞重疊或間隔出現，如：形形色色、十全十美。頂真修辭的成語：首尾蟬聯效果的修辭法，如：自欺欺人。回文修辭的成語：來者不善，善者不來。「來」跟「善」構成回文的修辭。鑲嵌修辭的成語：如：乃武乃文、七上八下。倒裝修辭的成語：之無不識，是「不識之無」倒裝。

### (四) 誇飾、婉曲、析形、節縮、互文、轉品、藏詞：

誇飾修辭的成語：如：萬家燈火、三人成虎。婉曲的成語修辭：天下無雙，意即「天下第一」的意思。析形的成語：離合字形的修辭方法，如：止戈為武。節縮修辭的成語：班輪之巧，「班」與「輪」皆是人名節縮。互文修辭的成語：前後句語義相互補足，如：猛將如雲，謀臣如雨。轉品修辭的成語：轉換詞性的修辭方法，如：十風五雨，「風」與「雨」在這邊是名詞轉換為動詞。藏詞修辭的成語：將某個文字藏省的方法，如：司空見慣，將「渾閑事」藏省了。

以上，從何永清（2005）的成語修辭技巧分類中可以了解到，成語的修辭方法便有十九種之多，若能將此十九種的修辭方法通徹了解，相信在使用及了解成語方面，一定有所助益，說不定還能達到事半功倍之效用呢！因此，要想能夠把成語用得靈活，記得清楚而不容易混淆，就必須把成語的實質精神，真正地吸收、消化，才能融會貫通，左右逢源。

## 第二節 目前語文教學的相關研究

語文在文化傳承中，扮演著不可或缺的角色。在相關的理論探討中，整理了 Vygotsky、Cole、Saljo 等三位學者的重要理論，再來是分析目前的語文教學現況，最後再就與本研究最為相關的作文與成語的部份，作更深入的整理說明。

### 壹、語文的重要性

語文是民族文化的累積，有文化，才有語文，而文化更是藉由語文來流傳（王建華，2000）。中華民族五千年的文化精華所衍生出的語文結晶，是重要也是必備的工具。

教育部（2003）所公布的「國語課程綱要」中指出國語學習領域所要培養學生的五項基本能力：

- 一、 具備基本的聽、說、讀、寫、作能力
- 二、 使用語文，解決問題的能力
- 三、 應用語文，拓展多元視野
- 四、 欣賞文學，體認中華文化精髓
- 五、 培養學生自學的能力

由這五項能力可以發現，在國語領域中，不再是強調記憶背誦等能力，而是學習如何應用語文，並進而達成自我的提昇。

Gutlip, G. W. 曾指出：一切學習的基礎在於「語文」，學習語文的目的不僅只是利用語文來表達情意及看法，而是用它來作為其它學科學習的重要工具（朱孝利，1999）。語文能力好，學習能力自然加分；語文能力差，光是花在解讀文字的理解上，相對的定要花費比一般人更多時間。

## 貳、 相關的理論與研究

幸曼玲（2003）指出進行語文教學如果忽略舊有經驗且又被動學習的話，語文的學習將是生硬無趣的。以下就幾位學者的觀點說明如下：

### 一、 Vygotsky（1978）

Vygotsky 是以社會建構的觀點來看待學童的學習並認為人的心智是存在社會多層次的情境中。傳遞歷史文化內涵的主要工具是語言符號，但 Vygotsky 認為語言並不僅是傳遞工具，同時也是人類心理功能的建構工具，人類在學習語言的歷程中，亦建構了高層次的心理功能。

### 二、 Cole（1998）

Cole 認為第一、孩子是透過文化來學習，並非模仿，再經由人際間的互動而進一步地內化至個人內，最後成為個人的心理工具。第二、透過多次的使用心理工具，使個體具備了高層次的思考能力，也造就了歷史文化的傳承。由此可見，語言並非單純的工具而已。第三、進行語言學習時，必須在實際的情境中進行，如此才能讓此學習富有意義。

### 三、 Saljo（1998）

Saljo 提出語言是人類最重要的心理工具。因此，學習語文不應只是單向傳輸，而是應在有意義的情境中互動，藉由不斷的澄清與思考，加強本身的內涵並能將其運用抒發表達於外。

綜上所述，語文的學習不僅是為了溝通，更是強化內在思考，進而傳承文化的重要工具。教學者應讓孩子在有意義的情境中互動學習，讓孩子懂得使用，並更進一步地加以思考內化才是語文學習的主要目標。

## 參、 語文教學現況

語文是一切學科之母，也是傳遞文化最根本的一股力量，近來國小國語文教學的缺失分述如下（張巧琳，2004）：

### 一、 教學未能把握重點

很多的語文教學仍停留在生字新詞的解說，對於文章的剖析及寫作技巧，仍缺少適度的引導和解說。

### 二、 缺少指導學生說話的技巧

教學活動仍是以老師單向灌輸居多，雖已漸漸注意學生說話表達的機會，卻也時常忽略對於說話技巧的指導。

### 三、 作文教學，長於批評，短於指導

作文批改，教師應多提出具有指導性、輔助性和鼓勵性的評語，而避免流於一味的批評，而使學生喪失寫作文的信心。

### 四、 加強「寫字」教學

寫字教學著重於讓學生了解字形字義，並能寫出一手端正的字體，但不少教師認為寫字課就是書法課，或許有本末倒置的疑慮。

### 五、 未能善用社會資源

如何將提昇學生語文能力由原本「學校的事」轉變為「大家的事」，促使語文教育更加蓬勃發展，是未來相當重要的課題。

綜上所述，可知目前語文教學仍有些困境仍須突破。而語文素養的高低代表一個國家的競爭力，因此如何加強語文教育是刻不容緩的事。提升語文教育的方法，許多學者也提出許多相關的策略，茲就與本論文所要探究最為相關的「作文教學」與「成語」的部分做進一步的說明。

## 肆、 作文教學與成語

卓惠美（2004）指出現在學生語文程度的低落：錯別字多、亂用成語、語彙不多、文章無重點等。因此如何引發學生對語文的興趣，進入有趣的語文世界，在小學如此奠定基礎的階段上是非常重要的。

作文是語文教學的綜合活動，由作文的表現可以考驗其語文能力的高下（張巧琳，2004）。在學習語文的過程中，能將平日所讀訴諸文字再加以修飾表達出來，是語文能力的最高表現之一。要寫好一篇好作文，平日需藉著大量閱讀來儲存語彙，多做語法練習，再加以教師適切的引導，才能提升學生的作文能力（葉興華，1999）。

張月娟（2007）提出在作文能力的提升中，「成語教學」具有極大的功效。成語的教學不宜貪求快速，也不需選擇艱深冷門的成語，否則將會抹殺學生學習成語的信心。

對於成語教學與寫作，張月娟（2007）提出策略如下：

### 一、 介紹常用成語

進行成語教學時，教師宜介紹生活中常用之成語，除典故、釋義的介紹外，更重要的是，也要做成語的造句練習及相近成語、相反成語的統整。

### 二、 一日以兩則為宜

教師在課堂上花五分鐘講解或造句，讓學生回家做練習，每天累積二則，長久日子下來，累積的成語數量也是很可觀的！

### 三、 進行成語評量

進行成語評量，分為配對遊戲與造句練習，主要是讓學生能夠理解成語意義，並具備正確使用成語的能力。

對於學生寫作能力的培養，不必強求能立即寫出長篇文章，而是要循序漸進地累積能力，從基礎奠定起。經由以上的成語教學策略，學生的句子會寫得更加有內涵與豐富，逐一儲存於實力中，對於將來書寫長篇作文，定有極大的幫助與意想不到的效果！

由本節所述，可以得知，語文是一切學科的基礎，但它不僅只是溝通的工具，也是思考轉化為內在的心理工具。而現今語文能力的低落，慢慢也會影響一個國家的競爭力。因此，如何提升語文能力為當務之急。其中作文是語文教學的綜合表現，藉由作文就能看出一個人知識的廣度與深度，以及判斷是非與遣詞造句的能力（蔡榮昌，2006）。藉由成語教學中成語知識的累積，更是作文能力提昇的最大效力，作文能力增強，語文的表達更有強力的後盾支持，語文能力自然跟著提升。

### 第三節 國小成語教學的相關研究

本節分為三部份來說明，第一部份為目前國小成語教學的現況整理，第二部份為成語教學的原則與方法，第三部份為目前成語教學的相關研究。茲分別整理如下：

#### 壹、 國小成語教學現況

了解國小成語教學現況，從國小教科書著手最為直接，依據研究者（2007）所整理的「國小教科書成語教材之內容分析－以翰林版國語課本為例」的文章中，歸納結論如下：

一、 年級數越低，所教的成語較少；年級數越高，所教的成語則越多。

低年級小朋友，由於剛接觸國語，且需從注音符號開始學起，所以課本上教給低年級小朋友的成語有 9 個。中年級隨著小朋友能力和所學字彙的增加，所學到的成語有 61 個。高年級小朋友的智力及記憶力更為成熟，且需提供作文及說話的需要，從課本可以學到的成語增加為 159 個。由此可見，伴隨能力的增加，課本也提供了更多成語的學習。

二、 所教的成語隨年級數增加，分配大致均勻，少數有幾課較為集中。

由結論一可知，成語所學隨年級數增加。不過從研究者整理中可以發現，成語雖然大致上逐冊增加，分配上大致均勻，不過仍有少數幾課較為集中，學生學到那課時，會覺得較備感吃力。而這也有待老師，多詳加解釋及複習，來幫助學生達到事半功倍的效果。

三、 年級數越低，所教的成語越淺顯；年級數越高，所教的成語則越加深。

從各年級整理出來的成語看來，越低年級所教的成語較淺顯易懂；越高年級則成語用字較艱難，涵義也較深。因此，課本所教的成語，有因能力的增加而給予加深加廣。

四、 課本教材中所出現的成語，大部份都很常用也很口語化。

課本所教的成語，幾乎都很常用也很生活化，不僅淺顯也很接近日常生活的實用性。這樣的成語，頗能符合學生的需要，又能實際運用於生活環境中。

五、 由課文中學成語，可以學到與課文主題相關的成語。

每篇課文，皆有主題與其所要闡述的相關道理，學生藉由學習課文，可以從中學到與課文中主題相互佐證的成語，日後說話或寫文章時，便可互相參照應用，增加學生成語應用的能力。

六、 課文教材成語，重複性不高。

課文教材裡出現的成語，雖有重複，但重複性不高。這樣學生便可學到更多更新的成語，而不會在原地踏步，浪費太多的精力和時間。

## 七、 所學成語略嫌少，國中銜接？

依整理，小學六年，只學了 229 個成語，對於現今升學主義的需求，似乎明顯不足。等到學生在國中為基測做準備時，恐負擔更重。這一點，實有必要再做加強。

## 八、 小學成語較受忽略，無編排進程目標及妥善規劃。

小學課本教材，以學習語文為主，雖說成語也是語文的一部份，但仍是備受忽略的一環。既無進程目標的規劃，也無教材的編纂，所學到的成語，均只是伴隨所選課文而來，沒有真正奠定成語能力的培養。

## 九、 教科書版本不一，成語銜接問題？

目前教科書版本眾多，學生所學成語不一，程度也不一，升上國中後，勢必有落差，有些成語早已學過，有些成語或許根本就沒看過，對於學生的學習會有一定的影響，在未來成語教學的規劃中，這也是不容忽視的一點。

綜上所述，在每冊的國語教科書中均內含成語的教學內容，嚴格說來，仍缺乏有系統的規劃，內容也不足。王月鳳（2004）也認為歷年來的國小課程規劃都未對成語部分加以規定。但成語是語文的精髓，是值得我們研究和教導下一代的寶藏，要想提昇學生的口語表達及作文書寫能力，可以在最貼近學生的國語教科書上再多費心設計，定能對學生的語文能力助益良多，也能繼續將中華文化的精粹流傳長遠。

## 貳、 成語教學的原則與方法

何永清（2005）認為成語教學的原則，需具有：

一、生動化：成語教學成功的保證就是多樣的教學法與趣味的教材。教師不能僅僅只是將成語解釋出來或淪為抄寫等制式的教學方式，而應推陳出新，多多運用語文的小遊戲來引起學生學習成語的動機，當小朋友對成語產生興趣，他們的學習就會日趨主動。

再來二、統整化：結合說話、讀書、作文等等的成語教學，能將成語的運用變得更加熟練，甚至可以統整到數學、社會及藝術與人文等學科的相關領域上。

其三、實用化：成語的教學，必須使學生在作文上能運用得宜，否則成語的教學，只完成了一半。

最後，四、多樣化：運用豐富的教學資源，除了成語辭典外，電腦媒體或漫畫等方式也都是教師可以用來教學的教學媒介。

另外根據以上原則，何永清（2005）也提出了幾個具體可行的成語教學方法，分享如下：

一、成語尋寶：教師可以讓學生運用成語辭典或上網分別找出成語中的動物名、植物名、花草名、人名、數字等，讓學生在搜尋的過程中，對成語印象深刻。二、成語接龍：運用成語作頂真式的連環串接，也可運用於小組成語競賽之中。三、成語填字：教師可設計將成語中的一些字空下，讓學生來練習填寫。四、成語配對：運用成語來對仗。五、說成語故事：讓學生可以對成語產生的背景有初步的了解，將來對於此成語也較容易認識。六、成語造句：運用成語來造句，可以提昇學生運用成語的能力。七、用數字猜成語：與數學結合的一種成語教學。八、成語形象畫圖：結合藝術與人文，將成語意象進行推廣。

以上，這麼豐富的成語教學活動，實是有趣，有別於傳統教學之不同，若教師想在成語這番領域有所教學突破的話，相信以上的這些活動，定能帶給小朋友另一番不同的體驗，也或許這些活動也能作為拋磚引玉之效，引領更多成語創新的教師，激發出更多不同的想法，讓成語教學的天地，更添繽紛美麗的色彩，蘊育更多成語智慧的結晶。

## 參、成語教學研究

了解小學課本及成語教學的原則方法後，再進一步討論學界對於成語教學的一些建議：

在語文領域中，成語的運用是其中非常重要的一環，能善加利用的話，對於作文、說話等方面將有畫龍點睛之效果。在劉淑貞（2003）利用成語教學方案對自閉症學生閱讀及溝通能力增進之研究結果指出，成語教學方案對於自閉症學生的閱讀及溝通能力是具有增進效果的。想當然爾，對於正常智力、學習能力正常的學生而言，成語教學若施行得宜，相信對於學生語文能力的提昇，一定也會有相當的助益。學生的成語能力，若單只知成語故事而不知如何應用的話，是無法

達到效果的（蕭嫻慈，2002）。那如何進行有效的成語教學呢？江惜美（1994）提出：依據不同的類型的成語，有需注意的教學技巧。

## 一、以歷史故事產出的成語

要請小朋友清楚記下：

- （一） 故事在哪個時代發生的
- （二） 故事中主要的人物有哪些
- （三） 故事發生的前因後果
- （四） 故事中的原義及後來引申義
- （五） 如何應用於生活中
- （六） 以事物原理產出的成語

這類成語往往特別彌足珍貴，因其為先民觀察事理的體驗。在教導這類的成語時，可以從事物最根本道理說起，再漸漸衍生與現實的人事結合。

## 二、以寓言、神話、詩詞用語呈現的成語

這類的成語富有教育意味，但由於某些用詞較艱深，因此學生較不易學習。如果教師能找到以漫畫方式呈現的補充教材，是最為理想的。

除了依據成語的類型，有不同的教學技巧外，張君松（2005）還提出若想讓學生能更加有效的學習成語，可以讓學生先對中國的文化與歷史有初步的認識，再將成語的字義、語義及隱喻的意義加以說明，最後說明在何種情境或造句時可以運用到，其中的說明要多多的舉例並給予學生作業的練習。教學要由淺入深，學生才能經由作業練習而達到精熟。

徐明珠（2008）也談到，有了成語基本的蹲馬步外，這些可以作為寫作的材料，仍需藉由構思和選擇，才能將架構和素材鋪陳經營起來。撰寫文章首先要確立核心價值，也就是寫作的立意，是想要警世，或是傳達個人感受，抑或撩撥讀者愁緒，喚起共鳴，又或是教忠教孝，激發愛國意識等，如此字裡行間才能充滿作者想要表達的觀點。

其次，構思內容。這是考驗作者肚子有多少墨水的階段，能夠從自己的知識庫中提取何種知識，其知識性的程度不只透露一個人的文采，更能彰顯智慧的深度和廣度。

再者，組織材料。此項階段是邏輯能力的展現，作者可以將知識庫的材料做巧妙的佈局和運用，有將事物化繁為簡的歸納能力，也有將簡化之原則推演成為放諸四海皆準之道理。

進一步，確立選材之內容。一般而言，主題類型舉凡修身勵志、政治經濟、社會文化、軍事科技、文學藝術、怡情養性、時事議題等皆可入題。

最後，則進入實質創作階段。將以上之基本功融入於文章中，並藉由緊實的布局及修辭的運用，把材料鋪陳為草稿，再逐步檢視、評析、修正，創作即告完成。

綜上所述，教師在進行成語教學時，能先針對成語教材做教學前的分析，再佐以配合孫世民（1995）的研究結論指出應善用生動活潑的教學法來激發學生學習動機，使學生了解成語的形、音、義，活用於作文及說話中，再輔以情境式引導，樹立成語正確價值概念，若能如此兼顧到認知、技能、情意的成語教學才是理想的。成語的文化千錘百鍊，積累不易，因此自當珍惜源於歷史故事、古書章句之成語典故，並加以活用之，以使成語能成為現代文學中歷久彌新的主流文化。

#### 第四節 電腦測驗的相關研究

現今各式考試場合中的主要施測方式仍是以傳統的紙筆測驗為主，但這種傳統測驗方式不僅費時費力也容易產生相當大的測量誤差和無效率性（余民寧，2002）。幸好在目前資訊科技日益發達下，測驗的計分、回饋及迷思概念的診斷藉由電腦化測驗的發展愈來愈有效率，使得電腦化測驗成為最佳的測驗選擇（陳新豐，2005）。因此依目前科技迅速發展的情況下，電腦發展和測驗運用的相結合是不可抗拒及無法避免的趨勢（何榮桂，1997）。隨著網際網路的興起，測驗不僅是與電腦結合而已，更與網路科技結合，除了電腦化測驗已經成為事實之外，邁向網路線上測驗的時代更是未來測驗發展的趨勢（余民寧，2004）。Bunderson、Inouvy 和 Olsen 於 1989 年指出電腦測驗有四代的發展類型，其中現今最備受重視與討論的為電腦化適性測驗。以下分述電腦適性測驗的意義、類型與相關理論。

## 壹、電腦測驗的意義

電腦測驗根據其背後所持測驗理論基礎的不同，分成適性與非適性兩類。非適性的電腦化測驗（Computer-Based Testing，簡稱為 CBT），主要還是以古典測驗理論為依據，是一種將傳統紙筆測驗改成以電腦螢幕或網路當呈現介面，逐一或全部呈現電腦輔助施測的方式，學生在作答時，只須使用鍵盤或滑鼠輸入即可，作答完畢即可立即知道自己的得分和作答情況（余民寧，2004）。

結合試題反應理論的則是電腦化適性測驗（Computerized Adaptive Testing，簡稱為 CAT），是指考生藉由操作電腦進行測驗，而測驗題目的選擇則是依據考生的能力來適性挑選（李茂能，2001）。在張俊欽（2005）的研究結果也指出，電腦化適性測驗可以快速正確地分析學生在試題上的作答情形，並達成受試者能力值準確的估計。因此這樣的測驗方式不僅符合經濟效益，且測驗結果精確、誤差最小。

目前適性測驗的發展，均使用不同的分支策略，作為彈性呈現不同難度試題給不同能力程度的考生施測。這些不同的分支策略雖然在考生作答第一道試題後，採用不同的分支策略呈現下一道試題給考生作答，但都有個共同點：那就是先從題庫中挑選一道難易適中的試題作為起始題，讓每位考生都作答，在作答完畢之後，下一道應該呈現的試題是依據考生在上一道試題的作答結果來決定。這個選題的過程會繼續下去，一直到事先預期施測的題數測完、達到可以容忍的最小測量誤差、或題庫中已無適當的試題可以再呈現為止（Warm，1978）

## 貳、電腦測驗發展的類型

依陳新豐（2004）電腦測驗四種類型整理如下：

第一代是電腦化測驗，基本上，此類測驗是將傳統的紙筆測驗題目改用電腦來呈現，其試題內容同於紙筆測驗。

第二代是電腦化適性測驗，此類測驗結合試題反應理論及適性測驗的優點，可針對不同能力和不同程度的考生及其不同的作答速度，在電腦上提供適合其能力作答的適當難度的試題，以最少的題目達到精確估計考生能力的效能。

第三代是連續性測驗，類似電腦化適性測驗，同樣根據受試者前一題的表現來決定往後的測驗流程，也就是將已量尺化題目參數的試題安排在課程中，以持續而不干擾的方式，估計學生成就表現或學習剖面的動態改變歷程。基本上，連續性測驗是將測驗融入於教學中，可使教學與測驗不再區別，也就是在測驗的過程中，連續性測驗乃根據前一題的答題結果，來決定下一個教學內容。

第四代則是智慧性評量，是根據知識基礎及推論程序，結合人工智慧，利用電腦提供智慧性的記分、受試者學習剖面的解釋、對結構複雜的答案精密計分，自動化解釋個人能力的側面圖，並提供智慧型的建議，例如：針對個別差異情形，了解學生是認字的困難導致閱讀文章有所困難，則應加強認字的基礎技能才能有助閱讀理解。

本研究所採取的為第三代連續性測驗，內容是「適性測驗」的題目。陳新豐（2002）指出適性測驗是根據受試者先前作答的經驗，再挑選適合受試者能力的題目繼續施測，作答後立即評分，以利選擇下一題施測的題目，以此類推，直到完成預定的終止標準為止。

適性測驗可分為二段式、金字塔式、彈性水準測驗及分層式適性測驗（王寶墉，1995）。本研究所採取的為金字塔式適性測驗，意義如下（陳新豐，2002）：

金字塔式測驗又稱分枝、多階段測驗，此種施測方式採固定分支方式，選題方式以單一試題為選擇單位，如圖 2-1 中，位於金字塔尖端的題目（第 1 題）是難度最中間者，受試者由該題開始作答，如果答對，接下來就答第 3 題（比第 1 題難），反之，如果第 1 題答錯，那接下來就答第 2 題（比第 1 題簡單），答完第二階段的題目之後，繼續用相同的分枝策略答其餘各階段的題目。舉例來說，一個含括 55 題的測驗而言，每位受試者只需要答 10 題（吳裕益，1997）。金字塔測驗是適性測驗中最常用且最受重視的一種彈性的測驗方式，因分支功能的不同又可分為固定分支與變化分支，因此很受測驗學者的青睞。綜言之，在多種適性測驗中，金字塔式測驗具有如下優點，包括：充分運用貝氏策略、便於新增鑑別力參數及猜測參數和適用於團體測驗（王寶墉，1995）。

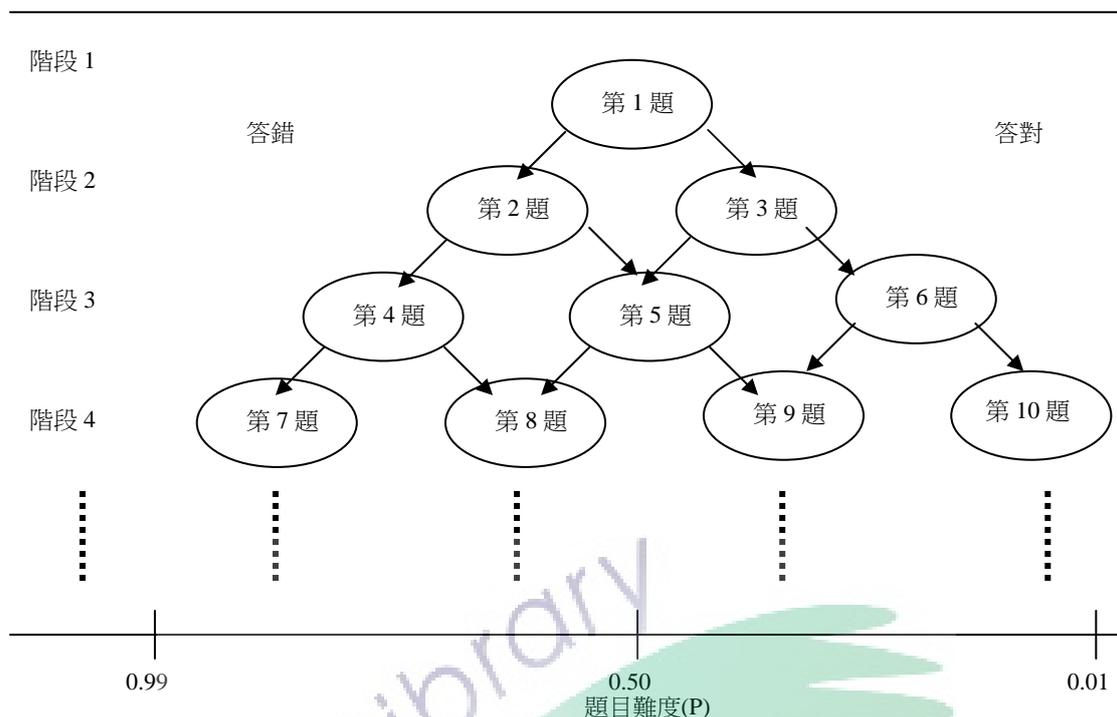


圖 2-1：金字塔式適性測驗示意圖

## 參、電腦測驗相關理論

在電腦測驗最為相關的重要理論，便是「古典測驗理論」及「試題反應理論」，茲說明整理如下：

### 一、古典測驗理論

#### (一) 理論說明

古典測驗理論 (Gulliksen, 1987) 是最早的測驗理論也是目前測驗學界使用與流通最廣的理論依據。其主要是以真實分數模式為理論架構，依據弱勢假設 (weak assumption) 而來，其理論模式的發展已為時甚久，且具有相當的規模，所採用的計算公式簡單明瞭、淺顯易懂，適用於大多數的教育與心理測驗資料，以及社會科學資料的分析 (余民寧, 2004)。

當某位受試者接受一份測驗的施測後，他在該測驗上的實得分數，即代表他在某些特定的情境下在這些試題樣本上的能力。施測時當然會有許多因素會影響受試者在測驗上的表現。因此即使在內容範圍相同但試題樣本不同的條件下，或在不同的時間、主測者、與施測情境的條件下，受試者的表現可能都會不一樣。所以，如果我們在所有可能的施測情境下、在所有可能的不同時間範圍內、或盡可能使用不同試題，來針對同一位受試者進行同樣的測驗多次，則我們就可以獲得許多有關該受試者的實得分數。這些實得分數的平均數，即代表該受試者能力的不偏估計值，該估計值即被定義為「真實分數」。

因此，所謂的真实分數模式，即是指一種直線關係的數學模式，用來表示任何可以觀察到、測量到的實得分數皆由下列兩個部分所構成的一種數學函數關係。這兩個部份分別是：一為觀察不到，但代表研究者真正想要去測量的潛在特質部份，叫作「真實分數」；另一為觀察不到，且不代表潛在特質，卻是研究者想要極力去避免或設法降低的部份，叫作「誤差分數」。這兩個部份合併構成任何一個真實的測量值，且彼此之間具有並延伸出多種基本假設，能符合這些基本假設的測量問題，即為真實分數模式所探討的範疇（余民寧，2004；Lord & Novick，1968）。

根據古典測驗理論的假設，每位受試者都具有某種潛在特質，該潛在特質無法單由一次測驗的實得分數來表示，必須由受試者在無數次測驗上所得的實得分數，以其平均數來表示，該數值即為受試者所得的真实分數。真實分數並不受測量次數的影響，它是代表長期測量結果永恆不變的部份。然而實際上，單獨一次測量所得的實得分數，總會與真實分數間產生一段的差距，這段差距即為「誤差」。若以數學公式來表示為  $X = T + E$ ，即觀察分數等於真實分數與誤差分數之和。古典測驗理論即是建立在上述這種真實分數模式及其假設的基礎上，針對測驗資料間的實證關係，進行有系統解釋的一門學問（余民寧，1992）。

陳新豐（2002）指出雖然古典測驗理論在心理與教育測驗上的應用已具有相當的成效，但由於考生能力的不同及能力估計值會受試題的難度及鑑別度的影響，因此在實施古典測驗理論時會有五個限制：

1. 試題的難度及鑑別度，受限於樣本的依賴
2. 必須依賴平行測驗來比較不同受試者的能力
3. 要建構平行測驗實有困難
4. 試題訊息缺乏
5. 測驗誤差同質的假設並不適切

由此可見，古典測驗理論的缺失很多，也無法有效地處理測驗的設計及能力評估的問題，而為了克服古典測驗理論的眾多問題，遂有「試題反應理論」的產生。

## (二) 試題難度的計算

一個題庫的良窳，取決於其試題的品質，而一個試題的好壞會有其量化的分析，也就是試題難度與鑑別度的計算（高啟洲，2005）。余民寧（2004）提到難度指標的分析主要以答對百分比法最為簡單實用，而本研究所使用的 Tester 2.0 軟體也是採用此方法來計算難度指標。

在全體學生當中，答對每道試題的人數佔總人數的百分比值，稱為答對百分比法。而所得的百分比值就是這道試題的「難度指標」，其計算公式如下（余民寧，2004）：

$$P_i = \frac{R_i}{N} \times 100\%$$

其中， $P_i$  為第  $i$  個試題的難度指標， $R_i$  為答對第  $i$  個試題的人數， $N$  為總人數。 $P$  值的指標是介於 0 到 1 之間，一般而言， $P$  值越大，即表示該試題越容易，越多數的學生對於該試題都是答對的；反之， $P$  值越小，即表示該試題越困難，越少數的學生答對該試題。

## (三) 試題鑑別度的計算

鑑別度指標的高低，能表示該試題對於區別學生能力高低的準確力。試題鑑別度的分析方法，也有兩種，主要還是以本研究所使用的 Tester 2.0 軟體進行的內部一致性分析方法來作介紹，其公式如下（余民寧，2004）：

$$D_i = P_{iH} - P_{iL}$$

$$i = 1, 2, \dots, n$$

其中， $D_i$  是表示第  $i$  個試題的鑑別度指標， $P_{iH}$  和  $P_{iL}$  則分別表示高、低分組的人在第  $i$  個試題上答對人數百分比值。一般來說，試題的  $D$  值 = 0，代表此試題無鑑別力，而試題的  $D$  值  $\geq 0.20$ ，則表示該試題具有較佳的鑑別力（林妙香，2000）。

## 二、 試題反應理論

試題反應理論（IRT）在測驗的情境之下，可以藉由受試者特質的了解而達到預測或解釋受試者的能力。試題反應理論認為，受試者與特質之間存有關係，受試者經由測驗試題表現出其特質或能力，而潛在特質或能力亦透過測驗試題而表達出來。因為此種關係是一種數學函數的關係，所以試題反應模式也是一種數學模式（陳新豐，2002；Hambleton & Swaminathan，1985；Van Der Linden & Glas，2000）。

陳新豐（2002）說明試題反應理論就是把受試能力與試題難度、鑑別度及猜測等參數間的關係以數學符號來表示。試題反應理論的假定是屬於強假定，古典測驗理論則是弱假定，兩者是截然不同的。雖然在試題反應理論中的測驗資料很難完全符合單向度的假定，但試題反應理論允許某種程度的違反假定，因此具有很大的優勢，以下就試題反應理論所符合的理論整理如下：

### （一） 單向度

試題反應理論符合單向度的標準，根據此理論而產出的試題能集中在鑑定某一種特定的能力。

### （二） 局部獨立

試題反應理論同時也假定受試者對某一個試題作答的好與壞，並不會受其它試題的影響。

### (三) 試題特徵曲線

試題的特徵曲線即是該題答對機率對受試者能力的迴歸線。而此條曲線只跟受試者的能力高低有關，對於受試者是否成常態分布並不構成影響，此種不變性也是試題反應理論的一大特色。

### (四) 答題速度

在試題反應理論下，受試者在答題時並沒有時間的限制，所以受試者沒有作答的題目是代表他不會而不是沒有足夠的時間作答。

試題反應理論近來普遍受到重視與運用，不僅適用於精熟式適性測驗，也適用於適性成就測驗或人格測驗。具有三大特色如下（李茂能，2001）：

- (一) 依據不同的能力使用不同的估計標準誤
- (二) 難度估計值與能力使用共同量尺
- (三) 參數估計值具有不變性

另外，由研習資訊的期刊中，余民寧教授整理出對試題反應理論的發展有實際貢獻的代表性作者和著作如表 2-1 所示：

表 2-1：對試題反應理論發展有實際貢獻的代表性作者和著作

作者（年代）	代表作及其貢獻
Tucker(1946)	第一位提出試題特徵曲線概念的人。
Lord(1952)	第一位導出兩個參數常態肩形模式的參數估計公式，並考慮試題反應理論應用性的人。
Rasch(1960)	試題反應理論中 Rasch 模式的創始者，影響深遠。
Lord & Novick(1968)	第一本介紹古典與當代測驗理論模式的經典作品，引發學者對「潛在特質」概念的重視與研究。
Wright & Panchapakesan (1969)	美國地區第一篇介紹 Rasch 模式的參數估計法，並發展有名的 BICAL 電腦程式的代表作品。
Samejima(1969)	她的一系列作品描述新的試題反應模式及其應用，其中包含處理多分法與連續性資料的模式，甚至擴展到多向度的試題反應模式，為一艱澀難懂的重要著作。
Bock(1972)	提供許多估計模式參數的新概念。
Andersen(1973)	歐洲地區談論測驗模式的重要著作。
1976	Lord 等人創作第一版有名的電腦程式：LOGIST。
1977	Journal of Educational Measurement 第四季出版一冊專門探討試題反應理論的專輯。
Baker(1977)	第一篇評論試題反應模式參數估計法的文獻探討。
Wright & Stone(1979)	第一本描述各種 Rasch 模式理論及其應用的專書。
Lord(1980)	第一本以試題反應理論命名的專書，是當代測驗理論發展的里程碑。
Weiss(1980)	第一本編輯成的論文輯，專談試題反應理論的實際應用課題——電腦化適性測驗。

（續下頁）

(續表 2-1)

Andersen(1980)	對測量模式參數估計法有貢獻的方法學專論。
Bock & Aitkin(1981)	提出邊緣的最大近似值估計法——EM 估計程序，對參數估計法的改進貢獻不少。
Masters(1982)	第一位發表部份知識計分模式，對改進 Likert 式評定量表的計分與次序反應資料的計分貢獻不小。
Wright & Masters(1982)	闡述 Rasch 模式的各種模式成員，證明皆與部份計分模式相通，對 Likert 式評定量表與次序反應資料的計分方式改進不少。
Mislevy & Bock(1982)	發表另一有名的電腦程式：BILOG。
1982	Applied Psychological Measurement 第四季出版一冊專門探討試題反應理論及其應用的進階專輯。
Wainer & Messick(1983)	編輯而成的論文集，以表揚 Lord 一生對試題反應理論的貢獻，並兼論該理論的應用與未來。
Weiss(1983)	編輯而成的論文集，專談試題反應理論的應用與未來，並介紹它在電腦化適性測驗上的應用。
Hambleton(1983)	編輯而成的論文集，專談試題反應理論的模式與應用。
Hulin, Drasgow, & Parsons(1983)	為一本試題反應理論的教科書，增加對「適合度測量」概念的說明與應用。
Embretson(1985)	編輯而成的論文集，專談試題反應理論的未來發展。
Baker(1985)	為一本導論性的試題反應理論教科書，專為沒有數學訓練基礎的讀者而作，並附有 CAI 的電腦教學磁片。
Hambleton & Swaminathan(1985)	為一本進階的試題反應理論教科書。

(續下頁)

(續表 2-1)

Crocker & Algina(1986)	談論與比較古典與當代測驗理論的導論性教科書。
Wainer & Braun(1988)	專談有關效度方面的論文集，也談試題反應理論在效度上的應用。
Linn(1989)	負責主編第三版的「教育測量」(Educational Measurement)，其中增加一章專門介紹並評論試題反應理論。
Freedle(1990)	專談人工智慧及其在當代測驗理論上應用之論文集。
Suen(1990)	介紹各種測驗理論方面的教科書。
Wainer 等人(1990)	專談電腦化適性測驗方面的入門書，也談試題反應理論在電腦化適性測驗上的應用。
Hambleton, Swaminathan, & Rogers(1991)	試題反應理論方面的入門書，適用於非數學主修的初學者閱讀。

由此表的發展趨勢來看，試題反應理論將取代古典測驗理論是指日可待的事情，Sands, Waters & McBride 於 1997 年還曾提出電腦化適性測驗是主要未來發展的方向。首先是新題型的發展；其次是動態的題目參數估計；再來是線上題庫的等化；最後，則為多向度的適性測驗。因此電腦與試題反應理論所結合的電腦化適性測驗無疑是未來測驗發展的趨勢。

## 第三章 研究方法

本章共分六節，依次說明「研究架構」、「研究對象」、「研究設計」、「研究工具」、「資料分析」等六方面。

### 第一節 研究架構

本研究首先針對「成語的探究」、「目前語文教學的相關研究」、「國小成語教學的相關研究」及「電腦測驗的相關研究」等作文獻方面的探討，來作為開發測驗的重要基礎。同時利用 Tester 2.0 的分析來做成語題目預試的難度與鑑別度的計算分析。接著再根據分析後的結果，將題目排列成金字塔式的適性排列，並開始著手測驗介面的設計。測驗介面的開發工作分為「測驗的設計」、「測驗的建置」、「測驗的測試」及「測驗的修正」等，在此期間，測驗的開發會與專家諮詢討論作不斷的修正以趨完成。測驗開發完成後會由專家及使用者來作測驗的效能及滿意度的評估，以檢驗測驗的適切性。最後，並根據研究結果來完成本研究的論文報告。綜上所述，本研究之架構圖，如圖 3-1 所示。

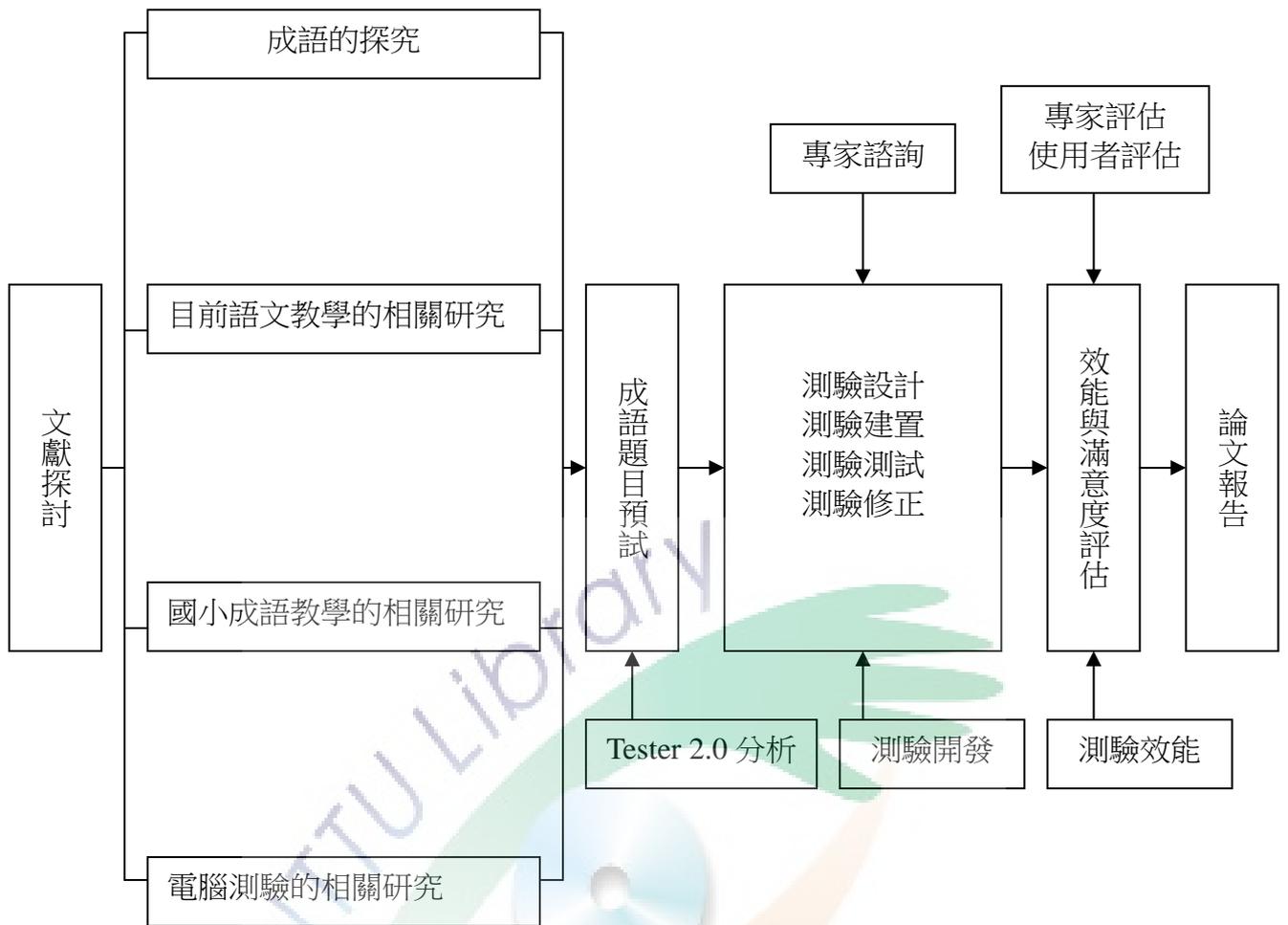


圖 3-1：研究架構圖

## 第二節 研究對象

以立意取樣的方式，抽取高雄縣五所國民小學的五、六年級學童，進行成語電腦適性測驗的施測。抽取後的學校及人數如下：

第一間為「高雄縣鳳山市文山國民小學人文藝術暨雙語實驗學校」五年三班 32 人、六年二班 31 人；第二間為「高雄縣中山國民小學」五年四班 34 人、六年四班 35 人；第三間為「高雄縣後庄國民小學」五年三班 31 人、六年三班 29 人；第四間為「高雄縣五甲國民小學」五年一班 25 人；六年四班 21 人；第五間為「高雄縣金潭國民小學」五年一班 23 人、六年二班 23 人，總計十個班級，共 284 名學生。

### 第三節 研究設計

本研究的進行，研究者以本身所整理的一至六年級翰林版國語課本中的成語及坊間可收集到的成語題庫等資料來作為試題來源之參考，設計成語題庫試卷，並以高雄縣文山國小一至六年級的學生，每個年級各抽一班來做試卷預試。所得的預試樣本，再以 **Tester 2.0** 軟體，完成題目難度、鑑別度的計算後，依每個題目的難度及鑑別度整理排列成適性測驗金字塔階層的出題順序。

完成適性測驗題庫的建立後，研究者研擬使用 **Flash** 多媒體的呈現方式讓學生使用電腦來完成此一電腦適性測驗。

施測完成後，研究者會將傳統測驗的分數與電腦適性測驗的分數作相關性的比較，並且對使用者做滿意度的問卷調查分析。

### 第四節 研究工具

本研究所使用之研究工具有：成語題目預試試卷、成語電腦適性測驗、成語傳統紙筆試卷及使用成語電腦適性測驗滿意度調查問卷。

#### 壹、成語題目預試試卷

成語題目預試試卷之題目來源為研究者本身所整理的一至六年級翰林版國語課本中的成語及坊間可收集到的成語題庫等資料來作為試題來源之參考，所設計而成的成語題庫試卷。其中剔除較為艱深冷門不常見的成語內容，以常用常見能應用的成語為主要挑選內容。所選試題如有題目敘述不清、誘答選項不佳等情形，則進行逐題修正。修正後，共選擇五十題。

#### 貳、成語電腦適性測驗

成語電腦適性測驗為根據研究目的、文獻探討等設計而成的。內容由預試的五十題成語題目中刪除及修改為二十八題，並作金字塔階層的適性排列，施測介面以 **Flash** 作為呈現的方式。

## 參、 成語傳統紙筆試卷

內容等同於成語電腦適性測驗，只是沒有作金字塔階層的適性排列，但二十八題的排列順序，則是按照難度由易而難來排列，越前面的試題越簡單，越後面的試題則越困難。

## 肆、 使用成語電腦適性測驗滿意度調查問卷

此滿意度之問卷，主要是針對研究者所開發的成語電腦適性測驗做滿意度的調查，與專家討論後，主要內容有：學生是否喜歡做成語電腦適性測驗、學生是否喜歡做成語紙筆測驗試卷、學生是否對成語電腦適性測驗感到有趣、學生是否有受到成語電腦適性測驗之歡迎動畫的吸引、學生是否覺得成語電腦適性測驗的整體版面設計美觀、學生是否覺得成語電腦適性測驗的題目文字清楚明確、學生是否覺得成語電腦適性測驗用滑鼠點選答案很方便、學生是否覺得成語電腦適性測驗的題目動畫很搭配、學生是否覺得成語電腦適性測驗做起來很新奇、學生之後是否還會想再做其它類似的電腦適性測驗等。

## 第五節 資料分析

本研究的資料分析，是針對成語題目預試試卷、成語電腦適性測驗、成語傳統紙筆試卷及使用成語電腦適性測驗滿意度調查問卷所獲得的資料進行統計分析。

- 壹、以 Tester 2.0 做成語題目預試試卷之題目難度及鑑別度的分析。
- 貳、以 Spss 12.0 做成語電腦適性測驗及成語傳統紙筆試卷的皮爾森積差相關的分析。
- 參、以 Spss 12.0 做使用成語電腦適性測驗滿意度調查問卷的次數分配表分析。

## 第六節 研究流程

本研究之「研究流程圖」如圖 3-2 所示，首先進行成語題目的預試，針對預試結果作難度分析及鑑別度分析，依其分析結果，將題目排列成金字塔階層的出題順序，結合 Flash 開發建立一個成語電腦適性測驗，將開發而成的成語電腦適性測驗，立意抽取五所學校做施測，再觀察其與傳統測驗效能的相關性及使用者對其施測後的滿意程度。

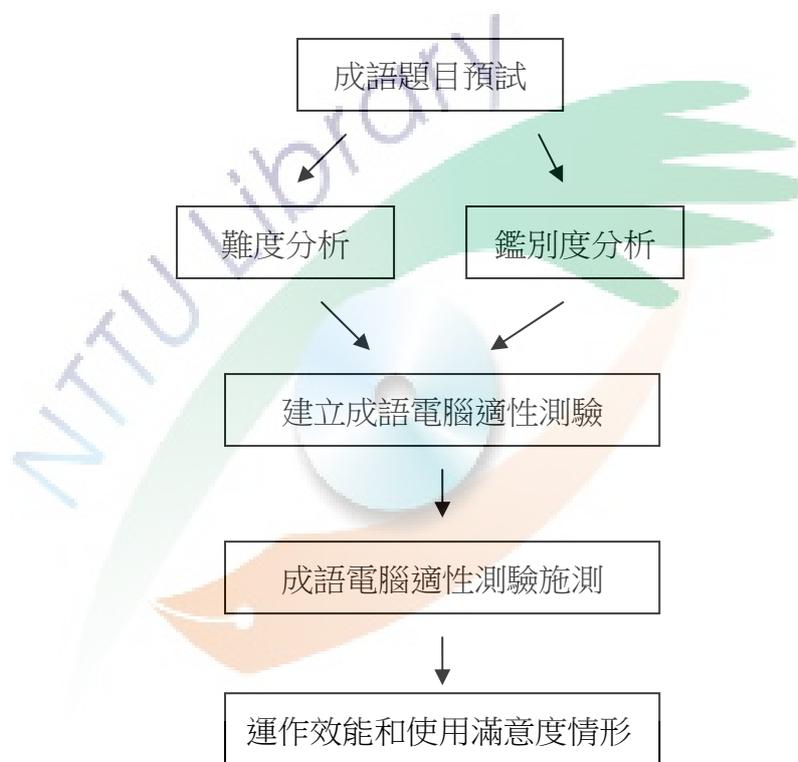
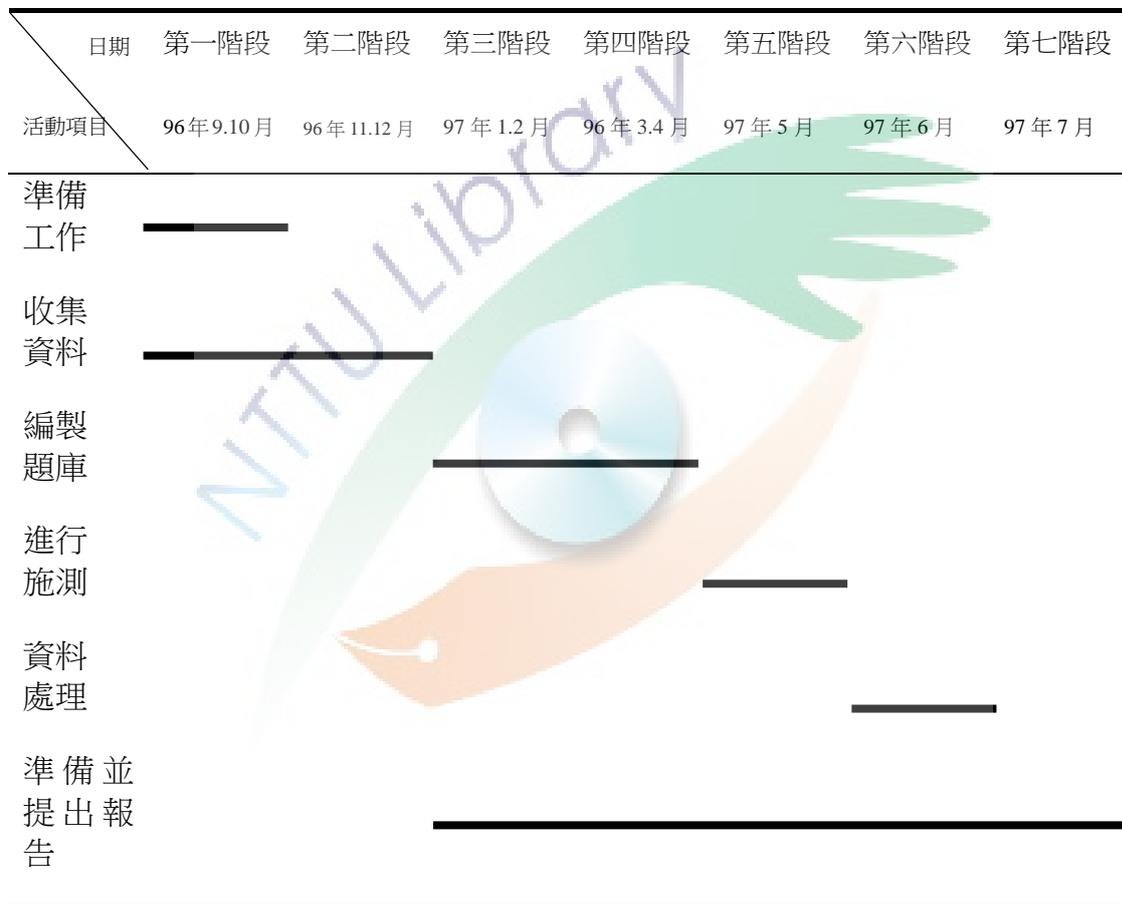


圖 3-2：研究流程圖

研究時間進度表的甘特圖如表 3-1 所示。第一階段的準備工作，預定於民國九十六年九月至十月間完成；第二階段的資料收集，預定於民國九十六年九月至十二月間完成；第三階段、第四階段的編製題庫，預定於民國九十七年一月至四月間完成；第五階段的進行施測，預定於民國九十七年五月完成；第六階段的資料處理，預定於民國九十七年六月完成；第七階段的準備並提出報告，則是一直持續直至完成。

表 3-1：研究時間進度表



## 第四章 結果與討論

本研究旨在發展一個「成語電腦適性測驗」，以提供教師一個電腦網路輔助測驗的環境，並能針對學生提供量身訂做的適性測驗。因此，如何將一些零散的成語題目，篩選排列成符合學生能力的適性題庫，也是本章所要呈現的研究結果。其次，測驗效能與架構是否充實完備，便關係著本研究的成敗，故開發題庫結合多媒體的 Flash 介面實為一重要的工作。最後，則是有關 Flash 題庫測驗的效能評估與使用情形，以了解本開發測驗之功能。故本章分為「題庫建置」、「測驗開發」與「效能評估」這三方面加以說明並討論。

### 第一節 成語電腦適性測驗建置過程與結果

首先，題庫測驗的建置，必須要有具代表性的成語題目來提供使用者線上施測。因此，本節旨在呈現題庫建置的過程，並進一步利用 Tester 2.0 的軟體來進行篩選分析，最後再與適性測驗整合，成為符合學生能力的適性測驗題庫。

#### 壹、預試題庫建置過程

從文獻上可知，試題品質的良窳決定題庫建置品質的優與劣。因此，在建立題庫的試題方面，研究者以本身所整理的一至六年級翰林版國語課本中的成語及坊間可收集到的成語題庫等資料來作為試題來源之參考，且為對試題及適合施測的年級有更精確的掌握，本研究有先進行預試，茲將過程與結果說明如下：

本研究最初所建立的題庫是依國小一至六年級學生而建立，測驗內容為成語。故在建立紙筆題庫時，試題來源參考包括研究者所整理的一至六年級翰林版國語課本中的成語及坊間可收集的成語題庫等。

由於本研究一開始設定的對象為國小一至六年級的學生，故根據一至六年級翰林版國語課本中的成語及坊間可收集的成語題庫等試題中，剔除較為艱深冷門不常見的成語內容，以常用常見能應用的成語為主要挑選內容。所選試題如有題目敘述不清、誘答選項不佳等情形，則進行逐題修正。修正後，共選擇五十題。

## 貳、 紙筆題庫預試結果分析

分為以下「受試樣本」、「分析軟體」及「分析結果」等三部分，說明如下：

### 一、 受試樣本

紙筆題庫預試的樣本為研究者立意抽取，高雄縣鳳山市文山國民小學人文藝術暨雙語實驗學校一至六年級，每個年級各抽取一班，共選取一年五班，計 33 名；二年一班，計 35 名；三年三班，計 28 名；四年三班，計 31 名、五年三班，計 31 名及六年四班，計 31 名，共計六個班級，189 名學生。每班皆做相同的試卷，施測時間為一節課，計四十分鐘。

### 二、 分析軟體

本預試所使用的分析軟體為一 TESTER for Windows 程式 2.0 版（版權所有 ©2002 年，余民寧教授 國立政治大學教育學系）。

簡介如下：

TESTER for Windows 程式 2.0 版主要用來執行古典測驗理論下的試題分析、測驗分析、和學生問題表分析等各種方法，以落實各種實證指標的建立工作。

本程式使用的最大容量範圍限制：學生數在 10000 人以內、試題數在 250 題以內、5 個選項以內的單選之選擇型試題、且每題的輸入資料格式必須是數值型（numerical）的作答反應資料（例如：即使選項為 A、B、C、D 的文字格式，也必須轉成 1、2、3、4 的數學格式來輸入）。此外，本程式的 S-P 表繪圖功能，僅限於學生數在 60 人、試題數在 50 題以內的資料，才能啟動使用。

### 三、 分析結果

#### (一) 年級分析

在測驗初步建立之後，製成五十題的預試測驗如附錄一。研究者立意抽取高雄縣鳳山市文山國民小學人文藝術暨雙語實驗學校共六個班級，189 名學生為預試對象，本預試分析結果作為適性題庫題目修改及刪除的依據。

預試結果採用 Tester 2.0 軟體進行試題難度、鑑別度、優劣率等相關的分析。茲將分析結果整理如表 4-1 所示：

表 4-1：國小成語測驗題庫分析一覽表

	良質試題 (A 和 B)	劣質試題 (A'和 B')	試題良質率
一年級	16	34	32%
二年級	21	29	42%
三年級	25	25	50%
四年級	29	21	58%
五年級	35	15	70%
六年級	36	14	72%

由 Tester 2.0 軟體執行後的報表中，被標註為 A 和 B 的題目，我們歸類為良質試題，表示可以藉由這些題目測出學生的能力；反之被標註 A'和 B'的題目，我們歸類為劣質試題，表示這些題目無法測出學生的成語能力。其中的試題良質率為良質試題佔全部試題的百分比。所以由表 4-1 可知這份試卷，施測於一年級時，約有 32% 的良質試題；施測於二年級時，約有 42% 的良質試題；施測於三年級時，約有 50% 的良質試題；施測於四年級時，約有 58% 的良質試題；施測於五年級時，約有 70% 的良質試題；施測於六年級時，約有 72% 的良質試題。

良質試題率，隨著年級的增加而增加，代表這份試卷對於中低年級的小朋友來說，有些試題可能過難還未學過，以致於缺乏鑑別度。但對於高年級的小朋友，這份試卷都約有七成的良質率，所以這份試卷對於高年級的小朋友來說是較具有鑑別度的。因此本研究的正式施測對象便設定為以國小高年級的五、六年級小朋友為主。

## (二) 試題分析

五、六年級預試利用 Tester 2.0 的施測結果分析整理表如附錄二及附錄三。其中內部一致性係數和差異係數整理如表 4-2 所示：

表 4-2：五、六年級預試之內部一致性係數和差異係數整理表

	內部一致性係數	差異係數
五年級	0.85	0.38
六年級	0.91	0.35

信度係數在 .70 至 .93 之間，斂聚效度及區辨效度在文獻上是相當肯定的。所以這份試卷施測於五、六年級上是具有顯著性的。差異係數  $> .40$  及  $< .60$  是最好的（余民寧，2002）。不過施測結果出來的差異係數分離程度較小，代表學生的學習成就差異過於明顯，而造成此一現象最大的原因，便是試題有待修改的空間，因為 Tester 2.0 的分析結果，也能顯示出試題的難度及鑑別度，因此有些鑑別度  $< .20$  的題目，便需要加以修改或刪除。

## 參、Flash 題庫與適性測驗的整合

依據 Tester 2.0 對五、六年級預試施測的分析，首先扣除劣質試題（A'和 B'）及鑑別度低於 .20 的題目，再修改題意不清或調整較缺乏誘答力的選項，最後按難度由易而難的排列後共 28 題如表 4-3，所完成正式施測的傳統紙筆測驗如附錄四。

表 4-3：正式施測題庫按難易度排列及編號表

難度	鑑別度	題目	編號
0.92	0.65	(4) 指高興得過了頭，轉而招致悲傷的事出現①苦中作樂②生不如死③悲從中來④樂極生悲	Q1
0.85	0.71	(4) 百貨公司開幕，大把大把的紅包從屋頂灑下來，真是「□□□□」呀！①禍從天降②天旋地轉③異想天開④喜從天降	Q2
0.83	0.41	(3) 看他們夫婦兩人含情脈脈、「□□□□」的樣子，有多幸福啊！①心血來潮②心如止水③心心相印④心安理得	Q3
0.80	0.44	(4) 「五體投地」①比喻人神通廣大，力強可畏②形容心情起伏不定，驚惶不安③比喻花樣繁多、變化多端④比喻非常欽佩	Q4
0.77	0.59	(4) 比喻針對問題所在，採取有效的措施①藥到病除②鐵口直斷③妙手回春④對症下藥	Q5
0.75	0.52	(2) 形容非常突出①五窮六絕②數一數二③三三兩兩④九死一生	Q6
0.72	0.61	(3) 「冷眼旁觀」①平靜無事②心氣平和，不急躁、不發怒③形容漠不關心 ④形容非常謹慎，不敢疏忽	Q7
0.70	0.56	(3) 災區居民生活困頓，請大家發揮「□□□□」的精神，有錢出錢，有力出力。①雪上加霜②風花雪月③雪中送炭④冰雪聰明	Q8
0.68	0.41	(4) 光明___落①累②雷③壘④磊	Q9
0.65	0.43	(2) 「有目共睹」①比喻才思敏捷，談吐風雅②極為清楚明顯，大家都看得見③比喻得到眾人的讚美④一句話就能說中事情的要點或真相	Q10
0.61	0.62	(3) 比喻對事物的情況捉摸不透；也形容人的心機極深①高不可攀②一望無際③深不可測④遙不可及	Q11

(續下頁)

(續表 4-3)

0.54	0.54	(1) 犯罪集團為了讓青少年「□□□□」地效忠，往往利用毒品控制他們，手段十分卑劣。①死心塌地②死有餘辜③九死一生④死不瞑目	Q12
0.53	0.79	(4) 人類因能完成夢想而偉大，但夢想的實現，絕非「□□□□」可達成。①朝不保夕②命在旦夕③朝令夕改④一朝一夕	Q13
0.51	0.77	(2) 無微不____①致②至③置④制	Q14
0.50	0.48	(1) ____步自封①故②固③雇④顧	Q15
0.48	0.66	(4) 「迎刃而解」①對於問題不求了解②還有餘力去處理更多的問題③問題難以釐清頭緒④解決問題很順利	Q16
0.46	0.65	(2) 比喻空談理論，不能解決實際問題①空中樓閣②紙上談兵③白紙黑字④千軍萬馬	Q17
0.44	0.81	(3) 相隔數十年後，他們雖然「□□□□」，但恍如隔世。①鏡花水月②明鏡高懸③破鏡重圓④目光如鏡	Q18
0.40	0.55	(1) 他在送舊晚會上「□□□□」，博得滿堂彩。①引吭高歌②引人注目③引以自豪④引以為榮	Q19
0.39	0.61	(1) 「不可開交」①沒法解開或擺脫②不聽也不問③不分是非④指一件事不可預測和估量	Q20
0.38	0.67	(3) _____相關①習習②昔昔③息息④襲襲	Q21
0.36	0.83	(2) 有話不便說①一言難盡②有口難言③苦不堪言④百口莫辯	Q22
0.35	0.71	(2) 小心_____①亦亦②翼翼③意意④奕奕	Q23
0.30	0.70	(3) 不可思____①異②義③議④意	Q24
0.29	0.60	(2) 「有口皆碑」①比喻嚴守秘密，不輕易告訴別人②比喻人人稱讚、頌揚③事情的問題所在④以懇切真摯的態度，有耐心的勸告他人	Q25

(續下頁)

(續表 4-3)

0.25	0.56	(3) 形容一無所有了①一言難盡②面目全非 ③別無長物④孤掌難鳴	Q26
0.24	0.51	(3) 人____亦____①雲②耘③云④芸 (1)「風起雲湧」①聲勢盛大②變化很多， 無法預測③雙方一觸即發的緊張態勢④對事 物情況捉摸不透	Q27
0.19	0.51		Q28

表 4-3 的 28 題成語題目中，屬於判讀「成語字音字形」類型的題目有 7 題；屬於挑選「符合解釋的成語」類型的題目有 7 題；屬於判別「成語的解釋」類型的題目有 7 題；屬於「選擇適當的成語填入句子」類型的題目有 7 題，題目類型平均分配。

而這份傳統紙筆測驗上的 28 題題目，同時也是成語電腦適性測驗所要施測的內容，所不同的是傳統紙筆測驗是按照難度的排列，而成語電腦適性測驗主要是與適性測驗結合，呈現七層金字塔式的適性排列如圖 4-1 (題目編號同表 4-3)。

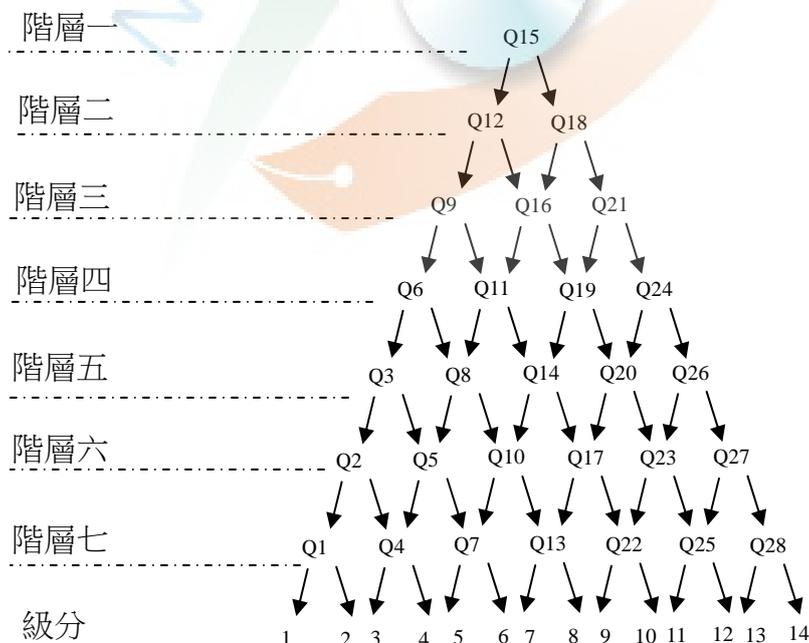


圖 4-1：七層金字塔式適性排列的線上 Flash 題庫

如圖 4.1 所示，此一線上 Flash 題庫共分為七個階層。階層一：Q15，共 1 題；階層二：Q12、Q18，共 2 題；階層三：Q9、Q16、Q21，共 3 題；階層四：Q6、Q11、Q19、Q24，共 4 題；階層五：Q3、Q8、Q14、Q20、Q26，共五題；階層六：Q2、Q5、Q10、Q17、Q23、Q27，共六題；階層七：Q1、Q4、Q7、Q13、Q22、Q25、Q28，共 7 題，總計 28 題。Q15 為初始題，難度為 .50 的中等試題，試題答對往左邊，反之往右邊繼續作答，其餘試題以此類推，直到測出最後的級分為止。一般傳統紙筆測驗需將 28 題的題目完全做完才能得到分數，本研究中的成語電腦適性測驗，學生只需做 7 題，即能測出學生的成語能力。因此若傳統紙筆測驗測出的分數能與成語電腦適性測驗所測出的級分得到高度相關，那麼本研究的成語電腦適性測驗即達成目標。而施測的結果與討論則留待第三節繼續探討。

## 第二節 成語電腦適性測驗之發展模式

本節將成語電腦適性測驗之開發過程分為「測驗規劃」、「測驗設計」、「測驗建置」、「測驗運行與支援」等四個階段。以下茲就四個階段來說明本整合 Flash 題庫測驗之開發過程。

### 壹、測驗規劃

在測驗規劃的階段裡，選擇一個良好適當的測驗施測介面，是關係著本研究所發展的題庫測驗是否可行且易被國小學童所接受。

目前常被大家用來做線上測驗的程式很多，例如：ASP、PHP+SQL 資料庫、C++、JavaScript、Macromedia Dreamweaver 製作最基本的 HTML 及 Flash 等等不勝枚舉。然而考量到本研究者的科技技術背景且施測對象為國小學童，因此結合了音效、影像、動畫，且具備了高品質、體積小、互動性高等優點的 Flash，便成為了本研究用來製作題庫測驗的主要介面程式。

Flash 的特色大略分述如下列幾點：

## 一、 高畫質

由於 Flash 使用的是向量繪圖方式，因此可以製作出畫質佳的動畫，不論縮放視窗的大小，都不改其品質。

## 二、 檔案體積小

Flash 所製出的動畫，檔案體積很小，相當適合透過網路媒體傳播。

## 三、 簡易的製作環境

製作環境「面板化」，可縮放摺疊，更省空間，管理效果更方便。

## 四、 可匯入 / 匯出多種檔案格式

點陣：BMP、JPEG、GIF 動畫、PNG

向量：AI (Adobe Illustrator)、WMF、FreeHand、Flash SWF

聲音：WAV

視訊：AVI、QuickTime、MPEG、DV

## 五、 影音整合

Flash 可以在播放動畫時加入音效與背景音樂，這是以往使用 GIF 動畫無法做到的，Flash4 之後的版本在聲音的壓縮上支援 MP3 格式，可以進一步縮小動畫檔的體積，FlashMX 更可匯入視訊檔（例如 QuickTime）。

## 六、 威力強大的 ActionScript 語言

Flash 有專屬的程式語言 ActionScript，可以讓動畫產生互動效果，網路上時常可以看到 Flash 設計的遊戲或應用程式（例如留言板、繪圖板、線上訂票訂房系統）。

## 七、 外掛程式隨處可見

利用搜尋引擎打關鍵字搜尋，均可搜尋上網路上很多免費且實用的 Flash 外掛小程式。

## 八、 標準播放器

可將.swf 動畫轉換成.exe 可執行檔。

## 九、 模擬動畫下載情況

Flash 動畫採串流傳輸模式，不用等到 Flash 完全下載，可以邊下載邊播放，避免訪客漫長等待。

## 十、 容易與網頁整合

Flash 輸出檔 (.swf) 是目前所有瀏覽器當可以撥放的檔案格式。

## 十一、 高品質的列印功能

Flash 以向量圖方式處理內容，圖形精美且無鋸齒。

綜合上述的特點，Flash 是一套簡單並且功能強大的製作軟體，除了動畫、網頁、甚至遊戲都能夠使用此套軟體製作。現在的瀏覽器都內建 Flash Player 就可以知道其普及的程度，尤其在跨平台上的優點和生動的動畫，更讓大眾有很高的接受度。所以研究者選擇這套軟體來製作線上題庫測驗的施測介面，希望其高畫質、顯示速度快及多媒體的高度整合能力，能為這次的成語電腦適性測驗製作更有加分的效果，且易被國小學童所接受。

## 貳、 測驗設計

研究規劃以電腦為工具的資訊處理，主要的效能是可以提高資訊的品質與工作效率（饒瑞文，1993）。因此本研究所設計的成語電腦適性測驗旨在利用電腦作為施測工具，Flash 為多媒體的施測介面，以成語適性測驗題庫為內容，期能達到更準確地評估學生成語能力並能節省施測及批改試卷的時間。

結構化分析中最主要的工具就是資料流程圖，它是利用圖形來描述各部份之間的資料流情況。一般來說，完整的資料流程圖是一個網路圖，用來描述每個部份的資料流動情形，資料流程圖是由以下四個基本元件構成（宋麗麗、曹延傑，1990；陳新豐，2002；Shelly、Cashman & Rosenblatt，2001）：

### 一、 資料流

資料流的方向是以箭頭來表示，主要用來追縱資料中的流程。

### 二、 處理程序

每一個處理程序是以圓圈來表示，並將處理程序的名稱寫在圓圈內。處理程序是從事資料處理的部份。

### 三、 檔案

在結構化分析的資料流程圖中是以並行雙橫線來表示檔案，檔案名稱是寫在雙橫線之間。檔案主要是用來儲存資料的地方。

### 四、 源頭或終點

源頭是所使用的資料來源；終點則是接收所產生的資料之接收處。

根據上述四個基本元件，研究者提出本測驗之資料流程圖如圖 4-2 所示。從圖 4-2 可知，本測驗主要的使用對象為國小高年級學生、教師及專家學者，學生可以經由本測驗，並且得到適當的測驗結果；教師主要在提供適性題庫來豐富測驗內容，而測驗則能有效節省教師施測及批改時間；專家學者則可以提供測驗建置時、測驗操作與運行後之結果的評估及考量。

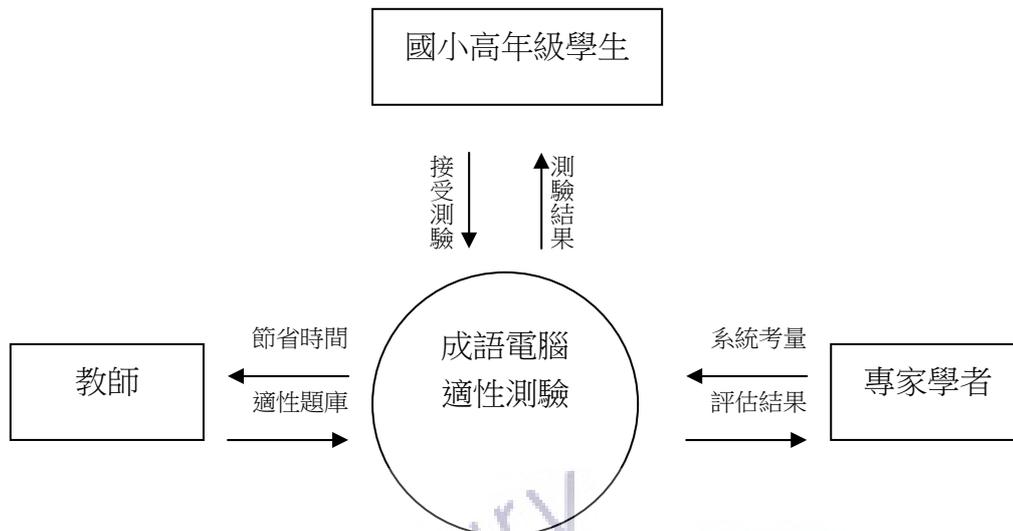


圖 4-2：「成語電腦適性測驗」之資料流程圖

## 參、測驗建置

「測驗建置」旨在說明「測驗發展」、「測驗測試」及「適性測驗建置相關問題」等三個部分，依序說明如下：

### 一、測驗發展

經由測驗規劃與設計的步驟，確定本測驗所應具備的內容、資料流程之後，接下來便要進入到測驗實際製作的階段，完成所需要建置的測驗。以下分別就「硬體規格」、「軟體系統」及「程式分析」等三部份加以說明。

在硬體方面，研究者預設輸出畫面為 800×600，但因為 Flash 的向量圖形不失真且高畫質，因此不管使用者的系統輸出入畫面預設的解析度如何，畫面都是精細且無鋸齒狀。另外，最普及的 IE 瀏覽器，都內建有 Flash Player 的播放程式，因此使用者不需額外下載其它的播放器，線上下載即可使用。

在軟體系統方面，使用的是 Macromedia Flash 8。成語電腦適性測驗的歡迎動畫、版面設計、題目文字、題目動畫及整體美觀、使用的介面，都盡量能以使用者感到最友善、最親切為考量來製作。

在程式分析分方面，則是利用 **Tester 2.0** 的程式來進行題目的分析，包括試題的良質率、難度、鑑別度、內部一致性係數、差異係數等。再經由每題的難度及鑑別度來做題庫金字塔式的適性排列，最後會依據 **Flash** 測驗所得的級分與傳統測驗所得的分數來做相關等效能的評估。

## 二、 測驗測試

首先在整合測試階段，都是由研究者，直接輸入修改編輯，來測試整個 **Flash** 適性測驗執行上的正確性。再經由專家學者的修正建議直到完成。驗收測試階段則是開放線上測驗後，請到台東大學教學科技碩士班一年級的學生一一驗證無誤後，成語電腦適性測驗終於在九十七年四月上旬左右完成所有測試工作，使測驗運行無誤，才於民國 97 年 4 月 15 日起，開始進行本研究的正式施測。

以下將以成語電腦適性測驗之開頭歡迎動畫、注意事項說明、測驗畫面及最後得分結果圖示如下：



- (一) 如圖 4-3，成語電腦適性測驗的開頭歡迎畫面，按下右下角的按鈕之後，即可進入測驗。



圖 4-3：「成語電腦適性測驗」之開頭歡迎畫面

(二) 如圖 4-4，進入測驗前的注意事項說明頁。小朋友閱讀完注意事項後，接著按下「測驗開始」鈕後，即進行施測。

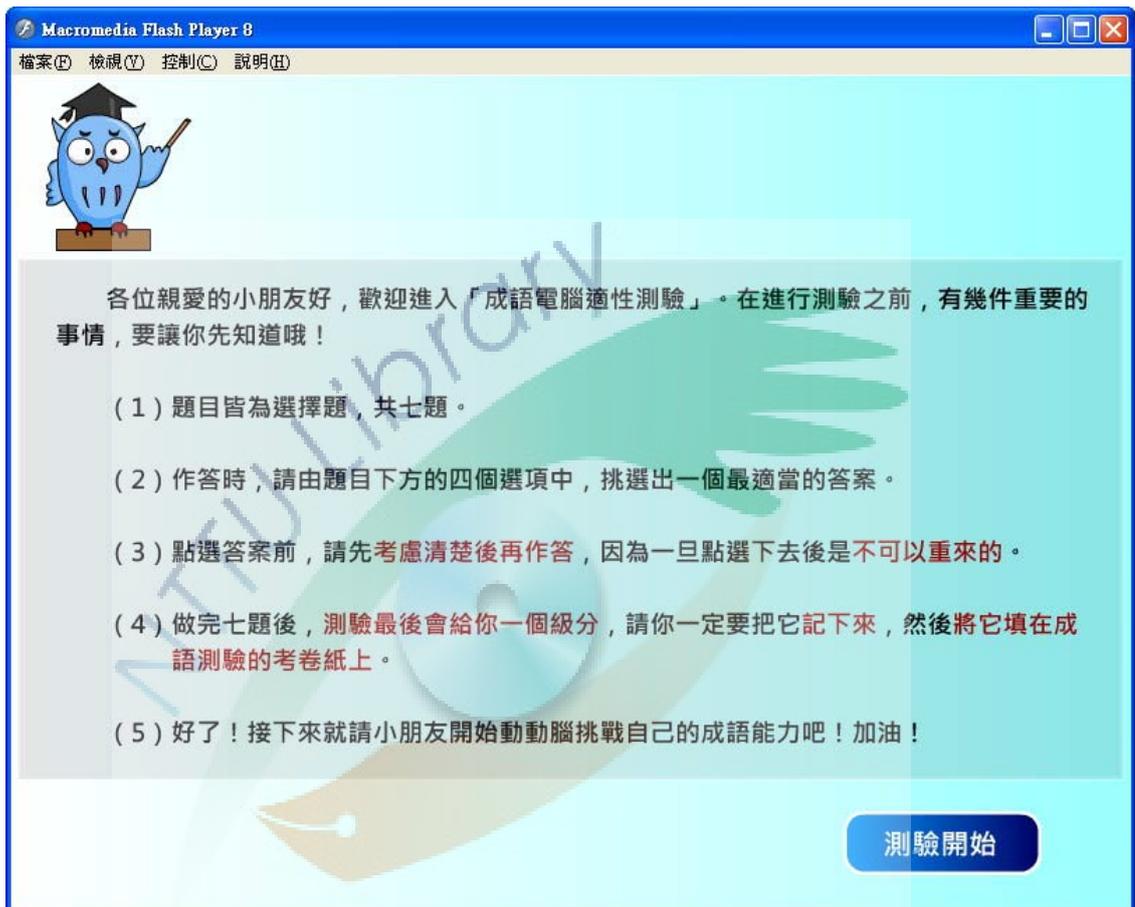


圖 4-4：「成語電腦適性測驗」之注意事項說明頁

(三) 如圖 4-5 所示，難度中等的測驗初始題。這題若答對會往難度更高的題目繼續作答，若答錯則往難度更低一點的題目繼續作答。



圖 4-5：「成語電腦適性測驗」之初始題

(四) 圖 4-6 所示為剩餘 27 題由易而難排列的適性題目之作答

畫面，每一題的適性題目都有依據其題意所搭配之動畫。



(續下頁)

(續圖 4-6)

<p>wkicon1_6</p> <p>請選出正確的答案：</p> <p>比喻針對問題所在，採取有效的措施</p> <ul style="list-style-type: none"><li><input type="radio"/> 藥到病除</li><li><input type="radio"/> 擱口直斷</li><li><input type="radio"/> 妙手回春</li><li><input type="radio"/> 對症下藥</li></ul> 	<p>wkicon1_7</p> <p>請選出正確的答案：</p> <p>形容非常突出</p> <ul style="list-style-type: none"><li><input type="radio"/> 五窮六絕</li><li><input type="radio"/> 數一數二</li><li><input type="radio"/> 三三兩兩</li><li><input type="radio"/> 九死一生</li></ul> 
<p>wkicon1_8</p> <p>請選出正確的答案：</p> <p>「冷眼旁觀」</p> <ul style="list-style-type: none"><li><input type="radio"/> 平靜無事</li><li><input type="radio"/> 心氣平和，不急躁、不發怒</li><li><input type="radio"/> 形容漠不關心</li><li><input type="radio"/> 形容非常謹慎，不敢疏忽</li></ul> 	<p>wkicon1_9</p> <p>請選出正確的答案：</p> <p>災區居民生活困頓，請大家發揮「□□□□」的精神，有錢出錢，有力出力。</p> <ul style="list-style-type: none"><li><input type="radio"/> 雪上加霜</li><li><input type="radio"/> 風花雪月</li><li><input type="radio"/> 雪中送炭</li><li><input type="radio"/> 冰雪聰明</li></ul> 
<p>wkicon1_10</p> <p>請選出正確的答案：</p> <p>光明___落</p> <ul style="list-style-type: none"><li><input type="radio"/> 累</li><li><input type="radio"/> 雷</li><li><input type="radio"/> 霹</li><li><input type="radio"/> 霹</li></ul> 	<p>wkicon1_11</p> <p>請選出正確的答案：</p> <p>「有目共睹」</p> <ul style="list-style-type: none"><li><input type="radio"/> 比喻才思敏捷，談吐風雅</li><li><input type="radio"/> 極為清楚明顯，大家都看得見</li><li><input type="radio"/> 比喻得到眾人的讚美</li><li><input type="radio"/> 一句話就點出事情的要點或真相</li></ul> 

(續下頁)

(續圖 4-6)

<p>wkicon1_X</p> <p>請選出正確的答案：</p> <p>比喻對事物的情況捉摸不透；也形容人的心機極深</p> <ul style="list-style-type: none"><li><input type="radio"/> 高不可攀</li><li><input type="radio"/> 一望無期</li><li><input type="radio"/> 深不可測</li><li><input type="radio"/> 遙不可及</li></ul> 	<p>wkicon1_Y</p> <p>請選出正確的答案：</p> <p>犯罪集團為了讓青少年「□□□□」地效忠，往往利用毒品控制他們，手段十分卑劣。</p> <ul style="list-style-type: none"><li><input type="radio"/> 死心塌地</li><li><input type="radio"/> 死有餘辜</li><li><input type="radio"/> 九死一生</li><li><input type="radio"/> 死不瞑目</li></ul> 
<p>wkicon1_M</p> <p>請選出正確的答案：</p> <p>人類因能完成夢想而偉大，但夢想的實現，絕非「□□□□」可達成。</p> <ul style="list-style-type: none"><li><input type="radio"/> 朝不保夕</li><li><input type="radio"/> 命在旦夕</li><li><input type="radio"/> 朝令夕改</li><li><input type="radio"/> 一朝一夕</li></ul> 	<p>wkicon1_X</p> <p>請選出正確的答案：</p> <p>無微不至</p> <ul style="list-style-type: none"><li><input type="radio"/> 致</li><li><input type="radio"/> 至</li><li><input type="radio"/> 實</li><li><input type="radio"/> 制</li></ul> 
<p>wkicon1_7</p> <p>請選出正確的答案：</p> <p>「迎刃而解」</p> <ul style="list-style-type: none"><li><input type="radio"/> 對於問題不求了解</li><li><input type="radio"/> 還有餘力去處理更多的問題</li><li><input type="radio"/> 問題難以釐清頭緒</li><li><input type="radio"/> 解決問題很順利</li></ul> 	<p>wkicon1_Q</p> <p>請選出正確的答案：</p> <p>比喻空談理論，不能解決實際問題</p> <ul style="list-style-type: none"><li><input type="radio"/> 空中樓閣</li><li><input type="radio"/> 紙上談兵</li><li><input type="radio"/> 白紙黑字</li><li><input type="radio"/> 千軍萬馬</li></ul> 

(續下頁)

(續圖 4-6)

<p>請選出正確的答案：</p> <p>相隔數十年後，他們雖然「□□□□」，但恍如隔世。</p> <ul style="list-style-type: none"><li><input type="radio"/> 鏡花水月</li><li><input type="radio"/> 明鏡高懸</li><li><input type="radio"/> 破鏡重圓</li><li><input type="radio"/> 目光如鏡</li></ul> 	<p>請選出正確的答案：</p> <p>他在送舊晚會上「□□□□」，博得滿堂彩。</p> <ul style="list-style-type: none"><li><input type="radio"/> 引吭高歌</li><li><input type="radio"/> 引人注目</li><li><input type="radio"/> 引以為豪</li><li><input type="radio"/> 引以為憾</li></ul> 
<p>請選出正確的答案：</p> <p>「不可開交」</p> <ul style="list-style-type: none"><li><input type="radio"/> 沒法解開或擺脫</li><li><input type="radio"/> 不聽也不聞</li><li><input type="radio"/> 不分是非</li><li><input type="radio"/> 指一件事不可預測和估量</li></ul> 	<p>請選出正確的答案：</p> <p>_____ 相關</p> <ul style="list-style-type: none"><li><input type="radio"/> 習習</li><li><input type="radio"/> 音音</li><li><input type="radio"/> 息息</li><li><input type="radio"/> 曉曉</li></ul> 
<p>請選出正確的答案：</p> <p>小心 _____</p> <ul style="list-style-type: none"><li><input type="radio"/> 亦亦</li><li><input type="radio"/> 翼翼</li><li><input type="radio"/> 意意</li><li><input type="radio"/> 奕奕</li></ul> 	<p>請選出正確的答案：</p> <p>有話不便說</p> <ul style="list-style-type: none"><li><input type="radio"/> 一言難盡</li><li><input type="radio"/> 有口難言</li><li><input type="radio"/> 苦不堪言</li><li><input type="radio"/> 百口莫辯</li></ul> 

(續下頁)

(續圖 4-6)



圖 4-6：「成語電腦適性測驗」其餘 27 題之作答畫面

(五) 圖 4-7 為測驗結果之所得級分畫面。級分共分為 1-14 級分，級分越高代表學生成語能力程度越高，反之則學生成語能力程度較低。



圖 4-7：「成語電腦適性測驗」測驗結果之所得級分畫面

### 三、 適性測驗建置相關說明

本研究所建置的成語電腦適性測驗之發展研究，雖然可以適合用於多種學科，但是因為人力與物力的關係，所以以語文科中與作文方面很有相關的成語部分作為題庫的內容，而在題型部分，因為研究主題的焦點乃在於傳統紙筆測驗與電腦適性測驗的相關探討，因此目前只是以選擇題型為唯一的題型。本部份主要的目的即是呈現本測驗之六大要素，整理統整如下：

### (一) 金字塔式適性測驗模式

金字塔型適性測驗是一種題目間分支策略 (Vale, 1981; 1982)。測驗的進程是讓所有的受試者皆從塔尖的的初始題 (亦即第一題) 開始, 若在某一階段的反應是答對的, 則分支至下一題較難的題目; 反之, 則接受下一題較簡單的題目。每一位受試者在任一階段皆只對一個題目做出反應 (蔡振昆, 2001)。這樣的測驗模式能利用較少的題數, 做出更適性的能力評估。

### (二) 題庫

題庫的建置及其性質部分已在先前的章節中詳細說明, 所以此處只是簡單地予以介紹。本研究之「成語電腦適性測驗」所採用的題庫是利用研究者本身所整理的一至六年級翰林版國語課本中的成語及坊間可收集到的成語題庫等資料來作為主要的試題內容, 另外再加上研究者經紙筆預試的篩選修改, 測驗軟體 Tester 2.0 的統計分析, 而成為新的題庫。

### (三) 初始題的設定

經由 Tester 2.0 軟體分析出題目的難度及鑑別度後, 再由金字塔式的適性排列, 決定由測驗難度中等 ( $P = .50$ ) 的題目為初始題。

### (四) 題數的決定

一方面由 Tester 2.0 篩選出的良質試題, 再配合金字塔式適性測驗的階層題數分配, 最後再參考專家學者的建議, 本研究之「成語電腦適性測驗」的題庫總題數為 28 題。

### (五) 中止標準的設定

成語電腦適性測驗 Flash 的總題數為 28 題, 按照金字塔的適性分配, 共可分為七個階層, 雖說總共有 28 題的題目, 但施測者只需作答 7 題, 即可完成測驗。

## （六） 能力估計結果

做完 7 題的成語電腦適性測驗之後，隨著金字塔式的路徑作答，將會引領施測者至最後的級分估計階層，此次的七個階層作答後，將有 1-14 個級分的結果，成語能力越高者，級分數越高；反之，則越低。

## 肆、 測驗運行與支援

以下說明本研究所發展「成語電腦適性測驗」運行時之「軟硬體之規格需求」及其功能說明。

### 一、 硬體部份

#### （一） 研究者端

1. CPU：AMD Athlon(tm) 64 Processor 3200+
2. RAM：1.00 GB
3. HDD：WDC WD2500KS-00MJB0 250GB
4. 網路卡：10/100 Mbps

#### （二） 使用者端

1. CPU：Intel Pentium II 350 MHz（含）以上
2. RAM：128 MB（含）以上
3. HDD：2 GB（含）以上
4. 網路卡：10/100 Mbps
5. 顯示卡：可顯示 256 色以上，及解析度至少達 800×600（含）以上

## 二、 軟體部份

### (一) 研究者端

1. Macromedia Flash 8
2. Tester 2.0
3. 作業系統：Microsoft Windows XP
4. 成語電腦適性測驗：研究者利用 Macromedia Flash 8、Tester 2.0 等程式，建置在網際網路的環境，讓使用者連結下載，完成線上施測。

### (二) 使用者端

使用者端由於網際網路為開放性架構，因此本研究所建置的成語電腦適性測驗其使用者端之軟體，使用者可以自行依其習慣偏好調整以配合不同的工作平臺，以下為本研究對於使用者端建置之軟體部份。

1. 作業系統：Microsoft Windows 95/98/ME/2000/XP
2. 全球資訊網瀏覽器：Microsoft Internet Explorer 4.0 (含) 以上

## 三、 測驗功能

### (一) 測驗立體化

由圖 4-1 題庫的金字塔適性測驗排列的路徑圖可以看出題目間彼此連結的關係，研究者透過 Flash 的介面將彼此的階層結構、整個測驗的前因後果清楚的呈現出來。

## (二) 作答線上化

本研究的受試者可隨時上線作答，受試者可在任意時間與任意地點，只要透過網路的連結，便可直接在網路上回答本適性測驗的試題，以評量出自己成語的能力。

## (三) 結果自動化

受試者在線上作答完成後，透過金字塔的路徑進行，最後至會進入級分的階層，讓受試者知道此次測驗的程度級分，此過程完全可在線上自動完成，不需教師額外評分。

以上，為本研究所發展「成語電腦適性測驗」之功能，至於使用效能與使用者的滿意程度則會在本章第三節中繼續討論。



### 第三節 成語電腦適性測驗之滿意度評估

本節旨在評估本研究開發的「成語電腦適性測驗」的使用效能，茲從「使用結果與傳統測驗的相關性」和「使用者滿意程度」等二方面加以說明。

#### 壹、使用結果與傳統測驗的相關性

##### 一、受試樣本

正式施測樣本為研究生立意抽取，共抽取高雄縣五所學校之五、六年級各一班學生。第一間為「高雄縣鳳山市文山國民小學人文藝術暨雙語實驗學校」五年三班 32 人、六年二班 31 人；第二間為「高雄縣中山國民小學」五年四班 34 人、六年四班 35 人；第三間為「高雄縣後庄國民小學」五年三班 31 人、六年三班 29 人；第四間為「高雄縣五甲國民小學」五年一班 25 人；六年四班 21 人；第五間為「高雄縣金潭國民小學」五年一班 23 人、六年二班 23 人，總計十個班級，共 284 名學生，如表 4-4 所整理。每班皆做 7 題的成語電腦適性測驗、28 題的傳統紙筆測驗及 10 題的滿意度問卷。施測時間為一節課，計四十分鐘。施測日期於民國 97 年 4 月 15 日起，開始陸續進行本研究的正式施測。

表 4-4：正式施測之抽樣學校及班級數、人數一覽表

編號	縣市	校名	班級數	人數
S1	高雄縣	文山國小	2	63
S2	高雄縣	中山國小	2	69
S3	高雄縣	後庄國小	2	60
S4	高雄縣	五甲國小	2	46
S5	高雄縣	金潭國小	2	46
合計			10	284

## 二、 測驗內容

本測驗的「成語電腦適性測驗」的題庫為本章第一節表 4-3 所整理出共 28 題的成語題目，同時也是傳統紙筆測驗的題目，如附錄三。另外為了解使用者在使用「成語電腦適性測驗」後的態度與滿意程度，在與專家討論之後，設計出十題問卷如下表 4-5，問卷形式如附錄五。

表 4-5：正式施測之十題問卷題目

問卷題目
1.我比較喜歡做成語電腦適性測驗。
2.我比較不喜歡做成語紙筆測驗試卷。
3.我覺得做成語電腦適性測驗很有趣。
4.我認為成語電腦適性測驗的歡迎動畫很吸引我。
5.我覺得成語電腦適性測驗的整體版面設計很美觀。
6.我覺得成語電腦適性測驗的題目文字清楚明確。
7.我覺得成語電腦適性測驗用滑鼠點選答案很方便。
8.我覺得成語電腦適性測驗的題目動畫很搭配。
9.我做完成語電腦適性測驗後的感覺很新奇。
10.如果有其它類似的電腦適性測驗我會很想試試看。

### 三、 測驗實施方式

測驗實施的規劃流程如表 4-6：

表 4-6：正式施測時間規劃流程表

時間（分）	測驗內容
5	測驗說明
7	成語電腦適性測驗
18	傳統紙筆測驗
10	問卷

測驗範圍為成語，測驗內容共分為三個部份：「成語電腦適性測驗」、「傳統紙筆測驗」、「問卷」，測驗時間為一節課，四十分鐘。首先，會有五分鐘的測驗說明，說明之後，會讓學生經由網路（<http://psa.idv.tw/~sunny/welcome.swf>）連結「成語電腦適性測驗」的 Flash，雖然「成語電腦適性測驗」的題庫有 28 題，但經由金字塔適性測驗的分支理論，學生只需做 7 題，即能得出學生的成語級分；做完 7 題的成語電腦適性測驗後，會要求學生將所得的級分結果寫在傳統問卷上，接著再讓學生做完整 28 題的傳統紙筆測驗。最後等二份測驗做完之後，便給予學生做 10 題的回饋問卷。這樣的施測方式，施測在五所學校，共十個班級下，都很順利地在一節課四十分鐘內完成。

#### 四、 分析結果

施測共 284 名學生，回收也有 284 份資料，因此有效率為 100%，本次分析利用 Tester 2.0 軟體針對各校之五、六年級學生回收資料進行分析。茲將分析的結果摘要整理如表 4-7。

表 4-7：正式施測利用 Tester 2.0 分析結果摘要表

學校編號	五年級			六年級		
	r	K-R	D*	r	K-R	D*
S1	0.48	0.74	0.45	0.59	0.85	0.44
S2	0.49	0.79	0.54	0.56	0.84	0.40
S3	0.43	0.62	0.58	0.72	0.75	0.49
S4	0.76	0.84	0.45	0.55	0.81	0.50
S5	0.59	0.73	0.59	0.48	0.91	0.42

表 4.7 中，S1—S5 為學校之編號，編號之學校內容可查閱表 4.4。其中 r 為成語電腦適性測驗與傳統紙筆測驗的相關係數；K-R 為庫李信度(內部一致性係數)；D\*則為差異係數。

相關係數 > .50，就屬於高度正相關，所以 S1 六年級的 .59；S2 六年級的 .56、S3 六年級的 .72；S4 五年級的 .76、六年級的 .55；S5 五年級的 .59 均為高度正相關，而 S1 五年級的 .48；S2 五年級的 .49；S3 五年級的 .43；S5 六年級的 .48，也都很接近 .50，均屬於正相關，代表本研究所發展之「成語電腦適性測驗」與傳統的紙筆測驗的評量結果相關性高。

信度係數在 .70 至 .93 之間的資料是具有顯著性的，所以 S1 五年級的 .74、六年級的 .85；S2 五年級的 .79、六年級的 .84；S3 六年級的 .72；S4 五年級的 .84、六年級的 .81；S5 五年級的 .73、六年級的 .91，均在此斂聚效度與區辨效度間，代表這些資料是具有很高的可信度的，而 S3 五年級的 .62，也很接近 .70，因此也具有不錯的信度。

差異係數介於 .40 與 .60 之間是最好的，本次施測結果 S1 五年級的 .45、六年級的 .44；S2 五年級的 .54、六年級的 .40；S3 五年級的 .58、六年級的 .49；S4 五年級的 .45、六年級的 .50；S5 五年級的 .59、六年級的 .42，全都介於 .40 與 .60 之間，代表此次施測的成語題目，皆具有非常好的效度。

本研究旨在了解「成語電腦適性測驗」與「傳統紙筆測驗」的相關。因此以 Pearson 積差相關法，求出此二種測驗的相關情形，如表 4-8 所示。

前述是以 Tester 2.0 進行各校各班「成語電腦適性測驗」與「傳統紙筆測驗」的相關分析。這裡將以五所學校，共十個班級的「成語電腦適性測驗」與「傳統紙筆測驗」的資料利用 Spss 12.0 進行 Pearson 相關分析。

表 4-8：「成語電腦適性測驗」與「傳統紙筆測驗」之相關分析

	成語電腦適性測驗	傳統紙筆測驗
成語電腦適性測驗	1	0.57***
傳統紙筆測驗	0.57***	1

\*\*\* $p < .001$

由表 4.8 可知，「成語電腦適性測驗」與「傳統紙筆測驗」之間，呈現高度正相關 ( $r = .57$ ,  $p < .001$ )，而二測驗間之相關達顯著水準。

## 貳、使用者滿意程度

本研究旨在發展一個「成語電腦適性測驗」的 Flash 題庫測驗，以提供教師一個電腦輔助測驗的環境，並能針對學生提供量身訂做的適性測驗。本研究是以國小所學的成語內容為命題範圍，施測對象為國小五、六年級的學生，為能真確收到使用者的使用感受，研究者設計當學生授受連續性的測驗後，便給予一份評估感受的十題問卷讓學生填答，填答內容如附錄四。以下茲從這十題問卷內容加以分析整理如表 4-9。



表 4-9：使用者評估結果次數分配表

編號	題目	非常同意	同意	沒意見	不同意	非常不同意	平均數
Q1	我比較喜歡做成語電腦適性測驗。	119 (41.9%)	73 (25.7%)	80 (28.2%)	8 (2.8%)	4 (1.4%)	4.04 (.97)
Q2	我比較不喜歡做成語紙筆測驗試卷。	61 (21.5%)	47 (16.5%)	117 (41.2%)	33 (11.6%)	26 (9.2%)	3.30 (1.19)
Q3	我覺得做成語電腦適性測驗很有趣。	113 (39.8%)	87 (30.6%)	60 (21.1%)	17 (6.0%)	7 (2.5%)	3.99 (1.04)
Q4	我認為成語電腦適性測驗的歡迎動畫很吸引我。	91 (32.0%)	60 (21.1%)	95 (33.5%)	22 (7.7%)	16 (5.6%)	3.66 (1.17)
Q5	我覺得成語電腦適性測驗的整體版面設計很美觀。	80 (28.2%)	81 (28.5%)	102 (35.9%)	10 (3.5%)	11 (3.9%)	3.74 (1.03)
Q6	我覺得成語電腦適性測驗的題目文字清楚明確。	126 (44.4%)	86 (30.3%)	57 (20.1%)	9 (3.2%)	6 (2.1%)	4.12 (.98)
Q7	我覺得成語電腦適性測驗用滑鼠點選答案很方便。	162 (57.0%)	68 (23.9%)	40 (14.1%)	8 (2.8%)	6 (2.1%)	4.31 (.96)
Q8	我覺得成語電腦適性測驗的題目動畫很搭配。	88 (31.0%)	76 (26.8%)	94 (33.1%)	17 (6.0%)	9 (3.2%)	3.76 (1.06)
Q9	我做完成語電腦適性測驗後的感覺很新奇。	95 (33.5%)	78 (27.5%)	83 (29.2%)	18 (6.3%)	10 (3.5%)	3.81 (1.08)
Q10	如果有其它類似的電腦適性測驗我會很想試試看。	116 (40.8%)	64 (22.5%)	82 (28.9%)	8 (2.8%)	14 (4.9%)	3.92 (1.12)

在本研究的正式施測時，計有 284 位國小五、六年級的學生，上網連結成語電腦適性的 Flash 測驗，回收的問卷也為 284 份，回答率為 100%。

本問卷量表採五點量表，換算各題平均數為 3 分，若此問題層面的平均超過 3 分者，代表使用者滿意程度良好，反之，若此問題層面的平均低於 3 分者，則代表使用者滿意程度較不佳。茲就十題問卷題目分析如下：

### 一、 Q1：我比較喜歡做成語電腦適性測驗。

在此問題層面裡，其所得的平均數為 4.04，高於量表平均的 3 分，且同意的人有 73 人，非常同意的人有 119 人，同意的人數共 192 人，約佔全部的 67.7%，接近七成；而不同意的人有 8 人，非常不同意的人有 4 人，不同意的人數共 12 人，只佔 2.37%，代表絕大部份的學生都喜歡做成語電腦適性測驗，滿意程度很高。

### 二、 Q2：我比較不喜歡做成語紙筆測驗試卷。

在此問題層面裡，其所得的平均數為 3.03，只略高於量表平均的 3 分，且同意的 47 人與不同意的 33 人，人數也都各佔 16.5%、11.6%，代表此一問題層面，學生偏向於中間地帶，對於傳統的紙筆測驗模式，有人喜歡，也有人不喜歡，人數約佔各半，沒有特別偏向哪一邊。

### 三、 Q3：我覺得做成語電腦適性測驗很有趣。

在此問題層面裡，其所得的平均數為 3.99，高於量表平均的 3 分，且同意的人有 87 人，非常同意的人有 113 人，同意的人數共 200 人，約佔全部的 70.4%；而不同意的人有 17 人，非常不同意的人有 7 人，不同意的人數共 24 人，只佔 8.5%，代表絕大部份的學生都覺得做成語電腦適性測驗很有趣，滿意程度很高。

### 四、 Q4：我認為成語電腦適性測驗的歡迎動畫很吸引我。

在此問題層面裡，其所得的平均數為 3.66，高於量表平均的 3 分，且同意的人有 60 人，非常同意的人有 91 人，同意的人數共 151 人，約佔全部的 53.1%；而不同意的人有 22 人，非常不同意的人有 16 人，不同意的人數共 38 人，僅佔 13.3%，代表有一半以上的學生對於成語電腦適性測驗的歡迎動畫，滿意程度高。

## 五、 Q5：我覺得成語電腦適性測驗的整體版面設計很美觀。

在此問題層面裡，其所得的平均數為 3.74，高於量表平均的 3 分，且同意的人有 81 人，非常同意的人有 80 人，同意的人數共 161 人，約佔全部的 56.7%；而不同意的人有 10 人，非常不同意的人有 11 人，不同意的人數共 21 人，僅佔 7.4%，代表一半以上的學生對於成語電腦適性測驗整體版面的設計，滿意程度是偏高的。

## 六、 Q6：我覺得成語電腦適性測驗的題目文字清楚明確。

在此問題層面裡，其所得的平均數為 4.12，高於量表平均的 3 分，且同意的人有 86 人，非常同意的人有 126 人，同意的人數共 212 人，約佔全部的 74.7%；不同意的人有 9 人，非常不同意的有 6 人，不同意的人數共 15 人，僅佔 5.3%，代表絕大部份的學生都認為成語電腦適性測驗的題目文字清楚明確，滿意程度很高。

## 七、 Q7：我覺得成語電腦適性測驗用滑鼠點選答案很方便。

在此問題層面裡，其所得的平均數為 4.31，高於量表平均的 3 分，且同意的人有 68 人，非常同意的人有 162 人，同意的人數約佔全部的 80.9%，接近九成；不同意的人有 8 人，非常不同意的人有 6 人，不同意的人數共 14 人，僅佔 4.9%，代表幾近全部的學生認為成語電腦適性測驗用滑鼠點選答案很方便，滿意的程度非常好。

## 八、 Q8：我覺得成語電腦適性測驗的題目動畫很搭配。

在此問題層面裡，其所得的平均數為 3.76，高於量表平均的 3 分，且同意的人有 76 人，非常同意的人有 88 人，同意的人數共 164 人，約佔全部的 57.8%；不同意的人有 17 人，非常不同意的人有 9 人，不同意的人數共 26 人，僅佔 9.2%，代表有一半以上的學生都認為成語電腦適性測驗中的動動畫跟題目很搭配，滿意程度高。

## 九、 Q9：我做完成語電腦適性測驗後的感覺很新奇。

在此問題層面裡，其所得的平均數為 3.81，高於量表平均的 3 分，且同意的人有 78 人，非常同意的人有 95 人，同意的人數共 174 人，約佔全部的 61.0%；而不同意的人有 18 人，非常不同意的人有 10 人，不同意的人數共 28 人，約佔 9.8%，代表絕大部份的學生對於做成語電腦適性測驗是感到很新奇，滿意程度很高的。

## 十、 Q10：如果有其它類似的電腦適性測驗我會很想試觀看。

在此問題層面裡，其所得的平均數為 3.92，高於量表平均的 3 分，且同意的人有 64 人，非常同意的人有 116 人，同意的人數共 180 人，約佔全部的 63.3%；而不同意的人有 8 人，非常不同意的人有 14 人，不同意的人數共 22 人，僅佔 7.7%，代表絕大部份的學生喜歡這樣的測驗方式，對於若有其它類似的電腦適性測驗，他們想再繼續參與的意願很高。

由以上的分析結果，各問題層面的滿意程度，其平均皆高於 3 分，顯示本研究所發展的「成語電腦適性」，在學生的使用感受上，滿意程度非常良好。

## 第五章 結論與建議

本研究旨在發展一個「成語電腦適性測驗」，以提供教師一個在網際網路環境下的輔助教學評量，並能針對學生提供量身訂做的適性測驗。

首先在理論驗證方面，本研究探討電腦測驗、金字塔適性測驗、試題品質及實際驗證建置題庫等相關層面；其次，開發「成語電腦適性測驗」作為本研究的主要研究工具；最後在測驗評估方面，本研究評核測驗的運作效能及使用者的滿意度評估等向度。

綜上所述，本研究乃是結合編擬題庫、工具發展、效能及滿意度評估的研究。茲說明本研究的「主要發現」與「結論」，並根據研究結論提出「建議」。最後，再說明本研究之「研究限制」，提供教學及研究人員參考。

### 第一節 主要發現

如前言所述，本研究乃是結合編擬題庫、工具發展、效能及滿意度評估的研究。茲分述主要成果如下：

#### 壹、在編擬題庫方面

在編擬題庫方面，本研究是屬於電腦測驗的第三代－連續性測驗。連續性測驗是根據受試者前一題的表現來決定往後的測驗流程。測驗內容則是依據「金字塔式」分支結構的適性測驗。金字塔式測驗具有充分運用貝氏策略、便於新增鑑別力及猜測參數和適用於團體測驗的優點。最後研究者再利用 Tester 2.0 軟體，分析出各題目之難度、鑑別度，依照金字塔式的分支排列成「成語電腦適性測驗」之題庫。

## 貳、在工具發展方面

研究者將「成語電腦適性測驗」開發過程分為「測驗規劃」、「測驗設計」、「測驗建置」、「測驗運行與支援」等四個階段，茲依序說明四個階段的研究成果。

### 一、測驗規劃

在測驗規劃方面，研究者首先從文獻探討的方式提出「成語電腦適性測驗」之架構初稿，並透過成語題目的預試，篩選修改成正式施測的題庫，並徵詢專家學者的意見，確立本研究之可行性。同時，也釐清了本研究的本質及界限。本研究發展的測驗主要是提供學生一套量身訂做的適性測驗，以國小五、六年級為適用對象，測驗的學科範圍為成語部分，以結合了音效、影像、動畫，且具備了高品質、體積小、互動性高等優點的 Flash，作為製作題庫測驗的施測介面。

### 二、測驗設計

「測驗設計」主要的效能是提高資訊品質與工作效率，本研究所設計的「成語電腦適性測驗」，會以能達到更準確地評估學生成語能力並節省教師施測及批改試卷的時間為主要考量方向。因此本研究提出「成語電腦適性測驗」之資料流程圖，可知本測驗主要的使用對象為學生、教師及專家學者。學生經由此測驗得到適當的結果；教師主要提供產生適性測驗題庫之試題；專家學者則是提供此測驗建置及運作時的考量及評估。

### 三、測驗建置

測驗建置包含了「測驗發展」和「測驗測試」兩個部份。其中，測驗發展部分，本研究發展的「成語電腦適性測驗」，硬體方面，研究者預設輸出畫面為 800×600，其實使用者的輸出畫面解析度可以不拘，因 Flash 為向量圖形，具有不失真的特性；軟體方面，作業系統為 Windows XP，建構施測介面則是使用 Macromedia Flash 8 來撰寫具有多項優點及高度整合能力的 Flash 程式，最後則是利用 Tester 2.0 來進行適性題目分析與排列。

在測驗測試方面，首先都是由研究者親自修改及執行，來測試整個成語電腦適性測驗的正確性。再經由專家學者的修正建議直到完成。在開放線上測驗驗收階段，有請到台東大學教學科技碩士班一年級的學生進行施測驗證無誤後，便確立了測驗建置的完成性。

#### 四、 測驗運行與支援

測驗運行時之「軟硬體之規格需求」，硬體方面，在研究者端的規格如下：(一) CPU：AMD Athlon(tm) 64 Processor 3200+ (二) RAM：1.00 GB (三) HDD：WDC WD2500KS-00MJB0 250 GB (四) 網路卡：10/100 Mbps。在使用者端，則為：(一) CPU：Intel Pentium II 350 MHz (含) 以上 (二) RAM：128 MB (含) 以上 (三) HDD：2 GB (含) 以上 (四) 網路卡：10/100 Mbps (五) 顯示卡：可顯示 256 色以上，及解析度至少達 800×600 (含) 以上。

至於在軟體部分，在研究者端：計有 (一) Macromedia Flash 8 (二) Tester 2.0 (三) 作業系統：Microsoft Windows XP (四) 成語電腦適性測驗：研究者利用 Macromedia Flash 8、Tester 2.0 等程式，建置在網際網路的環境，讓使用者連結下載，完成線上施測。而使用者端由於網際網路為開放性架構，因此本研究所建置的成語電腦適性測驗其使用者端之軟體，使用者可以自行依其習慣偏好調整以配合不同的工作平臺，以下為本研究對於使用者端建置之軟體部份。(一) 作業系統：Microsoft Windows 95/98/ME/2000/XP (二) 全球資訊網瀏覽器：Microsoft Internet Explorer 4.0 (含) 以上。

#### 參、 在效能及滿意度評估方面

在效能及滿意度評估方面，旨在說明本研究所發展的「成語電腦適性測驗」的使用成效，以下依「使用結果與傳統測驗的相關性」、「使用者評估」等二個方面說明：

##### 一、 使用結果與傳統測驗的相關性

正式施測，研究者抽取五所學校，共 284 名五、六年級的學生，進行成語電腦適性測驗及傳統紙筆測驗。利用 Tester 2.0 測驗出來的結果，每間學校五、六年級學生所分析出來的數據，在信度方面，除了後庄國小五年級學生的 .62 頗接近 .70 外，其餘的九班學生的測驗信皆是高於 .70，因此顯示本題庫測驗具有很高

的信度。在效度方面，五間學校所分析出來的差異係數全都介於 .40 與 .60 之間，代表此題庫測驗具有非常的效度。而在相關性方面，有四個班級的數據是頗接近 .50，其餘六個班級皆大於 .50，由此可知，本研究所發展的「成語電腦適性測驗」與「傳統紙筆測驗」的施測結果，呈現高度正相關。

## 二、 使用者評估方面

在使用者做完「成語電腦適性測驗」與「傳統紙筆測驗」後，會讓使用者填答有關「成語電腦適性測驗」使用滿意度的問卷，計有 284 位學生填答問卷，回收問卷數也為 284 份，回答率為 100%。研究主要發現在整體滿意度方面平均值皆大於 3，亦即表示使用者對於本測驗是抱持著滿意的態度，其中「我覺得成語電腦適性測驗用滑鼠點選答案很方便。」選項的滿意程度最高 (M=4.31)。

## 第二節 結論

根據上述主要發現，本研究結論如下：

### 壹、 題庫建置方面

以「Tester 2.0」軟體的試題分析及「金字塔式」適性測驗，說明如下：

#### 一、 利用 Tester 2.0 軟體進行試題分析具便利性

研究者利用 Tester for Windows 程式 2.0 版，進行試題預試的分析，只要鍵入學生作答的資料，Tester 2.0 軟體即能算出問題注意係數表、內部一致性係數、差異係數及各試題之難度、鑑別度等一些重要的分析訊息，就連學生作答後的分數，也不需研究者一一批改，Tester 2.0 軟體自動能將各學生的作答資料跟正確答案一一比對算出學生所作答後的分數。

在正式施測後的試題分析，研究者同樣採用 Tester 2.0 軟體，所不同的是，在此次的分析中，研究者加入效標關聯的項目，主要是用來作「成語電腦適性測驗」與「傳統紙筆測驗」相關性的分析，因此從 Tester 2.0 軟體分析的報表中，即能很清楚地了解到「成語電腦適性測驗」與「傳統紙筆測驗」在各校的五、六年級學生作答上，均呈現高度的正相關。

## 二、 利用金字塔式適性測驗便於達成學生能力的估計

適性測驗中最常用且最受重視的一種彈性的測驗方式就屬金字塔適性測驗（陳新豐，2002）。一份包括 55 題的測驗內容，每位受試者只需要作答 10 題，即可達成學生能力的估算（吳裕益，1997）。而本研究所發展的「成語電腦適性測驗」內容包括 28 題的測驗內容，每位受試者只需要作答 7 題，即能得出學生能力的估計。

## 貳、 工具研發方面

在工具研發方面以「整合」與「功能」二部份，說明整理如下：

### 一、 Flash 施測介面與金字塔適性測驗題庫內容可整合為一

目前測驗的發展，已逐漸邁向與電腦化的結合，許多電腦化測驗的研究，更是如雨後春筍般不勝枚舉。可是，金字塔適性測驗與 Flash 介面所整合的題庫測驗，目前並不多見。本研究首先嘗試將金字塔適性測驗與 Flash 介面整合在一起，就「成語」這個部分，設計相關試題，統合成一個結合影音、互動性高，又能以設計好的適性試題，即時達到學生成語能力估計的 Flash 題庫測驗。

### 二、 成語電腦適性測驗具有多項功能

本研究所發展的「成語電腦適性測驗」，結合網際網路，可提供受試者隨時上網作答，不受時空的限制，作答完成後，電腦會自動產生回饋訊息給受試者。因此，本測驗的功能有測驗結構立體化（金字塔式分支）、作答線上化及結果自動化。

## 參、效能及滿意度評估方面

在滿意度評估方面，分為「估算效能」及「滿意度」等二部份，說明整理如下：

### 一、 成語電腦適性測驗能力估算效能良好

傳統的紙筆測驗仍有其測驗效力的優勢存在，本研究所發展的「成語電腦適性測驗」所估計的效力若能與傳統紙筆測驗所估計的等同，代表本研究發展的測驗是具有測驗效度的。除了利用 **Tester 2.0** 軟體，分析出各校五、六年級的「成語電腦適性測驗」與「傳統紙筆測驗」的分數均呈現高度正相關外，同時研究者也將五校共 284 名學生的「成語電腦適性測驗」與「傳統紙筆測驗」的分數，利用 **Spss 12.0** 作 Pearson 積差相關分析，也達數據 .57 的高度正相關。由此可知，本研究所發展的「成語電腦適性測驗」是可行的，不僅測驗效力與傳統紙筆測驗呈正相關，而且測驗題數、測驗時間及測驗批改所花費的精力都要比傳統紙筆測驗節省很多，受試者不必做全部的試題，即能在短短做數題的時間裡，得到適性的能力估計，施測者也不需再花額外的時間，做每份試卷的事後批改工作，因此本題庫測驗的發展是具有實質上的效果的。

### 二、 使用者對成語電腦適性測驗感到滿意

本測驗經使用者使用後，在整滿意度方面平均值皆大於 3，表示使用者對於本測驗抱持著滿意的態度。其中「我覺得成語電腦適性測驗用滑鼠點選答案很方便」(M=4.31) 滿意度最高，接下來分別是「我覺得成語電腦適性測驗的題目文字清楚明確」(M=4.12)、「我比較喜歡做成語電腦適性測驗」(M=4.04)、「我覺得做成語電腦適性測驗很有趣」(M=3.99)、「如果有其它類似的電腦適性測驗我會很想試試看」(M=3.92)、「我做完成語電腦適性測驗後的感覺很新奇」(M=3.81)、「我覺得成語電腦適性測驗的題目動畫很搭配」(M=3.76)、「我覺得成語電腦適性測驗的整體版面設計很美觀」(M=3.74)、「我認為成語電腦適性測驗的歡迎動畫很吸引我」(M=3.66)、「我比較不喜歡做成語紙筆測驗試卷」(M=3.30)。

## 第三節 建議

茲依據前述主要成果和結論提出建議，並分為三方面說明如下：

### 壹、對建置題庫的建議

針對建置題庫，主要以豐富題庫內容方向為主，建議如下：

#### 一、可多參考各家版本之語文教科書

目前研究者只針對一家版本（翰林版）的國小語文教科書做成語部分的整理，或許整理出來的成語內容，對於使用此版本的學校較能涵蓋到，而對於使用別家版本的學校，可能就會有疏漏涵蓋不全的情況產生。不過研究者當初在設計成語題庫時，除了參考翰林版的語文教科書外，也有收集坊間的成語題庫評量，應能使此次所發展的「成語電腦適性測驗」內容較不失偏頗。因此，若有後續的研究發展，想讓題庫內容能更加全面化的話，各家版本的語文教科書，如：南一、康軒等，均可再做更深入的整理探討。

#### 二、擴增題庫內容

本研究發展的「成語電腦適性測驗」之題庫內容是由當初預試的 50 題，分析再修正為 28 題，經由金字塔分支結構排列後，於正式施測時，學生只需作答 7 題，即能達成成語能力的估計。若想使學生能力的估計達到更精準，建議可以在題數的方面作擴增，金字塔分支越細，學生能力的估算也能更加準確。

#### 三、題庫內容以隨機呈現

本研究所發展的「成語電腦適性測驗」之題目，每個能力的估算都僅有一個題目為代表，倘若在日後的研究上，能在每個能力的估算方面都能有多題為代表，題目的內容就能以亂數抓取的方式來呈現，不僅能讓此測驗更加豐富多元，也不易有固定制式的呆板印象。

## 貳、 對研發工具的建議

對於研發工具方面的建議，以下有二點建議：

### 一、 可採用 PHP 和 MySQL 等作為開發測驗的工具

本研究所發展的「成語電腦適性測驗」是以 Flash 作為開發測驗的工具。若想把此測驗功能作更有效率的提升與控管，可採用 PHP 來製作動態網頁並以 MySQL 作資料庫管理的動作，這兩種軟體不僅具有免費的優點外，也能讓題庫測驗的功能更趨完備。

### 二、 結合各類專業人士共同參與開發測驗

本研究所發展的「成語電腦適性測驗」，研究者從構思架構與功能，到實際擔任發展人員、測試人員、美工人員、施測人員及評估人員等角色，加上題庫內容的增建，也需要更多機制來掌控試題的品質，深感研究時間之不足。因此，研究者建議未來在開發類似之測驗時，應結合各類專業人士的共同參與，尋求各方面技術層面上的支持，如測驗開發者，著重於程式介面的設計與功能的發展；美工人員美化版面，加強受試者參與作答之動機；測驗人員專注於題庫方面的建置，強化試題的代表性，如此，應能使測驗的發展與成果更加地成功與完善。

### 三、 Flash 動畫輔助測驗，提昇學生成語能力

測驗的目的，不在於考倒學生，或是只是測出學生現階段的能力而已，而是學生能否經由測驗，學習到更多不同的知識，轉而再次提昇本身的能力。

本研究所發展的測驗，其中的 Flash 動畫，只是與題目略為相關，並作為版面美化之效用。若想藉由動畫的輔助來引領學生達到成語能力的提昇，研究者建議在之後的研究發展上，Flash 動畫的設計應可再做動畫內容的設計與延伸，如此一來，學生可藉由動畫的引領，達到望「圖」生義，明瞭每一成語的內容，這樣的測驗也才會更完整，更富有意義。

## 參、對未來研究的建議

對於未來研究方面的建議，二點建議如下：

### 一、電腦化適性測驗

為因應科技進步，教學評量與電腦科技化整合的趨勢，已是未來一股研究潮流。Bunderson、Inouvy 和 Olsen（1989）便曾指出電腦發展的第四代—電腦化適性測驗，為現今最受重視與討論的電腦測驗（陳新豐，2002）。

電腦化適性測驗是結合試題反應理論，兩者相輔相成應運而生。要產生具備有試題反應理論的題庫，則有下列特點（王寶壟，1995；吳裕益，1997；余民寧，1994；陳麗如，1999；陳新豐，2002；Van der Linden & Glas，2000）：

- （一）題庫的試題參數應包括難度、鑑別度及猜測度等三個參數
- （二）難度參數應能涵蓋所有受試的能力
- （三）鑑別度參數應在 0.8 以上
- （四）題庫的題數至少要在 100 題以上
- （五）對於題庫中之試題所測的能力或特質要相當程度的確定

電腦化適性測驗的產生，使得測驗能更配合考生的程度、減少測量的誤差而能更精準地估算受試者之能力，無疑是未來此方面未來研究發展的走向。

### 二、應用於不同學科上

除了上述讓測驗的發展走向電腦化適性測驗及更注重試題的安全性外，研究者也建議在未來更多元的發展上，可以讓此測驗與其它更多不同的學科相結合。本研究所發展的測驗，是以與「成語」部分結合為主，未來也可以應用在其它例如：國語、英語、數學及社會等不同學科上，便得以了解整合測驗在不同學科的應用情形和效能。

## 第四節 研究限制

本研究的研究限制，主要在測驗發展及施測等相關議題方面，說明如下：

首先，是有關題庫測驗的安全問題。由於本研究的題庫測驗是建置在網路上面，答題者只要連線便能進入測驗，若答題者有心想存取題目內容，只需利用重複試答的方式，便能將題目一一找出，再刻意背誦或是利用抄寫就能將題目記錄下來，而本測驗並無安全性的配套措施來禁止答題者不適當行為，因此題庫測驗之安全性確受威脅。所以在未來若要將此成語電腦適性測驗題庫測驗繼續擴大推廣，應做好題庫測驗安全性的考量，例如：答題者的身份確認或是測驗設定只能施測一次等。不過，因為本研究的施測，只在施測當日才在班級上公布網址，故不會有提早竊取試題的可能，且研究者全程於電腦教室監考，不會讓答題者有重複試答的機會，所以本研究題庫測驗的安全性威脅度較低。

再來是城鄉學校的數位落差問題。研究者所取樣的學校裡，有屬於都市型的學校，也有屬於鄉村型的學校，在此二種不同類型的學校班級裡施測時，研究者可以感受到明顯數位落差的存在。在都市型學校班級施測時，研究者一公布測驗連結網址，學生幾能自行輸入來連結測驗進行施測，只是有些速度快些、有些慢些。而在鄉村型學校班級施測時，一旦公布測驗連結網址後，很多學生連鍵盤數字、英文字母及符號的位置都搞不太清楚，因此需要較多的協助與說明。而研究者施測時，該班的電腦老師都會在場給予幫忙，所以這方面的問題，得以快速解決，不致於耽誤太多施測的時間。不過，關於城鄉數位落差的層面，仍有待政府發展更多的教育配套措施，來縮短其差異性。

在「成語」及「電腦測驗」這二個區塊裡，研究者試著將這二部份作結合，並將自己本身發展題庫測驗與施測的相關經驗，整理於此論文中，希冀後來若有陸續的相關研究，本研究也能提供未來發展之參考，期待日後有更健全完整的題庫測驗產生，來造福更多的教師、學生、研究的專家學者及一切會使用到此題庫測驗的相關人員！

最後，依據本研究目的，綜合本研究後的成果與結論，重要性如下：

- (一) 結合金字塔式適性測驗，發展更符合學生能力的題庫內容。
- (二) 開發一個適性測驗的施測環境，提供受試者立即的結果報告。
- (三) 測驗的測驗效力與滿意度皆高，可作為日後研究發展之參考。

# 參考文獻

## 壹、中文部份

- IRT 台灣題測中心。IRT 電腦適性測驗介紹。2007 年 7 月 16 日，取自：  
<http://www.irt.org.tw/index.php?mod=01>。
- 方金雅、陳新豐、黃秀霜等人（2004）。語文領域之電腦化測驗與評量。課程與教學，3，17-31。
- 王月鳳（2004）。國民小學本國語文審定本成語內容分析之研究。國立新竹教育大學語文教學碩士班碩士論文，未出版，新竹。
- 王建華（2000）。國語文教學的基本概念與指導。人文及社會學科教學通訊，4，217-230。
- 王寶墉（1995）。現代測驗理論。臺北：心理出版社。
- 史遷（2006）。猜猜謎學成語。臺北：文房文化。
- 朱孝利（1999）。國小高年級國語科閱讀指導之教學研究。臺北：文景。
- 朱麗芳（2007）。成語語法功能和運用的初步考察。現代語文，9，47-48。
- 江惜美（1994）。成語故事教學指導。國語文教育通訊，8，31-35。
- 何永清（2005）。成語的語法與修辭及其教學探究。台北市立師範學院學報，1，1-24。
- 余民寧（1991）。試題反應理論的介紹（一）—測驗理論的發展趨勢。研習資訊，6，13-18。
- 余民寧（1992）。次序性資料的內容效度係數和同質性信度係數之計算。測驗年刊，40，199-214。
- 余民寧（1994）。試題反應理論的介紹（11）—題庫的建立。研習資訊，4，11-16。
- 余民寧（2004）。教育測驗與評量。臺北：心理出版社。

- 余桂林 (2001)。《現代漢語詞典》四字成語的意義及釋義特徵。萍鄉高等專科學校學報，3，88-90。
- 何榮桂 (1997)。從「測驗電腦化與電腦化測驗」再看網路化測驗。測驗與輔導，144，2972-2974。
- 吳裕益 (1996)。題目反應理論。上課講義 (未發表)。
- 吳霞雲 (主編) (1999)。新編成語辨正。臺北：武陵。
- 呂杰 (2007)。漢語成語的修辭功能。現代語文，1，73-74。
- 宋麗麗、曹延傑 (1990)。結構化分析。臺北：資訊工業策進會。
- 李茂能 (2001)。電腦化適性測驗的過去、現在與未來。臺灣教育，604，52-61。
- 李書瑩 (2007)。我的第一本成語故事。臺北：三采文化。
- 卓惠美 (2004)。談「全語文」作文教學上的應用及策略。南投文教，20，75-77。
- 幸曼玲 (2003)。語文教學的困境與省思。教師天地，26，61-64。
- 林妙香 (2000)。入學測驗簡易化之目標與實際：以北區高中入學測驗為例。國家科學委員會研究彙刊，2，135-147。
- 柯光霖 (2006)。網路多元化教學評量與適性學習研究。義守大學資訊管理研究所碩士論文，未出版，臺北。
- 孫世民 (1994)。談成語教學。中國語文，3，63-65。
- 徐明珠 (2008)。成語之應用與寫作技巧。國政評論，教文 (評) 097-031 號。
- 徐康 (2005)。漢語成語和英語成語的異同。US-China Foreign Language，3，5-8。
- 涂淑遠 (2006)。國小高年級國語教科書中成語之內容分析及教學研究。國立花蓮教育大學語文科教學碩士班碩士論文，未出版，花蓮。
- 高啟洲 (2005)。互動式學習系統之設計。南大學報，1，111-132。
- 張心怡 (2003)。英語線上適性測驗於企業徵才之探討。義守大學資訊管理學系碩士班碩士論文，未出版，高雄。
- 張月娟 (2007)。寫作教學七八九—好用的策略介紹。國民教育，3，49-56。

- 張巧玲（2004）。國小國語文教學現況及其改進的策略。中國語文，561，38-43。
- 張君松（2005）。泰國學生學習中文成語的困難及教學補救策略研究。國立臺灣師範大學華語文教學研究所碩士論文，未出版，臺北。
- 敖幼祥（2006）。敖幼祥的漫畫中國成語。臺北：時報文化。
- 教育部（2003）。國語文領域課程綱要。2007年10月6日，取自：  
<http://itseed.trps.cyc.edu.tw/resource/file/256.doc>。
- 陳季佳（主編）（2006）。國語 1-12 冊。臺南：翰林。
- 陳桂成（2005）。成語界說與成語詞典立目。辭書研究，1，23-27。
- 陳新豐（2002）。線上題庫與適性測驗整合系統之發展研究。國立政治大學教育學系教育心理與輔導組博士論文，未出版，臺北。
- 陳新豐（2005）。傳統紙筆測驗與線上電腦化測驗試題參數估計差異之比較。教育研究與發展期刊，3，123-145。
- 陳麗如（1998）。電腦化適性測驗之題庫品質管理策略。台灣師大資訊教育研究所碩士論文，未出版，臺北。
- 賀非雨（主編）（2006）。成語 300 首。合肥：安徽少年兒童出版社。
- 董國峰（1990）。成語故事解說。臺北：作文出版社。
- 葉興華（1999）。國語科教學之檢討與改進。課程與教學季刊，4，119-134。
- 維基百科。成語和成語的使用。2008年7月2日，取自：  
<http://zh.wikipedia.org/w/index.php>。
- 劉淑貞（2003）。成語教學方案對自閉症學生閱讀及溝通能力增進之研究。國立臺東大學教育研究所碩士論文，未出版，臺東。
- 蔡榮昌（2006）。從讀寫結合談仿寫。屏縣教育季刊，27，5-19。
- 鄭培秀（2005）。成語句法分析及教學策略研究。國立中山大學中國文學系碩士論文，未出版，高雄。
- 鄭博真（2003）。為國語文教學尋找創新之路。人文及社會學科教學通訊，5，29。

學力 e 試網。現代測驗 IRT。2008 年 1 月 5 日，取自：<http://www.edutest.com.tw/>。

蕭嫻慈（2002）。關於學生成語練習的策略。中國語文，1，71-73。

顏崑陽（主編）（1992）。故鄉實用成語辭典。臺北：牛頓出版。

饒瑞文（1993）。系統開發導論。臺北：松岡。



## 貳、西文部份

- Andersen, E. B. (1973). *Conditional inference and models for measuring*. Copenhagen: Mentalhygiejnisk Forlag.
- Andersen, E. B. (1980). *Discrete statistical models with social science applications*. Amsterdam: North-Holland.
- Baker, F. B. (1985). *The basics of item response theory*. Portsmouth, NH: Heinemann.
- Baker, F.B. (1977). Advances in item analysis. *Review of Educational Research*, 47,151-178.
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37, 29-51.
- Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika*, 46, 443-459.
- Bunderson, C. V., Inouye D. I., & Olsen, J. B.(1989). The four generations of computerized educational measurement. In R. L. Linn(ED.) *Educational Measurement*(3<sup>rd</sup> ed., pp.360-407). New York: American Council on Education.
- Cohen, R. j., Montague, P., Nathanson, L. S., & Swerdlik, M. E. (1988). *Psychological testing: An introduction to tests and measurement*. Mountain View, CA: Mayfield.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Embretson, S. E. (Ed.) (1985). *Test design: Developments in psychology and psychometrics*. Orlando, FL: Academic.
- Freedle, R. (Ed.) (1990). *Artificial intelligence and the future of testing*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Guilford, J. P. (1954). *Psychometric methods*. New York: McGraw-Hill.

- Guion, R. M., & Ironson, G. H. (1983). Latent trait theory for organizational research. *Organizational Behavior and Human Performance*, 31, 54-87.
- Gulliksen, H. (1987) . *Theory of mental test*. Hillsdale, NJ : Lawrence Erlbaum Associates. ( Originally published in 1950 by New York : John Wiley & Sons )
- Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, 9,139-150.
- Hambleton, R. K. & Swaminathan, H. (1985) . *Item response theory : principles and applications*. Boston : Kluwer Nijhoff Publishing.
- Hambleton, R. K. (1989). *Principles and selected applications of item response theory*. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 147-200). New York: Macmillan.
- Hambleton, R. K. (Ed.) (1983). *Applications of item response theory*. Vancouver, BC: Educational Research Institute of British Columbia.
- Hambleton, R. K., & Cook, L. L. (1977). Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement*, 14, 75-96.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: SAGE.
- Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Application to psychological measurement*. Homewood, IL: Dow Jones-Irwin.
- Linn, R. L. (Ed.) (1989). *Educational measurement (3rd ed.)* . New York: Macmillan.
- Lord, F. M. (1952). A theory of test scores. *Psychometric Monograph*, 7.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbawn Associates.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.

- Mike Cole (1998). *Cambridge Journal of Education*. Volume 28, 2, 235 - 238.
- Mislery, R. J., & Bock, R. D. (1982). *BILOG: Maximum likelihood item analysis and test scoring with logistic models for binary items*. Chicago: International Educational Services.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: The Danish Institute for Educational Research.
- Richardson, M. W. (1936). The relationship between difficulty and the differential validity of a test. *Psychometrika*, 1, 33-49.
- Saljo, R (1998). *Thinking with and through artifacts in Faulkner, D, Littleton, K & Woodhead, M Learning relationships in the classroom*, 57.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 17.
- Sands, William A., Waters, Brian K. & McBride James R. (1997) . *Computerized Adaptive Testing : From Inquiry to Operation*. Washington, DC : American Psychological Association.
- Shelly Gray B. Cashman Thomas J. & Rosenblatt Harry J. (2001). *System Analysis and Design*. Thomson Learning, Inc.
- Suen, H. K. (1990). *Principles of test theories*. Hillsdale, NJ: Lawrence ErlbaumAssociates.
- Terman, L. M. (1916). *The measurement of intelligence*. Boston, MA: Houghton Mifflin.
- Thurstone, L. L. (1929). Theory of attitude measurement. *Psychological Bulletin*, 36, 222-241.
- Tucker, L. R. (1946). Maximum validity of a test with equivalent items. *Psychometrika*, 11, 1-13.
- Vale, C. D. (1981) . Design and implementation of a macrocomputer-based adaptive testing system. *Behavior Research Methods and Instrumentation*, 13, 399-406.
- Vale, C. D. (1982) . *Design of a macrocomputer-based adaptive testing system*.

Proceeding of the 1982 computerized Adaptive Testing Conference, Department of Psychology, University of Minnesota.

Van der Linden W. J. & Glas, C. A. W. (2000) . *Computerized adaptive testing : theory and practice*. Dordrecht; Boston : Kluwer Academic.

Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in Society: The Development of Higher Psychological Process*. Cambridge, MA: Harvard University Press.

Wainer, H. et al. (1990). *Computerized adaptive testing: A primer*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wainer, H., & Braun, H. I. (Ed.) (1988). *Test validity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wainer, H., & Messick, S. (Ed.) (1983). *Principals of modern psychological measurement: A Festschrift for Frederic M. Lord*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Warm, T. A. (1978) . *A primer of item response theory*. Springfield, VA : National Technical Information Services.

Weiss, D. J. (Ed.) (1980). *Proceedings of the 1979 computerized adaptive testing conference*. Minneapolis: University of Minnesota.

Weiss, D. J. (Ed.) (1983). *New horizons in testing: Latent trait test theory and computerized adaptive testing*. New York: Academic.

Wright, B. D., & Panchapakesan, N. (1969). A procedure for sample-free item analysis. *Educational and Psychological Measurement*, 29, 23-48.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA.

# 附錄一

## 成語預試題目

### 國小學童成語能力測驗

\_\_\_\_年\_\_\_\_班\_\_\_\_號

性別：\_\_\_\_

姓名：\_\_\_\_\_

#### 選擇題（每題 2 分，共 100 分）

- |  |  |
|--|--|
| 1 ( ) 光明____落①累②雷③疊④磊                      | 1 2 ( ) 形容一無所有了①吞雲吐霧<br>②吞吞吐吐③別無長物④吞<br>聲忍氣                    |
| 2 ( ) 不____下問①齒②尺③恥④侈                      | 1 3 ( ) 形容神槍手的射擊技術高超<br>①百發百中②千軍萬馬③十<br>全十美④一五一十               |
| 3 ( ) 小心____①亦亦②翼翼③<br>意意④奕奕               | 1 4 ( ) 指事物變化很多，無法預測<br>①不解風情②變幻莫測③七<br>手八腳④難言之隱               |
| 4 ( ) 不可思____①異②義③議④意                      | 1 5 ( ) 比喻針對問題所在，採取有<br>效的措施①藥到病除②鐵口<br>直斷③妙手回春④對症下藥           |
| 5 ( ) 莫____其妙①名②冥③明④銘                      | 1 6 ( ) 比喻空談理論，不能解決實<br>際問題①空中樓閣②紙上談<br>兵③白紙黑字④千軍萬馬            |
| 6 ( ) 無微不____①致②至③置④制                      | 1 7 ( ) 比喻對事物的情況捉摸不<br>透；也形容人的心機極深①<br>高不可攀②一望無際③深不<br>可測④遙不可及 |
| 7 ( ) _____相關①習習②昔昔③<br>息息④襲襲              |  |
| 8 ( ) 不____示弱①干②甘③杆④乾                      |  |
| 9 ( ) _____步自封①故②固③雇④顧                     |  |
| 1 0 ( ) 迫不____待①急②即③疾④<br>及                |  |
| 1 1 ( ) 形容變化無窮①口若懸河②<br>口是心非③口不擇言④千變<br>萬化 |  |

- 18 ( ) 指高興得過了頭，轉而招致悲傷的事出現①苦中作樂②生不如死③悲從中來④樂極生悲
- 19 ( ) 形容一個人對事情毫不關心①不聞不問②不聲不響③不分皂白④不辨東西
- 20 ( ) 指因意見不合、言語衝突而不愉快地各自離開①不告而別②不歡而散③曲終人散④依依不捨
- 21 ( ) 比喻事物發展得迅速①快馬加鞭②刻不容緩③一刻千金④一日千里
- 22 ( ) 用來表示多餘、沒有必要的舉動①五體投地②手舞足蹈③多此一舉④隻手遮天
- 23 ( ) 指意外的災禍或事故；也特指人的死亡①七上八下②不三不四③一五一十④三長兩短
- 24 ( ) 形容非常突出①五窮六絕②數一數二③三三兩兩④九死一生
- 25 ( ) 比喻做事踏實、認真①不翼而飛②腳踏實地③一飛沖天④落井下石
- 26 ( ) 「十拿九穩」①比喻很有把握②形容既高且遠的地方③形容專心學習的樣子④形容驚嚇過度的樣子
- 27 ( ) 「約法三章」①泛指事先約好或規定的事②形容分散而不完整、不團結③指能夠達到目標的唯一方法④泛指旅途平安無阻或事情進展順利
- 28 ( ) 「目不轉睛」①形容自傲自大，瞧不起他人②形容因驚嚇或錯愕而神情呆滯的樣子③以魚眼睛混充珍珠。比喻以假亂真④形容凝神注視的樣子
- 29 ( ) 「五體投地」①比喻人神通廣大，力強可畏②形容心情起伏不定，驚惶不安③比喻花樣繁多、變化多端④比喻非常欽佩
- 30 ( ) 「冷眼旁觀」①平靜無事②心氣平和，不急躁、不發怒③形容漠不關心④形容非常謹慎，不敢疏忽
- 31 ( ) 「有目共睹」①比喻才思敏捷，談吐風雅②極為清楚明顯，大家都看得見③比喻得到眾人的讚美④一句話就能說中事情的要點或真相
- 32 ( ) 「有口皆碑」①比喻嚴守秘密，不輕易告訴別人②比喻人人稱讚、頌揚③事情的問題所在④以懇切真摯的態度，有耐心的勸告他人
- 33 ( ) 「風平浪靜」①徹底改正以往的過錯②心平氣和③水面平靜無風浪④痛恨、怨恨到極點
- 34 ( ) 「迎刃而解」①對於問題不求了解②還有餘力去處理更多的問題③問題難以釐清頭緒④解決問題很順利
- 35 ( ) 「不可開交」①沒法解開或擺脫②不聽也不問③不分是非④指一件事不可預測和估量

- 36 ( ) 這次的旅遊活動，不但使我忘卻了工作的煩惱，而且增進了同事間的情誼，真是「□□□□」。①一擲百萬②一錢不值③一舉兩得④一瀉千里
- 37 ( ) 他「□□□□」的陷害你，你卻依然待他如親兄弟，實在太不值得了。①千辛萬苦②千言萬語③千方百計④千里迢迢
- 38 ( ) 既然當初是你一時衝動造成的錯誤，而今以『□□□□』，你只有自己承擔後果了。①雞犬不寧②雙管齊下③離鄉背井④覆水難收
- 39 ( ) 求求你救救我兒子！只要你救得了他，就是替你做牛做馬我也「□□□□」。①一廂情願②心甘情願③事與願違④心灰意冷
- 40 ( ) 每年除夕夜，餐桌上總是擺滿了「□□□□」等著我們去品嚐。①如數家珍②珍奇異寶③山珍海味④珍禽異獸
- 41 ( ) 犯罪集團為了讓青少年「□□□□」地效忠，往往利用毒品控制他們，手段十分卑劣。①死心塌地②死有餘辜③九死一生④死不瞑目
- 42 ( ) 看他們夫婦兩人含情脈脈、「□□□□」的樣子，有多幸福啊！①心血來潮②心如止水③心心相印④心安理得
- 43 ( ) 百貨公司開幕，大把大把的紅包從屋頂灑下來，真是「□□□□」呀！①禍從天降②天旋地轉③異想天開④喜從天降
- 44 ( ) 你不要遇到一點挫折就「□□□□」，還有很多機會可以把握啊！①冷冷清清②心灰意冷③冷暖自知④冷嘲熱諷
- 45 ( ) 災區居民生活困頓，請大家發揮「□□□□」的精神，有錢出錢，有力出力。①雪上加霜②風花雪月③雪中送炭④冰雪聰明
- 46 ( ) 相隔數十年後，他們雖然「□□□□」，但恍如隔世。①鏡花水月②明鏡高懸③破鏡重圓④目光如鏡
- 47 ( ) 老虎「□□□□」的樣子真是嚇人！①七嘴八舌②齜牙咧嘴③油嘴滑舌④人多嘴雜
- 48 ( ) 人類因能完成夢想而偉大，但夢想的實現，絕非「□□□□」可達成。①朝不保夕②命在旦夕③朝令夕改④一朝一夕
- 49 ( ) 他在送舊晚會上「□□□□」，博得滿堂彩。①引吭高歌②引人注目③引以自豪④引以為榮
- 50 ( ) 好朋友就是一旦有難，能挺身而出，「□□□□」的人。①笑裡藏刀②借刀殺人③刀光劍影④拔刀相助



## 附錄二

### 五年級成語預試分析表

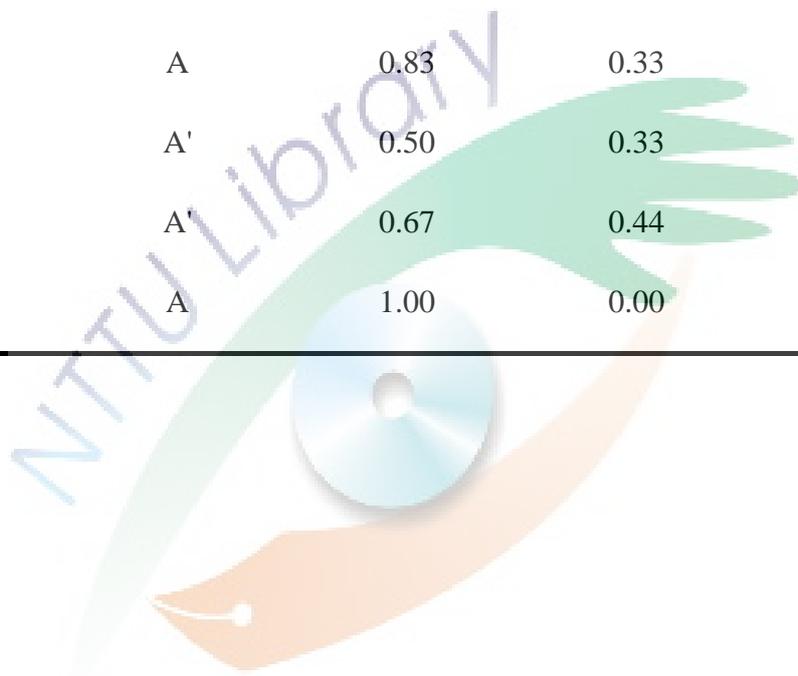
題號	判定類別	難度	鑑別度	是否刪除
1	A	0.72	0.56	
2	A'	0.61	0.33	是
3	A	1.00	0.00	是
4	B	0.17	0.33	
5	A'	0.67	0.00	是
6	B'	0.33	0.00	是
7	A'	1.00	0.00	是
8	A	0.50	0.56	
9	B'	0.56	0.22	是
10	B	0.50	0.56	
11	A	0.83	0.33	
12	A	0.83	0.33	
13	A	0.94	0.11	是
14	A	0.89	0.22	
15	A	0.56	0.89	
16	B'	0.61	0.33	是
17	A'	0.72	0.33	是

18	A	1.00	0.00	是
19	A'	0.94	0.11	是
20	A	0.61	0.78	
21	B	0.17	0.33	
22	A	0.94	0.11	是
23	A'	1.00	0.00	是
24	B	0.44	0.67	
25	A	1.00	0.00	是
26	A	0.89	0.22	
27	B'	0.22	0.00	是
28	A'	0.72	0.33	是
29	A	0.61	0.78	
30	A'	0.83	0.11	是
31	A'	0.72	0.33	是
32	B'	0.11	0.22	是
33	A'	0.61	0.56	是
34	A	0.61	0.78	
35	A	0.56	0.67	
36	A	0.89	0.22	
37	A	0.89	0.22	
38	A	0.83	0.33	
39	A	1.00	0.00	是

---

40	A	1.00	0.00	是
41	B'	0.56	0.44	是
42	A	0.89	0.22	
43	A	0.94	0.11	是
44	A	0.89	0.22	
45	A	0.61	0.78	
46	A	0.61	0.78	
47	A	0.83	0.33	
48	A'	0.50	0.33	是
49	A'	0.67	0.44	是
50	A	1.00	0.00	是

---





## 附錄三

### 六年級成語預試分析表

題號	判定類別	難度	鑑別度	是否刪除
1	A	0.65	0.50	
2	A'	0.70	0.20	是
3	A'	0.85	0.30	是
4	B	0.35	0.70	
5	A'	0.75	0.30	是
6	A	0.60	0.60	
7	A'	0.85	0.30	是
8	A'	0.55	0.30	是
9	B	0.60	0.80	
10	B'	0.35	0.10	是
11	A	0.85	0.30	
12	A	0.75	0.30	
13	A	1.00	0.00	是
14	A	0.85	0.30	
15	A	0.75	0.50	
16	A	0.50	0.60	
17	A	0.65	0.70	

---

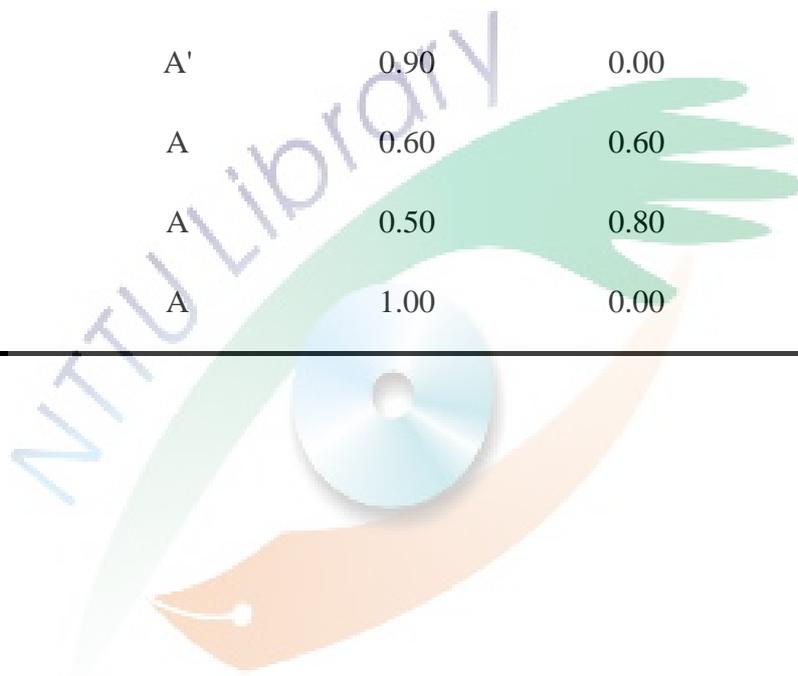
18	A	0.85	0.30	
19	A'	0.85	0.10	是
20	A	0.90	0.20	是
21	B'	0.35	0.30	是
22	A	0.95	0.10	是
23	A	0.80	0.40	
24	A	0.75	0.50	
25	A	0.85	0.30	
26	A	0.90	0.20	是
27	B	0.05	0.10	是
28	A	0.85	0.10	是
29	A	0.80	0.40	
30	A	0.70	0.60	
31	A	0.65	0.70	
32	B	0.25	0.50	
33	A'	0.65	0.30	是
34	A	0.55	0.50	
35	A'	0.55	0.10	是
36	A	0.90	0.20	是
37	A'	0.95	-0.10	是
38	A'	0.85	0.30	是
39	A'	0.90	0.20	是

---

---

40	A	0.95	0.10	是
41	A	0.60	0.40	
42	A	0.85	0.30	
43	A	0.85	0.30	
44	A	0.95	0.10	是
45	A	0.70	0.60	
46	A	0.50	0.60	
47	A'	0.90	0.00	是
48	A	0.60	0.60	
49	A	0.50	0.80	
50	A	1.00	0.00	是

---





## 附錄四

### 成語測驗

\_\_\_\_\_年\_\_\_\_\_班\_\_\_\_\_號

※剛才在「成語電腦適性測驗」中，我所得的**級分**是：\_\_\_\_\_分

※以下會出現幾題與電腦適性測驗重複的題目，請放心繼續作答。

1. ( ) 指高興得過了頭，轉而招致悲傷的事出現①苦中作樂②樂極生悲③悲從中來④生不如死
2. ( ) 百貨公司開幕，大把大把的紅包從屋頂灑下來，真是「□□□□」呀！①天旋地轉②禍從天降③喜從天降④異想天開
3. ( ) 看他們夫婦兩人含情脈脈、「□□□□」的樣子，有多幸福啊！①心血來潮②心如止水③心安理得④心心相印
4. ( ) 「五體投地」①比喻非常欽佩②比喻人神通廣大，力強可畏③形容心情起伏不定，驚惶不安④比喻花樣繁多、變化多端
5. ( ) 比喻針對問題所在，採取有效的措施①藥到病除②妙手回春③對症下藥④鐵口直斷
6. ( ) 形容非常突出①數一數二②三三兩兩③五窮六絕④九死一生
7. ( ) 「冷眼旁觀」①平靜無事②形容漠不關心③心氣平和，不急躁、不發怒④形容非常謹慎，不敢疏忽
8. ( ) 災區居民生活困頓，請大家發揮「□□□□」的精神，有錢出錢，有力出力。①雪上加霜②雪中送炭③風花雪月④冰雪聰明
9. ( ) 光明\_\_\_\_\_落①累②雷③磊④疊
10. ( ) 「有目共睹」①一句話就能說中事情的要點或真相②極為清楚明顯，大家都看得見③比喻得到眾人的讚美④比喻才思敏捷，談吐風雅
11. ( ) 比喻對事物的情況捉摸不透；也形容人的心機極深①一望無際②高不可攀③深不可測④遙不可及
12. ( ) 犯罪集團為了讓青少年「□□□□」地效忠，往往利用毒品控制他們，手段十分卑劣。①死不瞑目②死有餘辜③死心塌地④九死一生
13. ( ) 人類因能完成夢想而偉大，但夢想的實現，絕非「□□□□」可達成。①朝不保夕②朝令夕改③一朝一夕④命在旦夕
14. ( ) 無微不\_\_\_\_\_①至②致③制④置
15. ( ) \_\_\_\_\_步自封①固②故③雇④顧

16. ( )「迎刃而解」①對於問題不求了解②問題難以釐清頭緒③解決問題很順利④還有餘力去處理更多的問題
17. ( )比喻空談理論，不能解決實際問題①紙上談兵②空中樓閣③白紙黑字④千軍萬馬
18. ( )相隔數十年後，他們雖然「□□□□」，但恍如隔世。①目光如鏡②明鏡高懸③破鏡重圓④鏡花水月
19. ( )他在送舊晚會上「□□□□□□」，博得滿堂彩。①引人注目②引吭高歌③引以為榮④引以自豪
20. ( )「不可開交」①不分是非②不聽也不問③沒法解開或擺脫④指一件事不可預測和估量
21. ( )\_\_\_\_\_相關①昔昔②習習③息息④襲襲
22. ( )小心\_\_\_\_\_①亦亦②奕奕③意意④翼翼
23. ( )有話不便說①一言難盡②苦不堪言③有口難言④百口莫辯
24. ( )不可思\_\_\_\_\_①意②異③義④議
25. ( )「有口皆碑」①事情的問題所在②比喻人人稱讚、頌揚③比喻嚴守秘密，不輕易告訴別人④以懇切真摯的態度，有耐心的勸告他人
26. ( )形容一無所有了①一言難盡②別無長物③面目全非④孤掌難鳴
27. ( )人\_\_\_\_\_亦\_\_\_\_\_①云②芸③耘④雲
28. ( )「風起雲湧」①聲勢盛大②雙方一觸即發的緊張態勢③變化很多，無法預測④對事物情況捉摸不透

## 附錄五

### 使用成語電腦適性測驗滿意度調查問卷

小朋友，恭喜你順利完成成語電腦適性測驗和成語紙筆測驗試卷，關於測驗方式，老師最後有幾個問題想要問問你，請將符合你心中想法的答案於下列空格中勾選出來：

	非常同意	同意	沒意見	不同意	非常不同意
1.我比較喜歡做成語電腦適性測驗。	<input type="checkbox"/>				
2.我比較不喜歡做成語紙筆測驗試卷。	<input type="checkbox"/>				
3.我覺得做成語電腦適性測驗很有趣。	<input type="checkbox"/>				
4.我認為成語電腦適性測驗的歡迎動畫很吸引我。	<input type="checkbox"/>				
5.我覺得成語電腦適性測驗的整體版面設計很美觀。	<input type="checkbox"/>				
6.我覺得成語電腦適性測驗的題目文字清楚明確。	<input type="checkbox"/>				
7.我覺得成語電腦適性測驗用滑鼠點選答案很方便。	<input type="checkbox"/>				
8.我覺得成語電腦適性測驗的題目動畫很搭配。	<input type="checkbox"/>				
9.我做成語電腦適性測驗後的感覺很新奇。	<input type="checkbox"/>				
10.如果有其它類似的電腦適性測驗我會很想試試看。	<input type="checkbox"/>				

以上，謝謝小朋友完成最後的問卷部份，辛苦你囉！^\_^

