

Van Hiele 平面幾何思考層次測驗試題分析 —以高中生為例

馬秀蘭¹ 吳德邦² 吳順治³ 許天維⁴ 洪珮芬⁵

¹ 嶺東科技大學

² 國立臺中教育大學數學教育系

³ 聖約翰大學數位文藝學系

⁴ 國立臺中教育大學教育測驗統計研究所

⁵ 臺中縣文光國民小學

摘要

本研究的目的是以學生問題表 (S-P 表) 分析法針對「高中學生 van Hiele 平面幾何思考層次測驗」進行試題分析。研究樣本取自臺灣中部五個縣市之高中一至三年級學生,有效樣本 1254 名。採用 van Hiele 幾何思考層次測驗 (VHGT) 中文版為評量工具,並以 TESTER For Windows 程式 2.0 版軟體進行 S-P 表及試題分析。本研究分析的項目包括:內部一致性係數、平均答對率 (P)、差異係數、注意係數及鑑別度。經 TESTER For Windows 分析,獲得下列研究結論:(1) 此測驗的內部一致性係數為 0.81,信度良好。(2) 此測驗的平均答對率為 0.53,難度適中。(3) 差異係數為 0.35,顯示試題群和學生群兩者均具有相當的同質性。(4) 試題總數為 25 題,有 13 題被判定為 A (優良型試題),4 題被判定為 B (困難型試題),8 題被判定為 B' (拙劣型試題);其中層次二、三、四各有 1、1、2 題為 B,層次四、五各有 3、5 題為 B',幾何層次越高的題目,學生平均答對人數百分比越低,試題越困難,層次五的題目都被判定為 B'。(5) 此測驗除了有 4 題鑑別度差,3 題鑑別度尚可,2 題鑑別度優良以外,其餘的試題鑑別度非常優良。

關鍵字: S-P 表、van Hiele 平面幾何思考層次、高中學生、試題分析。

Item Analysis of the Chinese Version of the van Hiele Geometry Test of Senior High School Student

Hsiu-Lan Ma* Der-Bang Wu** Shun-Jyh Wu***
Tian-Wei Sheu**** Pei-fen Hung*****

*Lin-Tung University

**Department of Mathematics Education, National Taichung University
of Education

***Department of Digital Literature and Arts, St. John's University

****Graduate Institute of Educational Measurement and Statistics

*****Taichung County Weng-Kuang Elementary School

Abstract

The purpose of this research is to analyze the van Hiele geometry test of Senior high school student with student-problem chart analysis theory. Using the Chinese version of the van Hiele Geometry Test, tested a total of 1254 senior high school students from five counties and cities in the middle part of Taiwan. Analyze the test with S-P chart analysis theory by TESTER For Windows 2.0, developed by Yu (2002). The analysis includes correct ratio of the test, the internal consistency reliability coefficient, disparity index, caution index, and item analysis. After data processing, the following conclusions were drawn from this study: (a) The average of the test is 13.18 (the total grades are 25), the standard deviation is 4.70. (b) The internal consistency reliability coefficient of this test is 0.81, it is good. (c) The average of the correct ratio of the test is 0.53. The difficulty of the test is moderate. (d) The total amount of the questions is 25, there are 13 questions judged as A (fine question), 4 questions judged as B (difficult question), 8 questions judged as B' (clumsy question). Among them, the level two, three, four each has 1, 1, 2 questions as B, the level four, five each has 3, 5 questions as B'. The higher geometry level question, the lower correct ratio and the more difficult the question is. The questions of the level five are all judged as B'. (e) Except that 4 questions have bad item discrimination index, 3 questions are acceptable, 2 questions are fine, and the others are very fine. (f) Disparity index is 0.35, show that the test and the students are suitable homogeneity, and the content of testing are agree with the goal and content of learning.

Key words: senior high school student, van Hiele geometric thinking, Item analysis.

壹、緒論

一、研究動機與背景

荷蘭數學教育家 P. M. van Hiele & Dina van Hiele Geldof 伉儷，於西元1957年，為協助學生易於學習幾何，根據完形心理學之結構論及皮亞傑（J. Piaget）的認知理論（Molina, 1990; van Hiele, 1986），共同提出 van Hiele 學習理論。同年，van Hiele 於法國巴黎發表其研究成果，此理論立即吸引蘇聯教育家的注目（Wirszup, 1976），自1960年代起，更依據 van Hiele 的理論模式改革該國國民中小學之幾何課程（Burger & Shaughnessy, 1986; Fuys, Geddes, & Tischler, 1988; Hoffer, 1983; Usiskin, 1982）。van Hiele 的理論起初在國際間並未受到太多的注意，直到美國數學教育家 Wirszup(1976)首次將 van Hiele 的理論引入美國後才開始引起學界重視該理論，進行一序列相關的研究，應用於改革教學課程（Hoffer, 1983），更據其理論編製測驗，如：芝加哥研究計畫中，Usiskin(1982)等人編製的「van Hiele 幾何測驗」（van Hiele Geometry Test, VHGT）。

VHGT 已被國內外超過200多篇的研究論文暨碩博士論文（陳進春、吳德邦，2005；吳德邦，1996、1998、1999a、1999b；Usiskin, 1982; Wu, 1994; Ma, Wu, Wu, Chen, 2011）所採用，其重要性不可言喻（Usiskin & Senk, 1990）。但其信度不高，K-R 20信度從層次一至五分別為.31, .44, .49, .13, .10（秋季班）以及.39, .55, .56, .30, .26（春季班）（Usiskin, 1982）。信度低的原因是因為各分測驗的題數太少（Usiskin, 1982）。故擴充各分測驗的題數確有必要，由此誘發爾後吳德邦（1994~2011）畢生致力於 van Hiele 幾何思考層次測驗的開發與研究（吳德邦、馬秀蘭、吳順治、陳姿良、沈紀伶，2011）。

1994年吳德邦將 Usiskin 所編製 VHGT 翻譯為中文版，被應用於國內許多幾何思考層次的研究中，然而國內有關高中生之 van Hiele 幾何思考層次之研究甚少，亦缺乏相關主題之博碩士論文研究，對於 Usiskin 的 VHGT 之試題分析在國內也很少人去研究，故本文擬以高中生為例，探討 VHGT 之試題分析。

綜觀各方的研究（江仲翔，2003；吳婉嫻，2006；林孟嫻，2008；陳敏彥，2006），S-P 表分析法能有效的協助教師診斷學生試題的品質，作為改進教學、編製試題的參考，因此本研究打算以 S-P 表分析法對此測驗進行試題分析，以了解幾何測驗的品質，以作為教師進行補救教學或編製相關測驗的參考。

二、研究目的

根據前述的研究動機，本研究將利用 TESTER For Windows 對吳德邦（Wu, 1994）所翻譯的「van Hiele 測驗」進行 S-P 表分析及試題分析，以了解測驗的內部一致性係數、差異係數、注意係數及各試題的鑑別度。

三、研究限制

就研究對象而言，本研究僅抽取 1254 名臺灣中部五縣市一到三年級的高中生作為樣本，未將母群體擴大至全國，研究結論不宜做過度的推論。

就研究主題而言，本文僅以 S-P 表分析法對「van Hiele 測驗」進行試題分析，關於試題反應理論（Item response theory, IRT）中各項參數分析及試題關聯結構分析法（item relation structure analysis, IRS）等，因限於篇幅關係，擬另闢文討論之。

貳、文獻探討

一、van Hiele 幾何發展理論

van Hiele(1986)提出幾何發展理論包含五個幾何思考的層次，每一個層次都有其發展的特徵，茲將其概略特徵分述如下：

層次一，視覺的（visual）層次：此層次的學童是透過視覺觀察實物，由實物的輪廓來辨識形體或圖形，他們可以使用非標準的或標準的數學語言來描述圖形的形狀，但是卻無法理解這些數學語言的定義。

層次二，描述的（descriptive）層次：此層次的學童已經具有辨別圖形特徵的能力，他們能利用視覺來觀察組成圖形的基本要素與這些圖形之間的關係，可以開始尋找出某一類圖形的共同性質，可是卻無法說明這些圖形特徵之間有何關係存在。

層次三，理論的（theoretical）層次：此層次的學童已經能夠了解構成各種圖形的要素，並且能夠進一步探求各種幾何圖形的內在屬性以及各個圖形之間的包含關係，依據圖形的性質進行非正式的推演，但是還不能進行有系統的證明。

層次四，形式邏輯的（formal logic）層次：此層次的學生能用演繹邏輯證明定理，並建立相關定理的網脈結構關係。他們可以在一個公設的系統中建立理論，並且能夠證明與理解一個定理可以有很多不同證明的方法。

層次五，邏輯法則本質的（the nature of logical laws）層次：此層次的學生可以在不同的公設系統中建立定理，並分析比較這些定理的特性，他們能學習不同的幾何公設系統，了解抽象推理幾何，也能互相比較不同公設系統，甚至可以自創一套幾何公設系統。

根據 van Hiele 的理論，上述五個層次都有其次序性，學習者必須擁有前一個層次的概念與策略，才能有效地進行下一個層次的學習活動，而各層次與年齡不一定具有相關，通常國小低年級屬於視覺層次，國小中年級屬於描述層次，國小高年級屬於描述層次到理論層次的過渡期，國中屬於理論層次及形式邏輯層次，高中屬於形式邏輯的層次及邏輯法則的層次，大學屬於邏輯法則的層次。

二、試題分析

試題分析（item analysis）是評鑑試題品質優劣的方法，在整個測驗的編製過程中，扮演著一個相當重要的角色，它的功能在於瞭解題目的品質，刪掉或改寫品質不佳的題目，進而改善題目的品質。試題的質性分析可針對試題的內容與形式進行試題的邏輯審查，以確保試題具有一定的內容效度；量化分析主要在分析每個試題所具備的三大統計特徵，即：難度（difficulty）、鑑別度（discrimination）和誘答力（distraction）（王文中等人，1999；余民寧，2002）。

（一）難度指標的分析

難度指標是用來代表每個試題難易程度的指標，是指答對某題的人數佔總人數的百分比，其計算公式如下：

$$P = \frac{N_R}{N_T} \quad (1)$$

其中 P 為試題的難度指標， N_R 為答對試題的人數， N_T 為總人數。

此外，我們可以分別算出高分組和低分組學生在每道試題上的答對百分比（即 P_H 和 P_L ），再以下列公式來計算每道試題的難度指標：

$$P = \frac{P_H + P_L}{2} \quad (2)$$

P_H 表示高分組（全體受試者當中分數最高的 27%）答對該試題的百分比， P_L 表示低分組（全體受試者當中分數最低的 27%）答對該試題的百分比。難度指標其值越接近 0，即表示試題越困難；其值越接近 1，即表示試題越簡單。一般而言在常模參照測驗中，試題難度指標為 0.50 最為理想，Ahmanan & Glock(1981)就針對選擇題試題，主張應以 0.4 到 0.7 之間為難度值的範圍。

上述兩種難度指標，因前者易受極端值的影響，故後者為較多人常用的方法，故本文亦採用後者。

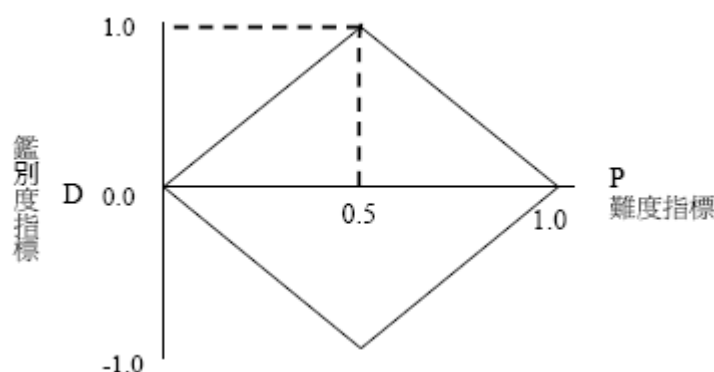
（二）鑑別度指標的分析

鑑別度分析的目的在于於想明瞭試題具有區別學生能力高低的功能如何。在試題鑑別度指標的分析方法上，採用下列公式來計算：

$$D = P_H - P_L \quad (3)$$

D 為鑑別度指標， P_H 為高分組在該試題的答對人數百分比， P_L 為低分組在該試題的答對人數百分比。美國測驗學者 Ebel 與 Frisbie (1991) 提出一套鑑別度判斷標準： D 為 .40 以上，表示試題非常優良； D 為 .30~.39，表示試題優良但可能需要修改； D 為 .20~.29 表示試題尚可但須作局部修改； D 為 .19 以下，表示試題低劣，需要刪除或修改。

從統計學觀點而言，試題的難度指標和鑑別度指標之間，具有密切的關係 (Ebel, 1967)。其間的關係可從圖一中清楚看出：當難度值趨向兩端（即 $P=0.00$ 或 $P=1.00$ ）時，鑑別度值亦將趨近於 0；但當難度值愈接近 0.50 時，鑑別度愈接近 1.00，達到最大。



圖一 難度指標與鑑別度指標的關係（引用余民寧，2002，p233）

（三）選項誘答力的分析

客觀測驗的選項誘答力分析可以提供試題分析的指標，作為判斷試題編製好壞與診斷學生作答反應組型的依據，協助教師改進編製試題的技巧與初步了解學生的作答情形，以便進一步調整或改變教學策略 (Haladyna, 1994)。

要分析選擇題的不正確選項是否具有誘答的功能，只要分析高、低分組學生在每道試題選項上選答的次數分配，再依據下列兩個原則加以判斷：

1. 每個不正確選項至少要有一位以上的低分組學生選擇它。
2. 低分組學生選擇不正確選項的人數要比高分組學生選擇不正確選項的人數還多。

若發現有違反以上兩項參考原則之一者，即表示試題的不正確選項缺乏誘答功能，必須加以修改、潤飾或刪除（余民寧，2002）。

三、S-P表分析技術

S-P表分析技術是由日本學者佐藤隆博（Takahiro Sato）於1970年代開發，針對每位學生與每個試題的作答反應組型所產生的注意係數，與整份測驗卷的差異係數做分析。

（一）差異係數

差異係數（disparity index）表示 S 曲線與 P 曲線分離的程度，其公式如下：

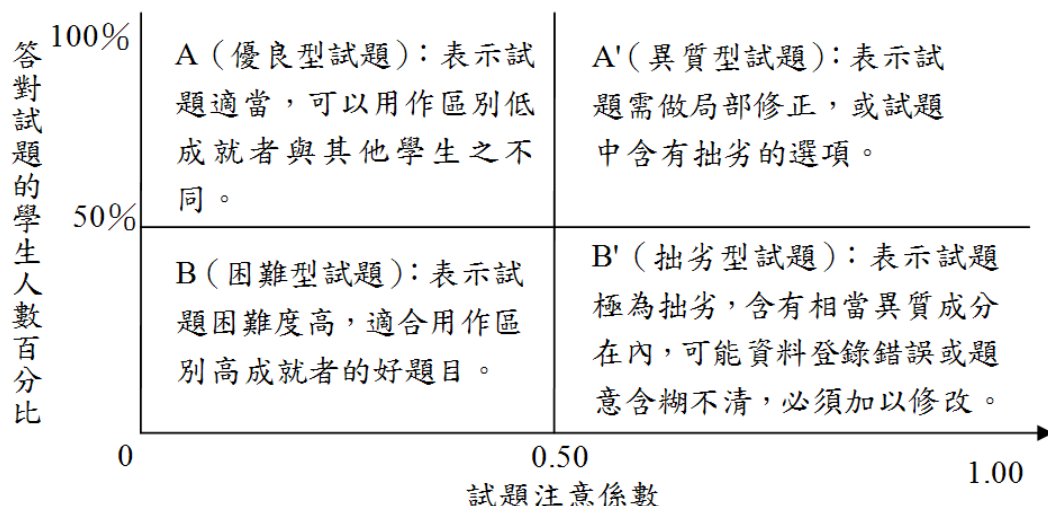
$$D^* = \frac{S(N, n, \bar{P})}{S_B(N, n, \bar{P})}, \quad \bar{P} = \frac{\sum_{i=1}^N X_i}{Nn} \quad (4)$$

$S(N, n, \bar{P})$ 表示在人數為 N，試題數為 n，平均答對率為 \bar{P} 的條件下，實際 S-P 表中兩曲線所包圍部分的面積； $S_B(N, n, \bar{P})$ 表示在人數為 N，試題數為 n，平均答對率為 \bar{P} 的條件下，兩曲線呈累加二項分配曲線所包圍部分之面積。正常的情況下，差異係數 D^* 會介於 0 到 1 之間。一般而言， $D^* = 0.5$ 左右為標準值，當 $D^* > 0.6$ 時，表示測驗含有異質因素，應對學生的反應組型予以注意，對試題加以檢討，並做適當的修改。因此差異係數可用來表示 S-P 表整體的非同質性程度指標（余民寧，2002）。

（二）注意係數

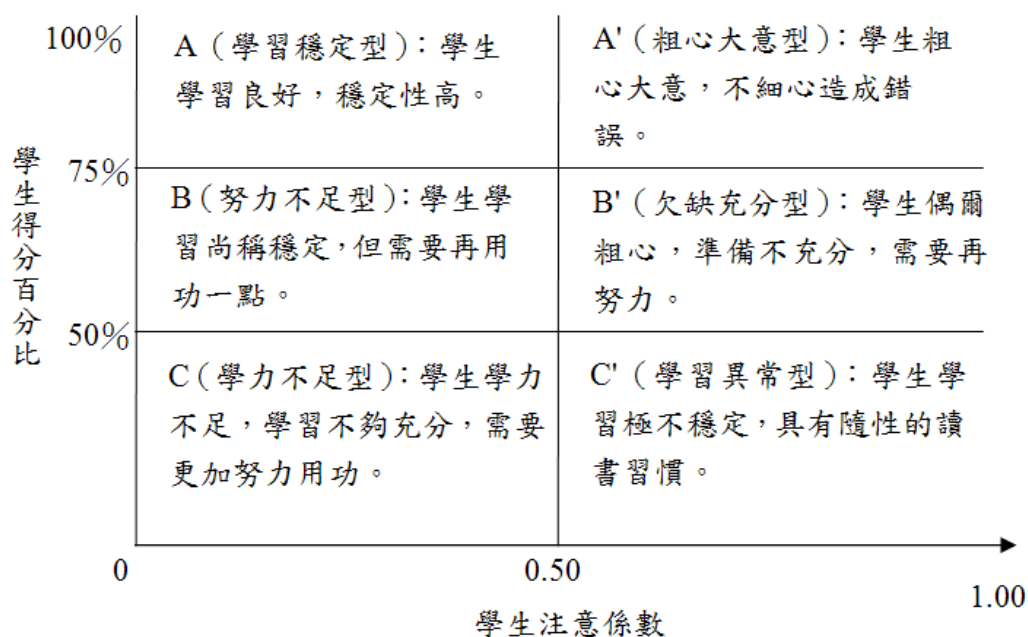
注意係數（caution index）是 S-P 表針對「個別」學生與試題所使用的另一類係數，可分為學生注意係數（caution index for students，簡稱 CS）與問題注意係數（caution index for items，簡稱 CP）兩種。當注意係數（CP 或 CS）介於 0~0.5 之間時，表示該試題或學生之反應組型的不尋常情況並不嚴重，屬於正常程度；當注意係數在 0.5~0.75 之間，則表示不尋常之情況已是嚴重狀況，教師應該加以注意；當注意係數大於 0.75 時，則表示不尋常之情況已是非常之嚴重，教師應更加特別注意。

S-P 表分析技術將試題診斷結果，依答對試題學生的百分比及 CP 分成四類：A（優良型試題）、A'（異質型試題）、B（困難型試題）、B'（拙劣型試題），如圖二所示：



圖二 試題診斷分析圖 (引自余民寧, 2002, p355)

S-P 表分析技術將學生學習診斷結果, 依學生得分百分比及 CS 分成六類: A (學習穩定型)、A' (粗心大意型)、B (努力不足型)、B' (欠缺充分型)、C (學力不足型)、C' (學習異常型), 如圖三所示:



圖三 學生診斷分析圖 (引自余民寧, 2002, p358)

由於利用同一份試題來分析試題 (CP) 和分析學生 (CS) 的學習, 似乎犯了邏輯上的謬誤。因為我們用這批學生來分析試題發現有 A', B' 的試題需要修改或刪除。我們又用這份試題來分析學生發現有 A', B', C' 的粗心大意等的學生問題。在邏輯上, 是否需先刪除或修改不合的試題, 再來分析學生的學習。也要刪除那些粗心大意等影響試題因素的學生再來分試題。如此循環下去, 本研究的試題和樣本都將出現一些矛盾的問題產生。因此, 本文只保留對試題的分析, 把學生的因素當做「必要且恒真」的公設。

四、高中幾何課程

高中階段幾何思考層次屬於形式邏輯的層次及邏輯法則的層次，幾何課程重在培養學生抽象的數學邏輯推理能力，其內容（教育部，2004）包含：

（一）三角函數的基本概念：銳角三角函數，三角函數的基本關係，簡易測量與三角函數值表，廣義角的三角函數，正弦定理，餘弦定理，基本三角測量。

（二）三角函數的性質與應用：三角函數的圖形，和角，倍角，半角公式，正、餘弦函數之疊合，複數的極式。

（三）向量：有向線段與向量，向量的基本應用，平面向量的坐標表示法，平面向量的內積。

（四）空間中的直線與平面：空間概念，空間坐標系，空間向量的坐標表示法，平面方程式，空間直線方程式，一次方程組。

（五）圓與球面的方程式：圓的方程式，圓與直線的關係，球面方程式，球面與平面的關係。

（六）圓錐曲線：圓錐曲線名詞的由來，拋物線（標準式），橢圓（標準式），雙曲線（標準式），圓錐曲線的光學性質。

參、研究方法與程序

一、研究對象

本研究之對象係隨機抽樣自中部五個縣市的高中一到三年級學生。筆試的樣本來自五個縣市（為避免影響到學生的心理，縣市別以代號取代之，其中台中市是指未升格為直轄市之台中市）。的高中一到三年級學生，刪除資料不完整的樣本後，共得有效樣本 1254 名。五個縣市各年級男女生樣本分配詳見表 1。

表 1 樣本分配表

縣市	性別	一年級	二年級	三年級	總和人數
縣市1	女	42 (3.3%)	51(4.1%)	8(0.6%)	101(8.1%)
	男	52(4.1%)	45(3.6%)	0(0%)	97(7.7%)
	總和人數	94(7.5%)	96(7.7%)	8(0.6%)	198(15.8%)
縣市2	女	50(4.0%)	52(4.1%)	51(4.1%)	153(12.2%)
	男	47(3.7%)	49(3.9%)	46(3.7%)	142(11.3%)
	總和人數	97(7.7%)	101(8.1%)	97(7.7%)	295(23.5%)
縣市3	女	47(3.7%)	39(3.1%)	43(3.4%)	129(10.3%)
	男	42(3.3%)	51(4.1%)	50(4.0%)	143(11.4%)
	總和人數	89(7.1%)	90(7.2%)	93(7.4%)	272(21.7%)
縣市4	女	41(3.3%)	34(2.7%)	39(3.1%)	114(9.1%)
	男	38(3.0%)	35(2.8%)	48(3.8%)	121(9.6%)
	總和人數	79(6.3%)	69(5.5%)	87(6.9%)	235(18.7%)
縣市5	女	48(3.8%)	35(2.8%)	39(3.1%)	122(9.7%)
	男	43(3.4%)	49(3.9%)	40(3.2%)	132(10.6%)
	總和人數	91(7.3%)	84(6.7%)	79(6.3%)	254(20.3%)
總計		450(35.9%)	440(35.1%)	364(29.0%)	1254(100%)

二、研究工具

本研究以 van Hiele 幾何思考層次測驗 (VHGT) 中文版為評量工具。此測驗共有五個層次，每個 van Hiele 層次各五題，共有 25 題選擇題。1 到 5 題屬於視覺層次；6 到 10 題屬於描述層次；11 到 15 題屬於理論層次；16 到 20 題屬於形式邏輯層次；21 到 25 題屬於邏輯法則本質層次。van Hiele 幾何思考層次測驗 (VHGT) 中文版由吳德邦 (Wu, 1994) 所翻譯，並應用在其博士論文的研究中，其效度及信度亦在其第一本博士論文的研究考驗過，各層次之再測信度分別為：.5549 ($P = .000$)，.8153 ($P = .000$)，.6492 ($P = .000$)，.4505 ($P = .000$)，和 .7902 ($P = .000$)。

本研究採用 TESTER For Windows 程式 2.0 版為試題分析的工具。TESTER For Windows 程式 2.0 版為余民寧 (2002) 所研發的軟體，它的目的在執行古典測驗理論下的試題分析、測驗分析及學生問題表 (S-P 表) 分析。

三、資料處理

本研究採用 TESTER For Windows 程式 2.0 版進行 S-P 表及試題分析，分析內容為：平均答對率 (P)、內部一致性係數、差異係數，以及試題分析。此軟體進行試題診斷結果判定類別分析的標準如下：

A：答對試題的學生人數百分比 $\geq 50\%$ ，CP < 0.50 。

A'：答對試題的學生人數百分比 $\geq 50\%$ ，CP ≥ 0.50 。

B：答對試題的學生人數百分比 $< 50\%$ ，CP < 0.50 。

B'：答對試題的學生人數百分比 $< 50\%$ ，CP ≥ 0.50 。

肆、結果與討論

一、利用 TESTER For Windows 進行測驗分析，結果如下：

(一) 問題注意係數分析結果，如表 2。

表 2 問題注意係數表

題號	答對人數	答對率	刪題後信度	注意指標	判定類別
1	1142	91.07	0.809	0.44	A
2	1112	88.68	0.805	0.27	A
3	1197	95.45	0.809	0.24	A
4	984	78.47	0.804	0.39	A
5	1024	81.66	0.802	0.28	A
6	839	66.91	0.795	0.23	A
7	1010	80.54	0.803	0.34	A
8	682	54.39	0.796	0.29	A
9	883	70.41	0.799	0.29	A
10	581	46.33	0.798	0.33	B
11	664	52.95	0.796	0.29	A

表 2 問題注意係數表 (續)

題號	答對人數	答對率	刪題後信度	注意指標	判定類別
12	782	62.36	0.796	0.28	A
13	650	51.83	0.798	0.33	A
14	370	29.51	0.801	0.36	B
15	640	51.04	0.797	0.31	A
16	385	30.70	0.814	0.69*	B'
17	589	46.97	0.801	0.39	B
18	546	43.54	0.799	0.35	B
19	321	25.60	0.817	0.77**	B'
20	361	28.79	0.811	0.61*	B'
21	258	20.57	0.816	0.76**	B'
22	210	16.75	0.813	0.66*	B'
23	545	43.46	0.814	0.60*	B'
24	303	24.16	0.819	0.86**	B'
25	451	35.96	0.812	0.63*	B'

(二) 測驗分析結果，如表 3。

表 3 測驗分析結果

總和 (Sum)	16529.00
平均每人得分 (Average)	13.18
最小值 (Minimum)	2.00
最大值 (Maximum)	24.00
全距 (Range)	22.00
變異數 (Variance)	22.11
變異數 (樣本)	22.10
標準差 (SD)	4.70
標準差 (樣本)	4.70
偏態 (Skewness)	-0.06
峰度 (Kurtosis)	-0.88
平均答對人數	661.16
受試者總數 (N)	1254.00
平均答對率 (P)	0.53
問題總數 (K)	25.00
內部一致性係數	0.81
差異係數	0.35

1. 經過 Tester 分析，此份測驗的平均答對率為 0.53，依據 Ahmanan & Glock(1981)難易度的標準，難易度適中。平均每人得 13.18 分，最小值為 2.00 分，最大值為 24.00 分，得分全距為 22.00 分，變異數為 22.11，標準差為 4.70，偏態為-0.66，峰度為-0.88，平均答對人數為 661.16 人。
2. 此測驗的內部一致性係數為 0.81，信度良好，具有使用的價值。
3. 試題 1~5（層次一）的內部一致性係數為 0.62，試題 6~10（層次二）的內部一致性係數為 0.70，試題 11~15（層次三）的內部一致性係數為 0.71，試題 16~20（層次四）的內部一致性係數為 0.33，試題 21~25（層次五）的內部一致性係數為 0.21。
4. 差異係數為 0.35，顯示試題群和學生群兩者均具有相當的同質性。
5. 試題總數為 25 題，從問題注意係數表得知，試題有 13 題（52%）被判定為 A，4 題（16%）被判定為 B，8 題（32%）被判定為 B'，其中試題 10（層次二），14（層次三），17、18（層次四）為 B，試題 16、19、20（層次四），21、22、23、24、25（層次五）為 B'，如表 4 所示。

表 4 試題診斷分析統計表

診斷結果	次數（題）	相對次數（%）
A（優良型試題）	13	52%
A'（異質型試題）	0	0%
B（困難型試題）	4 (10, 14, 17, 18) *	16%
B'（拙劣型試題）	8 (16, 19, 20, 21, 22, 23, 24, 25) *	32%
總計	25	100%

註一：*代表試題題號

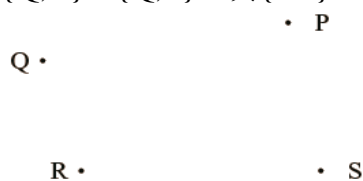
1. 依據 Ebel 與 Frisbie（1991）提出的鑑別度判斷標準，此測驗除了試題 3、19、22、24 鑑別度差，試題 1、16、21 鑑別度尚可，試題 20、25 鑑別度優良以外，其餘的試題鑑別度非常優良。
2. 試題 1~5（層次一）的平均答對率為 87.1%，試題 6~10（層次二）的平均答對率為 67.7%，試題 11~15（層次三）的平均答對率為 49.5%，試題 16~20（層次四）的平均答對率為 35.1%，試題 21~25（層次五）的平均答對率為 28.2%。

二、試題誘答力分析

如上一節所述，試題 1 到試題 5 屬於幾何思考層次一的題目，試題 6 到試題 10 屬於幾何思考層次二的題目，試題 11 到試題 15 屬於幾何思考層次三的題目，試題 16 到試題 20 屬於幾何思考層次四的題目，試題 21 到試題 25 屬於幾何思考層次五的題目。在此僅以難度指標較小的 3 題（試題 21、22、24）、難度指標較大的 3 題（試題 1、2、3）、鑑別度指標較高的 3 題（試題 8、11、18）、鑑別度指標較低的 3 題（試題 19、22、24）試題進行試題誘答力分析，其試題內容及試題選項誘答力分析結果如下：

(一) 第 21 題試題：

21、在與您過去所學之幾何學有所不同的 F-幾何學中，每一直線被定義為恰好只包含二點。若 P、Q、R、以及 S 為四個點，則可構成六條直線： $\{P,Q\}$ ， $\{P,R\}$ ， $\{P,S\}$ ， $\{Q,R\}$ ， $\{Q,S\}$ ，與 $\{RS\}$ 。



在 F-幾何學中，是這樣定義“相交”與“平行”這兩個術語的。

----兩直線 $\{P,Q\}$ 和 $\{P,R\}$ 相交在 P 點，因為 $\{P,Q\}$ 與 $\{P,R\}$ 有相同點 P。

----兩直線 $\{P,Q\}$ 和 $\{R,S\}$ 互相平行，因為 $\{P,Q\}$ 與 $\{R,S\}$ 沒有相同點。

由以上的資料得知，哪一個敘述是正確的？

- (A) $\{P,R\}$ 和 $\{Q,S\}$ 相交。
- (B) $\{P,R\}$ 和 $\{Q,S\}$ 平行。
- (C) $\{Q,R\}$ 和 $\{R,S\}$ 平行。
- (D) $\{P,S\}$ 和 $\{Q,R\}$ 相交。
- (E) 以上(A)-(D)都不正確。

題號:21

選 項	1	2*	3	4	5	其他
選項率	.40	.21	<input type="checkbox"/>	.06	.23	通過率:20.□7
			0			
高分組	.33	.□5	.□3	.03	.26	難 度:0.252
低分組	.30	.15	.19	.14	.21	鑑別度:0.205

試題 21 屬幾何思考層次五的題目，其通過率為 20.57，難度指標為 0.252，試題難度困難；鑑別度指標為 0.205，鑑別度尚可，CP 為 0.76，試題被判定為 B' (拙劣型)，評鑑結果為「修改」。從選項誘答力分析可知，試題的選項一、五對高分組較具誘答力。仔細審試題幹與選項的命題方式，並沒有任何缺點存在，原因可能是學生混淆了 F-幾何學與歐式幾何學中相交與平行的定義。所以經過試題分析後，教師宜加強各幾何學中相交及平行的定義與分辨，才能改善高分組學生作答情形。建議本題不必修改，但老師宜針對試題分析結果所顯示的意義進行補救教學。

(二) 第 22 題試題：

22、三等分角意指把一個角平分為三等分。在西元 1847 年，有一位數學家 (P.L. Wantzel) 證實，一般而言，只使用一個圓規與一個未標刻度的直尺是不可能三等分角的。根據他的證明，你能夠下什麼結論呢？

- (A) 一般而言，只使用一個圓規與一個未標刻度的直尺是不可能二等分角。
- (B) 一般而言，只使用一個圓規與一個有標刻度的直尺是不可能三等分角。
- (C) 一般而言，使用任何畫圖工具是不可能三等分角。
- (D) 將來有些人仍有可能發現用一般的方法，即只使用一個圓規與一個未標刻度的直尺去三等分角。
- (E) 永遠沒有人能發現用一般的方法，即只使用一個圓規與一個未標刻度的直尺去三等分角。

題號:22

選 項	1	2	3	4	5*	其他
選項率	.08	.15	.13	.45	.18	通過率: 16.75
高分組	.05	.06	.08	.54	.28	難 度: :0.197
低分組	.11	.23	.17	.37	.12	鑑別度: 0.163

試題 22 屬幾何思考層次五的題目，其通過率為 16.75，難度指標為 0.197，在這 25 題試題中，難度最難；鑑別度指標為 0.163，鑑別度非常低劣；CP 為 0.66，試題被判定為 B'，評鑑結果為「刪除」，刪除此題後，整份測驗的信度提高為 0.813。選項誘答力分析結果，選項四違反誘答力分析的參考原則，可能是學生不了解題目語意，所以建議宜修改或刪除本試題。

(三) 第 24 題試題：

24、兩本幾何學的書用不同的方法定義長方形。

哪一個敘述是正確的呢？

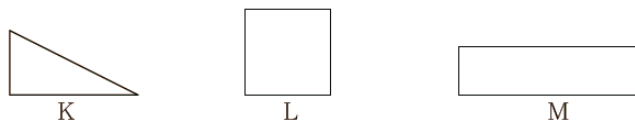
- (A) 其中一本書的定義寫錯了。
- (B) 其中一本書的定義是錯誤的，對長方形而言不能有二種不同的定義。
- (C) 在其中一本書裡，長方形必有不同於另一本書中之長方形的性質。
- (D) 在其中一本書裡，長方形必有與另一本書相同的性質。
- (E) 在兩本書裡，長方形的性質可能不同。

題號:24						
選 項	1	2	3	4	5*	其他
選項率	.04	.11	.21	.38	.25	通過率: 24.16
高分組	.01	.07	.13	.50	.29	難 度: 0.244
低分組	.08	.19	.30	.23	.20	鑑別度: 0.086

試題 24 屬幾何思考層次五的題目，其通過率為 24.16，難度指標為 .244，試題困難；鑑別度指標為 .086，鑑別度低劣；CP 為 .86，試題被判定為 B'，評鑑結果為「刪除」，刪除此題後，整份測驗的信度提高為 .819。從試題的選項誘答力分析資料中得知，選項四違反了選項分析的參考原則，仔細審試題幹與選項的命題方式，發現第一個選項和第二個選項有雷同之處，選擇第四個選項的學生比例較多，可能學生不了解題意，經過試題分析後建議宜刪除。

(四) 第 1 題試題：

1、下列各圖形中，哪些是正方形？



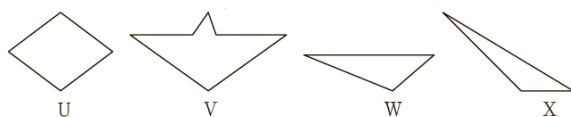
- (A) 只有 K。
- (B) 只有 L。
- (C) 只有 M。
- (D) 只有 L 和 M。

題號: 1						
選 項	1	2*	3	4	5	其他
選項率	.02	.91	.02	.05	.01	通過率: 91.07
高分組	.00	.99	.00	.00	.00	難 度: 0.882
低分組	.06	.77	.03	.12	.02	鑑別度: 0.217

試題 1 屬幾何思考層次一的題目，其通過率為 91.07，難度指標為 0.882，試題極容易；鑑別度指標為 0.217，鑑別度尚可；CP 為 0.44，試題被判定為 A（優良型），評鑑結果為「修改」。從試題的選項誘答力分析資料中得知，每個不正確選項都具誘答力，且符合誘答力分析的參考原則，仔細審試題幹與選項的命題方式並沒有任何缺點存在，建議予以保留。

(五) 第 2 題試題：

2、下列各圖形中，哪些是三角形？



- (A) 全部都不是三角形。
 (B) 只有 V。
 (C) 只有 W。
 (D) 只有 W 和 X。

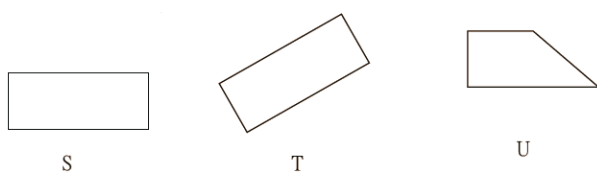
題號:2

選 項	1	2	3	4*	5	其他
選項率	.04	.02	.04	.89	.01	通過率: 88.68
高分組	.00	.00	.01	.99	.00	難 度: 0.828
低分組	.13	.07	.09	.66	.04	鑑別度: 0.332

試題 2 屬幾何思考層次一的題目，其通過率為 88.68，難度指標為 0.828，試題極容易；鑑別度指標為 0.332，鑑別度優良；CP 為 0.27，試題被判定為 A，評鑑結果為「保留」。從試題的選項誘答力分析資料中得知，每個不正確選項都具誘答力，且符合誘答力分析的參考原則。

(六) 第 3 題試題：

3、下列各圖形中，哪些是長方形？



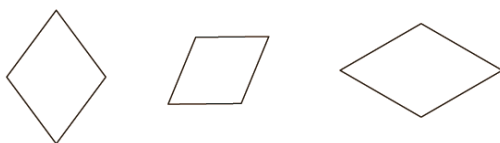
- (A) 只有 S。
 (B) 只有 T。
 (C) 只有 S 和 T。
 (D) 只有 S 和 U。
 (E) 全部都是長方形。

題號:3						
選 項	1	2	3*	4	5	其他
選項率	.02	.01	.95	.01	.01	通過率: 95.46
高分組	.00	.00	1.00	.00	.00	難 度: 0.925
低分組	.05	.02	.85	.05	.03	鑑別度: 0.144

試題 3 屬幾何思考層次一的題目，其通過率為 95.46，難度指標為 0.925，試題極容易；但鑑別度指標為 0.144，鑑別度低劣；CP 為 0.24，試題被判定為 A，評鑑結果為「刪除」，但刪除此題後，整份測驗的信度降低為 0.809。從試題的選項誘答力分析資料中得知，試題的不正確選項都具誘答力，仔細審試題幹與選項的命題方式並沒有任何缺點存在，所以建議予以保留。鑑別力低落的原因是因為此試題難度過於簡單，無法有效鑑別高分組及低分學生。

(七) 第 8 題試題：

8、菱形是一個四邊都等長的四邊形。圖形如下，



對任一菱形而言，下列 (A) 到 (D) 中，哪一個性質是錯誤的？

<若 (A) 到 (D) 的性質都成立，則選 (E)。>

- (A) 對角線長度相等。
- (B) 每一對角線平分菱形的兩個角。
- (C) 兩條對角線互相垂直。
- (D) 對角相等。
- (E) 對每一個菱形而言，以上 (A) 到 (D) 都成立。

題號:8						
選 項	1*	2	3	4	5	其他
選項率	.54	.07	.09	.06	.25	通過率: 54.39
高分組	.89	.00	.01	.00	.09	難 度: 0.513
低分組	.13	.16	.17	.14	.39	鑑別度: 0.764

試題 8 屬幾何思考層次二的題目，其通過率為 54.39，難度指標為 0.513，試題難度適中；鑑別度指標為 0.764，鑑別度非常優良；CP 為 0.29，試題被判定為 A，評鑑結果為「保留」。從試題的選項誘答力分析資料中得知，每個不正確選項對低分組都有不錯的誘答力，尤其是第五個選項，是個優良的題目。

(八) 第 11 題試題：

- 11、若已知圖形 F 是一個多邊形。下列哪一個敘述是正確的？
- (A) 若圖形 F 是一個長方形，則圖形 F 必定是一個三角形。
- (B) 若圖形 F 不是一個長方形，則圖形 F 必定是一個三角形。
- (C) 圖形 F 不可能同時是一個長方形，而且又是一個三角形。
- (D) 圖形 F 一定是一個長方形，或是一個三角形。
- (E) 以上 (A) - (D) 都不正確。

題號:11						
選 項	1	2	3*	4	5	其他
選項率	.04	.06	.53	.06	.31	通過率: 52.95
高分組	.00	.01	.94	.00	.05	難 度: 0.545
低分組	.12	.15	.15	.13	.45	鑑別度: 0.796

試題 11 屬幾何思考層次三的題目，其通過率為 52.95，難度指標為 0.545，試題難度適中；鑑別度指標為 0.796，鑑別度非常優良；CP 為 0.29，試題被判定為 A，評鑑結果為「保留」。從試題的選項誘答力分析資料中得知，每個不正確選項對低分組都具有的誘答力，是個優良的題目。

(九) 第 18 題試題：

- 18、以下有兩個敘述。
- I. 若一個圖形是長方形，則它的對角線互相平分。
- II. 若一個圖形的對角線互相平分，則它是一個長方形。
- 下列哪一個敘述是正確的？
- (A) 若要證明 I 為真，則只要證明 II 為真即可。
- (B) 若要證明 II 為真，則只要證明 I 為真即可。
- (C) 若要證明 II 為真，則只要找到一個對角線互相平分的長方形即可。
- (D) 若要證明 II 為偽，則只要找到一個對角線互相平分但不是長方形即可。
- (E) 以上 (A) - (D) 都不正確。

題號:18						
選 項	1	2	3	4*	5	其他
選項率	.12	.14	.17	.44	.13	通過率: 43.54
高分組	.03	.04	.01	.87	.05	難 度: 0.497
低分組	.14	.19	.32	.12	.23	鑑別度: 0.751

試題 18 屬幾何思考層次四的題目，其通過率為 43.54，難度指標為 0.497，試題難度適中；鑑別度指標為 0.751，鑑別度非常優良；CP 為 0.35，試題被判定為 B (困難型)，評鑑結果為「保留」。從試題的選項誘答力分析資料中得知，每個不正確選項無論是對高分組或低分組都具有的誘答力，是個非常優良的題目，具有區別高成就者的功能。

(十) 第 19 題試題：

19、下列有關幾何學中的敘述，何者正確？

- (A) 每一個術語都能被定義，而且每一個真實的敘述都能被證明是正確的。
- (B) 每一個術語都能被定義，但假設某些特定的敘述恆為真則是必要的。
- (C) 某一些術語必須當作是未定義名詞，但每一個真實的敘述都能被證明是正確的。
- (D) 某一些術語必須當作是未定義名詞，而且假設某些特定的敘述恆為真是必要的。
- (E) 以上 (A) - (D) 都不正確。

題號:19

選 項	1	2	3	4*	5	其他
選項率	.21	.25	.16	.26	.12	通過率: 25.60
高分組	.22	.26	.09	.35	.09	難 度: 0.267
低分組	.23	.24	.18	.19	.17	鑑別度: 0.163

試題 19 屬幾何思考層次四的題目，其通過率為 25.60，難度指標為 .267，試題困難；鑑別度指標為 .163，鑑別度低劣，CP 為 .77，被判定為 B'，評鑑結果為「刪除」，刪除此題後，整份測驗的信度提高為 .817。從試題的選項誘答力分析資料中得知，選項二違反了選項分析的參考原則，對高分組更具誘答力，可能語意不清，學生不了解「術語」、「定義」等意思，或是學生從未學習過建立幾何學公設系統需要何種條件而影響學生作答，因此建議必須加以修改或刪除。

小結：

上述層次五的 5 個題目 (Q21~Q25) 均被判定為 B' (拙劣型試題)，我們認為其可能原因是學生不瞭解題意，或是有其它因素，值得進一步探究。有些題目甚至違反誘答力分析 (例如：第 22 和 24 題)，此研究結果和 Usiskin(1982)所提出的論調 “Level five either does not exist or is not testable. All other levels are testable” (p. 79) (對高中學生而言，van Hiele 幾何思考層次五是不存在的或不可測的) 是一致的。由此誘發爾後吳德邦 (1994) 撰寫其博士論文，驗證了「對大專學生而言，van Hiele 幾何思考層次五是存在的而且是可測的。」

伍、結論與建議

一、結論

從以上所分析的資料，可得到下列的結果：

- (一) 此測驗的內部一致性係數為 0.81，信度良好。
- (二) 此測驗的平均答對率為 0.53，難度適中。
- (三) 差異係數為 0.35，顯示試題群和學生群兩者均具有相當的同質性。
- (四) 試題總數為 25 題，有 13 題被判定為 A (優良型試題)，4 題被判定為 B (困難型試題)，8 題被判定為 B' (拙劣型試題)；其中層次二、三、四各有 1、1、2 題為 B，層次四、五各有 3、5 題為 B'。
- (五) 此測驗除了試題 3、19、22、24 鑑別度差，試題 1、16、21 鑑別度尚可，試題 20、25 鑑別度優良以外，其餘的試題鑑別度非常優良。

二、建議

根據上述的發現，提出下列的建議：

- (一) 受試者總數為 1254 人，學生診斷分析統計結果，有 12.7% 屬於 A，1.4% 屬於 A'，36.4% 屬於 B，4.0% 屬於 B'，28.3% 屬於 C，17.1% 屬於 C'，這些學生的幾何思考層次發展情形如何呢？未來可以針對試題關聯結構進行分析。
- (二) 研究結果試題 16~20 (層次四) 的平均答對率為 35.1%，試題 21~25 (層次五) 的平均答對率為 28.2%，顯示學生在演繹邏輯證明定理及在不同的公設系統中建立定理並分析比較這些定理特性的能力低落，建議學校未來教學應多加強幾何證明，教材也應多引入不同的公設系統幾何內容，以提升學生幾何思考層次。
- (三) 本研究採 S-P 表分析法屬古典測驗理論，雖能有效的協助教師診斷試題的品質，但卻有古典測驗理論的缺點，即分析結果會因施測的學生不同而有不同的結果，有樣本相依的情形，因此未來建議可以進行進一步探討「不同試題分析法」、試題反應理論 (Item response theory, IRT)、試題反應結構 (IRS)、以及模糊測度、灰色理論等各項參數分析來了解試題的品質。

致謝

本文係「行政院國家科學委員會科學教育處」補助專題研究計畫之部分研究成果，特此申謝！

(計劃編號：NSC 87-2511-S-142-003、NSC 92-2521-S-142-004)。

參考文獻

中文部份

- 王文中、呂金燮、吳毓瑩、張郁雯、張淑慧 (1999)。教育測驗與評量—教室學習觀點。臺北：五南。
- 江仲翔 (2003)。應用 S-P 表分析高級中學數學科測驗試題。國立中山大學應用數學研究系碩士論文，未出版，高雄市。
- 余民寧 (2002)。教育測驗與評量—成就測驗與教學評量。臺北：心理。
- 吳德邦 (1996)。范析理 (van Hiele) 模式對我國師範學院學生在非歐幾何學的學習成就與幾何思考層次之研究。臺中師院學報，9，443-474。
- 吳德邦 (1998)。國中學生 van Hiele 幾何思考層次之研究。八十七學年度數學教育專題研究計畫成果討論會，臺北市行政院國家科學委員會科學教育發展處，臺北市，民國 87 年 11 月 28-29 日、12 月 6 日。
- 吳德邦 (1999a)。學士後國小職前與在職進修專班教師 van Hiele 幾何思考模式之研究。數理教學與師資培育學術研討會，國立彰化師範大學，彰化市，民國 88 年 12 月 18-19 日。
- 吳德邦 (1999b)。幼教職前與在職進修教師 van Hiele 幾何思考模式之研究。跨世紀幼教師資培育：回顧與展望學術研討會，國立臺中師範學院，臺中市，民國 88 年 5 月 27-28 日。
- 吳德邦、馬秀蘭、吳順治、陳姿良、沈紀伶 (2011)：編製 van Hiele 幾何思考層次四測驗之歷程。第三屆科技與數學教育學術研討會論文集 (pp. 276-287)。台中市：國立台中教育大學數學教育學系。
- 吳婉嫻 (2003)。利用 S-P 表分析高中地圖技能—以一個班級為個案研究。國立台灣大學地理環境資源研究所碩士論文 (未出版)。
- 林孟嫻 (2008)。國小學童小數加減表現之研究—以 S-P 表與次序理論分析為例。國立屏東教育大學數理教育研究所碩士論文 (未出版)。
- 教育部 (2004)。普通高級中學課程暫行綱要。臺北：教育部。
- 陳進春、吳德邦 (2005)。醫護專科學校學生 van Hiele 幾何思考層次之研究。測驗統計年刊，13 (2)，230-260。
- 陳敏彥 (2006)。應用 S-P 表與次序理論分析原住民學生在分數乘法之認知診斷。國立臺中教育大學數學教育學系碩士論文，未出版，臺中市。

英文部份

- Ahmanan, J. S., & Glock, M. D.(1981). *Evaluating student progress: Principles of tests and measurement*(6th ed.). Boston, MA: Allyn, & Bacon.
- Burger, W. F., & Shaughnessy, J. M.(1986). *Assessing children's intellectual growth in geometry*(Final report of the Assessing Children's Intellectual Growth in Geometry project) . Corvallis, OR: Oregon State University, Department of Mathematics.
- Ebel, R. L.(1967).The relation of item discrimination to test reliability. *Journal of Educational Measurement*,4,125-128.
- Ebel, R. L., & Frisbie, D. A.(1991). *Essentials of educational measurement, 5th ed.* Englewood Cliffs, NJ: Prentice-Hall.
- Fuys, D., Geddes, D., & Tischler, R.(1988). *The van Hiele model of thinking in geometry among adolescents.* Reston, VA: The National Council of Teachers of Mathematics, Inc.
- Haladyna, T. M.(1994). *Developing and validating multiple-choice test items.* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hoffer, A.(1983). van Hiele based research. In R. Lesh, & M. Landau(Eds.), *Acquisition of mathematical concepts and processes.*(pp. 205-228). New York, NY: Academic Press.
- Ma, H. L., Wu, D. B., Wu, S. J., & Chen, T. L.(2011). A study of the van Hiele model of geometric thinking to in-service kindergarten teachers in Taiwan. In Ubuz, B.(Ed.). *Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education, Vol. 1*, pp. 477. Ankara, Turkey: PME. ISBN: 978-975-429-294-7.(NSC 87-2511-S-142-003).
- Molina, D. D.(1990). *The applicability of the van Hiele theory to transformational geometry.* Unpublished doctoral dissertation, The University of Texas at Austin.
- Usiskin.(1982). *van Hiele levels and achievement in secondary school geometry*(Final Report of the Cognitive Development and Achievement in Secondary School Geometry Project). Chicago, IL: University of Chicago, Department of Education.(ERIC Reproduction Service No. ED 220 288).
- Usiskin, Z., & Senk, S.(1990). Evaluating a test of vanHiele levels: A response to Crowley and Wilson. *Journal for Research in Mathematics Education*, 21(3), 242-245.
- Van Hiele, P. M.(1986). *Structure and insight: A theory of mathematics education.* Orlando, FL: Academic Press.
- Wirszup, I.(1976). Breakthroughs in the psychology of learning and teaching geometry. In J. L. Martin, & D. A. Bradbard(Eds.), *Space and geometry: Papers from a research workshop*(pp. 75-97). Columbus, OH: ERIC Center for Science, Mathematics and Environmental Education.(ERIC Document Reproduction service No. 132033).
- Wu, D. B.(1994). *A study of the use of the van Hiele model in the teaching of nonEuclidean geometry to prospective elementary school teachers in Taiwan, the Republic of China.* Unpublished Doctoral dissertation, University of Northern Colorado, Greeley.