

國立臺灣師範大學教育心理學系
教育心理學報，民74，18期，85—96頁

國中學生測驗智巧和冒險取向的實驗研究*

吳 鐵 雄

本研究的目的是在探討測驗編製原則教學效果與年級和測驗智巧與冒險取向的關係，並研究學生自我概念與此兩種測驗行為的關係。受試者取樣自臺北市二所國中，共 787 人，其中一年級 274 人，二年級 255 人，三年級 248 人。受試者隨機分派為實驗組和控制組，實驗組教以測驗編製原則，而控制組則無。得到結果如下：(1)教導測驗編製原則，有助提高學生測驗智巧的行為，但不一定會影響冒險取向。(2)實驗組學生的測驗智巧有隨年級而發展的趨勢。(3)學生的自我概念和測驗智巧與冒險取向間亦無實質關係存在。

就心理測驗的實施而言，一般認為題目的內容、隨機誤差會影響測驗的結果，但除此之外，心理學者們認為尚有其他因素也會影響測驗的分數 (Ebel and Damrin, 1960; Thorndike, 1949; Vernon, 1962)。測驗智巧 (test wiseness) 和測驗的冒險取向 (risk taking) 便是其中的兩個因素。在國內，心理測驗學者們談論測驗的編製時，也一再的強調，如果測驗編製的技巧不好，編製出來的題目常易對受測者產生暗示作用，致受測者猜答而得分。Millman, Bishop 和 Ebel (1965) 將測驗智巧 (TW) 定義為：「一個受試者能利用測驗的特性 (characteristics) 和形式 (format)，以及測驗的情境而獲取高分的能力」。而 Slakter (1967) 則認為測驗冒險取向 (RT) 是指受試者在知道答錯會有倒扣情形下，依然進行猜答的行為。過去有些研究者比較中國學生與美國學生的測驗智巧和冒險取向，均發現中國學生在這兩方面都比美國學生為低 (Lo and Slakter, 1973; Wu and Slakter, 1978)。到底是何因素造成這種結果，至今未見實證性研究加以探討，因此，本測驗的目的想了解教導學生一些測驗編製的原則，是否會提高學生的智巧，並進而提高其冒險取向。

Dimond 和 Evans (1972) 發現學生智力和 TW 兩者之間並無顯著的關係存在。吳鐵雄 (民 73) 利用中國學生為對象，則發現 TW 和智力測驗分數有正相關存在，似乎有顯示 TW 與智力發展的關係。Slakter, Koehler 和 Hampton (1970a) 以美國五年級到十一年級學生為對象，研究 TW 與年級之關係，發現兩者有顯著關係，而且在十個年級中，呈直線趨向，他們發現年級愈高，TW 的分數也愈高。但是同樣用美國學生為對象，Ardiff (1965) 研究三年級至六年級學生的 TW 行為，都發現 TW 和年級沒有關係。

以中國小學四年級到國中三年級學生為對象，吳鐵雄 (民 73) 發現年級和部分測驗智巧的行為有顯著關係存在。但是一般研究都發現，TW 和性別並無關係 (吳鐵雄，民 73; Lo and Slakter, 1973; Slakter, Koehler, and Hampton, 1970a)。在 TW 的泛文化研究方面，Vernon (1962) 比較 108 位英國學生和 75 位美國學生在字彙和閱讀的測驗分數，他發現英國學生在字彙測驗上的得分比美國學生為高，但是在選擇式的閱讀測驗上，兩者却無差異。在比較 TW 與不同文化背景的關係，Millamam 和 Setijadi (1966) 發現美國學生在客觀式測驗上的得分確實比印尼學生的得分為高。

* 筆者特別感謝王瑪麗、劉英台兩位老師在實驗過程中的協助，同時也要感謝師大教心系夜間部朱雅玲和石兆蓮等同學在資料整理與建檔的幫助。

尤其當測驗題目中包含了明顯錯誤選項時，美國學生的得分更高，顯示出美國學生的TW比印尼學生為高。但是印尼學生在論文式測驗上的得分却比美國學生為佳。Lo 和 Slakter (1973) 比較中美兩國高二學生的TW，發現中國學生在四種TW策略上都比美國學生為低。利用國小五年級、國中二年級和高中二年級學生為對象，Wu 和 Slakter (1978) 亦發現各年級的中國學生其TW分數亦比同年級的美國學生為低。

在測驗冒險取向 (RT) 方面，Slakter 等人 (1971) 以客觀式的測驗為工具，研究年級和RT之關係，在利用美國五年級和十一年級學生為對象，他們發現無論是紐約州或密西根州學生，低年級學生的RT分數均比高年級學生為高。但這種年級的影響只能解釋1%的變異量。在以國小五年級、國中二年級和高中二年級學生為研究對象，Wu 和 Slakter 却發現受試者在RT測驗的得分有隨年級增高的趨勢。吳鐵雄 (民73) 的研究亦發現同樣現象，且呈三次曲線趨向。RT在性別差異的研究，研究者曾利用各種工具，所得到的結果却甚為不一致。有的研究發現RT與性別並無固定的關係 (Wallach and Kogan, 1961)；有的却發現男孩子的冒險行為高於女孩子 (Kass, 1964)。而Slovic (1966) 的研究却發現在低年齡組裏，男女孩的冒險行為無差異。但在高年齡組裏，則男孩子表現較高的冒險性。利用客觀式的測驗為工具，Slakter (1967), Slakter 和 Cramer (1969), Slakter 和 Koehler (1968) 均發現大學男女學生的RT得分並無顯著差異。但是當利用八年級學生為對象時，Slakter (1969) 却發現女生比同年級的男生在冒險取向的得分為高。吳鐵雄 (民73) 却發現在中國學生中，性別與RT測驗的得分無關。在對不同文化背景學生的研究中，Slakter (1969) 發現加拿大八年級學生與美國十二年級學生的RT的平均得分相似。Lo 和 Slakter (1973) 比較中國學生和美國學生在RT測驗的平均得分，發現兩者相當接近。Wu 和 Slakter 的研究亦有類似發現，但是發現中美兩國不同年級學生在RT的平均得分上，却有不同的趨勢。在中國學生中，高年級學生RT分數較高，而美國學生則是低年級學生的得分較高。

研究者曾探討測驗智巧是否與經驗或教學有關。在1920年代和1930年代裏，一系列的研究顯示出：如果利用類似的材料來練習，則學生在比西量表上的得分會提高 (Rugg 和 Colloton, 1921)。而在英國使用很多的 Morary House Test，學生在此測驗上的結果也比用其他測驗的結果，其IQ提高了9個單位 (Vernon, 1954)。但是在SAT測驗上，這種練習的效果卻沒有這麼明顯，研究者認為可能是由於受試者本身的原因 (Levine and Angoff, 1961)。Vernon (1962) 在比較不同文化背景學生的TW時，發現英國學生在經過練習之後，其在客觀式題目的閱讀測驗上得分有顯著的進步，而美國學生則沒什麼進步，他將原因歸諸於受測者對測驗型態的熟悉。這種結果似乎暗示了練習經驗對測驗分數的影響，與文化背景之不同有關。另外，很多研究的發現，認為TW行為可以學習 (Gibb, 1964; Moore, Schutz and Baker, 1966; Slakter, Koehler and Hampton, 1970b)，在我國則尚未有類似研究。

因而，本研究根據一般測驗編製的重要原則，編一個小小的教材，教導學生這些原理原則，用以了解這種教學是否對學生的TW產生影響，以探討下列的問題：

1. 國中學生在教學測驗編製原理原則之後，是否會提高其測驗智巧？
2. 國中學生了解測驗編製的基本原則之後，會不會因而提高其冒險取向的得分？
3. 探討測驗智巧與冒險取向與學生自我概念是否有關係存在？

方 法

一、受 試 者

本研究受試者取自臺北市芳和國中和弘道國中。共取國中一年級至三年級男女學生共 787 人，其

中男生 384 人，女生 403 人。從上述文獻中可知，過去的研究雖然均顯示性別與測驗智巧和冒險取向並無穩定關係存在，但為了儘可能避免性別因素對研究結果的影響，取樣時已考慮到男女人數的分配。其人數分配如表一所示。同時，因為此兩校二、三年級均採能力分班，所以在取樣時，所有學生均來自中段或中段上下班級。在選取樣本學校中，芳和國中為男女合班，因此每年級取樣二班，然後隨機分派一班為實驗組，一班為控制組；而弘道國中則為男女分班，故每年級取樣四班，男女班級各半，且隨機分派男女生班各一班為實驗組、另各一班為控制組。

表一 取樣受試人數

		一 年 級	二 年 級	三 年 級	小 計
實 驗 組	男	66	62	57	185
	女	71	63	76	210
控 制 組	男	67	70	62	199
	女	70	60	63	193
小 計		274	255	258	787

二、研究工具

(一)標準教育智力測驗 (Standard Educational Intelligence Test, SEIT)

本測驗是 Slakter 在1967年所編製。整個測驗共分兩部分，第一部分測驗冒險取向 (RTOOE)，第二部分測量測驗智巧。RTOOE 包括10組同義字或反義字，其中五組為有意義的字，而另五組則為無意義字所組成。根據 Slakter 等人 (1971) 的報告，原測驗之年級信度分別為 .68到.88。這部分測驗因中英文文字結構不同，由筆者將之改為中文，其信度與原測驗大略相近 (Wu and Slakter, 1978)。下面為其中的一個題目：

謀——闖

因為中文字中並無闖字，因此本題實際上並無答案。在作答說明中強調如果答錯的話，將倒扣分數；如果不作答則不給分也不扣分。因此，無論受試者對本題答為「同義」或「反義」，都算他顯示出測驗的冒險取向。本部分的分數便是受試者對無意義題目作答的比率，即對無意義題目作答的題數除以 5。測量測驗智巧的題目共有44個四選一的選擇題，其中28題是有意義的題目，16題則為無意義題目，以測量下列四種不同的測驗智巧行為 (Millman 等, 1965)，每種行為四題：

1. 題幹選項 (Stem-Options, ST)：由題幹的暗示，而從選項中選答。
2. 荒謬選項 (Absurd-Options, AO)：剔除顯然是錯誤的選項，而從剩餘的選項中選答。
3. 類似選項 (Smilar-Options, SO) 剔除意義相似的選項，而選答只剩下的選項。
4. 特殊限定 (Specific-Determiners, SD)：剔除選項中包含有暗示性的字詞如「絕不」、「一定」等，而猜對答案。

Millman 等人(1965)認為上述四種的前後兩種屬於利用線索的策略 (cue-using strategies)；而中間兩種則屬於演譯推理 (deductive reasoning)。下面是「類似選項」的一個題目：

33. 算術的結變原則是用在

1. 算 加 法
2. 算 除 法
3. 算 和 數
4. 算 總 計

因為算術上並無「結變原則」，這題本無答案，但是，在四個選項中因為第 1、3 和 4 項的意義

一樣，所以本題的答案便是 2。四類 TW 的分數便是受試者「答對」這些題目的題數。受試者如果具備上述四種 TW 行為，則他的得分必高。

(二) 田納西自我概念量表

本量表為林邦傑（民 69 年）根據 Fitts (1965) 所編的 Tennessee Self Concept Scale 所修訂而成。全測驗包括 70 個自我描述的句子。每個題目都有「完全不相同」、「大部分不相同」、「部分相同、部分不同」、「大部分相同」和「完全相同」五個程度不同的答案。如果受試者答「完全不相同」，便給 1 分，答「完全相同」則給 5 分。但其中有部分題目為反向記分。測驗結果可分別計算出下列十個分數：

1. 生理自我 (Physical Self, PH)：用以表示受試者對其身體、健康狀態、身體外貌、技能與性方面的感覺。
2. 道德倫理自我 (Moral-Ethical Self, ME)：用以表示受試者對其道德價值、宗教信仰、好壞人之看法。
3. 心理自我 (Personal Self, PER)：用以表示受試者對其個人價值及能力等的評價。
4. 家庭自我 (Family Self, FA)：用以表示受試者對自己作為家庭中的一份子的價值感及勝任感。
5. 社會自我 (Social Self, SO)：它反映受試者在與他人交往中的價值感及勝任感。
6. 自我認同 (Identity, ID)：它反映着受試者對自我現況的描述。
7. 自我滿意 (Self Satisfaction, SA)：它反映着受試者對自己現況的滿意或接納的程度。
8. 自我行動 (Behavior, B)：它反映受試者在接納自己或拒絕自己後，實際所採取的應對行動，或表現在外的行為。
9. 自我總分 (Total Positive, TOT)：它是上述第 6、7、8 項的和，表示受試者對整個自我的看法，得分越高，表示愈喜歡自己、信任自己，認為自己是個有價值的人，並且依此行事。
10. 自我批評 (Self Criticism, SC)：用以測量受試者自我防衛或自我批評。

根據林邦傑（民 69）的報告，本測驗之穩定係數為 .67 至 .85，而折半相關則為 .66 至 .95。

三、研究程序

筆者自一般學者（簡茂發和郭生玉，民 69 年；陳英豪和吳裕益，民 71 年）認為編寫選擇題應注意之原則中，選取較重要的六種，包括選項中避免使用絕對性字句、題幹與選項之間應有邏輯性存在、各選項之間必須互相獨立、題幹應避免暗示答案、各選項為正確答案之次數應大約相同、和正確答案應隨機排列等。利用這些原則，編寫成一份簡單的材料。當樣本選定之後，用這份教材教實驗組，而控制組則不接受教學。經過約一星期後，兩組學生分別實施標準教育智力測驗和田納西自我概念量表。所有測驗工作均在一週內完成。教學與測驗之實施均由取樣的兩所學校受輔導與測驗專業訓練的輔導老師負責。

四、資料處理

測驗結果之記分全部由電腦處理，其中尤其田納西自我概念量表，記分過程相當繁複，用電腦記分應能使人為的錯誤減至最低程度。資料的分析採用獨立樣本雙因子變異數分析和相關係數分析。在變異數分析中，如交互作用達顯著水準，則進行單純主要效果考驗。全部分析均以 .05 為顯著水準。

結果與討論

一、測驗智巧

雖然在資料中包括性別，但由於過去的研究均發現男女學生在測驗智巧上並沒有顯著差異，因此

在統計分析時並不考慮性別因素。利用年級和組別為自變數，所得各組的平均數與標準差呈現在表二。

表二 四種測驗智巧分數的平均數與標準差

	一年級		二年級		三年級		全部	
	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD
ST								
實驗組	2.11	.99	2.19	.96	2.06	1.00	2.12	.99
控制組	2.04	1.01	1.68	.98	2.00	.99	1.91	1.01
AO								
實驗組	1.85	1.05	2.00	.90	1.77	.92	1.87	.96
控制組	1.71	.99	1.89	1.13	1.76	1.06	1.78	1.06
SO								
實驗組	1.18	.88	1.37	1.05	1.22	.97	1.25	.97
控制組	.92	.78	.89	.85	1.10	.99	.97	.87
SD								
實驗組	1.34	1.15	1.67	1.46	2.00	1.50	1.67	1.40
控制組	1.08	1.04	.81	.85	1.05	.98	.98	.97

人數不等變因子變異數分析後，其結果呈現在表三，其分佈圖則如圖一所示。表三顯示：在ST分數上，年級與組別交互作用與組別均達到顯著水準，因此進行組別的單純主要效果分析，結果顯示只有在二年級學生中，實驗組的平均數高於控制組的平均數 ($F=17.15, P<.01$)。在AO分數上

表三 年級×組別測驗智巧之變異數分析結果

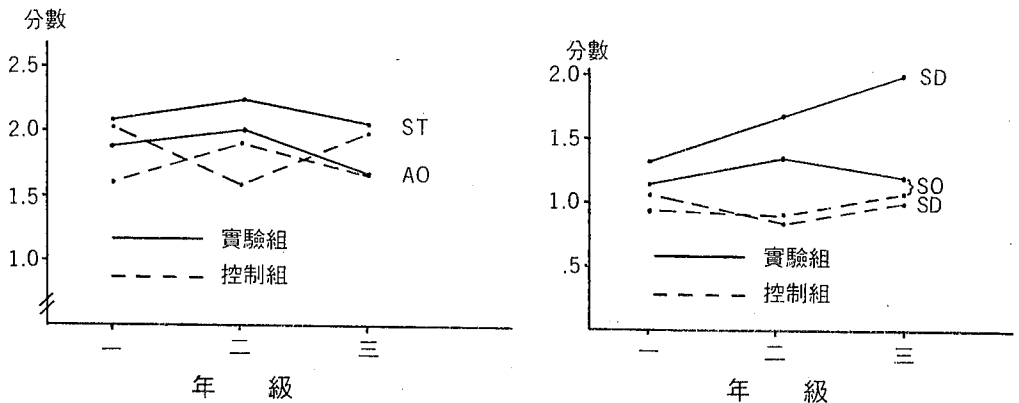
	ST		AO		SO		SD	
	F	P	F	P	F	P	F	P
年級	1.45	.235	2.40	.091	.98	.375	5.69**	.004
組別	8.75**	.003	1.66	.198	18.57**	.000	64.98**	.000
年級×組別	4.46*	.012	.30	.738	2.41	.090	6.67**	.001

* $P<.05$

** $P<.01$

，年級、組別和兩者的交互作用均沒有顯著差異。至於SO分數，只有組別的差異達到顯著水準，由表二可以看出實驗組的平均數高於控制組的平均數。在SD方面，年級與組別交互作用及兩個主要效果考驗均達到顯著水準。單純主要效果考驗結果，實驗組和控制組學生的平均數在國中一年級 ($F=3.75, P<.05$)、二年級 ($F=33.67, P<.01$)、和三年級 ($F=41.30, P<.01$) 均有顯著差異，且都是實驗組高於控制組。至於年級差異的考驗，則只有在實驗組達到顯著水準 ($F=10.30, P<.01$)。從表二可見，實驗組學生的SD平均分數有隨年級而呈直線發展之趨勢。

由變異數分析的結果，可知測驗智巧與年級之間並無明顯的關係。這種結果與過去一些研究並不



圖一 各年級測驗智巧的趨向

相符。無論在美國的研究 (Slakter 等人, 1970a), 或是以中國學生為對象 (吳鐵雄, 民 73; Wu and Slakter, 1978), 均發現學生的測驗智巧有隨年級而發展的現象。但是, 這些研究的對象, 年級的分佈都比較寬。例如 Slakter (1970a) 與 Wu 和 Slakter (1978) 的研究, 都是以國小至高中學生為對象, 而吳鐵雄 (民 73) 之研究對象也包括國小四年級到國中三年級的學生。這種由於受試者年級分佈寬狹之不同與 T W 結果有差別之現象, 似乎說明了測驗智巧的發展需要較長的時間才能顯示出差別, 也就是說在國中階段, 學生的測驗智巧並沒有顯著的發展。

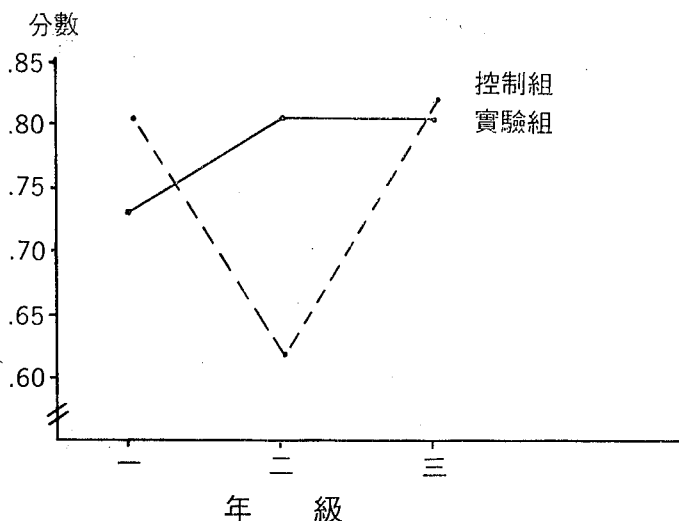
上面所引述的研究均發現: 在四種測驗智巧的行為中, S T 和 A O 的平均數均高於 S O 和 S D 的平均數。顯示出一般學生較具備從題幹與選項的相似以及從剔除荒謬選項中選取正確答案的測驗智巧, 但是却較無法從剔除類似選項及剔除含特殊字詞選項的測驗智巧。本研究之控制組學生亦表現出類似的現象, 他們的 S O 和 S D 平均數低於機率的平均值。但是, 實驗組學生的 S D 得分却顯著的提高, 似乎又顯示出國中學生容易自原則的教學中獲利。在實驗組與控制組差異的統計分析結果中, 只有 A O 的得分兩組學生沒有差異, 這種結果似可歸因於教材的內容。在教學中, 對於剔除荒謬選項原則的說明較其他三種測驗智巧行為不明顯, 因此我們可以了解, 只要針對測驗編製原則教導學生, 學生便能因而較能猜對答案。Bloom 和 Droder (1950) 的研究亦發現訓練能提高學生在問題解決測驗的得分, 這種發現印證了本研究的結果。

二、測驗的冒險取向

在三個年級中, 實驗組和控制組學生 R T 的平均分數與標準差如表四所示, 其現象如圖 2。由於各組人數並不完全相等, 因此進行人數不等雙因子變異數分析, 其結果顯示組別與年級的交互作用達到顯著水準 ($F = 13.10, P < .01$), 單純主要效果的考驗結果顯示出: 除了三年級外, 在國中一、

表四 測驗冒險取向的平均數與標準差

	一 年 級		二 年 級		三 年 級		全 部	
	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD
實 驗 組	.72	.33	.80	.31	.80	.34	.77	.33
控 制 組	.80	.29	.62	.37	.81	.31	.74	.34
全 部	.76	.32	.70	.35	.80	.32		



圖二 各年級學生的冒險取向

二年級學生中，實驗組與控制組有顯著的差異。一年級學生控制組高於實驗組 ($F=4.79, P<.05$)，但二年級學生則是實驗組高於控制組 ($F=23.34, P<.01$)；至於年級的差異只有在控制組達到顯著水準 ($F=16.27, P<.01$)，Scheffé 事後比較的結果發現一年級和三年級均高於二年級。

一般而言，本研究的結果顯示出年級間RT得分的差異較不明顯，此結果與Slakter等人(1971)、Wu和Slakter(1978)、及吳鐵雄(民73)之研究不盡相似。這種現象似也是由於年級範圍不同所產生的結果。因為對照本研究受試者RT的平均數與吳鐵雄(民73)研究中國中學生的結果，兩者相當接近，而且比該研究中國小學生的平均數為高，似乎支持筆者在民國73年研究中的推論，我國國中學生面對激烈的升學競爭，養成了傾向冒險的測驗行為。有趣的是，測驗編製原則的教學效果，在不同的年級有不同的結果，對一年級學生而言，這種教學降低了冒險取向，但對二年級學生却有助長冒險取向的現象，而對三年級學生則無影響，這是什麼原因，無法了解，有待進一步的研究。

三、測驗智巧和冒險取向與自我概念的關係

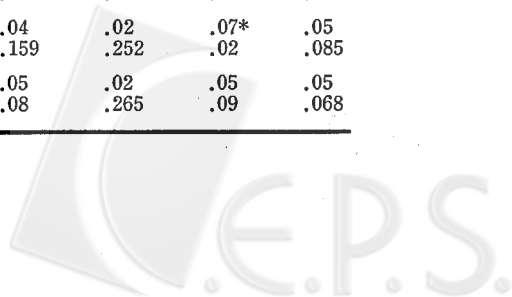
表五呈現TW和RT與自我概念的相關係數。由表中可見，雖然四種測驗智巧行為與部分自我概念之間有正相關存在，却找不出一個固定的結論，這種結果似在預料之中。而且，即使相關係數達到顯著水準，其數值一致偏低，自我概念所能解釋TW的變異量都不足1%，這種相關實不足取。

表五 測驗智巧與冒險取向和自我概念之相關係數

		ID	SA	B	PH	ME	PER	FA	SO	TOT
RT	r	.06*	.04	.07*	.04	.04	.04	.06*	.06*	.06*
	P	.039	.156	.025	.115	.112	.149	.045	.042	.045
ST	r	.06*	.08*	.08*	.07*	.09**	.08*	.03	.07*	.08*
	P	.049	.013	.013	.033	.007	.017	.238	.02	.013
AO	r	.02	.03	.03	-.01	.07*	-.003	.04	.02	.03
	P	.259	.229	.242	.345	.03	.467	.114	.308	.223
SO	r	.05	.04	.04	.05	.01	.04	.02	.07*	.05
	P	.06	.131	.122	.067	.340	.159	.252	.02	.085
SD	r	.07*	.02	.05	.06	.04	.05	.02	.05	.05
	P	.02	.259	.064	.053	.131	.08	.265	.09	.068

* $P<.05$

** $P<.01$



另一方面，冒險取向與家庭自我、社會自我、自我認同、自我行動和自我總分之間存在着顯著的正相關，此種結果與筆者的預料相當接近，顯示了學生愈是喜歡自己、信任自己，則愈傾向於在測驗時喜歡冒險，因為他相信自己的能力，認為猜對答案的機會較大。但是，相同的，自我概念解釋冒險取向變異量均甚低，這種相關可能與人數有關，實沒有太多的實質意義，因此比較可靠的結論有待進一步探討。

結論與建議

本研究以臺北市兩所國中一至三年級學生為對象，主要在探討教導學生有關測驗編製的原則，是否有助在不知答案的選擇題中猜對答案，並研究學生自我概念與測驗智巧和冒險取向的關係，結果可歸納為下列幾點：

1. 如果學生了解心理學者所提編製選擇式試題的原則，則有助於他們在不確定答案時猜中答案而得分的機會，尤其剔除含有絕對字詞選項的教學，效果更大。
2. 學生在接受測驗編製原則的教學之後，並不一定會改變其冒險取向。
3. 實驗組學生的測驗智巧有隨年級而發展的趨勢，唯此種現象並不存在於控制組學生。
4. 一般而言，學生的自我概念和測驗智巧及冒險取向之間，並無實質關係存在。

本研究結果似乎肯定了編製測驗原則的教學對國中階段學生的效果，這種發現支持教育行政單位提倡改進教師命題技巧的政策。Dimond 和 Evans (1972) 認為測驗智巧並不是一種普通能力，而是一種特殊的能力，只有在試題中含有若干線索時才能發生效力。因此，教師在編製測驗時，應時時注意避免提供學生任何答案的暗示。相反的，如果國中教師想幫助他們的學生在考試中得較高的分數，則可以在平時多教導學生有關編製測驗原則的知識。

參考文獻

- 吳鐵雄 (民73年) 國民中小學男女學生測驗智巧和冒險取向。師大教育心理學報，17期，121—130頁。
- 林邦傑 (民69年) 田納西自我概念量表指導手冊。臺北市，正昇教育科學社。
- 陳英豪、吳裕益 (民71年) 測驗的編製與應用。臺北市，偉文圖書公司。
- 簡茂發、郭生玉 (民69年) 測驗的編製。楊國樞等：社會及行為科學研究法 (上)。臺北市，東華書局。
- Ardiff, M. B. (1965) *The relationship of three aspects of test-wiseness to intelligence and reading ability in grades three and six*. Unpublished master's thesis. Cornell University.
- Bloom, B. S. & Broder, L. J. (1950) *Problem-solving processes of college students: An exploratory investigation*. Chicago: The University of Chicago Press.
- Dimond, J. J. & Evans, W. J. (1972) An investigation of the cognitive correlates of test-wiseness. *Journal of Educational Measurement*, 9, 145-150.
- Ebel R. L. & Damrin, D. E. (1960) Tests and examinations. In Harris, C. (ed.) *Encyclopedia of Educational Research*. New York: The Macmillan Company, 1502-1517.
- Gibb, B. G. (1964) *Test-wiseness as secondary cue responses*. Unpublished Doctoral

- Dissertation, Stanford University. Ann Arbor, Michigan: University Microfilms, No.64-7643.
- Kass, N. (1964) Risk in decision making as a function of age, sex, and probability preference. *Child Development*, 35, 577-582.
- Levine, R. S. & Angoff, W. H. (1961) The effects of practice and growth on scores on the Scholastic Aptitude Test. *Research and Development Report*, No. 58-6/SR-586. Princeton, N. J: Educational Testing Service.
- Lo, M. I. & Slakter, M. J. (1973) Risk taking and test-wiseness of chinese students. *The Journal of Experimental Education*, 5, 56-59.
- Millman, J., Bishop, H. & Ebel, R. (1965) An analysis of test-wiseness. *Educational and Psychological Measurement*, 25, 707-726.
- Millman, J. & Setijadi. (1966) A comparison of the performance of American and Indonesian students on three types of test items. *The Journal of Educational Research*, 59, 273-275.
- Moore, J. C., Schutz, R. E. & Baker, R. L. (1966) The application of a self-instructional technique to develop a test-taking strategy. *American Educational Research Journal*, 3, 13-17.
- Rugg, H. & Colloton, C. (1921) Constancy of the Stanford-Binet I. Q. as shown by retests. *Journal of Educational Psychology*, 12, 315-322.
- Slakter, M. J. (1967) Risk taking on objective examinations. *American Educational Research Journal*, 4, 31-43.
- Slakter, M. J. (1969) Generality of risk taking on objective examinations. *Educational and Psychological Measurement*, 29, 115-128.
- Slakter, M. J. & Cramer, S. H. (1969) Risk taking and vocational or curriculum choice. *Vocational Guidance Quarterly*, 18, 127-132.
- Slakter, M. & Koehler, R. A. (1968) A new measure of risk taking on objective examinations. *California Journal of Educational Research*, 19, 132-137.
- Slakter, M. J., Koehler, R. A. & Hampton, S. H. (1970a) Grade level, sex, and selected aspects of test-wiseness. *Journal of Educational Measurement*, 7, 119-122.
- Slakter, M., Koehler, R. A. & Hampton, S. H. (1970b) Learning test-wiseness by programmed test. *Journal of Educational Measurement*, 7, 247-254.
- Slakter, M. J., Koehler, R. A. & Hampton, S. H. (1971) Sex, grade level, and risk taking on objective examinations. *The Journal of Experimental Education*, 3, 65-68.
- Slovic, P. (1966) Risk taking in children: Age and sex differences. *Child Development*, 37, 169-176.
- Thorndike, R. L. (1949) *Personnel Selection: Test and measurement techniques*, New York: Wiley.
- Vernon, P. E. (1954) Symposium on the effects of coaching and practice in intelligence tests: Conclusions. *British Journal of Educational Psychology*,

24, Part 2, 57-63.

Vernon, P. E. (1962) The determinants of reading comprehensions. *Educational and Psychological Measurement*, 22, 269-286.

Wallach, M. A. & Kogan, N. (1961) Aspects of judgement and decision-making: Interrelationship and changes with age. *Behavior Science*, 6, 23-26.

Wu, T. H. & Slakter, M. J. (1978) Risk taking and test wiseness of chinese student by grade level and residence area. *Journal of Educational Research*, 71, 167-170.



Bulletin of Educational Psychology, 1985, 18, 85—96.
Taiwan Normal University, Taipei, Taiwan, China.

EXPERIMENT STUDY ON JUNIOR HIGH STUDENTS' TEST-WISENESS AND RISK TAKING

TIEH-HSIUNG WU

ABSTRACT

The purpose of this study is to investigate the relationships between teaching of the principles of item construction grade and test-wiseness (TW) and risk taking (RT). And this research also explores the relationships between self concept and the two testing behaviors. The subjects of 787 students were selected from two junior high schools in Taipei City. There were 274, 255, and 248 first, second, and third year students, respectively.

Subjects were randomly assigned to the experiment group and control group. The students of the experiment group were given a short instruction of the material related to the principles of item construction, while students of the control group were not. About three days after the instruction, both groups were given Standard Educational Intelligence Test for testing students' TW and RT.

The main findings of the study are:

1. Teaching of the principles of item construction improves students' score on TW, but does not necessarily increase risk taking.
2. For the experiment group, there is evidence that the higher the grade, the higher the average TW score.
3. There is no substantial relationship between self concept and students' TW and RT.

