

表 4.20 是台灣區及各分區五年級甲式測驗與六年級乙式測驗兩式測驗題目參數連結的斜率及截距。表內數據是用「平均數、標準差法」所得到的轉換常數，轉換方法與前述高雄市與台南市之相互轉換相同。

表4.20 常數樣本五(甲式)、六(乙式)年級兩式測驗題目參數連結的斜率及截距

地區	連結方式	斜率	截距
台灣省	六(乙)→五(甲)	1.2184	.5295
	五(甲)→六(乙)	.8208	-.4346
台北市	六(乙)→五(甲)	.8578	.3468
	五(甲)→六(乙)	1.1658	-.4043
高雄市	六(乙)→五(甲)	.9219	.5580
	五(甲)→六(乙)	1.0848	-.6054
台灣區	六(乙)→五(甲)	.8903	.5056
	五(甲)→六(乙)	0.1232	-.5678

## 五、常模研究

### (一)高雄市與台南市常模

本研究第一年僅建立台南市及高雄市五、六年級常模。甲、乙兩式雖然只分別有高雄市及台南市常模，但可透過等百分等級等化法，使兩式測驗均分別有兩市的常模。表 4.21 是兩式測驗高雄市及台南市標準化樣本的平均數及標準差。百分等級，T 分數（直線轉換）及 DIQ 常模請參見附錄所附資料。DIQ 的平均數是 100，標準差是 15，與魏氏智力測驗相同。

表4.21 兩式測驗次數分佈的平均數及標準差

測 驗	年級	N	M	S D
甲 式 (高雄市)	五	1066	27.701	8.873
	六	1078	32.510	10.125
乙 式 (台南市)	五	1163	25.864	7.979
	六	1220	29.983	9.332

## (二)台灣區及各地區常模

本研究第二年以較具代表性的常模樣本來建立台灣區及台灣省、台北市及高雄市三個地區性常模。建立全台灣區常模時，台北市與高雄市的受試樣本沒有全部採用，而是依據該二區佔全台灣區之比率，從新抽取適當人數與台灣省全體樣本共同組成全台灣區常模樣本。表 4.22 及表 4.23 分別是全台灣區及各地區國小六年級（乙式）及五年級（甲式）的得分之基本統計及內部一致性信度。由於不同地區之平均數及標準差均不同，因此，除了建立全台灣區之百分等級、T分數（直線轉換）及 DIQ（ $\mu = 100$ ， $\sigma = 15$ ）常模表外，也建立分區常模，這些常模表均列在附錄部份。

另外，依據常模樣本所得到的內部一致性信度均在 .90 附近，大致尚稱理想。

表4.22 各地區國小學術性向測驗六年級乙式之基本統計及內部一致性

統計數	台灣省	台北市	高雄市	全台灣區
題數	60	60	60	60
平均	28.695	34.963	33.696	29.372
標準差	10.776	9.704	10.845	10.888
Cronbach $\alpha$ 係數	.903	.882	.906	.905
測量標準誤	3.356	3.331	3.321	3.353
統計數	男生		女生	
題數	60		60	
平均	30.224		29.218	
標準差	11.334		10.376	
Cronbach $\alpha$ 係數	.912		.896	
測量標準誤	3.355		3.346	

表4.23 國小學術性向測驗傳統題目之信度(五年級甲式)

統計數	台灣省	台北市	高雄市	全台灣區
題數	60	60	60	60
平均	23.939	31.271	28.276	25.134
標準差	8.938	9.319	9.272	9.463
Cronbach $\alpha$ 係數	.860	.870	.871	.874
測量標準誤	3.361	3.364	3.327	3.360

  

統計數	男生	女生
題數	60	60
平均	25.470	24.747
標準差	9.951	8.869
Cronbach $\alpha$ 係數	.885	.858
測量標準誤	3.371	3.344

## 六、題目相對偏失 (bias) 研究

題目偏失的定義很多，研究的方法也很多，但卻一直沒有那一種定義或方法能普遍被接受。就 IRT 的觀點來說，能力相同的不同群體受試，如果答對某一題的概率 (P) 不一樣，那便是有題目偏失。此定義顯然相當合理，但卻沒有辦法真正用來找出那些題目是有偏失的，因為受試的真正能力是無法知道的。受試的能力只能用所要評鑑有無偏失的試題來估計，如果那些試題本身是有偏失的，那所估計的能力也必然有系統性的偏差，即使在評鑑某一試題時，只使用其他的試題來估計受試的能力，其他的試題仍有可能有偏失，因此，仍然無法得到不同群體受試能力的不偏估計值。可見題目偏失的研究除了定義上有爭議外，研究方法本身也有難以克服的問題。

本研究是以不同群體所估計的題目特徵曲線之差異情形，作為評鑑題目之相對偏失的指標。以題目的性別偏失而言，研究者分別以男生及女生受試樣本來估計每個題目的  $a, b, c$  三個參數。估計題目參數時是將受試的能力標準化 ( $\mu = 0, \sigma = 1$ )，所得到的題目參數會受到受試樣本能力