

第四章 研究結果

第一節 傳統題目統計數與測驗信度

本研究以 ITEMAN 軟體進行各分測驗各式別的傳統題目分析，此軟體可以得到各題目的難度(P)及兩種鑑別度(r_{bis} 及 r_{pb})指數， P 是答對百分比， r_{bis} 及 r_{pb} 均是題目得分與總分之相關係數， r_{pb} 將題目得分視為自然的二分變項(答對得 1 分，答錯得 0 分)， r_{bis} 則將題目得分視為連續的常態分佈變項，只是用人為的方式界定為二分變項。 r_{pb} 受 P 之影響較大， P 較極端時， r_{pb} 與 r_{bis} 之差距較大。此兩種鑑別度係數雖然不是完全正相關，但二者之相對地位是一樣的。全測驗 P 的平均值在 .393 至 .741 之間，「加減法」分測驗最高，「圖形與空間」最低。 r_{pb} 的平均值在 .326 至 .547 之間， r_{bis} 的平均值在 .455 至 .762 之間。各分測驗 Cronbach α 係數在 .621 至 .938 之間。表 4.1 至 4.10 為各年級受試學童在各分測驗的測驗分析摘要表。

表 4.1 國小數學診斷測驗傳統題目參數摘要表

第一分測驗：認數						
統計數	一年級	二年級	三年級	四年級	五年級	六年級
題數	28	26	37	41	40	43
受試人數	519	546	542	540	566	593
平均數	18.118	18.357	21.181	28.102	27.564	29.889
標準差	7.064	5.039	6.015	8.288	8.863	9.506
α 係數	0.922	0.867	0.834	0.910	0.923	0.930
測量標準誤	1.974	1.841	2.453	2.489	2.455	2.507
P 的平均數	0.647	0.706	0.572	0.685	0.689	0.695
r_{pb} 的平均數	0.564	0.495	0.380	0.472	0.506	0.513
r_{bis} 的平均數	0.759	0.730	0.535	0.659	0.692	0.708

表 4.2 國小數學診斷測驗傳統題目參數摘要表

第二分測驗：分數				
統計數	三年級	四年級	五年級	六年級
題數	24	24	40	52
受試人數	540	532	528	581
平均數	13.302	14.942	24.189	33.983
標準差	4.131	4.410	8.089	11.845
α 係數	0.786	0.798	0.894	0.941
測量標準誤	1.911	1.983	2.635	2.880
P 的平均數	0.554	0.623	0.605	0.654
r_{pb} 的平均數	0.410	0.424	0.441	0.501
r_{bis} 的平均數	0.578	0.587	0.586	0.667

表4.3 國小數學診斷測驗傳統題目參數摘要表

第三分測驗：圖形與空間						
統計數	一年級	二年級	三年級	四年級	五年級	六年級
題數	9	28	38	41	34	37
受試人數	527	555	551	545	567	633
平均數	4.875	10.995	19.305	21.646	16.259	18.964
標準差	2.300	3.796	5.587	7.038	5.813	6.360
α 係數	0.696	0.621	0.772	0.838	0.797	0.817
測量標準誤	1.269	2.336	2.667	2.830	2.619	2.722
P 的平均數	0.542	0.393	0.508	0.528	0.478	0.513
r_{pb} 的平均數	0.538	0.395	0.326	0.366	0.354	0.360
r_{bis} 的平均數	0.697	0.489	0.437	0.477	0.455	0.467

表4.4 國小數學診斷測驗傳統題目參數摘要表

第四分測驗：加減法						
統計數	一年級	二年級	三年級	四年級	五年級	六年級
題數	35	41	51	54	50	40
受試人數	565	588	590	610	618	632
平均數	20.396	25.544	33.373	40.685	40.335	31.896
標準差	7.099	8.976	9.609	10.024	9.085	7.177
α 係數	0.899	0.925	0.919	0.931	0.933	0.909
測量標準誤	2.256	2.452	2.737	2.638	2.348	2.165
P 的平均數	0.583	0.623	0.654	0.753	0.807	0.797
r_{pb} 的平均數	0.472	0.496	0.444	0.467	0.495	0.477
r_{bis} 的平均數	0.652	0.675	0.619	0.684	0.762	0.728

表4.5 國小數學診斷測驗傳統題目參數摘要表

第五分測驗：乘除法					
統計數	二年級	三年級	四年級	五年級	六年級
題數	34	49	57	59	65
受試人數	582	602	602	617	639
平均數	18.495	25.121	36.955	32.323	38.310
標準差	6.446	9.432	10.351	11.251	13.200
α 係數	0.869	0.909	0.922	0.919	0.938
測量標準誤	2.331	2.838	2.887	3.200	3.276
P 的平均數	0.544	0.513	0.648	0.548	0.589
r_{pb} 的平均數	0.432	0.428	0.437	0.420	0.452
r_{bis} 的平均數	0.590	0.574	0.619	0.559	0.606

表4.6 國小數學診斷測驗傳統題目參數摘要表

第六分測驗：四則運算			
統計數	四年級	五年級	六年級
題數	21	30	43
受試人數	434	422	445
平均數	11.763	13.706	27.609
標準差	3.774	5.497	8.655
α 係數	0.746	0.805	0.903
測量標準誤	1.901	2.425	2.693
P 的平均數	0.560	0.457	0.642
r_{pb} 的平均數	0.408	0.386	0.447
r_{bis} 的平均數	0.561	0.500	0.600

表4.7 國小數學診斷測驗傳統題目參數摘要表

第七分測驗：量						
統計數	一年級	二年級	三年級	四年級	五年級	六年級
題數	12	12	29	31	44	41
受試人數	556	542	539	539	506	562
平均數	6.930	8.893	17.336	15.870	23.698	24.066
標準差	2.791	2.326	4.761	5.063	8.486	8.815
α 係數	0.753	0.699	0.785	0.783	0.888	0.910
測量標準誤	1.386	1.275	2.208	2.358	2.845	2.640
P 的平均數	0.577	0.741	0.598	0.512	0.539	0.587
r_{pb} 的平均數	0.521	0.489	0.379	0.362	0.409	0.466
r_{bis} 的平均數	0.703	0.724	0.523	0.489	0.536	0.617

表4.8 國小數學診斷測驗傳統題目參數摘要表

第八分測驗：時間與計算						
統計數	一年級	二年級	三年級	四年級	五年級	六年級
題數	11	17	22	16	27	27
受試人數	541	547	541	587	584	611
平均數	6.854	10.543	13.985	8.560	16.731	17.982
標準差	2.747	3.344	3.479	3.039	5.222	5.860
α 係數	0.762	0.755	0.712	0.670	0.826	0.871
測量標準誤	1.340	1.655	1.868	1.746	2.176	2.108
P 的平均數	0.623	0.620	0.636	0.535	0.620	0.666
r_{pb} 的平均數	0.547	0.456	0.383	0.411	0.430	0.481
r_{bis} 的平均數	0.727	0.634	0.560	0.544	0.580	0.651

表4.9 國小數學診斷測驗傳統題目參數摘要表

第九分測驗：統計與圖表						
統計數	一年級	二年級	三年級	四年級	五年級	六年級
題數	16	26	46	33	35	30
受試人數	533	571	571	551	577	606
平均數	9.946	16.226	30.301	18.546	23.608	19.139
標準差	3.039	4.130	8.664	5.512	6.726	6.411
α 係數	0.728	0.753	0.904	0.797	0.883	0.882
測量標準誤	1.585	2.053	2.691	2.485	2.296	2.202
P 的平均數	0.622	0.624	0.659	0.562	0.675	0.638
r_{pb} 的平均數	0.452	0.385	0.437	0.366	0.449	0.477
r_{bis} 的平均數	0.632	0.560	0.621	0.486	0.627	0.640

表4.10 國小數學診斷測驗傳統題目參數摘要表

第十分測驗：輔助計算器與解題策略				
統計數	三年級	四年級	五年級	六年級
題數	13	22	24	43
受試人數	511	545	552	585
平均數	7.763	14.470	16.190	26.468
標準差	2.504	4.594	4.130	7.940
α 係數	0.599	0.832	0.781	0.884
測量標準誤	1.586	1.884	1.931	2.707
P 的平均數	0.597	0.658	0.675	0.616
r_{pb} 的平均數	0.420	0.475	0.411	0.415
r_{bis} 的平均數	0.556	0.646	0.580	0.560

第二節 題目反應理論模式題目分析

本研究所發展的數學診斷測驗，除了進行傳統的試題及測驗分析之外，也進行單參數 IRT 模式題目分析。由於各年級受試均不夠多（在 434 至 639 之間），較不適合進行二、三參數 IRT 模式題目分析。研究者選擇進行單參數題目分析，分析時選擇將受試能力標準化，可得到各題的難度值及所有題目的平均鑑別力。所有受試的能力估計值均在同一量尺上，可以相互比較。

有了 IRT 題目參數，且均在同一量尺上，那未來就可利用 IRT 的優點，發展成電腦化適性測驗（CAT），也可以較有效評估受試在每一內容領域之進展過程。

表4.11至表4.20 是各分測驗題目參數之聯合估計摘要表。單參數 IRT 的 a, b, c 分別代表題目的鑑別力、難度及擬似猜對率 ($c=0$)。除了題目參數外，也提供各題模式適合度考驗 χ^2 值， χ^2 值受樣本數影響極大，只要人數夠多，一定會達到顯著水準，不能以有無達到顯著水準來刪除或保留題目。事實上，再理想的題目也不可能完全符合單參數 IRT 模式，只要大致接近所要採用的模式即可。 χ^2 未達到顯著水準，並不代表題目完全符合理論模式，只是代表還沒有足夠證據（人數不夠多）推翻虛無假設（符合理論模式）而已。模式適合度考驗的推理邏輯與統計考驗的邏輯不符，而且恰好相反，因此各題的 χ^2 值只能作為各題相對適合度的參考指標，如果某幾個題目的 χ^2 值明顯高出其他題目很多倍，那這幾個題目的模式適合度可能較差。