

出國報告（出國類別：其他）

第十屆海峽兩岸心理與教育測驗學  
會學術研討會  
參加報告

服務機關：國家教育研究院

姓名職稱：吳慧珉 助理研究員

派赴國家：中國

出國期間：101年9月20日~9月25日

報告日期：101年12月18日

## 摘要

第十屆海峽兩岸心理與教育測驗學術研討會由「中國教育學會教育統計與測量分會」以及「中國心理學會心理測量分會」主辦於2012年9月21日(六)~24日(一)在江西師範大學召開，主要議題包含測驗理論與技術新進展、教育心理統計方法與應用、教育考試與學業評價、人事測評理論研究與應用、電腦與網路測驗的發展與認知診斷理論與技術。海峽兩岸心理與教育測驗學術研討會是兩岸測驗學術與應用領域相當知名且重要的國際會議之一。研究者主要前往發表論文並聽取其他學者發表與進行交流。

## 目次

一、 目的.....	1
二、 參加會議經過.....	1
三、 與會心得.....	5
四、 建議.....	5
五、 攜回資料名稱及內容.....	5

## 一、目的

發表論文，論文題目條列如下：

吳慧珉、陳俊華、郭伯臣、楊智為（2012）。結合建構反應題的自動化診斷測驗系統。

陳俊華、吳慧珉、施淑娟、郭伯臣（2012）。多重解題策略Q矩陣設計之探究。大會  
**壹等獎論文**

陳俊華、郭伯臣、吳慧珉、白曉珊（2012）。認知診斷測驗選題策略之比較與探究。

吳慧珉、郭伯臣、許天維、葉昶成（2012）。NEAT垂直等化設計可能值方法估計效果之探討。

許天維、吳慧珉、郭伯臣、郭秀芬（2012）。納入輔助變項於多向度IRT模式估計之探討。

## 二、參加會議經過

海峽兩岸心理與教育學術研討會每兩年舉辦一次，今年是由江西師範大學舉辦，主要議題包含測驗理論與技術新進展、教育心理統計方法與應用、教育考試與學業評價、人事測評理論研究與應用、電腦與網路測驗的發展與認知診斷理論與技術。海峽兩岸心理與教育測驗學術研討會是兩岸測驗學術與應用領域相當知名且重要的國際會議之一。研究者主要前往發表論文並聽取其他學者發表與進行交流。

### 9月22日 論文發表

發表的論文簡介如下：

#### **論文題目：多重解題策略Q矩陣設計之探究**

認知診斷測驗使用 Q 矩陣表示試題與概念的關係，並藉由受試者的作答反應組型與 Q 矩陣，推估受試者的概念精熟狀態，進而瞭解受試者的學習狀況，有效實施補救教學。因此 Q 矩陣對認知診斷測驗相當重要，相關研究顯示 Q 矩陣會影響認知診斷模型的估計。目前研究大多探討單一解題策略的 Q 矩陣對特定認知診斷模型如 DINA 模型估計的影響；de la Torre,(2008)驗證多重解題策略情境下，MS-DINA 模型的估計比 DINA 模型估計較好。本研究目的在透過模擬研究方式，探討多重解題策略的 Q 矩陣設計對 MS-DINA 模型估計的影響，使用 EM 估計方法(de la Torre, 2009)進行認知屬性與試題參數的估計，研究結果顯示 Q 矩陣具備多重解題策略的試題數較少時，模型的概念估計精準度較高。概念數較少試題具備多重解題策略時，S-DINA 模型的概念估計精準度較低。不同多重解題策略 Q 矩陣設計，試題參數估計結果一致。不適合多重解題策略 Q 矩陣設計發生的試題數較多時，MS-DINA 模型的概念估計精準度較差。不適合多重解題策略 Q 矩陣設計發生在概念數較少試題時，對 MS-DINA 模型的概念估計精準度影響較大。不適合多重解題策略 Q 矩陣設計，對 MS-DINA 模型的試題猜測度估計影響較大。

### 論文題目：結合建構反應題的自動化診斷測驗系統

建構反應題能詳細記錄學生解題過程，提供更多測驗訊息以診斷學生的學習狀況，並降低測驗不確定性如選擇題的猜測度；紙筆測驗雖然能夠得到學生詳細解題歷程，卻會增加教師的閱卷時間及負擔。因此本研究提出結合建構反應題的自動化診斷測驗系統，透過系統詳細紀錄學生的解題歷程，依據計分規則，使用區塊分析法，自動化診斷學生的錯誤類型與計分。本研究以數學科”分數的乘法”與自然科”電磁作用”兩個單元來評估系統成效，研究對象分別為 3 所學校共 6 個班級的 158 位六年級學生以及 4 所學校共 23 個班級的 569 位六年級學生。本研究的系統自動化診斷學生的數學科”分數的乘法”單元的錯誤類型正確率為 97.15%，診斷學生的自然科”電磁作用”單元的錯誤類型正確率為 98.18%，整體的錯誤類型正確率為 96.67%，研究結果顯示系統能正確且有效的診斷學生的錯誤類型。

### 論文題目：認知診斷測驗選題策略之比較與探究

古典測驗理論(classical test theory, CTT)與試題反應理論(item response theory, IRT)能提供受試者能力在團體中所佔相對位置的訊息，卻無法顯示出受試者是否精熟某種概念或技能。受到美國政府於 2002 年實施「沒有落後的孩子」(No Child Left Behind Act, NCLB)法案的影響，認知診斷模型(cognitive diagnostic models, CDMs)近年來在心理計量研究領域中備受重視，能反應受試者的概念或技能精熟狀態，因此結合認知診斷模型的測驗是個新興的議題。認知診斷模型中以 DINA(deterministic input, noisy “and” gate model; Junker & Sijtsma, 2001)模型最簡單且容易解釋，因此本研究使用 DINA 模型估計試題 CDI 訊息量(cognitive diagnostic index)，提出基於 CDI 訊息的組型層次與屬性層次加權 CDI 法(pattern and attribute level weighted CDI, PAW-CDI)，透過模擬研究方式探討在不同試題參數設定、不同概念數下、不同題數下，隨機選題(Random)、CDI、PAW-CDI 選題法成效。

### 9 月 23 日 論文發表

發表之論文簡介如下：

#### 論文題目：NEAT垂直等化設計可能值方法估計效果之探討。

本研究係以單向度單參數試題反應理論為基礎，以定錨不等組 (non-equivalent groups with anchor test design, NEAT) 垂直等化設計進行模擬施測，探討單向度納入輔助變數之可能值方法 (PV)、加入輔助變數之期望後驗估計法 (EAP\_AV)、期望後驗估計法 (EAP)、最大概似估計法 (MLE) 以及加權最大概似估計法(WLE)五種不同估計方法對於群體參數的估計效果。研究結果顯示：不管是高年級還是低年級，PV 及 EAP BV 方法估計群體能力平均數的 RMSE 皆遠優於 EAP、MLE、WLE 等三種估計方法。在施測題數為 18 題的情境中，就 PV 及 EAP BV 此二種估計方法而言，施測人數 10920 的群體能力平均數的 RMSE 皆略低於 16128 人的情境，但差異不大。而在 EAP、MLE、WLE 等另外三種方法在估計群體能力平均數方面，就低年級群組方面，此 3 種估計方法的 RMSE 亦是施測人數 10920 的群體能力平均數

的 RMSE 皆略低於 16128 人的情境，但是在高年級組部份，則是 16128 人的情境的 RMSE 略低，但差異皆不大。在施測題數為 36 題的部份，五種估計方法的 RMSE 皆是施測人數 10920 人的情境優於 16128 人的情境，除了在低年級群組中 EAP、MLE、WLE 此三種方法的 RMSE 在 2 種施測人數情境間差異較大外，其餘情境差異不大。就群體能力標準差的估計結果而言，PV 的估計結果遠優於其它四種估計方法。而不管是在題數為 18 題或是 36 題的情境中，不同施測人數間相對之各種估計法的 RMSE 均無明顯差異。

### 論文題目：納入輔助變項於多向度 IRT 模式估計

本研究採用模擬資料進行分析，探討不同估計方法在多向度試題反應理論下對群體參數估計之差異。本研究以多向度隨機係數多項洛基模式(MRCMLM)為基礎，使用模擬資料探討最大概似估計法(MLE)、加權概似估計法(WLE)、納入背景變項的期望後驗估計法(EAPBV)與可能值方法(PV)對於回覆群體參數的精確性，瞭解納入背景變項的估計方法對群體參數估計的效果。本研究結果顯示：在群體平均數估計方面，以 EAPBV、WLE、和 PV 的方法較好，在群體標準差之估計方面，則以 PV 的方法最好，顯示納入輔助變數有助於提升能力參數之估計成效，特別是對於標準差之估計。

研究者所發表的議題主要是針對大型測驗與認知診斷測驗，目前大型測驗的理論探討在中國仍屬新興領域，有多位學者針對等化與次級資料之分析等議題與研究者進行交流。而認知診斷測驗是中國目前相當熱門的研究議題，針對多重策略之認知診斷測驗與認知診斷測驗之等化等議題，中國於此方面已有不錯的研究成果，可供我們學習參考。



與大陸著名測驗學者張厚燦教授討論交流



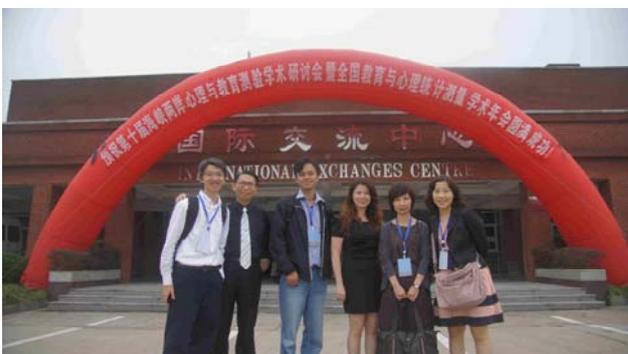
研究者進行論文發表



研討會開幕典禮



研討會閉幕典禮



與其他學者在研討會門口合影



美國學者張華華教授與臺灣學者郭伯臣教授發表演講進行交流

### 三、 與會心得

此次會議是自費參加，感謝會議主辦單位中國教育學會教育統計與測量分會的用心籌畫，使我有機會與大陸學者專家藉此學術場合得以進行深入交流。大陸近幾年於測驗領域之發展進步神速，研究主題也逐漸符合國際潮流，掌握國際評量發展趨勢，且積極與美國著名的測驗學者合作，截長補短，發展評量領域之理論與實務。形成一股不容小覷的力量。或許是競爭激烈，大陸的學生學習意願非常積極強烈，主動問問題勤作筆記，令我感觸良多！國內的學生或許是學習資源過於豐富，學習態度反倒需要加強，而另一方面也督促我自己以更嚴謹更積極的態度面對學術研究！

### 四、 建議

參與國際會議是增進專業領域之良好方式，大陸這幾年學術領域發展快速，且積極舉辦國際研討會，增加國際知名度，參加此種學術研討會有助於年輕學者擴展視野，期望相關單位能在法律範圍內，放寬補助標準，提高補助經費，特別是對於年輕之學者之補助，讓新進之年輕學者之相關研究成果能宣揚於國際，也帶動國際交流，提高台灣之國際知名度。

### 五、 攜回資料名稱及內容

第十屆海峽兩岸心理與教育測驗學術研討會會議手冊，內容包含所有發表論文摘要。另外，也獲得許多後續將舉辦的相關研討會之宣傳資訊。