第七節　辦理工作坊

壹、背景

面對全球化之浪潮，各國越來越重視自己國家的學生是否具有國際競爭力，國際大型測驗在這一方面提供一良好參考指標，因此像 TIMSS、PISA 及 PIRLS 等測驗愈來愈受到重視，這些測驗經過長久的發展與修正，能提供良好的範本，國內實施 TASA 時，如能參考這些大型測驗的實施流程與分析模式，將能讓整個研究計畫更嚴謹。

本次工作坊是邀請在大型測驗設計與分析均有深入探討與研究的學者 Dr. Margaret Wu，介紹大型測驗設計時的誤差來源，透過控制誤差來源，提升測驗參數估計的精準度。Dr. Margaret Wu 亦是著名的測驗軟體 ConQuest 的作者，ConQuest 的理論基礎是試題反應模式（Item Response Modelling, IRM），工作坊的內容涵蓋基礎理論的介紹與軟體的使用，透過這些課程，將有助於提升大型測驗資料分析之實務能力。

貳、目的

一、瞭解大型測驗設計時之誤差來源。

二、ConQuest 軟體介紹與實際操作。

三、在實際測驗如何運用可能值作次級資料分析。

四、透過可能值計算，估計 TASA 群體分數資料，包含全體受試者及縣市群體分數可能值估計。

參、工作坊時間：

2009 年 11 月 18 日（星期三）至 11 月 20 日（星期五），共三天。

肆、工作坊地點：

國立臺中教育大學，臺中市民生路 140 號

伍、工作坊內容與日程

「大型測驗暨 ConQuest 工作坊」

日程表（**Workshop Schedule**）

時間 Time		活動內容 Activity	地點 Place
第一天：2009 年 11 月 18 日 (星期三) DAY1：2009/11/18（Wed）	08:30-09:00	報到 Registration	臺中教育大學求真樓四樓會議廳(K401)
	09:00-09:10	開幕及貴賓致詞 Opening Words & Welcome Remarks	
	09:10-10:30	Sources of error in large-scale assessment(1)	
	10:30~10:50	Tea Break	
	10:50-12:00	**Computer Lab**	
	12:00-13:30	Lunch Break	
	13:30-15:00	**Sources of error in large-scale assessment(2)**	
	15:00-15:20	Tea Break	
	15:20-16:30	**Computer Lab**	

時間 Time		活動內容 Activity	地點 Place
第二天： 2009 年 11 月 19 日 (星期四) DAY2： 2009/11/19(Thu)	08:30-09:00	報到 Registration	臺中教育大學 求真樓四樓會 議廳(K401)
	09:10-10:30	Plausible Values and Imputation	
	10:30~10:50	Tea Break	
	10:50-12:00	Computer Lab	
	12:00-13:30	Lunch Break	
	13:30-15:00	Plausible Values for Secondary(or Subscale) Analysis	
	15:00-15:20	Tea Break	
	15:20-16:30	Computer Lab	
時間 Time		活動內容 Activity	地點 Place
第三天： 2009 年 11 月 21 (星期五) DAY3： 2009/11/21(Fri)	09:00-12:00	Private discussion Issue: using TASA data to construct conditioning variables, estimate population parameters, and students' plausible value	臺中教育大學

陸、工作坊內容

大型測驗暨 ConQuest 工作坊主要包含三大議題：大型測驗設計時的誤差來源、可能值(plausible values)的應用與實務分析，主講者是在大型測驗設計與分析均有深入探討與研究的學者 Dr. Margaret Wu，Dr. Margaret Wu 亦是著名的測驗軟體 ConQuest 的作者。第一天的議程是探討大型測驗設計時的誤差來源，包含測量誤差、抽樣誤差、等化誤差，依測驗目的不同決定可容忍的誤差範圍。影響測驗誤差的主要來源是測驗長度；在抽樣誤差的部分，必須考量到需抽樣到多大的樣本才能讓標準誤變小，而使得分數的比較是有意義的；等化誤差的來源主要是

實施測驗時的變異，或是試題具有差異（DIF）所造成，故在設計與實施測驗時，必須關注這三方面的誤差來源，才能對測驗結果作較正確的推論與解釋。第二天的議程是探討可能值的使用，目前國際大型測驗的技術報告都以可能值呈現學生的分數，主要是讓次級資料分析者，透過可能值能較準確回復群體參數估計的特性，如群體的平均數或變異數。第三天 Dr. Margaret Wu 分享她的實務工作經驗，如抽樣變異如何計算，三天工作坊的內容涵蓋基礎理論的介紹與軟體的使用，透過這些課程，讓與會者對於大型測驗資料分析有更深入的理解。