

「大型教育調查工作抽樣及量尺化分析工作坊」心得分享



任宗浩老師主講大型教育調查工作抽樣及量尺化分析工作坊

【文、圖／測驗及評量研究中心助理研究員 林奕宏】

近年來，PISA、TIMSS、PIRLS、NAEP 等大型的學生學習成就長期評量計畫日漸受到世界各先進國家的重視，我國也以其為目標發展了 TASA（臺灣學生學習成就評量資料庫）。為了提昇院內研究同仁對大型教育評量的專業知能，國民素養專案辦公室於 104 年 3 月 23 日至 25 日及 3 月 30 日至 4 月 1 日共計六個全天，在國教院三峽院區傳習苑 302 電腦教室，邀請了臺灣師大科學教育中心任宗浩副研究員，針對大型教育調查工作的抽樣設計、量尺化、資料分析等相關議題，進行一系列課程的主講。

任宗浩老師服務的臺灣師大科學教育中心負責規畫並執行 TIMSS 在臺灣的調查，歷經 1999 年、2003 年、2007 年、2011 年及即將開始進行的 2015 年 TIMSS 調查，任老師的實務經驗相當豐富，是國內少數具有實際執行大型教育調查經驗的學者。任老師在六天的工作坊中，詳細介紹了相關的理論概念，並搭配了實際上機操作，讓參加者可以把所學立即加以應用。六天的課程主題簡介如下：

- （一）3/23：抽樣設計、樣本估計、權重計算。實作：抽樣誤差的模擬。
- （二）3/24：抽樣誤差的估計。實作：誤差估計的模擬與真實資料的計算。
- （三）3/25：重複抽樣權重的產生。實作：相關軟體的介紹與操作。
- （四）3/30：大型教育測驗工具的信、效度與試題數的估計、題本配置。實作：試題反應理論的模擬資料產生及分析。
- （五）3/31：試題反應理論的能力估計方法（MLE、EAP、MAP）。實作：plausible values 的產生與測驗誤差的估計。

(六) 4/1：大型測驗調查標準誤的估計。實作：相關軟體的介紹與操作。

欲進行大型的教育調查工作，比如調查全國中、小學生的學習成就現況，第一步是先界定調查之目的為何，亦即須先訂出研究的問題。接著決定欲調查的內容，比如哪些學科有必要調查，各學科有哪些內容必須涵括在調查裡，此部份和課程綱要有直接的關係。其次則是決定各學科內容至少要有多少道試題。試題太多會讓調查的受試者樣本數增加，導致經費的提高；試題太少則會影響對學科內容評估的範圍，導致能推論的學科內容不足。試題數一旦確定，在適當的測量精準度（測量標準誤）前提下，即可據以預估調查的學生樣本總數，此部份直接影響到所需的經費，諸如試卷、監試人員、閱卷老師（若有非選擇題）等數量。其後即可依據調查目的擬定抽樣架構，包括目標母群、抽樣方式、抽樣單位、題本分配、權重計算等。其它工作尚包括：發展試題及預試、建置題庫、與被抽樣學校的行政連絡、監試人員與閱卷老師的招募與訓練、決定資料分析模式、資料回收與整理等等。

大型教育調查工作的結果，可提供給教育決策當局作為改進教育政策的重要依據，這正是本院的發展方向。本次為期六天的工作坊，已為本院在此方面的研究能量推進了一大步。