

估計，常以個別受試者的成績（能力值）平均值或變異數代表該群體之某一能力表現及其分散程度，更進一步進行各種假設檢定，例如：TASA 數學科即採用此方式（洪碧霞、林素微、林娟如，2006）。依據 Mislevy 等人（Mislevy, 1991; Mislevy, Beaton, Kaplan, & Sheehan, 1992; OECD, 2005; Lee, et al., 2007）之研究結果顯示，此種推論母群表現之方式容易造成偏誤。根據 Mislevy 等人之研究，可能值（plausible values）包含隨機誤差成分，不適合描述個體分數，但可能值具有良好群體估計一致性，適合描述群體之特性（Mislevy, 1991; Mislevy, et al., 1992）。因此，目前國際上大型測驗皆以此種技術進行群體統計特性描述（OECD, 2005; Lee, et al., 2007），本研究擬深入探討此一技術之使用方式，運用於臺灣學生學習成就評量資料庫（TASA），進而介紹國內相關研究領域使用。

綜合上述可知，目前國際上較知名的大型標準化測驗在評量架構、試題與測量模式之配合上仍有不一致與不足之處。因此，本計劃擬探討 NAEP、TIMSS 及 PISA 等測驗之資料分析步驟與方法，進而提出適用於國內 TASA 之標準化資料分析步驟與方法。

## 貳、研究目的

針對上述背景及動機，特提出本研究計畫，研究目的如下：

1. 藉由國外大型標準化測驗（NAEP、TIMSS 及 PISA）之文獻蒐集，以建立一套適合 TASA 之標準化流程，包含抽樣設計與抽樣權重、測量模式、試題特性與背景變項資料分析、量尺化程序等。
2. 針對某一些議題進行實徵資料實驗，以評估所研發標準化測驗之成效。