

【文／測驗及評量研究中心助理研究員 謝佩蓉】

對大多數的人而言，一份考卷最重要的是「考了什麼題目」。考題所涵蓋的內容範疇當然也是「大規模學業成就表現調查」（以下簡稱大規模調查）高度關心的議題。大規模調查所關心的對象是全體學生的表現情形而不是個別的表現，譬如，學生能力國際評量計劃（PISA）分析的是各國 15 歲學生在閱讀、數學以及科學的表現，而不是針對個別學生提供測驗結果。

為了能夠測量到更周延的內容範疇，大規模調查的題目數量通常都很多，而每位學生體力能負荷的作答題數有限，「心理計量領域」學者會採用的方式是：將幾個題項組成一個區塊（想像是一個積木），再透過某種排列方式，讓各個區塊輪流組成一份考卷（想像考卷是各個積木輪流排列組成）。

根據這樣的原則，當我們要進行大規模的追蹤調查時，又該怎麼組卷呢？所謂追蹤調查指的是，針對特定的同一群人觀察或蒐集資料好幾次。譬如，2012 年抽樣，針對新移民家庭子女進行大規模調查，在 2014 年再繼續針對同一群樣本，調查他們的學業成就表現變化。為了良好地設計大規模追蹤調查組卷，日前特別邀請來自澳洲 PISA 團隊的吳博士和臺灣 TIMSS 團隊的任博士提供國際級的专业意見。會中訂出的幾項原則如下：

（一）連結 2012 年和 2014 年兩次調查的共同試題數量一定要夠多，最好要占 2014 年試題數量的一半，才能夠準確地測量學業成就表現進展。

（二）2014 年試題區塊排列組合方式，和 2012 年相同。

（三）2012 年保留的試題，就留在它既定的試題區塊之中；加入新題時，則插入該區塊被刪除試題的位置。

（四）挑選試題時，以良好鑑別度為最優先考量，其次是難度。連結 2012 年和 2014 年兩次調查的共同試題，可以由 2012 年試題中挑選難度略高且鑑別度良好的試題；2014 年再加進的新題則可以是難度中等且鑑別度良好的試題。

（五）試題所能涵蓋的內容範疇，亦需一併考量。

橫斷性調查研究設計，將所有的變項都在同一個時間點調查，僅能闡述相關性，難以說明因果關係。為能進一步了解研究對象所處之環境脈絡和時間變化對於學習表現之影響，提供更具信服力的證據，籌劃大規模追蹤調查計畫以了解同一群研究對象長時間之成長變化，是值得投資的方向。組卷原則既經確立，建構大規模追蹤調查的測量工具便能收事半功倍之效。