

壹、緒論

2001年美國推行 No Child Left Behind (U.S. Department of Education, 2001) 對全國學生進行普測，低於通過標準的學生需接受補救教學，聯邦政府更會依各州表現情形給予不同補助 (NCLB; P. L. 107110)，故通過標準之訂定茲事體大，牽動學生的自信與身心，更觸及家長與社會的敏感神經，且隨著國際大型評比之建立，透過建立標準來了解學習表現與各國概況，已為世界潮流，標準設定在教育評量儼然佔有一席之地。

標準設定的模式約可分為兩種，一為常模參照 (norm-reference)，例如國際數學與科學教育成就趨勢調查 (Trends in International Mathematics and Science Study, TIMSS)、國際學生成就評比計畫 (Programme for International Student Assessment, PISA)，將學生能力等化到同一量尺後，以特定的百分比將學生分級，百分比對應的分數即決斷分數，這種模式建立在相對標準上，同一學生會隨同儕之不同，而可能被定義在不同級別 (Mullis, Erberber, & Preuschoff, 2008)。另一標準設定模式為標準參照 (criterion-reference)，例如美國國家教育進展評估 (National Assessment of Educational Progress, NAEP)，先將學生成就分為基礎、精

熟、進階三個水平，定義三個水平的學生應具備的能力後，再將學生進行歸類，由此而得的學生表現情形，已成為美國解釋學業成就的重要指標之一 (National Center of Education Statistics, 2008; Zenisky, Hambleton, & Sireci, 2007)。此類標準設定建構在絕對標準上，學生不因族群或班級不同而被歸類到不同水平，且在定義各個水平應有表現時，往往會呼應課程綱要和評量目的，將學生分級後，可了解學生精熟哪些課程、哪些能力需要提升，這是常模參照較難達到的。

臺灣學生學習成就評量 (Taiwan Assessment of Student Achievement, TASA) 為國內以課程改革為基礎，探討學生成就表現的大型評量資料庫，其建置目的為了解學生學習成就的表現，分析學生學習變遷之趨勢，進而檢視當前國家教育體制與政策實施之成效 (臺灣學生學習成就評量資料庫, 2011)，而欲檢視學生學習是否達到政策制定者所預期的程度，就需進行效標參照型的標準設定。臺灣自2001年起將英語課程從國中延伸至國小中年級，英語課程逐漸受到重視，國內也有研究探討2005年 TASA 英語科學生表現，以書籤標定法 (Bookmark) 對聽力與閱讀各建立通過／不通過的決斷點，發現約有七成以上的學生在通過的水準，但其訂出的二分標準，距今已不合時宜，且隨著課程

的調整，過去的標準已無法詮釋學生表現（吳毓瑩、陳彥名、張郁雯、陳淑惠、何東憲、林俊吉，2009），故興起重新設定標準之舉。

TASA 2009六年級英語科標準設定在2010年8月舉行，第一天採行Yes/No Angoff標準設定法，標準設定成員皆要判斷各水平邊緣受試者（borderline students），在每一個題目的答對情形，因為標準設定是一個判斷決斷分數的測量過程，在解釋設定結果時，要將影響整個過程的誤差因素加以考量（Cizek & Bunch, 2007）。Yin和Sconing（2008）曾以概化理論探討Angoff和Mapmark的變異誤，得到沒有特定的因素會影響兩種方法的決斷分數標準誤；另有研究檢視NAEP 2005年標準設定的效度表現情形（Sireci, Hauger, Wells, Shea, & Zenisky, 2009）。國內也陸續對TASA數學科標準設定提出有力的支持證據，標準設定效度相關的實務性研究逐漸產生（吳宜芳、鄒慧英、林娟如，2010）。

綜合以上，本研究將根據Kane（1994）與Pitoniak（2003）提出的過程、內部及外部等多元效度證據的觀點，評估TASA 2009國小六年級英語科標準設定的效度，並提出數項建議，供未來標準設定之參考。

貳、文獻探討

一、Angoff標準設定法

Angoff-based為一種以測驗為中心、試題評定取向的標準設定方法，也被NAEP運行多年（Hurtz & Auerbach, 2003; Reckase, 2000），此法廣為使用在美國各州政府評量（Council of Chief State School Officers [CCSSO], 2001）、瑞典的數學考試（Näsström, & Nyström, 2008），也用在醫學等證照考試，為最常使用的標準設定法之一。

最原始的Angoff標準設定法由Angoff（1971）所提出，標準設定成員要評估邊緣學生在每一題的答對機率，再將各題的答對率平均，找出決斷分數，如今也可運用到多元計分評量。Angoff標準設定法簡單易懂，也可在學生施測前執行，由於是逐題判斷，被判別較高決斷分數的試題可中和較低決斷分數的試題，有補償（compensatory）決斷分數的作用（Hambleton & Pitoniak, 2006），但當題數過多，容易造成標準設定成員的負擔，較不適宜使用，故後續有許多改良版的Angoff標準設定法，Yes/No Angoff即為其中之一，此方法在運用上，標準設定成員仍然要逐題進行判斷，不同之處在於判斷邊緣學生是／否答對即可，簡化原先評估答對率的做法（Impara & Plake, 1997）。

標準設定成員對邊緣學生的定義，以及估計學生的答對情形，是一項