

壹、緒論

教育政策研究常牽涉到方案 (program)，方案基本上可視為是一組活動或計畫，其目的是要達成某些預先設定而且對公眾有益處的結果。通常與政策有關的方案大都是經由政府單位資助而執行，所以實質上它是依靠納稅人的稅捐才能進行的，因此，由政府資助的方案都必須接受評鑑，檢驗其實際效果是否值得接受政府的資金援助，以符合績效責任制度 (accountability) 的原則。方案評鑑在公共政策辯論以及監督政府所資助的計畫上扮演重要角色，它可以影響到有關計畫的設計、運作、資金運用等方面的決定。

方案評鑑有許多不同的技術，其中一項是隨機化對照試驗 (Randomized Controlled Trials, RCT)。近年來，採用該方法來評鑑教育方案的效果，有愈來愈受政府決策單位以及教育研究人員重視的趨勢，此外，與其相似的集群隨機化試驗 (Cluster Randomized Trials, CRT) 研究方法亦愈來愈受到注意。事實上，該等研究方法在不少領域已行之有年，例如，Greenberg 及 Robins (1986) 的論文中即曾指出，社會科學家應用隨機化試驗已有數十年的歷史，對於評量社會政策的效果方面已累積了不少經驗。其他研究領域如醫學及公共衛生 (Borman, 2009; Donner & Klar, 2000)，在過去 50 年間，透過大量採用隨機化臨床試驗的研究方式，成功撲滅了美國境內的麻疹及小兒麻痺症等疾病 (Borman, 2009)。又例如美國的勞工局於上世紀 1980 年代期間，即已開始資助採用集群隨機化試驗的研究，以探討一些就業訓練方案的成效 (Raudenbush, Martinez, & Spybrook, 2007)。教育研究應用隨機化對照試驗和集群隨機化試驗雖然起步較晚，但近年來也有一些呼喚教育改革的報告，文中亦提及在教育研究方面應進行隨機化對照試驗，例如，由美國國家數學諮詢委員會所提出之頗具爭議的《成功的根基》(Foundation for success) 彙報，除了針對各級數學教學方面提出很多明確的建議，並鼓勵多執行高品質且能夠探討出因果關係的研究計畫，還特別強調進行隨機化對照試驗，可以對國家層級的政策提供有用的資訊 (National Mathematics Advisory Panel [NMAP], 2008)。此外，美國聯邦教育部在 2002 年通過《教育科學革新法案》(Education Sciences Reform Act)，正式成立了一個名為教育科學院 (Institution of Educational Studies, IES) 的機構，專事教育方面的研究工作，並以推動隨機化試驗為其首務之一 (Education Sciences Reform Act, 2002)，該機構自 2002 年至 2006 年這 5 年的期間，即資助了五十五個集群隨機化試驗的研究計畫，美國 IES 促進集群隨機化試驗的決心，於此可見一斑。

在臺灣方面，雖然一般研究方法的教科書都會對隨機分派及實驗法有所著墨，但教育研究期刊中採用隨機化試驗的論文為數不多，筆者曾按一般途徑做初步的搜尋，只搜尋到 3 篇有用隨機分派的實驗研究法論文 (方金雅、蘇姿云, 2005; 洪麗瑜、黃冠穎, 2007; 章勝傑、李冠蓉, 2003)，而且研究對象人數非常少。再者，究竟何謂隨機分派？何謂隨機化對照試驗與集群隨機化試驗？雖然從某一個角度而言，它們可以說是在教育方案評鑑與教育政策等領

域中較為新興的方法，但為何需要應用此類研究方法？它們有什麼優、缺點？在設計方面應該留意什麼事項？本文旨在對隨機化試驗做綜合的介紹，包括隨機分派與因果關係的建立；並介紹兩種隨機化試驗的形式，包括隨機化對照試驗與集群隨機化試驗，以及進行時考量的要點。然而，本文目的並非要對隨機化試驗做全面性的回顧，而是從較廣的層面，整理相關的背景知識，讓有興趣的研究者對此方法在概念上能有初步但較全面的瞭解，並能留意到在教育環境中執行隨機化試驗時，實務上並不簡單，需要有彈性做計畫上的調整，因此本文不列舉制式化的進行步驟，如讀者對隨機化試驗大致上的進行方式有興趣，可參考本文所報導的三個隨機化試驗實例。

貳、隨機化試驗備受注意的緣故

事實上，隨機化試驗並不是一個新的概念，隨機化的觀念可上溯至二十世紀二、三〇年代英國著名統計學家 Ronald A. Fisher (1925) 所從事的研究，他認為當要分派 (assign；亦有翻譯為分配或指派) 研究對象到實驗處理 (treatments) 的不同組別時，分派這些研究對象的方式，必須是與研究對象的特質沒有任何關係的，例如，並不會因為某研究對象比較高或者是比較敏捷而被分派至實驗組，因此分派的過程必須是隨機的，這樣的安排被認為對實驗設計是既重要而且是相當可取的。Campbell 與 Stanley (1963) 的名著 *Experimental and quasi-experimental designs for research* 一書中，除了向教育界介紹前實驗、準實驗與真實驗等設計的差別之外，並提供內在效度以及外在效度的各類威脅作為一個架構，以比較不同研究設計的限制，尤其是在建立因果關係方面。該書所提出眾多研究方法的專業術語，時至今日，仍為教科書所廣為沿用，Campbell 與 Stanley 是忠實倡導教育研究應推行隨機化試驗的學者，他們認為要面對教育實施 (educational practices) 方面的爭議，以及執行某項教育實施是否真的帶來進步，只有透過隨機化試驗才能得到解決。

教育研究雖然行之有年，但歷來外界對很多教育研究的品質有不少批評，包括研究是否嚴謹、研究方法是否合宜，以及研究結果的應用性等方面。例如，美國教育部發表的 2002-2007 年政策計畫書中，就曾指出有別於醫學及農業學等領域，教育界長久以來依賴意識形態或者是專業人士所達成的共識來運作，很容易受到一時流行風尚的影響，因此不能像科學研究那樣，藉由科學方法的應用，以及系統化的蒐集與應用客觀的資料以制定政策，從而積累出進步的成果 (U.S. Department of Education, 2002)。由此可見，至少當時美國教育部的官方立場，會認為很多教育研究的可信度並不高。

另一方面，美國國會對於隨機化試驗之所以備受注意亦扮演一個重要角色，而且影響了不只一個面向。首先，Borman (2009) 及 Slavin (2003) 指出在 1998 年間，美國國會核准 1 億 5,000 萬美元作為全面學校改革示範方案 (Comprehensive School Reform Demonstration,