

方案理論導向評鑑模式之探究

■ 文／林素卿·國立彰化師範大學教育研究所教授

方案理論導向評鑑 (program theory-driven evaluation, PTDE) 或稱為方案理論評鑑 (program-theory evaluation) 或稱方案理論本位評鑑 (program theory-based evaluation)，此類型的評鑑模式係透過方案理論 (program theory) 建構方案評鑑的架構，並據以提出評鑑問題、驗證方案的因果機制及評估方案的成效。因為傳統的評鑑模式常以績效評估為主要目的，忽略方案為何有效或無效原因之探討，易使評鑑成為一種「黑箱評鑑」，其評鑑結果恐難被決策者或利害關係人所應用，亦可能對方案成敗因素產生錯誤的歸因 (Chen & Rossi, 1983)，基於對黑箱評鑑的批評與反省，PTDE模式繼而興起。

PTDE緣起與發展

PTDE的相關概念可追溯到1930年代Ralph W. Tyler之系統闡述評鑑的目的和測試方案的理論。1960年代起已有學者對PTDE提出討論，主張引用理論作為分析方案運作機制的依據。1970年代PTDE開始受到評鑑學界的重視，有些學者嘗試運用理論建立假設，考驗方案之運作及成果。1980年代已有更多的評鑑學者投入PTDE的探究，提出許多觀點，有關PTDE的學

術論文亦陸續出版。此時期PTDE已經受到評鑑學界相當多的關注，逐漸成為受歡迎的評鑑思維。1990年Huey T. Chen出版《理論導向評鑑》(Theory-driven evaluations) 一書後，造成很大的迴響，引發更多關於PTDE學理和方法論的探究 (Coryn, Noakes, Westine, & Schröter, 2011)。

之後，Stewart I. Donaldson受到PTDE先行者著作的影響，致力於PTDE模式的再建構、擴充與改革，於2007年出版《方案理論導向評鑑科學：策略和應用》(Program theory-driven evaluation science: Strategies and applications)，深入說明PTDE的核心概念且提供具體的實踐範例，其主張PTDE是一種評鑑科學，此論點已經超越前述先行者之論點，發展出更嚴謹的評鑑方法。近十幾年來，PTDE因其嚴謹的評鑑設計，企圖揭開方案運作機制的「黑箱」，因此被廣泛應用於各領域的評鑑中，成為當代方案評鑑的主流模式之一。

PTDE核心概念

PTDE具有三種基本的特性：

(1) 詳細描述輸入；中介過程；短、中、長期結果及三者間的預期關係，以解釋方案介入處理之

情形。

(2)測量方案理論中所有假設性的構念。

(3)分析數據及資料，確定假設性的關係與實際出現的程度(Shadish, Cook, & Campbell, 2002)，其核心概念主要包括方案理論、評鑑設計步驟、評鑑的範疇，分述如下：

一、方案理論

PTDE的方案理論是用來引導評鑑的設計和執行，其本質是方案有效運作的合理性和可行性的模式建構，關注方案所預期的效果及從輸入到輸出之黑箱過程中，如何轉換的一系列命題，因此，PTDE的方案理論既是描述性、說明性；也是規範性、實踐性(Chen, 1990)。PTDE方案理論分為二類：

(1)概念性理論

概念性理論是行動的基本信念，是建構方案計畫時所依據的假設，用來引導評鑑的設計和執行(Weiss, 1998)，例如規範性理論(prescriptive theory或normative theory)旨在建議方案應該是什麼或應該如何做才能更好，其中包括行動導向(action orientation)、介入處理設計(treatment design)、實施的概念化(conceptualization)和結果成效標準的選擇等(Chen, 1990)；又如方案影響理論(program impact theory)是由一系列行動及概念假設所組成的，描述方案所期望產生改變的因果關係(Rossi、Freeman & Lipsey, 2004)。

(2)經驗性理論

經驗性理論如Weiss(1998)所稱的執行理論(implementation theory)，詳細說明方案如何運作和可能產生的結果，例如因果性理論(causative theory)乃詳細說明方案如何運作，藉由界定某些過程中將發生的情況和可能產生的結果，說明介入處理和產出間之因果關係(Chen,

1990)；又如方案過程理論(program process theory)說明方案如何與標的群體(targeted group)產生互動，如何運用資源和組織的活動維持方案的發展及預期的結果(Rossi、Freeman & Lipsey, 2004)。

二、評鑑設計步驟

PTDE評鑑設計的主要步驟，包括發展方案理論、系統性闡述評鑑問題並排出優先順序、回應評鑑問題(Donaldson, 2003, 2007)，分述如下：

(1)發展方案理論

PTDE所建構的方案理論(含概念性和經驗性)。在方案理論發展的過程中，評鑑人員和利害關係人之間必須有高度的互動，才能順利發展出符合雙方需求的方案理論。一旦發展完成方案理論後，則利用這些理論作為確認評鑑問題，選擇評鑑方法的依據，但必要時評鑑人員必須重新評估是否調整部分的方案理論，或者全面性改變方案理論。

(2)系統性闡述評鑑問題並排出優先順序

評鑑人員和相關的利害關係人通常會列出一系列評鑑問題的清單，但是不同利害關係人對於每個評鑑問題的價值，看法可能不一，因此必須詳盡的闡述每個評鑑問題。然後，試圖將這些問題排出優先順序，使評鑑人員很清楚知道在評鑑過程中，哪些問題是合理的、適切的、有價值的和可以回應的。

(3)回應評鑑問題

在評鑑過程中，評鑑人員首先應該對下列狀況提出清楚的界定：缺失(breakdowns)(例如不良的實施、不合適環境脈絡、錯誤的理論等)、非意圖的結果(含正面和負面)、副作用、方案的有效性(或效力)。

其次，測量方案理論中所假設的過程構念、成果構念及環境脈絡構念。最後，根據所蒐集

的資料，解釋理論構念間的因果關聯，包括關係的方向和強度的差異、調節其他構念間關係的程度等，以回應利害關係人所關心的評鑑問題。

三、評鑑範疇

Chen (1990, 2005) 根據方案理論，將PTDE範疇分為六類：

(1) 介入處理

介入處理 (treatment) 是使方案產生預期改變的基本要素。評鑑人員將根據研究架構對介入處理的過程和產出的結果進行系統性的評鑑，關注的議題包括：介入處理的概念化和設計（例如介入處理的性質、要素、要素間的組織及測量等）、測量場域中實際介入處理之情形、評估設計和實際執行之間的一致性。

(2) 實施環境

方案可以透過不同的方法來實施，而方案如何被實施可能影響方案的過程及產出。因此實施環境 (implementation environment) 評鑑旨在檢視及評估方案介入處理實施的實際環境與所規劃的環境是否一致，關注的議題包括：介入處理是否針對標的群體 (targeted group)？執行人員是否具備所需之專業能力與資格？傳遞的模式和組織間的協調是否適切且充分？

(3) 成果

成果 (outcome) 即是方案預期目標達成的程度，對利害關係人而言是非常重要的，因為方案目標常被用來引導利害關係人的後續的活動和決定資源的分配，且經常被用來作為判斷方案成效的標準。不過，除了方案目標和預期成果外，方案亦可能產生非預期的結果，但是非預期的結果經常被利害關係人所忽略。因此，成果評鑑應關注預期成果和非預期成果二部分的評估。

(4) 影響

影響 (impact) 評鑑是評估介入處理對成果產出的影響。利害關係人如果想知道介入處理是否有效或方案應朝哪個方向發展，就必須進行影響評鑑，否則很難判斷方案是否有效。此範疇涉及因果的推論，因此必須有充分且明確的證據，作為評估方案介入處理對意圖成果或非意圖成果的影響。

(5) 中介機制

中介機制 (intervening mechanism) 評鑑旨在了解連結實施的介入處理到成果之間的因果過程，亦即，藉由介入處理的過程，探討如何產生預期的成果。中介機制的評鑑將提供利害關係人，有關為何方案有效或無效的資訊，診斷方案的優點和缺點，作為改善方案依據 (Chen, 1990)。

(6) 概括化

概括化 (generalization) 評鑑是方案的成果如何類推到未來目標之系統性評估。有時候決策者或利害關係人希望評鑑的結果，可以應用於未來或特定的人、環境或機構。就此需求，評鑑人員就必須對方案進行概括化評估。

PTDE共同的迷思及回應

PTDE模式發展至今，評鑑學界對此類模式出現一些共同的迷思 (Donaldson, 2003)，分述如下：

一、方案理論等同社會科學理論

有些評鑑學者，例如Stufflebeam (2001) 認為很少有發展成熟的社會科學理論可做為評鑑的基礎，質疑基於理論導向的PTDE是不可行的。Donaldson (2003) 認為此斷言無疑混淆了PTDE所使用的方案理論和社會科學理論的性質。

PTDE之方案理論是實驗處理或是介入方法之

理論，其典型的定義如Bickman（1987）所界定的：「方案理論是方案如何運作的合理和明智模式的建構」或如Lipsey（1993）所定義的：「方案理論是指黑箱運作中有關輸出到輸入的命題，亦即，如何透過介入處理改善現有的狀況」。PTDE中的方案理論是基於多種資訊的來源，包括先前的理論和研究、最接近於方案執行的暗示性理論（implicit theories）、方案行動過程中的實際觀察、測試方案性質的重要假設之探索性研究等（Donaldson, 2007）。

二、探討方案成敗之因果關係不是評鑑的主要任務

Scriven（2007）認為評鑑的主要任務是判斷方案的優缺和價值，而不是解釋方案如何有效。解釋方案為何或如何有效通常需要有關方案的專業知能，且已超出大多數評鑑人員的能力範圍，不應該列入評鑑的任務之中。

Chen（1994）則針對Scriven的質疑作出回應，認為對方案評鑑而言，不論是判斷方案的優缺和價值或解釋方案的因果關係，二者同等重要，否則無法提供利害關係人改進方案或作決策所需之訊息或知識。

Chen以醫學界開發新藥為例，說明解釋性知識的重要性：「如果黑箱評鑑顯示新藥可能治癒疾病，但沒有提供有關新藥治療機制的訊息，內科醫生將很難開此新藥為處方，因為不知道此新藥在哪些條件下可能有效，哪些條件下可能會產生副作用」。

三、PTDE等同邏輯模式

Weiss（2007）發現很多PTDE的評鑑人員，雖聲稱採用方案理論，但仍只集中於實施理論，通常以框框和箭頭圖示出輸入、活動、結果、外部因素和影響等元素之線性關係，這種模式應屬於方案邏輯（program logics）或邏輯模式（logic

models）而不是真正的方案理論，因為它們不涉及方案因果機制的考驗，也沒有顯示足夠的因果線索，因此無法為複製方案或改進方案提供有用的訊息。

PTDE的方案理論是一個評鑑實踐的綜合理論，具體說明評鑑實務可行的做法，提供改善方案之有價值的知識。而邏輯模式僅是PTDE用來發展方案理論的工具，不屬於評鑑實務的綜合理論，僅是PTDE的一部分，因此邏輯模式不等於PTDE。

四、單純的成果評鑑不屬於PTDE的範疇

有些評鑑學者揚言PTDE之範疇不包括單純的成果評鑑。Donaldson（2003）指出這是另一種貶低PTDE價值的迷思，事實上，PTDE並沒有將單純的結果評鑑排除在外，例如在評鑑過程中，縱使方案理論已經發展完成，但是如果因為評鑑的資源有限，或受到實務運作的限制，PTDE當然還是以回答重要性評鑑問題為重，其中最適當的設計，當然就是單純的成果評鑑了。

五、PTDE忽略方案副作用的評估、成本高且費時

有些批評者指出PTDE的方案理論中並不包括非預期結果和副作用的假設，因此忽略方案副作用的評估。Chen（1990）對此批評提出回應，指出發展PTDE之方案理論時，評鑑人員和利害關係人會同時考量在什麼情況下如何評估方案潛在的副作用，且在方案評鑑過程中也會思考和調查是否有負副作用的產生。

另外，Scriven（1998）批評PTDE較其他取向評鑑來得昂貴且費時，因為PTDE在方案理論建構過程中，通常會有兩個（含以上）的競爭性理論（competing program theories）接受檢驗，藉由小規模的試驗，檢視哪個方案理論較為適切，

實施活動是否需要調整，再根據實際狀況決定或改變方案理論，以符合利害關係人的需要。Donaldson (2003) 對Scriven的批評提出反駁，指出如果一個方案沒能作好準備就大規模付諸於實施，結果肯定是無效的。藉由小規模的試驗調整，可以節省大量的經費及減少大規模評鑑的時間和資源。

藉由PTDE 提升方案的品質

方案評鑑旨在改善方案的缺失，提升方案品質。藉由評鑑的實施與回饋，成為各項教育方案永續發展的動力（吳清山、王湘栗，2004）。然而，目前方案評鑑常僅止成效的評估，缺乏探討方案為何有效或無效的原因，評鑑結果往往無法

作為改善方案的依據。

PTDE的主要特色是透過方案理論建構方案評鑑的架構、提出評鑑問題、驗證方案的因果機制及評估方案的成效；最大的優勢是解釋方案所潛藏的因果機制，從中建構新的知識，產生有效的模組，作為複製方案或者改進方案的重要依據，這是其它類型的評鑑模式所不及的。

筆者發現國內對PTDE模式尚不熟悉，應用有限，期待未來學界有更多關於PTDE的探究，經由PTDE理論與實務不斷的精煉與成熟，可望為當前方案評鑑的困境解套，提供值得嘗試的一條明徑，讓方案評鑑發揮實質的意義，促使方案永續發展，發揮更大的效益。📖

◎參考文獻

- 吳清山、王湘栗（2004）。教育評鑑的概念與發展。*教育資料集刊*，29，1-26。
- Bickman, L. (1987). The functions of program theory. *New Directions for Program Evaluation*, 33, 5-18.
- Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review*, 7, 283-302.
- Chen, H. T. (1990). *Theory-driven evaluations*. Newbury-Park, CA: Sage.
- Chen, H. T. (1994). Theory-driven evaluation: Needs difficulties and options. *Evaluation Practice*, 15, 79-82.
- Chen, H.T. (2005). *Practical Program Evaluation: assessing and improving program planning, implementation, and effectiveness*. Thousand Oaks, CA: Sage Publications Inc.
- Coryn, C. L. S., Noakes, L. A., Westine, C. D., & SchrÖter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32(2), 199-226.
- Donaldson, S. I. (2003). Theory-driven program evaluation in the new millennium. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp.109-141). Mahwah, NJ: Lawrence Erlbaum.
- Donaldson, S. I. (2007). *Program theory-driven evaluation science*. New York, NY: Lawrence Erlbaum.
- Lipsey, M. W. (1993). Theory as method: Small theories of treatments. *New Directions for Program Evaluation*, 57, 5-38.
- Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (2004). *Evaluation: A systematic approach* (6th ed.). Thousand Oaks, CA: Sage Publications Inc.
- Scriven, M. (1998). Minimalist theory: The least theory that practice requires. *American Journal of Evaluation*, 19, 57-70.
- Scriven, M. (2007). The logic of evaluation. In H. V. Hansen et al. (Eds.), *Dissensus and the search for common ground* (pp. 1-16). Windsor, ON: OSSA.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Stufflebeam, D. L. (2001). Evaluation models. *New Directions for Evaluation*, 89. San Francisco, CA: Jossey-Bass.
- Weiss, C. H. (1998). *Evaluation*. Upper Saddle River, NJ: Prentice Hall.
- Weiss, C. H. (2007). Theory-based evaluation: Past, present and future. *New Directions for Evaluation*, 114, 68-81.