

第五代政策評估的興起：複雜性科學的公共治理新典範

張世賢*

摘要

面對 21 世紀氣候變遷與能源轉型等高度錯綜複雜的「棘手問題」，傳統依循牛頓機械論的線性評估工具已顯露疲態，而強調多元協商的第四代評估亦易陷入主觀建構的溝通內耗。在此脈絡下，全面導入「複雜性科學」思維的「第五代政策評估」應運而生。本研究採取文獻研究法與批判性論述分析法，系統性探討最新第五代政策評估的理論演進、核心途徑與方法論質變。研究發現：第五代評估透過「解釋性現實主義」與「動態演化評估」兩大核心途徑，將公共政策視為「複雜適應系統」，成功深化了「脈絡—機制—結果（CMO）」的客觀因果解釋。在方法論上，由傳統「化約論」（reductionism）工具轉向「行動者模型」（ABM）與「社會網絡分析」（SNA）等系統科學工具，促使評估者角色從場外裁判轉化為系統共同進化者。本研究進一步檢閱英國國家政策指南與台灣本土能源轉型之治理困境，驗證第五代框架在處理「非線性」湧現效應時的實踐價值。最後，本研究針對數據壁壘與傳統「關鍵績效指標」（KPI）問責體制的衝突進行詰辯，並提出法制化複雜性指南、建置系統模擬平台及推動實驗型問責文化等具體建議，期能為當代政府建構具備動態適應力與治理韌性的新典範。

關鍵詞： 第五代政策評估、複雜性科學、現實主義評估（CMO）、行動者模型（ABM）、第四代政策評估

* 國立臺北大學公共行政暨政策學系退休教授。本論文經兩位雙向匿名審查通過。E-mail: chang.shihhsien@gmail.com

本論文經兩位雙向匿名審查通過。收件：2026/9/10。同意刊登：2026/9/28。

壹、前言

公共政策的推行在 21 世紀面臨前所未有的嚴峻挑戰。隨著全球化、數位化浪潮席捲，以及氣候變遷、跨國流行病、人工智慧（AI）治理等議題的浮現，公共政策所處的環境已化為高度錯綜複雜的生態系統。傳統政策科學長期依循「牛頓式機械論」，將政策干預視為線性的「投入→產出→效果」因果關係，並依賴事後靜態的「關鍵績效指標」（key performance indicators，簡稱 KPI）進行審查。然而，此種還原論工具在面對動態演化的「棘手問題（Wicked Problems）」時，往往顯得捉襟見肘。

為回應傳統評估的侷限，古巴（Egon Guba）與林肯（Yvonna Lincoln）於 1989 年提出「第四代政策評估」，強調回應性與建構主義取向，試圖透過利害關係人的協商與多元價值對話來尋求共識。然而，第四代評估過於沉溺於各方利害關係人的主觀建構與政治協商，實務上常陷入無共識的口水戰（唯我論），且忽略了政策因果機制在客觀現實中的複雜交互作用。

在此脈絡下，當前政策科學界進一步催生了「第五代政策評估（The Fifth Generation of Evaluation）」的理論框架。第五代政策評估的核心宗旨，在於深化對客觀因果機制的實質解釋，並全面導入「複雜性科學」（Complexity Science）思維，將政策視為一個「複雜適應系統」（Complex Adaptive Systems, CAS）。本研究旨在系統性評析第五代政策評估的理論演進、核心途徑、方法論與角色轉變，並透過英國與台灣本土的實例應用，探討其在當代公共治理中的實踐價值、面臨之挑戰與未來發展建議。

貳、文獻探討

一、第四代政策評估的檢視

古巴（Egon G. Guba）與林肯（Yvonna S. Lincoln）提出第四代政策評估（Fourth Generation Evaluation），其評估方法的核心在於評估者不再扮演至高無上的裁判，而是轉為「諮商者」與「促成者」的角色。其運作起點是去蒐集並回應所有利害關係人（Stakeholders）的「主張（Claims）」、「關切（Concerns）」與「議題（Issues）」，透過多元對話與協商來共同建構政策價值。這三個核心概念的具體內涵如下：

（一）主張

主張指利害關係人對某政策問題所持有的看法、訴求或利益要求。這是利害關係人基於自身立場，所提出的政策主張，可以「做得很好」、「有利」或「符合期待」。從利害關係人的主張，可以進一步了解其關切甚麼。

（二）關切

關切指利害關係人對於其政策主張所關心的事項，認為與他有利害關係，必須關心，需要除去其擔憂、焦慮、或所認為的潛在風險。

（三）議題

議題指不同的利害關係人之間，因為價值觀、利益或立場不同，對同一政策問題有不同的主張、關切事項，而產生爭議、衝突；該政策問題即成為政策議題。

第四代政策評估的學理基礎是相對主義（**Relativism**）與建構主義（**Constructivism**）。相對主義主張：不存在客觀且唯一正確的絕對真理，社會現實是由不同利害關係人（**Stakeholders**）依據自身的脈絡、經驗與價值觀所共同建構出來的「**多重現實（Multiple Realities）**」。因此，政策評估的目的不再是尋找單一的客觀事實，而是透過對話與協商，理解並融合這些多元的觀點。建構主義認為現實與知識是人們在特定社會文化背景下共同建構的，並不存在單一客觀的絕對真理。強調協商與回應，重視多元利害關係人的主張、關切與意見，透過溝通與協商來達成共識。超越實證技術，打破前三代過分依賴計量與實驗的技術取向，轉而關注價值判斷和社會互動。

第四代政策評估的精髓，就是將這些利害關係人的主張、關切與議題，當作評估的指引，引導各方在協商過程中互相理解，最終達成集體的共識建構。

第四代政策評估雖然成功打破了前三代「管理主義」與「過度崇尚科學客觀」的盲點，轉而重視多元利害關係人的內心感受，但在學術界與實務推行上，它也面臨了許多嚴峻的挑戰與缺點。

二、第四代政策評估的缺點

第四代政策評估（**Fourth-Generation Evaluation**）由林肯與古巴（**Lincoln & Guba**）提出，強調回應性建構主義（**Responsive Constructive Evaluation**），主張

透過利害關係人的對話與協商來建構政策價值，評估焦點完全放在「多元對話」、「社會建構」、循證（Evidence-Based）上，其核心在於透過客觀、量化的科學證據來引導政策制定與評估（Guba & Lincoln, 1989），儘管此模式具備民主與包容的理想，但在實際推行時卻面臨深刻的現實限制。其缺點：「證據」的定義過於狹隘、協商成本高、共識困難、權力不平等，但這些缺點不等於「主觀口水戰」或「唯我論」。

（一）「證據」的定義過於狹隘

政策本質上是資源分配與價值的選擇，「證據」無法自動決定哪種價值更重要（Head, 2008）。在政治攻防中，決策者常落入「政策引導證據」（Policy-based evidence）的陷阱，選擇性使用對自己有利的數據（Sanderson, 2002）。雖然強調人人平等對話，但在實際的協商會議中，弱勢族群（如偏鄉居民、移工）往往缺乏論述能力或資源，最終話語權仍可能被相對強勢的團體壟斷，弱勢聲音可能再度被淹沒。

循證往往過度依賴可量化的數據或隨機對照試驗（RCT），容易忽略無法被量化的脈絡（如文化價值、歷史背景）。這種由上而下的菁英模式，常將住民經驗或在地智慧視為「非科學」而加以排除（Greenhalgh & Russell, 2009）。且證據缺乏脈絡可複製性（脈絡盲點），在 A 地（或實驗環境）證明有效的政策，直接複製到 B 地時，常因制度、法律、民情不同而宣告失敗，科學證據往往難以做到「去脈絡化」的移植（Pawson, 2006）。

（二）協商成本高

第四代政策評估要求所有利害關係人進行反覆的對話與「赫密士辯證 circle」（Hermeneutic Circle）。這種「參與式」的設計，意味著評估不再是專家填填表格、跑跑數據就結束，而是需要大量的協商會議、工作坊與溝通成本。在現代政府追求「效率」與「即時回應」的脈絡下，冗長的協商程序往往導致決策延宕，甚至錯失解決政策問題的最佳時機。嚴謹的科學證據需要長期的追蹤與研究，要納入所有利害關係人的主張、關切與議題（Claims, Concerns, Issues），並透過反覆論證，達成共識，過程極其漫長且成本高昂（Guba & Lincoln, 1989）。

(三) 共識困難

當政策涉及高度價值衝突（如核能發電、同婚政策、文資保存與開發）時，各方的核心價值往往是「不可共量」的。在多元社會中，各方利益和世界觀根深蒂固，對話最終往往只能走向「政治上的相互妥協」或「各自表述」，難以達成林肯與古巴理想中「融合大眾意見的更高層次共識」。面對當前氣候變遷、流行傳染病等需要即時做出決策的公共危機，在短時間，很難產生共識。如以「慢工出細活」的對話模式，緩不濟急。但政治決策往往具有時效性，等證據出爐時，政策窗口可能早已關閉（Head, 2008）。

(四) 權力不平等

第四代評估假設人人皆可在平等地位上發言。然而在現實的政治場域中，權力是不對等的。資源豐富、擁有話語權或高知識水準的優勢團體（如資方、專家），往往比弱勢群體（如移工、偏鄉居民）更能系統性地表達自身利益。這導致最後的「共識」可能只是優勢階級將其意志合法化的工具，反而進一步剝奪了弱勢者的聲音。

(五) 其缺點不等於「主觀口水戰」或「唯我論」

雖然第四代評估充滿了溝通的政治角力與困境，但將其斥為「主觀口水戰」或「唯我論（Solipsism）」是對該理論的嚴重誤解。

1. 第四代政策評估是「互為主體性」，而非「主觀口水戰」。(1)口水戰的本質：流於情緒、毫無根據且拒絕聆聽的盲目爭吵。(2)第四代評估的本質：建立在互為主體性（Intersubjectivity）之上。它不否認每個人有不同的主觀世界，但強調透過「理性且有結構的對話」，讓不同的建構（Constructions）相互碰撞、修正。(3)證據的約束力：在第四代評估中，利害關係人提出的觀點不能只是憑空捏造的謊言，必須植根於實際的政策經驗、數據與生活脈絡中。評估者的角色就是「催化劑」與「調解者」，負責篩選、釐清並確保對話是基於彼此經驗的理解，而非無休止的口水戰。

2. 第四代評估是基於「社會建構」，而非「唯我論」。唯我論的本質：認為只有「我的意識」存在，其他人和外在世界都是不可信或不存在的，因此完全拒絕與外在世界妥協。第四代評估的本質屬於社會建構論（Social Constructivism），

主張「現實」是由群體在特定的社會與文化脈絡中，共同建構而成的。其超越個人的「共識追求」，雖然不承認有唯一、絕對客觀的物理真相（即反對實證主義的絕對真理），但它極力強調「群體」的重要性。每個人都必須把自己的建構拿出來與他人檢驗，並透過辯證 circle 不斷擴大視野。其目標是走向「主體間的交集」，這與唯我論那種「只有我說了算、縮回個人世界」的封閉態度完全背道而馳。

總之，第四代政策評估的真正問題，不在於它流於虛無的主觀口水戰，而在於它過於低估了現實政治中權力結構的頑固性，以及高估了人類透過對話達成共識的意願與能力。它是一個在哲學上精緻（強調互為主體與社會建構）、但在實務上高度艱難（高成本、難共識、受制於權力）的理想主義模型。

三、 第四代政策評估的轉向

(一) 轉向「知證」途徑（Evidence-Informed Approach）：

將「循證」（以證據為唯一導向）修正為「知證」（將證據視為決策的參考要素之一），承認證據的局限性，並在決策時將科學證據、預算限制、政治可行性與大眾輿論放在同等重要的天平上衡量（Head, 2008; Oxman et al., 2009）。

(二) 引進「現實主義評估」（Realistic Evaluation）：

受到現實主義科學觀的影響，評估不再盲目問「政策有沒有效？」，而是轉向追問「在什麼脈絡下（Context）、對誰（Who）、透過什麼機制（Mechanism）會產生什麼結果（Outcome）？」（Pawson & Tilley, 1997）。

(三) 轉向「商議民主」與「參與式評估」：

修正過去由專家壟斷證據的現象，透過公民會議、利害關係人工作坊等機制，讓民眾的「生活經驗」與實踐者的智慧轉化為合法的政策評估依據（Greenhalgh & Russell, 2009）。

(四) 採用「敏捷與適應性治理」

面對複雜與未知的社會環境，政策評估轉向結合大數據即時監測與滾動式修正，採取「小規模試點、即時回饋、快速修正」的學習型模式（Sanderson, 2002）。

四、政策評估受到複雜性科學的影響

(一) 政策評估的研究者提出複雜性科學的觀點

隨著全球化、數位化及社會議題的極度錯綜複雜（如氣候變遷、跨國流行病、AI 治理等），政策評估理論在 21 世紀已演進至全新的觀點，深深受到「複雜性科學」（Complexity Science）的逐漸影響。Sanderson (2000) 探討政策評估「複雜性轉向（complexity turn）」的早期經典文獻，最早論述政府政策作為開放且複雜的系統時，傳統實證主義評估方法為何失效，並提出政策應作為「實驗與持續適應學習」的觀點。

Patton (2010) 有系統性地運用「非線性」、「湧現性」及「不確定性」等複雜系統概念，建構出適合棘手問題與社會創新的「發展的評估」（DE）框架。這是將複雜性科學引入政策評估最具里程碑意義的著作。

Boulton 等人（2015）探討世界如何作為一個「複雜適應系統」（CAS）運作，並指出傳統線性預測的盲點，深入論述了複雜性科學如何從根本上改變策略制定、治理與政策成效評估的哲學。

Rutter 等人（2017）刊登於醫學權威期刊《刺絡針》的論文，聚焦於公共衛生政策，其核心立論在於批判傳統將政策孤立評估（還原論）的錯誤，極具代表性地指出評估複雜公共政策時，必須轉向「複雜系統方法（Complex Systems Approach）」。

HM Treasury (2020) 提出全球公共政策評估最具權威的官方指南之一，將「複雜性科學」正式納入國家政策評估的法定補充指南中，指導評估者如何處理非線性因果、反饋迴路與湧現效果。Barbrook-Johnson 等人（2021）系統性地將「複雜性科學」應用於食物、水資源、能源與環境等跨界政策評估時的創新方法論與實務挑戰，有別於第四代政策評估的觀點與方法。

(二) 複雜性科學促成政策評估理論的轉變

過去從第一代到第四代的政策評估，其根基是「牛頓式的機械論」，假定政策與結果之間是線性的「投入→產出→效果」因果關係。然而，複雜性科學（Complexity Science）引入了「網狀系統」與「動態演化」的思維，對政策評估理論帶來了顛覆性的影響：

1. 因果觀的轉變：從「線性因果」到「非線性與湧現（Emergence）」

傳統觀點：增加 A 預算，就會等比例增加 B 政策效果。

複雜性影響：系統具有湧現性（Emergence）與非線性。微小的政策擾動可能引發巨大的連鎖蝴蝶效應；或者投入大量資源，卻被系統內部的負回饋抵消。評估理論因此不再執著於尋找單一的「歸因（Attribution）」，而是評估政策對整個生態系統的「貢獻度（Contribution）」。

2. 目標觀的轉變：從「預設靜態目標」到「適應性與反饋迴路」

傳統觀點：依據政策出發前設定的 KPI 進行實質對照（事後評估）。

複雜性影響：公共問題多為動態的「棘手問題（Wicked Problems）」。複雜性科學促使評估理論發展出「動態演化評估」（Developmental Evaluation, 由 Michael Quinn Patton 提出），強調評估必須形成及時的反饋迴路（Feedback Loops），協助決策者在高度不確定性的環境中邊做、邊評、邊動態調整。

3. 方法論的轉變：從「還原論工具」到「系統科學工具」。傳統觀點：透過對照組、統計迴歸等方法，將變數孤立起來分析（還原論）。複雜性影響：為了看清全貌，評估引入了大量計算科學與系統分析工具：行動者模型（Agent-Based Modeling, ABM）：模擬成千上萬個體在政策介入後的互動與行為質變。參與式系統繪圖（Participatory Systems Mapping）：邀請利害關係人共同繪製政策系統結構圖，找出政策隱藏的槓桿點。社會網絡分析（Social Network Analysis, SNA）：評估政策執行過程中，跨部門與公私協力的網絡關係與演變。

4. 評估者角色的轉變：從「客觀裁判」到「系統共同進化者」

傳統觀點：評估者是站在場外的「獨立第三方專家」。

複雜性影響：評估者本身就是複雜系統的一部分，其評估行為會干擾系統（觀察者效應）。評估者現在被定位為「促動者」與「系統導航員」，職責是揭示系統的盲點、促進跨界對話，提升整個治理系統的適應力（Adaptability）與韌性（Resilience）。

複雜性科學透過英國 Centre for the Evaluation of Complexity across the Nexus（CECAN）等機構的推動，已從單純的「物理/生物理論」完全落地為「政策評

估的實務指南」，徹底改變了現代政府對於循證(evidence-based)決策的理解(HM Treasury, 2020)。

五、第五代政策評估名詞正式提出

「第五代政策評估」(Fifth-Generation Evaluation)名詞的正式提出，並不像第四代政策評估，由 Guba 與 Lincoln 於 1989 年奠定；而是由多位學者在 2010 年前後因應複雜性科學與社會建構論的發展分別提出。最主要的兩位先驅學者是 Korporowicz (2011)，他在波蘭教育界探討評估理論的演進時，正式發表了朝向第五代評估發展 (Evaluation as animation: Towards the fifth generation evaluation) 的專書章節。另一位是 Lund (2011, 2013)，她因應組織轉型與社會建構主義，撰寫了探討第五代評估 (Fifth generation evaluation) 概念的經典文章。

六、第五代政策評估之文獻探討

(一) 對 El Dessouky (2016) Public policy evaluation theory: From first to fifth generation 的文獻探討

該文確實嘗試建構並容納了第五代政策評估的核心內涵，將其定義為以「社會建構論 (Social Constructionist Paradigm)」為哲學基礎、以「行動研究 (Action Research)」為核心工具的評估典範。然而，若放在當代政策評估理論的整體脈絡下來看，該文對於「第五代」的定義與國際學術界最主流的共識（如 Brousselle 等學者提出的「解釋世代」）存在顯著的路徑差異。以下針對該論文有否完整容納第五代政策評估的重要內涵，進行深度的學術評論：

1. 理論與實踐的緊密結合：該論文所容納的「第五代」核心內涵 El Dessouky (2016) 在文中將第一代至第四代評估理論（承襲 Guba & Lincoln 的經典分類：測量、描述、判斷、回應性建構）進行了系統性梳理，並在此基礎上大膽提出了第五代政策評估 (Fifth-Generation Evaluation)。該文所容納的第五代重要內涵包含：本體論與認識論的轉移（從建構到社會建構）：論文明確指出，第五代評估必須從第四代的「建構主義 (Constructivist Paradigm)」走向「社會建構論 (Social Constructionist Paradigm)」，這意味著評估不再只是個別利益關係人主觀認知的集合，而是著重於知識如何在社會互動、語言交流與權力關係中被「共同生產」出來。方法論工具的創新(行動研究)：該文將行動研究 (Action Research) 視為第五代評估的核心工具。強調評估者不應只是旁觀的協調者(第四代角色)，

而必須將「理論與實踐」緊密結合，透過評估直接介入政策改善的循環。主客體關係的深化（社會建構對話）：第五代評估強調研究者與所有利益關係人之間，需建立更深層次的社會建構對話（**Social Constructive Dialogue**），以優化公共決策的品質與正當性。

2. 理論侷限與國際主流共識的「落差」：雖然 El Dessouky 成功勾勒出她所定義的第五代內涵，但從政策科學的視角來看，該文在容納「第五代政策評估重要內涵」上，存在以下關鍵性的侷限：1. 與國際主流的「解釋世代（**Explanation Generation**）」分道揚鑣在當代政策評估文獻中，多數公共政策學者（如 Brousselle & Buregeya, 2018; Pattyn 等人）所公認的第五代政策評估，其實是「解釋世代（**Explanation Generation**）」。主流第五代的內涵：重點在於解決複雜政策干預下的「因果機制」，核心工具是以理論為導向的評估（**Theory-Based Evaluation**），如巴森與提利的「現實主義評估（**Realist Evaluation**）」、梅恩的「貢獻分析（**Contribution Analysis**）」。其旨在回答政策「在什麼脈絡下、對誰、為什麼有效或無效？」該文的落差：El Dessouky 的第五代本質上更像是「第四代回應性評估的延伸與激進化版」（深化了參與和行動），反而忽略了 21 世紀解決「政策複雜性」與「機制解釋」這一主流評估學界最核心的第五代內涵。2. 概念仍流於「胚胎期」（**Embryonic Paradigm**）作者在論文的結論中也坦言，她所提出的第五代典範仍處於「胚胎階段（**Embryonic Paradigm**）」，仍需要被進一步詳細闡述」。文中對於「行動研究如何具體操作於高度政治化的公共政策環境」缺乏堅實的實證支持與框架建構，使得該論文的第五代內涵顯得理想色彩較重，操作化定義不足。

3. 總評：El Dessouky (2016) 的論文有容納她自身學術脈絡下的第五代政策評估內涵（即：社會建構論 + 行動研究 + 決策優化對話）。它成功為政策評估的「民主參與」與「實踐批判」開闢了新視野。

(二) 對 El Dessouky (2026). *Navigating Post-Pandemic Paradigms: Epistemological Insights and Emerging Trends in Public Administration and Policy* 的文獻探討

該論文雖然不是直接探討第五代政策評估。但深具第五代政策評估內涵。其探討公共行政與公共政策(PA/PP)學門在後疫情時代的知識論(Epistemological)變革與新興趨勢，引入「範式轉移」與「利害關係人動態(Stakeholder Dynamics)」兩大概念，將其視為引導政策變更，並進一步調節知識論發展與公共行政/政策新興趨勢的關鍵干擾變數(Moderating Factors)。在「後疫情時代」的大背景下，

作者提出三個要點，引申這篇論文的核心觀點：

1. 知識論的轉變：從「效率理性」走向「韌性與公共價值」傳統的公共行政（如新公共管理，NPM）深受經濟學與實證主義影響，高度強調效率、成本控制與市場化。然而，COVID-19 疫情與氣候變遷等全球性危機，暴露了過度追求「精實（Lean）」的政府在面對突發災害時的脆弱性。引申思考：論文中提到的「知識論變革」，意味著學術界與實務界正從過去追求「利潤與效率」的實證主義，轉向強調社會韌性（Resilience）、公共價值（Public Value）與社會正義的後實證主義或批判理論。研究公共政策不再只是算計 KPI，而是探討政府如何在不確定性中保障多元公民的福祉。

2. 利害關係人動態與範式轉移的干擾效應，將「範式轉移」與「利害關係人動態」列為干擾變數，這暗示政策變更不再是政府關起門來的單向輸出。數位科技具雙刃劍，後疫情時代加速了數位治理（AI、大數據）的普及，改變了民眾與政府的互動模式。公民、NGO、科技巨頭等「利害關係人」在政策制定中的權力槓桿正在重組。這種利害關係人權力的此消彼長，導致了傳統官僚體制的「範式」難以為繼。傳統的政策網絡變得更碎片化也更具衝突性，使得政策變遷的預測與管理變得高度複雜。

3. 「新政府角色」的具體實踐。敏捷與協同治理論文結論強調需要「新政府角色」來改革制度與服務，這呼應了近年國際上熱議的敏捷政府（Agile Government）與網絡治理（Network Governance）。新時代的政府不能再單打獨鬥，也不能只當一個冷漠的「守門人」或「管制者」。新角色更像是一個「平台促成者（Platform Facilitator）」或「生態系調度者」。政府需要打破部門壁壘（處置「煙囪式官僚」體制），並透過跨域協力（公私協力、跨國合作）來應對通膨、地緣政治、公共衛生與數位轉型等複合型危機（Polycrisis）。

參、研究方法

本論文研究方法採用「文獻回顧與概念比較分析」，以下說明其內容與方法、應用、以及操作的步驟。

一、文獻回顧與概念比較分析：內容與方法

1. 文獻回顧（Literature Review）的內涵：文獻回顧並非單純的文獻摘要累積，而是對特定研究主題既有學術成果的系統性梳理、批判與整合。其核心目的在於：確立研究定位：找出既有研究破口（Research Gap）。建構理論基礎：為研究主題尋找合適的解釋模型或變項。歷史脈絡梳理：理解該概念隨時間演進的發展歷程。

2. 概念比較分析（Conceptual Comparative Analysis）的方法：概念比較分析旨在釐清多個相似、相異或演進中概念的本質。常用方法包含：多維度比較法：設定統一的比較基準（如：哲學基礎、評估者角色、核心價值、主要方法），透視不同概念的異同。歷史演進分析：依據時間序列或範式轉移（Paradigm Shift），解析概念如何從第一代演進至第五代。矩陣分析法（Matrix Analysis）：利用圖表或表格將複雜的概念進行結構化對比，以提升研究的視覺化與思維清晰度。

二、 應用研究：探討第五代政策評估的興起

政策評估的發展歷程從早期強調客觀量化的實證主義，逐步演進至強調利害關係人參與的第四代評估（Guba & Lincoln, 1989）。而近年興起的第五代政策評估（Fifth-Generation Policy Evaluation），則是為了回應全球化、氣候變遷與大數據時代等「棘手問題」（Wicked Problems）而誕生。以下透過文獻回顧與概念比較，解析第五代政策評估的興起背景與核心特徵：

1. 歷史脈絡與範式轉移：政策評估學者將評估的演進劃分為不同世代。第一代至第三代普遍偏向實證主義與技術導向；第四代則引入了建構主義與回應性評估。然而，當代公共政策面臨高度不確定性，促使第五代評估的出現（Dunn, 2018; Patton, 2020）。

2. 世代概念比較分析：為了明確定義第五代政策評估，我們將其與前幾代（以最具有代表性的第三代、第四代為對比）進行比較。見表 1。

表 1
政策評估第三、四、五代之比較

比較項	第三代政策評估 （判斷導向）	第四代政策評估 （回應/建構導向）	第五代政策評估 （系統變革/發展導向）
哲學基礎	邏輯實證主義 （Positivism）	社會建構主義 （Constructivism）	系統思考與複雜科學 （Complexity Science）

比較項	第三代政策評估 (判斷導向)	第四代政策評估 (回應/建構導向)	第五代政策評估 (系統變革/發展導向)
評估者角色	獨立的客觀裁判者	促進對話的協調者、 談判者	共同演進的催化劑、變 革夥伴 (Patton, 2021)
核心價值	效率、效能與客觀科 學性	利害關係人多元觀點 與民主參與	系統適應力、永續性與 即時創新 (Williams, 2019)
主要方法	實驗設計、量化統計 分析	質性訪談、焦點團 體、焦點對話	大數據預測、動態系統 建模、發展性評估 (Dunn, 2018)
處理問題	線性、因果關係明確 的簡單問題	涉及多元價值衝突的 複雜問題	高度不確定、跨域的棘 手與生態系統問題

3. 第五代政策評估興起的關鍵核心

透過上述比較可知，第五代政策評估的興起並非否定前人，而是範式轉移。其核心在於「發展性評估」(Developmental Evaluation)，強調在政策環境高度變動時，評估者必須融入決策團隊，提供即時的反饋(Real-time Feedback)，以協助政策進行動態調整(Patton, 2020, 2021)。

本研究採取「文獻回顧與概念比較分析」，針對政策評估的典範演進進行質性對照研究。分析文本篩選自奠基性專書與權威期刊，包含 Patton (2010)、Boulton 等人 (2015)、Barbrook-Johnson 等人 (2021)、Sanderson (2000) 的文獻及英國財政部 *Magenta Book* 官方補充指南 (HM Treasury, 2020)。

三、操作化步驟

第五代政策評估(The Fifth Generation of Policy Evaluation)的核心精神，在跳脫傳統「線性因果」與「單一最優解」的靜態評估，改以複雜系統理論(Complex Systems Theory)為基礎，強調政策在異質行動者相互動態交織下所產生的「湧現(Emergence)」現象。在此範式下，基於行動者的模型(Agent-Based Modeling, ABM)與社會網路分析(Social Network Analysis, SNA)成為最關鍵的方法論組合。

為深化方法論的操作化（Operationalization），本研究特別將第五代政策評估的核心技術工具——行動者模型（ABM）與社會網絡分析（SNA）——之學術運作步驟規範如下：

（一）行動者模型（ABM）之學術運作步驟

ABM 旨在由下而上（Bottom-up）地模擬異質性個體在政策介入後的微觀行為與宏觀湧現結果。其標準學術操作流程包含五個核心步驟：

【步驟 1：行動者與環境定義】→【步驟 2：行為規則與演算法擬定】→【步驟 3：模型程式化與校準】→【步驟 4：數據驗證】→【步驟 5：政策情境模擬與湧現分析】

1.行動者與環境定義（Agent & Environment Definition）：界定系統中的核心「行動者」（Agents，如居民、企業、官僚）及其屬性（如預算限制、環境偏好），並定義行動者互動的「空間或網絡環境」。

2.行為規則與決策演算法擬定（Decision Rules & Behavioral Logic）：基於行為科學理論，撰寫「若……則……（If-Then）」的啟發式決策規則。例如：當鄰近社區有超過 30% 的住戶安裝綠能設備時（環境刺激），該行動者轉向綠能的意願將提高 15%。

3.模型程序化與校準（Model Programming and Calibration）：使用特定模擬軟體（如 NetLogo、Python Mesa 套件）將規則程式化。利用實證調查數據或微觀統計資料設定模型初始參數，確保模擬環境的異質性符合客觀現實。

4.經驗驗證（Validation & Sensitivity Analysis）：將模型執行的宏觀產出與歷史真實軌跡（Historical Data）進行比對，檢驗模型是否能準確「重現」現實系統。並透過敏感度分析，測試微調某一微觀參數時，宏觀系統是否會發生非線性跳躍。

5.政策情境模擬與湧現分析（Policy Scenario Simulation）：在模型中導入不同的政策干預方案（如不同的碳稅稅率、補貼退場機制），讓系統自動運行千百次。觀察系統在經歷非線性交互作用後，最終「湧現（Emergence）」出何種意料之外的集體行為或政策盲點。

(二)、社會網絡分析 (SNA) 之學術運作步驟

SNA 聚焦於政策執行網絡中，跨部門與公私協力關係的結構、密度及演變軌跡。其學術運作步驟如下：

【步驟 1：網路邊界界定與節點抽樣】→【步驟 2：關係數據蒐集與矩陣建構】→【步驟 3：網絡結構指標計算】→【步驟 4：個體節點中心度分析】→【步驟 5：縱貫性網絡演變分析】

1. 網絡邊界界定與節點抽樣 (Boundary Specification)：明確定義研究網絡的範疇。例如在「再生能源推動網絡」中，節點 (Nodes) 需包含經濟部能源署、地方政府、綠能開發商、地方 NGO 及社區發展協會等實體。

2. 關係數據蒐集與矩陣建構 (Relational Data Collection)：透過結構化問卷、深度訪談或官方公文往來紀錄，蒐集節點間的關係數據 (如資訊流、資金流、信任關係)。將關係數據轉化為二元 (0 或 1) 或加權的「鄰接矩陣 (Adjacency Matrix)」。

3. 網絡結構指標計算 (Whole-Network Metrics)：使用 UCINET 或 Gephi 軟體進行計算。分析整體網絡的密度 (Density) (評估跨部門協力的緊密程度)、集中度 (Centralization) (評估網絡是否過度依賴單一核心機關) 及凝聚子群 (Cohesive Subgroups) (找出內部是否存在本位主義的封閉派系)。

4. 個體節點中心度分析 (Egocentric Centrality)：計算個別行動者的點度中心度 (Degree Centrality) (誰掌握最多資源)、中介中心度 (Betweenness Centrality) (誰扮演跨界溝通的「橋接者」) 及接近中心度 (Closeness Centrality) (誰能最快傳遞或接收政策訊息)。

5. 縱貫性網絡演變分析 (Longitudinal Network Evolution)：蒐集多個時間點 (如政策規劃期、政策執行期) 的網絡數據，比對網絡結構的動態演變。檢視公私協力關係是否隨著時間流逝而深化，亦或在政策受挫時走向破碎化，以此評估治理系統的韌性。

肆、第五代政策評估的內涵

一、第四代政策評估如何轉化為「第五代政策評估」？

為了修正第四代的「過度主觀、缺乏效率與實據」的缺點，近年來公共行政學界與政策實務界進一步演進，發展出強調科技融合、大數據與永續治理的第五代政策評估。為了解決第四代政策評估的缺點，**第五代政策評估**在典範上全面升級，定位為「循證與智慧科技評估」（Evidence-Based & Intelligent Evaluation）。

1. 從「主觀建構」轉化為「知證」途徑（Evidence-Informed Approach）決策：第五代評估不再偏廢任何一方。它重新找回前三代的科學工具，但保留第四代的人文關懷，強調將「利害關係人的心聲」與「大數據/科學證據」進行交叉檢證（Triangulation）。其具體做法：評估者不只聽民眾的「關切」，同時也會調取公共數據（如財稅、醫療、交通等大數據）進行量化推估，讓評估結果既有「溫度」又有「數據」支撐。

2. 從「事後/階段性協商」轉化到「全生命週期與即時反饋（Real-Time）」：第四代評估多在政策告一段落或面臨衝突時才召開協商。第五代則強調政策的全生命週期管考，利用數位工具進行「動態監測」。其具體做法：透過數位公民參與平台、社群媒體文本分析（Sentiment Analysis），即時捕捉社會大眾最新的大眾情緒與意見，讓政策在執行過程中就能「滾動式修正」，大幅提升評估與回應的效率。

3. 從「在地利害關係人」轉化到「全球永續與跨世代正義」：第四代評估聚焦於「當前、在場」的利害關係人利益。而第五代政策評估則將評估視野擴大至永續發展目標（SDGs）與「未來世代」。其具體做法：在評估一項能源或建設政策時，不僅僅協商當代居民與商家的衝突（第四代），還必須透過 AI 與科學模擬，強制將「淨零碳排」與「50 年後的子孫權益」納入評估指標（第五代），以達成跨世代的公平正義。

二、第五代政策評估的理論脈絡

目前學術界對於第五代政策評估主要有兩種最具代表性的理論脈絡：一是 Brousselle 與 Buregeya（2018）提出的「解釋世代」（Explanation Generation）其代表理論包括 1. 雷波恩（Ray Pawson）與蒂利（Nick Tilley）的「現實主義評估（Realist Evaluation）」（Pawson & Tilley, 1997），以及 2. 梅恩（John Mayne）的「貢獻分析法（Contribution Analysis）」（Mayne, 2012；Mayne, 2019）。；二是 Lund（2011, 2013）基於社會建構論所發展的「反思對話與實踐趨向」。

(一) 脈絡一：以「解釋」為核心的第五代評估 (Brousselle & Buregeya 流派)

1. 產生背景：傳統的前三代評估側重於測量、描述與裁判，而第四代評估則強調多元利害關係人的協商與建構。然而，隨著現代公共政策日益複雜（例如大型社會福利、跨部門公共衛生或永續發展政策），評估者發現僅僅知道政策「有沒有效果」（傳統實證主義）或「各方的主觀認知為何」（建構主義）已不足以應付現實。政策決策者與實務界更迫切需要理解：政策是在什麼樣的「脈絡背景」下、透過何種「因果機制」、為什麼（Why）以及如何（How）產生了特定結果？為了彌補過去評估缺乏因果解釋力的缺憾，第五代「解釋世代」應運而生。

2. 學理基礎：此流派的核心學理基礎為「理論驅動評估」（Theory-Based Evaluation, TBE）與「批判現實主義」（Critical Realism）。它主張評估不能只看「投入」與「產出」的黑盒子，而必須去解構政策運行的內在邏輯。其核心方法論支柱包括：

(1) 現實主義評估 (Realist Evaluation)：現實主義評估由 Pawson 與 Tilley (1997) 提出，強調「脈絡—機制—結果」(Context-Mechanism-Outcome, CMO) 的配置分析，探討政策在何種環境下會啟動何種心理或社會機制。此途徑是對第四代評估「過度主觀建構」的核心批判。它轉向「以理論為基礎的評估 (Theory-based Evaluation)」，主張評估不應僅止於回答政策「有沒有效」，而必須聚焦於解釋「政策在什麼脈絡 (Context) 下、透過什麼機制 (Mechanism)、對誰、產生了什麼結果 (Outcome)」。

現實主義評估反對傳統評估中「某政策是否有效」的單純因果問法，主張評估應該回答：「在什麼情況下，哪些機制會對哪些人產生什麼影響？」其核心內涵：

A. 「脈絡-機制-產出」模型 (Context-Mechanism-Outcome, 簡稱 CMO 模型)。它認為政策或方案本身不會直接導致結果，而是政策提供了資源或機會，觸發了參與者的心理、認知或行為反應（機制），而這些反應又受到周遭環境（脈絡）的制約。

(a) 核心公式：「脈絡 (Context, C) — 機制 (Mechanism, M) — 結果 (Outcome, O)」：指方案實施的背景環境。包括地理位置、組織文化、參與者的社經背景、現有的法律制度等。相同的政策在不同的脈絡下，會產生完全不同的效果。機制

(Mechanism, M)：指方案如何發揮作用的隱藏邏輯。通常包含兩部分：方案提供的資源 (Resources)，以及參與者對該資源的推理與反應 (Reasoning)。機制通常是隱形、肉眼不可見的。產出 (Outcome, O)：方案實施後所產生的預期或非預期結果。

現實主義評估是第五代政策評估「解釋性取向」的哲學基石。該理論反對將政策視為「刺激—反應」的因果黑盒子，其核心心法在於建構並驗證由「脈絡 (Context, C) — 機制 (Mechanism, M) — 結果 (Outcome, O)」交互運作的 CMO 構念架構 (CMO Configurations)：

(a) 脈絡 (Context, C)：指政策推行時既存的社會結構、文化規範、法律制度、經濟條件或歷史背景。脈絡本身不直接產生結果，但它決定了政策的內部「機制」是否能夠被成功「激活 (Activated)」。相同的政策在都市 (脈絡 A) 可能高度有效，在偏鄉 (脈絡 B) 則可能因派系網絡與土地細碎化而失效。

(b) 機制 (Mechanism, M)：機制絕非政策活動本身，而是政策所提供的資源或機會 (Resources)，與行動者內在認知、推理、動機 (Reasoning) 之間發生交互作用的隱藏過程。「政策不直接改變行為，是政策改變了行動者的推理，進而引發行為質變」。

(c) 結果 (Outcome, O)：機制在特定脈絡下被激活後所產生的實質效應。現實主義評估所認知到的結果通常是「分化 (Differentiated)」且非線性的，致力於捕捉「對誰有效、對誰無效」以及意料之外的副作用 (Unintended Consequences)。

其動態交互方程式為：將 Outcome (結果) 視為 Context (情境) 與 Mechanism (機制) 交互作用後的函數結果：

$$O = f(C, M)$$

- **O**：表示結果 (Outcome)
- **f**：表示函數關係 (Function)
- **C**：表示情境變數 (Context)
- **M**：表示機制變數 (Mechanism)

B. 主要特點生成性因果觀 (Generative Causation)：不同於量化研究的「繼起性因果 (A 導致 B)」，現實主義評估認為因果關係是「生成」的——只有在特定的環境 (C) 下，方案引發了內在心理或結構的改變 (M)，才會帶出結果 (O)。理論導向 (Theory-driven)：評估的重點不是去測量最終數據，而是去檢驗、修正或細緻化關於方案如何運作的「方案理論 (Program Theory)」。

(2) 對現實主義評估的評論：現實主義評估在公共政策、社會福利及公共衛生領域 (如社區治安、醫療方案評估) 受到高度重視，但也存在實務操作上的挑戰。

A. 正面貢獻 (優點)，提升政策的可複製性 (Scalability)：傳統評估只告訴我們「這個在 甲 社區成功了」，但換到 乙 社區可能就失敗。CMO 模型能解釋「為什麼在 甲 成功、乙 失敗」，幫助決策者了解政策轉移時需要調整哪些脈絡條件。打開政策的「黑盒子」：它不只看投入與產出，更深入探討人的心理動機與社會結構的互動，讓評估結果更具深度與解釋力。具備高度包容性：在方法論上抱持務實態度 (Methodological Pluralism)，不排斥量化 (如統計數據)，也不排斥質性 (如深度訪談)，只要能拼湊出 CMO 關係的方法皆可使用。

B. 限制與挑戰 (缺點) 操作定義模糊且複雜：在實務上，研究者常常很難清晰界定什麼是「脈絡 (C)」，什麼是「機制 (M)」。

例如：參與者的「教育程度」在某些脈絡下是環境背景，但在某些方案中卻可能轉化為影響推理的機制。

耗費大量時間與成本：為了挖掘隱藏的機制與脈絡，評估者通常需要進行大量的質性訪談、文獻回顧與反覆的理論驗證，不適合需要快速出爐的短期評估。難以提供絕對的「黃金標準」：對於追求「單一、明確、可量化效果」的政治決策者而言，現實主義評估給出的結論通常是複雜且有條件的 (「如果...就會...」)，可能無法直接滿足「這筆預算該不該砍」的立即行政需求。

(3) 貢獻分析 (Contribution Analysis)：貢獻分析由 Mayne (2012) 提出，用以在無法進行嚴格隨機實驗的複雜環境中，推論政策對最終結果的貢獻程度。貢獻分析法的核心邏輯在於「建立一個令人信服的因果故事 (Performance Story)」。

它不追求測量精準的「歸因百分比」 (例如：本計畫佔了 60% 的功勞)，而是透過實證資料證明：在考慮了其他外部因素後，該計畫確實是導致改變發生的關鍵原因之一。

梅恩提出了執行貢獻分析法的六個核心步驟：步驟 1：確定所要探討的因果問題。明確界定計畫預期達成的改變是什麼、需要什麼證據，以及存在哪些相互競爭的解釋 (外部干擾因素)。

步驟 2：發展變革理論 (Theory of

Change, ToC)。繪製出詳細的邏輯模型，說明計畫投入、產出、短期成果到長期影響的因果鏈條，並明確指出其中的假設（Assumptions）與風險。步驟 3：收集既有的因果證據。優先使用現有的監測數據、過去的評估報告或學術文獻，來驗證變革理論中各個因果鏈條是否成立。步驟 4：組裝因果故事，並評估其挑戰。將證據填入變革理論中，檢視這個因果故事是否合理。此時必須辨識出「鏈條中的薄弱環節」以及「是否有其他外部因素（Alternative Explanations）也能解釋這個成果」。步驟 5：收集額外的新證據。針對步驟 4 發現的薄弱之處或主要的競爭性解釋，進行針對性的資料收集（如深度訪談、問卷、專家意見），以強化因果鏈。步驟 6：修正並建構最終的貢獻故事。整合所有證據，形成一個合乎邏輯、有實證支持且能回應質疑的「貢獻聲明（Contribution Claim）」。

(4) 對貢獻分析法的評論：a. 主要優點（Strengths）突破傳統實驗設計的限制：在公共政策與跨國發展計畫中，很難或甚至不道德去設置「對照組」（如不給予某地區醫療援助）。貢獻分析法提供了一套嚴謹的質性與量化混合路徑，在沒有對照組的情況下也能進行合理的因果推論。承認並包容「複雜性」：傳統評估常假設「A 導致 B」，但現實是「A+C+D 共同導致 B」。貢獻分析法主動將外部環境、其他利害關係人的行為納入分析，更符合現實世界的運作邏輯。促進組織學習與反思：透過對「變革理論」的反覆驗證與修正，管理團隊能清楚看到計畫在哪個環節斷裂、哪些假設不成立，對於計畫的修正與疊代極具價值。提高評估結果的可信度：它不美化成果，而是透過主動尋找「反面證據」與「競爭性解釋」來進行自我防禦，這反而讓最終的評估報告更具說服力。b. 主要缺點與挑戰（Weaknesses & Challenges）無法提供精確的量化歸因：對於習慣看「KPI 達成率」或「淨影響（Net Impact）數值」的決策者而言，貢獻分析法給出的「故事與論證」可能顯得不夠直觀，無法回答「每一塊錢投入換來多少具體產出」的效益問題。對評估者的能力要求極高：此方法極度依賴變革理論的建構。如果評估者對該領域的洞察力不足，漏掉了關鍵的外部因素或核心假設，整個「貢獻故事」就會流於形式或產生偏誤。時間與資源成本高昂：六個步驟涉及多次的證據檢視與補充調查，特別是當既有資料不足時，需要耗費大量時間進行質性訪談與多方論證（Triangulation）。主觀性的質疑：儘管梅恩強調證據的客觀性，但在決定「何種程度的證據算充分」以及「如何回應競爭性解釋」時，仍帶有評估者一定程度的主觀判斷。

(5) 總結：貢獻分析法是複雜方案評估(Complex Evaluation)的一座里程碑。它將評估的焦點從「證明(Proving)因果關係」轉向「解釋(Explaining)因果機制」。在實務應用上，它非常適合用來評估政策倡議、社會創新、社區營造等難以量化且受多方因素影響的長期計畫。

(二) 脈絡二：以「關係與反思對話」為核心的第五代評估

1. 產生背景：由丹麥學者 Lund (2011, 2013) 提出。她觀察到，過去數十年來公共組織過度濫用評估，導致大量的「非預期負面後果」(Unintended Negative Consequences)。例如：基層員工為了迎合評估指標而耗費大量行政資源，或為了達標而扭曲政策初衷(如應付故事實、造假數據)，反而損害了組織內部信任與公共服務品質。為了解決這種「評估異化」與複雜環境帶來的挑戰，第五代評估被設計為一種更具實踐導向、能將非預期衝擊降到最低的評估模式。

2. 學理基礎：此流派深刻植根於社會建構論(Social Constructionism)與複雜科學(Complexity Science)。

(1)社會建構論與關係轉向：認為組織和政策的現實是透過語言與持續的對話「共同建構」出來的。評估不應是外部專家前來考核的工具，而必須是組織內部進行反思性對話(Reflective Dialogue)的過程。評估不再只是事後的檢討與問責，而是一個推動組織與社會品質提升的動態過程(Evaluating for Quality)。評估者必須與社群共創「期望的未來」，透過持續的反思與集體行動，達到邊評估邊修正的互惠效果。

(2)欣賞式探詢(Appreciative Inquiry)與行動研究(Action Research)：評估的焦點從「挑錯與課責」轉向「發現組織既有的優勢」，並藉由評估過程直接促進組織成員的合作與共生。

(三)、 兩大流派之對比

表 2

解釋世代與反思世代政策評估對比

比較項	Brousselle & Buregeya (解釋世代)	Lund (反思對話世代)
產生背景	現代政策因果關係過於複雜，傳統評估無法合理解釋「為何/如何」有效。	傳統評估流於課責與指標考核，帶來嚴重的非預期行政負擔與扭曲。

比較項	Brousselle & Buregeya （解釋世代）	Lund （反思對話世代）
主要學理基礎	批判寫實主義、理論驅動評估（TBE）。	社會建構論、關係溝通、複雜理論。

三、第五代政策評估之四大取向

第五代政策評估（Fifth-Generation Policy Evaluation）是為了因應當代「棘手問題」（wicked problems）與全球化社會的極高不確定性而發展出的全新範式。它超越了第一至四代傳統評估偏重技術測量、實地描述、價值判斷及利害關係人建構的侷限，核心轉向將政策視為一個複雜適應系統（Complex Adaptive Systems, CAS）。第五代評估不再一味追求線性因果關係的實證與究責，而是強調複雜性科學、反思對話、民主治理及跨世代永續等四大取向的系統性整合，旨在協助政策在動態多變的環境中進行靈活的調適與協同演進（Guba et al., 2018）。分述如下，見表 3。

（一）複雜性科學取向（Complexity Science Approach）

傳統評估將政策視為線性的「投入—產出」模型，而複雜性科學取向則將政策情境視為動態且非線性的生態系統。

1. 動態網絡：政策利害關係人之間存在高度交織的互動，微小的行為改變可能引發巨大的連鎖效應（蝴蝶效應）。

2. 自組織與湧現（Emergence）：評估焦點從「預期結果的達成率」轉向觀測系統中自然產生的新模式與非預期後果。

3. 發展性評估：引進動態反饋機制，讓評估者成為政策團隊的戰略夥伴，邊評估、邊修正，以利政策靈活調適。

（二）反思對話取向（Reflexive Dialogue Approach）

此取向深受社會建構論影響，強調評估行為本身就會對組織與政策環境產生質變（Lund, 2018）。

1. 打破客觀中立迷思：評估者並非局外觀測者，其評估指標與方法會直接形塑受評者的行為與認知。

2. 持續性反思（**Reflexivity**）：要求評估者與決策者不斷審視自身的隱性假設、權力偏見及評估工具的潛在副作用。

3. 協同意義建構（**Sensemaking**）：透過定期的反思對話，協助政策社群對複雜現象建立共識，將評估轉化為組織學習的過程。

（三）參與式與民主治理取向（**Participatory and Democratic Governance Approach**）

本取向旨在扭轉由「技術專家」或「高層官僚」壟斷的評估權力，貫徹多元在地民主。

1. 多元主體參與：納入傳統上被邊緣化的弱勢群體，賦予其平等的評估話語權，不讓評估流於單一階級的工具。

2. 協商民主（**Deliberative Democracy**）：評估過程即是公民參與的論壇，透過利害關係人的充分對話，協商出彼此認可的評估指標。

3. 社會轉型與賦權：評估不再僅是為了行政究責，更是為促進社會正義、增進社區自主治理能力而服務。

（四）永續與跨世代評估取向（**Sustainable and Intergenerational Evaluation Approach**）

此取向將評估的時間與空間範疇拉長至「全球生態」與「未來世代」的宏觀視野（**United Nations Development Programme [UNDP], 2021**）。

1. 跨越當代框架：政策評估不能只看當前（例如 3 至 5 年）的經濟效益，必須將未出世世代的權益與環境承载力納入計分板。

2. 地球藍圖評估（**Blue Marble Evaluation**）：強調全球一體化思維，檢視地方政策是否符合全球永續發展目標（**SDGs**）。

3. 跨世代韌性：系統性評估政策在面對氣候變遷、資源枯竭等長遠挑戰時，能否為下一代留下足夠的調適彈性與生態資產。

表 3

四大取向之核心特徵對比

評估取向	核心關注點	評估者的主要角色	核心評估工具與方法
複雜性科學	系統適應力、非線性湧現、動態調適	系統思想家、調適性夥伴	系統動力學、發展性評估、多 Agent 模擬

反思對話	組織學習、假設審視、協同意義建構	促進者、共同學習者	行動研究、反思工作坊、敘事探究
參與式民主	社會正義、多元賦權、民主問責	賦權者、民主協商催化劑	參與式行動研究、焦點小組、利害關係人協商
永續跨世代	生態邊界、跨世代正義、全球韌性	地球公民守護者、長遠策略家	藍圖評估指標、生命週期評估、跨世代衝擊分析

伍、實例應用檢閱

一、國際經驗：英國國家政策指南與 CECAN 的跨界評估

為驗證第五代政策評估的實踐可行性，英國政府與相關研究機構的嘗試最具代表性。英國財政部於 2020 年特別與「國家級跨界複雜性評估中心（Centre for the Evaluation of Complexity across the Nexus, CECAN）」合作，頒布了《國家政策評估法定補充指南：處理解析政策評估中的複雜性》（HM Treasury, 2020）。這象徵著複雜性科學已正式落地為官方的政策評估指南。

在實務應用中，CECAN 將第五代評估框架全面導入至食物、水資源、能源與環境等跨界政策（Nexus Policy）的評估之中。以環境與能源跨界政策為例，傳統評估僅單獨計算「碳減排量」或「經濟補助成效」（還原論），卻常忽略補助政策可能導致農地過度轉作光電、進而衝擊糧食安全與水土保持的非線性連鎖反應。

透過第五代評估方法，評估團隊首先運用「**參與式系統繪圖（Participatory Systems Mapping）**」，邀請農業、經濟、環保團體與地方政府共同繪製政策系統結構圖，找出隱藏的反饋迴路與槓桿點。接著，利用「**現實主義評估（Realist Evaluation）**」，不再偏執於「政策是否全面成功」，而是去挖掘「在農村人口老化（脈絡）下，透過合理的綠能誘因（機制），對哪些特定農民（對象）產生了兼顧生計與轉型的正向結果」。最後，透過「**動態演化評估（DE）**」建立週期的反饋機制，讓決策者得以及時微調補助門檻，避免對生態系統造成不可逆的湧現傷害。

二、 在地實踐：台灣能源轉型的棘手治理與「環社檢核」機制

當代台灣推行的「2050 淨零轉型」與再生能源（特別是地面型光電與漁電共生）發展政策，堪稱公共政策界最典型的「棘手問題」。該政策涉及國土規劃、能源安全、產業轉型、生態保育及地方派系利益，其系統內部充斥著高度的不確定性、非線性反饋與利益衝突。本研究嘗試引進第五代政策評估框架，對台灣能源轉型政策進行模擬檢閱：

（一） 台灣能源轉型的 CMO 構念分析

若以現實主義評估的 CMO 框架剖析台灣光電政策的實務困境，可重構出以下兩組對立的 CMO 構念鏈：

1. 正向 CMO 鏈（預期因果）：

(1) 脈絡（C1）：台灣中南部日照充足、沿海地區存在大量因地下水超抽導致地層下陷、無法耕作的「不利農業經營用地」。

(2) 機制（M1）：政府提供躉購費率（FIT）的長期躉購保障（資源），激活了老邁農民追求穩定租金收益、免除耕作勞頓的理性推理（推理）。

(3) 結果（O1）：農地快速釋出，不利耕作區之光電裝置容量在短期內呈現爆發式線性增長。

2. 負向 CMO 鏈（湧現副作用）：

(1) 脈絡（C2）：台灣農村土地產權細碎、地方派系與光電掮客關係錯綜複雜、環境生態敏感區與農漁業生產基地高度重疊。

(2) 機制（M2）：龐大的商業利益（資源），激活了部分地方勢力進行土地炒作、尋租行為（推理）；同時也激活了環保團體與養殖漁民對「漁電共生」破壞原生生態與漁業生計的相對剝奪感與抗爭動機（推理）。

(3) 結果（O2）：系統湧現出社會不信任與在地抗爭等嚴重的非線性反饋，導致多項重大光電案場陷入停滯。

(二) 系統科學工具 (ABM 與 SNA) 在台灣能源轉型的應用設計

面對上述複雜系統的湧現危機，傳統事後 KPI 審計完全失效。第五代政策評估則以系統科學工具 (ABM 與 SNA) 落地實踐：

1. 基於行動者的模型 (ABM)：行動者規則與實證校準：ABM 則是扮演「政策沙盒」的角色。它在電腦中建立微觀行動者的行為規則，並用實證資料進行參數校準 (Calibration)，以模擬不同政策工具下，系統的宏觀湧現結果。

(1) 行動者規則設計 (Actor Rules)：以「民眾與企業裝設屋頂型太陽光電」的政策推動為例，模型中的主體 (Agents) 主要包含：住戶/企業主 (Households/Firms)。其決策規則由以下微觀邏輯驅動：

(2) 狀態屬性：

$$\text{Agent}_i = \{A_i, I_i, E_i, L_i\}$$

A ：屋頂面積。 I ：家庭收入。 E ：環境意識強度。 L ：所在縣市。

規則 1：經濟理性決策 (門檻規則)

行動者評估裝設光電的「淨現值 (NPV)」。當預期回報大於其心理門檻 (θ_i) 時，啟動裝設意願。

$$\text{NPV}_i = \sum_{t=1}^T \frac{\text{FIT}(L_i) \times \text{Generation}_i - \text{Maintenance}_t}{(1+r)^t} - \text{Setup Cost}$$

$$\text{If } \text{NPV}_i > \theta_i \rightarrow \text{Adopt Intent} = \text{True}$$

規則 2：社會網絡與鄰近效應 (Peer Effect)

行動者的決策受社會網絡鄰居影響。若其社交網絡 (或同社區) 中裝設光電的比例高於一定值，其環境意識或信任度會被觸發提升。

$$E_i(t+1) = E_i(t) + \beta \times \left(\frac{\text{Adopted Neighbors}(t)}{\text{Total Neighbors}} \right)$$

β ：社會網絡鄰居影響係數。

規則 3：政策適應規則

當政府調降躉購費率（FIT）或取消經濟補助時，行動者調整其決策門檻（ θ_i ）。

(3) 如何由實證資料校準（Empirical Data Calibration）

若 ABM 的參數純靠憑空想像，模型將失去政策評估價值。必須透過台灣當地的實證資料進行參數校準（Parameter Calibration）：

步驟一：微觀行為參數初始化（Micro-data Calibration）

a.空間與環境資料：利用內政部國土測繪中心、氣象署的日照資料，初始化不同地區（L）的屋頂面積（A）與實際發電潛能（Generation）。

b.行為意願普查：參考學術界或經濟部關於「民眾安裝綠能意願」的抽樣問卷。利用問卷中「費率降到多少錢你就不想裝？」的數據，擬合出風險厭惡門檻（ θ_i ）的統計分佈（如常態分佈或對數常態分佈）。

步驟二：巨觀歷史軌跡擬合（Macro-historical Calibration）

a.歷史數據比對：擷取台灣 2016 至 2024 年「各縣市歷年太陽光電實際累計裝設容量」作為歷史真實軌跡。

b.演化演算法優化（Evolutionary Algorithms / Reinforcement Learning）：將未知的微觀參數（如鄰近效應權重 β 、資訊傳播速率等）設為未知數。在虛擬環境中輸入 2016 年的初始政策條件（當時的 FIT 費率），讓模型開始運行。

利用電腦進行上萬次模擬，並透過遺傳演算法（GA）或強化學習 Heuristics，不斷修正微觀參數組，直到模擬出的「2016-2024 裝設成長曲線」與台灣經濟部統計的「實際歷史成長曲線」之間均方誤差（MSE）達到最小化。

步驟三：敏感度分析與政策逆向實驗（Policy Sandbox）

當模型成功校準、能精準重現歷史後，即可將其投入政策前瞻評估。例如：模擬「若 2027 年全面取消光電補貼，但改為強制新建物全面屋頂裝設」的政策。這時模型中異質行動者的連鎖適應反應，就能湧現出最接近真實社會的政策衝擊結果。

2. 社會網路分析 (SNA) 評估台灣能源轉型：研究設計

SNA 用於評估政策網絡的結構、核心推動者、以及政策傳遞與執行的瓶頸。本文以「台灣能源轉型（如：邁向 2050 淨零碳排、綠能推動、2030 能源配比等）」為主題設計的實證研究。

(1) 研究範圍與邊界 (Network Boundary)

a.分析焦點：評估台灣推動再生能源（離岸風電與太陽光電）政策執行網絡中，行動者間的「資訊流」、「政治聯盟」與「審查程序網絡」。

b.時間跨度：以蔡英文政府（2016-2024 年）的「20-30-50 能源配比公式」至賴清德政府（2024 年起）的淨零政策轉型為軸線。

(2) 節點定義 (Nodes: 異質行動者)

將台灣能源轉型網絡劃分為以下多層級治理節點：

a.政府核心機關：國家發展委員會、經濟部能源署、環境部氣候變遷署、地方政府（負責土地變更與發照審查）。

b.市場與國營事業：台灣電力公司（台電，掌握電網併聯併網權）、中油、民營光電/風電開發商、綠電購買企業（如 TSMC 等 RE100 企業）。

c.公民社會與知識媒介：台灣氣候行動網絡（TCAN）等環境 NGO、在地受影響社區/漁會、學術智庫。

(3) 關係線定義 (Edges: 連結強度與屬性)

定義三種網絡矩陣（以有向或無向矩陣呈現）：

a.資訊與諮詢線：政策公聽會、法規諮詢的互動頻率。

b.資源與合約線：綠電躉購（FIT）、企業綠電採購協議（CPPA）的簽署關係。

c.衝突/阻礙線：針對地方回饋金、環境影響評估（環評）爭議所產生的訴訟或抗議關係。

(4) 資料佐證與來源 (Empirical Data Sources)

為確保研究符合嚴謹的經驗研究，必須透過以下四大資料流交叉驗證 (Triangulation) 以建構網絡矩陣：

a.官方治理資料庫：至「臺灣博碩士論文知識加值系統」及行政院/經濟部公報，挖掘歷年能源政策白皮書、公聽會出席名單、環境影響評估會議紀錄，以此建構「共同出席網絡 (Affiliation Network)」。

b.新聞媒體與公眾輿論語料庫：利用台灣主要媒體資料庫（如聯合知識庫、中央社、PTT/Dcard 輿論）進行文本挖掘 (Text Mining)，提取行動者共同出現 (Co-occurrence) 的頻率與態度（正負向情緒），轉化為立場網絡。

c.NGO 倡議與產業報告：蒐集台灣氣候行動網絡 (TCAN) 的淨零評估報告、歐洲商會 (ECCT) 低碳白皮書，用以勾勒民間與外商對政府審查效率（如地方政府卡關光電許可）的壓力網絡。

d.深度訪談與問卷（第一手資料）：向風電/光電開發商、經濟部官員、能源學者發放結構式社會關係問卷（如「在遭遇電網併聯困難時，您最常尋求哪三個單位協助？」），將關係量化。

(5) 政策評估指標運用

a.點度中心性 (Degree Centrality)：評估發現台電與經濟部仍具絕對支配權；若地方政府的「入度」高但「出度」低，顯示地方審查程序成為綠能推動的實質瓶頸（結構洞阻礙）。

b.凝聚度 (Cohesion) 與派系分析 (Subgroup Analysis)：評估執政黨、非政府組織與環評委員之間是否存在高度碎片化，解釋為何近年綠能建置容量出現非預期性落後。

(三) 應用之效益

1.運用行動者模型 (ABM) 動態模擬政策演化：研究者可建構一個包含「農漁民、光電開發商、地方官僚、生態保育行動者」的 ABM 虛擬台灣農村系統。在模型中輸入不同的政策變數，例如：「若提高漁電共生中的漁業產量稽查頻率（政策干預）」，模擬個體行為在非線性交互作用後，宏觀系統會走向「漁民真

轉型、綠能真共生」，還是會湧現出「開發商集體撤資、導致國家綠能目標跳票」的極端後果。

2.運用社會網絡分析（SNA）破解地方治理僵局：研究者可針對台灣綠能審查體制建構 SNA 網絡模型。透過計算「中介中心度」，評估會發現，政策推行受阻的主因往往在於中央部會（核心節點）與地方環保 NGO、在地農漁會（邊緣節點）之間缺乏有效的資訊流橋接者，導致溝通成本過高。同時，利用「凝聚子群」分析可揭示出隱藏的地方利益共同體，協助政府重新設計審查流程。

（四）環社檢核機制法制化之批判性詰辯

為緩解綠能推動所引發的非線性衝突，民間團體與政府共創了「環境與社會檢核（環社檢核）」機制，從地面型漁電共生開始試辦。環社檢核本質上即是第五代政策評估中「行動研究取向」與「動態演化評估（DE）」的在地化體現，其核心功能在於選址與規劃階段進行「議題辨認」，排除高生態爭議區位，並與多元利害關係人進行深度對話。然而，該機制目前在實務推展上面臨嚴峻的「法制化」瓶頸：

現行法制困境：【再生能源發展條例】—— 缺乏上位母法明確授權 ——> 僅具地方政府審查「行政指導」效力無法有效拘束開發商與地方尋租體系

1. 母法授權之位階缺位：現行環社檢核主要依據經濟部所頒布之行政命令或作業要點（如漁電共生審查作業要點）進行。在台灣現行法制中，現行《再生能源發展條例》雖歷經多次修正，納入屋頂光電與地熱專章，卻始終缺乏對於「環社檢核」機制的上位母法明確授權。此一法源位階之低落，導致環社檢核在法律性質上偏向「行政指導」或地方政府審查的附款，缺乏實質的強制力與行政處分機關的課責依據。

2. 與《環境影響評估法》的雙軌脫鉤：依據台灣現行《開發行為應實施環境影響評估細目及範圍認定標準》，除位於重要濕地外，多數地面型太陽光電不予實施環境影響評估。環社檢核雖試圖填補此一法制空白，但因其並非法定環評程序，導致開發商在實務上面臨「雙軌審查」的程序混亂。環社檢核所辨認出的生態與社會對策，無法在《環境影響評估法》或《國土計畫法》中獲得剛性的法律拘束力，一旦開發商違反檢核承諾，主管機關難以依法對其進行裁罰或廢止再生能源設備登記。

3. 法律效果之非常規化不確定性：環社檢核強調「邊做邊評、滾動微調」的動態反饋迴路。然而，傳統行政法高度講求「法律確定性」與「信賴保護原則」。當環社檢核在「關注減緩區」要求案場動態修正設計時，常與業者已取得的容許使用或躉購費率權益發生衝突。在缺乏明確法制化規範的情形下，此類程序往往流於經濟部與地方派系之間的政治博弈，未能轉化為常規化的制度保障。

因此，從第五代政策評估的角度審視，台灣能源轉型的環社檢核機制若要真正發揮「系統導航員」的功能，必須透過《再生能源發展條例》或《氣候變遷因應法》進行實質修法，將環社檢核正名為法定的「前置選址與適應性治理程序」，方能化解法律真空引發的非線性治理危機。

伍、研究發現與討論

一、政策評估典範之演進對照

透過對文獻與實例的深入評析，本研究整理出政策評估典範演進的整合對照表（見表 4）。

表 4

政策評估的典範演進對照表

演進階段	核心焦點	方法論基礎	評估者角色	複雜性科學的定位
第一至三代	測量、描述、價值判斷	實證主義、定量、實驗設計	技術專家、裁判	忽略複雜性，視政策為機械封閉系統。
第四代	利害關係人協商、回應性	建構主義、定性、質性對話	協調者、談判者	注意到多元主體的複雜性，但缺乏系統機制的科學解釋。
第五代 (複雜性洗禮後)	因果機制解釋、系統動態、適應性進化	後實證主義、現實主義、計算機模擬與系統繪圖	系統導航員、批判性共創者	全面擁抱複雜性，將政策視為複雜適應系統（CAS）。

二、核心討論

(一) 超越「黑盒子」與「口水戰」的辯證融合

研究發現，第五代政策評估成功整合了前幾代的優點。它既不像第一至三代將政策因果簡化為冰冷的線性黑盒子，也不像第四代陷入極端相對主義的溝通內

耗（唯我論）。第五代評估透過「後實證主義」與「現實主義」的哲學基調，承認社會現象的客觀存在，同時包容主體理解的多元性，提供了一種更具包容性且科學化的動態解釋框架。

(二) 方法論工具的實踐門檻與「技術還原論」風險

儘管 ABM 模型與 SNA 分析等系統科學工具能提供極具價值的動態模擬，但在實務上面臨嚴峻的數據門檻與技術壁壘。複雜適應系統中的非線性數據往往難以標準化蒐集，且政府評估人員若缺乏計算科學素養，極易使這些新興工具淪為美化報告的「新技術還原論」，這是推行第五代評估時必須警惕的盲點。

(三) 政治問責體制的結構性衝突

傳統政府運作高度依賴確定性與短期 KPI 問責。然而，第五代評估核心在於擁抱不確定性與承認政策可能引發無法預測的湧現後果。當政策評估者從「場外裁判」轉變為「系統共同進化者」時，傳統文官體制往往難以接受評估報告中出現「動態調整」而非「達標與否」的彈性結論。這顯示出第五代政策評估的推廣，不僅是方法論的升級，更是政府治理文化的結構性革命。

陸、結論與建議

一、 結論

當代公共治理的環境已由機械式的封閉系統蛻變為網狀、動態且充滿不確定性的複雜適應系統。本研究評析顯示，最新第五代政策評估在複雜性科學的洗禮下，已成功突破傳統線性因果與第四代主觀建構的瓶頸。藉由轉向以理論為基礎的現實主義評估與動態演化評估，第五代評估賦予了現代政府「邊做、邊評、邊動態調整」的治理韌性，深刻重塑了「循證政策（evidence-based policy）」的科學內涵。儘管在技術門檻、數據壁壘與政治問責體制上面臨結構性挑戰，且台灣本土的「環社檢核」實踐亦遭逢法律位階缺位與非常規化之制度瓶頸，但其作為 21 世紀公共治理新典範的趨勢已然確定。

二、 建議

針對未來政府推動政策評估之實務，本研究提出以下三項具體建議：

(一) 上位母法修正，推進「環社檢核」機制法制化

建議政府積極推動《再生能源發展條例》或《氣候變遷因應法》之實質修法，明確增訂「環境與社會檢核專章」或授權條款，賦予環社檢核剛性的法律效果與裁罰依據。藉由法制化程序，將動態演化評估（DE）正式納入國家重大前瞻性政策的常規化法定程序中，脫離單一 KPI 的束縛。

（二）建構跨學門的系統科學工具庫，克服技術壁壘

政府應與學術機構深化合作，建置政策動態模擬平台，積極引進行動者模型（ABM）與社會網絡分析（SNA）。同時，應加強對公共行政文官體系的「系統思考（Systems Thinking）」培訓，將計算科學素養納入高階文官訓練課程，降低新技術工具的應用門檻，避免流於形式。

（三）推動「實驗型」的彈性問責文化與動態演化評估

建議審計與決策部門鬆綁傳統的事後靜態對照問責制。在高度不確定的棘手政策中，允許評估者扮演「系統導航員」角色，將「及時發現盲點、促進跨界溝通並滾動微調政策」本身，視為正向的治理績效指標，從根本上鼓勵政府組織發展出持續適應學習的韌性治理能力。

柒、研究限制

本研究雖透過文獻深化與國際、在地實例檢閱，系統性地建構了最新第五代政策評估的學術論述與應用藍圖，但在理論推導與實務操作的診斷上，仍存在以下三項不可規避的研究限制：

一、跨範疇數據整合之「次級資料偏誤」與「模擬不確定性」

在方法論操作上，本研究提出了行動者模型（ABM）與社會網絡分析（SNA）的學術運作步驟。然而，「複雜適應系統」（CAS）的核心在於微觀異質主體的非線性交互作用，其所需的數據多具備高動態性與隱蔽性。由於本研究屬於典範評析與架構建構，在實例應用檢閱中主要依賴次級文獻、官方指南（如英國 *Magenta Book*）與政策報告進行質性推論，缺乏第一手的微觀模擬數據驗證。未來研究若欲實際運行 ABM 或 SNA，將面臨大範圍異質數據蒐集的技術壁壘，且模擬模型本身的參數設定（Calibration）亦存在「過度簡化現實」的風險。

二、質性個案脈絡的「外推有效性（External Validity）」限制

本研究深化了現實主義評估的「脈絡—機制—結果（CMO）」理論架構，並以台灣能源轉型的「環社檢核」法制化困境作為在地實踐個案。然而，現實主義評估在本質上高度依賴特定政治、經濟與文化環境的「脈絡特殊性」。台灣農村土地產權細碎、特有的地方派系生態，以及《再生能源發展條例》與《環境影響評估法》的雙軌脫鉤，皆屬於高度本土化的特殊脈絡（Context）。因此，本研究基於台灣個案所推導出的負向 CMO 鏈與法制化修法建議，在面臨不同法律體系（如美歐的環評與土地信託制度）或不同政策範疇（如跨國 AI 治理、流行病防治）時，其「外推有效性」將受到限制，無法直接複製。

三、理論框架之「時間落後（Time Lag）」與實證期程限制

第五代政策評估理論全面擁抱複雜科學中的「湧現性」與「時間動態演化」。然而，政策效果的湧現往往需要長期的追蹤觀察。以台灣能源轉型政策為例，其目標遠景橫跨至 2050 年，而現行的「環社檢核」機制從試辦到目前的法制化爭議，實踐歷程僅約五至六年間，系統內部的長期生態效應與社會變遷尚未完全定型。本研究受限於當前的實證期程，僅能對既有的法制體制進行結構性斷面診斷，難以完整捕捉該政策在未來數十年間，因環境變遷與技術突破（如儲能技術或碳捕捉）所引發的深層二階或三階湧現效應。

參考文獻

- 行政院（2023）。**臺灣 2050 淨零轉型十二項關鍵戰略行動計畫**。行政院。
- Barbrook-Johnson, P., Proctor, A., & Bicket, M. (2021). Policy evaluation for a complex world: Practical methods and reflections from the UK Centre for the Evaluation of Complexity across the Nexus (CECAN). *Evaluation*, 27(1), 6–16.
- Boulton, J. G., Allen, P. M., & Bowman, C. (2015). *Embracing complexity: Strategic perspectives for an age of turbulence*. Oxford University Press.
- Brousselle, A., & Buregeya, J.-M. (2018). Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation*, 24(2), 153–168.
- Brousselle, A., & Buregeya, J. M. (2018). Theory-based evaluation: A pillar for the fifth generation of evaluation. *Evaluation*, 24(2), 242–256.
- Dunn, W. N. (2018). *Public policy analysis: An integrated approach* (6th ed.).

- Routledge.
- El Dessouky, N. F. (2016) Public policy evaluation theory: From first to fifth generation. *EPRA International Journal of Economic and Business Review*, 4(4),15-25.
- El Dessouky, N.F. (2026). Navigating Post-Pandemic Paradigms: Epistemological Insights and Emerging Trends in Public Administration and Policy. In: Cantafio, G., et al. *International Conference on Sustainability and Innovation Processes and Systems*. ICSIPS 2025. Lecture Notes in Networks and Systems, vol 1806. Springer, Cham. https://doi.org/10.1007/978-3-032-16389-9_46
- Gilbert, N. (2008). *Agent-based models*. SAGE Publications.
- Greenhalgh, T., & Russell, J. (2009). Evidence-based policymaking: A critique. *Journal of Health Services Research & Policy*, 14(3), 174–181.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. SAGE Publications.
- Guba, E. G., Lincoln, Y. S., & Lynham, S. A. (2018). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (5th ed., pp. 108–150). SAGE Publications.
- Head, B. W. (2008). Three lenses of evidence-based policy. *Australian Journal of Public Administration*, 67(1), 1–11.
- HM Treasury. (2020). *Magenta book supplementary guide: Handling complexity in policy evaluation*. Her Majesty's Stationery Office.
- Korporowicz, L. (2011). Ewaluacja jako animacja. W kierunku ewaluacji piątej generacji [Evaluation as animation: Towards the fifth generation evaluation]. In G. Mazurkiewicz (Ed.), *Ewaluacja w nadzorze pedagogicznym. Refleksje* (pp. 35–53). Ośrodek Rozwoju Edukacji.
- Lund, G. E. (2011). Femte generations evaluering [Fifth generation evaluation]. *EvalueringsNyt*, 27, 4–11.
- Lund, G. E. (2013). Fifth generation evaluation. *Taos Institute Online Organization Website*. haslebo-partnere.
- Lund, G. E. (2018). *Fifth generation evaluation: Developing new practice- oriented methodologies based on social constructionism*. ResearchGate.
- Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, 18(3), 270–280.

- Mayne, J. (2019). Revisiting contribution analysis. *Canadian Journal of Program Evaluation*, 34(2), 171–191.
- Oxman, A. D., Lavis, J. N., Lewin, S., & Fretheim, A. (2009). SUPPORT Tools for evidence-informed health Policymaking (STP) 1: What is evidence- informed policymaking? *Health Research Policy and Systems*, 7(Suppl 1), S1.
- Patton, M. Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. Guilford Press.
- Patton, M. Q. (2019). *Blue marble evaluation: Premises and principles*. Guilford Press.
- Pawson, R. (2006). *Evidence-based policy: A realist's perspective*. SAGE Publications.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. SAGE Publications.
- Rey, L., Brousselle, A., & Champagne, F. (2012). Logic analysis: Testing program theory to better evaluate complex interventions. *The Canadian Journal of Program Evaluation*, 26(3), 61–78.
- Rutter, H., Savona, N., Glonti, K., Syrett, M., & Cummins, S. (2017). The need for a complex systems approach in public health. *The Lancet*, 390(10112), 2602–2604.
- Sanderson, I. (2000). Evaluation in complex policy systems. *Evaluation*, 6(4), 433–454.
- Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Policy & Politics*, 30(1), 1–22.
- United Nations Development Programme. (2021). *Evaluation for a turbulent future: Incorporating complexity, agility, and sustainability into public policy assessment*. UNDP Independent Evaluation Office.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- Williams, K. (2019). *Introducing public policy: Platforms, politics, and practices*. Routledge.

The Rise of Fifth-Generation Policy Evaluation: A New Paradigm of Public Governance Formed by Complexity Science

Shih-Hsien Chang*

Abstract

Facing highly complex 21st-century "wicked problems" such as climate change and energy transition, traditional linear evaluation tools rooted in Newtonian mechanism have proven inadequate, while fourth-generation evaluation often suffers from the transaction costs of subjective, collaborative deliberation. Against this backdrop, fifth-generation policy evaluation has emerged by comprehensively integrating complexity science into public governance.

Employing a literature review and critical discourse analysis, this study systematically examines the theoretical evolution, core approaches, and methodological transformations of this new framework.

The findings indicate that through interpretive realism and dynamic evolutionary evaluation, the fifth-generation paradigm conceptualizes public policy as a complex adaptive system, thereby deepening objective causal explanations through the Context-Mechanism-Outcome framework. Methodologically, it shifts from reductionist tools to systemic methods such as Agent-Based Modeling and Social Network Analysis, redefining evaluators as system co-evolvers. By reviewing UK national policy guidelines and governance dilemmas in Taiwan's energy transition, this study validates the empirical value of the framework in managing non-linear emergent effects.

To address the conflict between data silos and traditional KPI accountability, this study proposes institutionalizing complexity guidelines, constructing systemic simulation platforms, and fostering an experimental accountability culture. Ultimately, this framework offers contemporary governments a new paradigm for building dynamic adaptability and governance resilience.

Keywords: fifth-generation policy evaluation, complexity science, realist evaluation, ABM, fourth-generation policy evaluation

* Professor Emeritus, Department of Public Administration and Policy, National Taipei University. E-mail: chang.shihhsien@gmail.com The paper was published under two double-blind reviews. Received: September 11, 2026. Accepted: September 23, 2026. The paper was published under two double-blind reviews. Received: September 10, 2026. Accepted: September 28, 2026.