

國際組織中長期目標檢視方法研析

林子立(東海大學政治系副教授)

前言：

APEC 於 2020 年發表「2040 太子城願景」(Putrajaya Vision 2040)，並於 2021 年要提出執行行動(Implementation Actions)，這即是今年 APEC 各級會議最重要的工作事項之一。

為了讓 APEC 後續在落實「2040 太子城願景」上能夠按部就班，不僅依循 APEC 的價值與原則持續推進工作，而在推進的過程中，也可透過評估機制來檢視工作進展的狀況；因此，本文透過檢視聯合國、經濟合作暨發展組織、東協、歐盟的中長程計畫為例，分析上述國際組織如何設定評估機制，相關的評估方法或工具分別為何，這將可作為 APEC 後續推進「執行行動」規劃方向之參考。

一、 聯合國永續發展目標

聯合國永續發展目標 (the Sustainable Development Goals，後簡稱 SDGs) 2015 年由聯合國大會批准，是建立在 2000 年批准的千禧年發展目標的基礎，完成沒有實現的目標。¹SDGs 繼續把重點放在減輕貧困和繁榮，以「沒有人落後」(No one is left behind) 的社會包容性需求。結合氣候變化與環境保護等議題，強化社區組織和更廣泛的參與。SDGs 重要的五個領域：人、地球、繁榮、和平與夥伴關係。

事實上，大部分學者同意 SDG 所強調的問題是刻不容緩²。而由個別國家所主導的政策其實是無法解決複雜的全球性問題³。一般而

¹ United Nations. 2015. "Transforming Our World: The 2030 Agenda for Sustainable Development".

² Sachs, J. D. (2005). The end of poverty: Economic possibilities for our time. New York: Penguin Press; Stiglitz, J. E. (2006). Making globalization work (1st ed.,,). New York: W.W. Norton & Co..

³ Easterly, W. (2006). The white man's burden: Why the West's efforts to aid the rest have done so much ill and so little good. New York: Penguin Press; Hoffmann, M. J. (2011).

言，組織內部和外部因素通常是目標是否達成的簡單分類。組織的內部流程、人力資源和組織發展影響 SDGs 的達成。同時，國際政治的外部因素也會導致目標是否達成，特別是大國高度競爭或是全球景氣榮枯，都會影響目標的達成。

後續行動和檢視過程是以國家主導的評估和數據為依據，以國家為中心推動 SDGs 主因在於成效涉及許多不同組織之間的協調，這些組織具有不同的觀點和優先順序不同，評估方法也都不盡相同。有些機構傳統上使用定量方法，而另一些機構則使用定性和參與性方法。此外，由於通訊技術的發達，利用智慧型手機或是平板電腦進行數據收集。大數據分析推文和社交媒體，有效而迅速便宜地收集信息。

從 2000 年的千禧年發展目標開始，評估與檢視的方法已經有重要的進步，並產生重大意涵。許多評估文獻都集中在評估發展項目的影響上，但對於 SDGs 評估的設計必須兼顧國家與部門政策。此外，評估通常涉及許多不同組成部分或項目的計畫。困難度在於涉及大量的執行機構和利益相關者，而且覆蓋範圍廣大，往往沒有關於每個組成部分的實施地點和實施方式的細節。

四種最常見的評估：

- (一) 政策評估：評估政策和廣泛計畫（例如國家計畫和捐助者合作計畫）的設計和實施情況。包含檢視其發展目標的完成度。這些評估側重於上游開發和規劃。許多評估是回顧性的，較多是在一個國家計畫週期結束時進行（持續四到五年）。這些評估中有許多使用經濟合作與發展組織（OECD）發展援助委員會（DAC）的評估標準（例如相關性，效率，功效，影響和永續性），但也可以使用許多其他政策評估方法。
- (二) 形成性評估：形成性評估的目的是向管理階層和其他利益相關者提供定期反饋，以幫助加強計畫和項目的實施。檢視與評估之間有著緊密的聯繫，以確保最大程度利用檢視作為好的管理工具，而不僅僅只是問責制。形成性評估結合了定量和定性方法，結合成一種混合方法。此外，評估作為一種學習工具，在整個計畫和項目週期中都使用形成性評估。該方法基於管理層與評估團隊之間的緊密合作，而檢視方法不同於許多總結性評估，後者常常強調「客觀性」，以保持管理者與評估者之間的距離來實現。形成性評估還包括採用定性和參與性方法，檢視

貧困和弱勢群體所表達聲音並促進社會正義。

- (三) 發展性評估：此種評估目的是，幫助管理者和其他利益相關者提高計畫績效，並學習選擇和設計未來計畫的經驗。但是，發展評估側重於創新計畫以及那些在複雜環境中運行的計畫。在這些環境中強調設計評估和實現目標。因評估而採取的干預措施會不斷發展和適應，而且往往沒有任何完成點。該方法基於管理人員和評估人員之間的密切合作，後者與計畫實施和調整密切相關。
- (四) 總結性評估：總結性評估的目的是，評估預期項目目標中觀察到的變化。可歸因於項目干預效果的程度。這些評估可以是定量的，估計變化的大小和統計意義，也可以採用更定性的方法。其中主要的證據來源之一是受影響人群和其他利益相關者的意見。傳統上，總結性評估用於問責制。⁴

SDGs 的中長期目標檢視是多元而且及其重要的檢視過程，經由追蹤 SDGs 的進程，建立全球指標是極其重要的。然而，這些指標也會因為國家地方及主題的變化而被錯誤引導。原因在於指標無法解釋變化為什麼發生，或是對於不同利益相關者的重要性。簡言之，在不同的範圍內，以何種數據作為成效指標可凸顯出國家政策的重點。

因此，必須理解這些指標的使用，檢視進程並加上評估才是完整的成效檢視。以「2030 年永續發展議程」第 72 條為例，後續行動和審查過程將嚴格的以證據為基礎，並由國家主導的評估和數據可及性，及時性，可靠性，並按收入，性別，年齡，種族，族裔，移民狀況，殘疾和地理位置以及在國家範圍內相關的其他特徵進行分類。

二、OECD

總部設在法國巴黎的經濟合作暨發展組織（Organization for Economic Cooperation and Development, OECD）是由 37 個市場經濟國家所組成的政府間國際組織，其宗旨是幫助會員國實現永續性的經濟增長、就業與生活水準上升。其中，OECD 的發展援助委員會（Development Assistance Committee, DAC）是以論壇形式探討對開發

⁴ Bamberger, M., Tarsilla, M., and S. Hesse-Biber. 2016. "Why so many 'rigorous' evaluations fail to identify unintended consequences of development programs. How mixed-methods can contribute". *Evaluation and Program Planning* 55 (2016): 155-162.

中國家進行援助、發展與解決貧窮等問題。在 DAC 之下，建立發展計畫之評估系統 (Evaluation of development programmes)，目的在於有更好的發展結果，利用評估建立強而有力的證據為基礎的決策。由此可知，評估與檢視不管在 UN 還是 OECD，都是備受重視。

可靠而獨立的評估有助於提高援助政策的效力，並有助於使捐助者和夥伴國政府對結果負責。DAC 評估資源中心 (the DAC Evaluation Resource Centre, DEReC) 從其所有成員的開發機構收集評估報告。DEReC 每週提供超過 3000 份報告，是一個廣泛而獨特的資源。⁵除此之外，亦有「評估網絡」(EvalNet) 範圍涵蓋所有會員國和 9 個多邊組織 (六間開發銀行，歐洲委員會，國際貨幣基金和開發計畫署) 的評估系統，進行跨組織管理評估系統。個別會員國也發展獨立評估功能，長期努力以改善整體複雜機構間的評估協調，並在整個發展合作領域提供獨立監督。舉例而言，英國成立援助影響獨立委員會 (the Independent Commission for Aid Impact, ICAI)，德國則是發展評估機構 (the German Institute for Development Evaluation, DEval)，瑞典則是設立援助研究專家團隊 (the Expert Group for Aid Studies, EBA)，作為監督或審查機構。不僅如此，歐盟也提出「更好法規」計畫改善績效和提高效率的重視。

進一步檢視 OECD 的評估系統，組織內部的基本原則相當清楚：能力、利益相關者的參與和溝通。不僅如此，能力建設、知識管理以及與捐助者和國伙伴的協調。⁶對於最基本的評估研究的文獻檢閱，會員國皆以公開的資料的方式提供政策文件與評估系統。OECD 也提供大量的工作文件，以及多邊、雙邊機構編寫的文件，像是「審慎地應用評估標準」(Applying Evaluation Criteria Thoughtfully) 等許多重要文獻。如此各方都能獲取有關其評估系統的基本評估政策、報告資料、工作計畫和年度審查，甚至是評估類型。

不僅如此，EvalNet 成員評估報告能夠彼此共享，學習其他會員國的成功與失敗經驗⁷。以「回應在發展中國家的難民危機」(Responding to Refugee Crises in Developing Countries: What Can We Learn From Evaluations?) 為例，運用會員國的評估工作，以改善難民危機的未來因應措施。不論是在彌合人道主義與發展之間的差距，或是加強對曠日費時危機的國際反應，都是從評估中獲取更好的結果。

⁵ OECD, 2016. Evaluation Systems in Development Co-operation: 2016 Review.

⁶ OECD, 2010. Quality Standards for Development Evaluation, DAC Guidelines and Reference Series.

⁷<https://www.oecd.org/dac/evaluation/dcdndep/membersofthedacnetworkondevelopmentevaluation.htm>

為了獲得最全面的信息，EvalNet 也採用問卷調查的方式，並用了開放式問題。再根據具體最新發展情況，與 EvalNet 秘書處合作，選擇了要調查的組織，僅可能使各種背景的機構都能參與評估。

不過，評估體系在不斷的發展之下，變得越來越分散。大多數組織報告結合使用集中式評估和分散式評估，目的是提高成本效益，對業務部門的評估具有相關性和自主權，並支持夥伴機構能力的發展。最後，分散式評估多使用在部門、重要主題或戰略性議題，而且大多數的會員國都有政策以規範對援助計畫的評估，因為不精準的評估會導致下次計畫的失敗或是失焦。

OECD 的評估程序

一般而言，評估工作的計畫編制與運營緊密相關，而評估和運營計畫則必須整合。利用諮詢機構、指導委員會、參考小組和特別顧問來監督援助計畫的工作和過程。評估的品質則取決於評估者的能力和知識。過去的評估經驗，以及援助當地的政治、制度、文化和傳統。在進行分散式評估時，一致性和嚴謹性就容易不夠。以圖一所呈現的就是評估過程的重要組成部分，從計畫開始，依序是補充評估工具、品質、能力建立、參與、討論、評估使用，以及最後的知識分享。



資料來源：OECD

圖一：評估程序

大體而言，若是利益相關者參與評估，則有助於援助計畫的永續發展，對於責任釐清也有幫助。儘管如此，只有少數機構在評估計畫的設計階段中，就鼓勵利益相關者參與。大約只有一半的 EvalNet 會員國鼓勵合作夥伴參加指導小組。可想而知，讓援助與受援助雙方能夠共同制定援助計畫，聯合評估就會較為切實可行。然而，這些評估的連貫性卻常常遭遇挑戰，這是由於援助方與受援助方有不同要求和優先順序，也是評估與檢視工作遭遇的困境。

評估結果不僅是提供透明度和滿足發布評估政策想要解決的問題。以 OECD 訪談報告可知，及時使用評價結果和建議是最重要的。同時，評估報告的傳播與被各方接受也同等重要。最後，評估結果和建議的採用與否取決於組織文化和管理層的支持，才能根據評估產生較好的決策。

三、ASEAN Masterplan 2025

在 2008 年全球金融風暴肆虐之下，東南亞國協（ASEAN）仍然有強勁的經濟增長。根據 IMF 統計數據，ASEAN 的 GDP 在 2012 年為 3,055 億美元，預計到 2025 年將達到 6,282 億美元。在此同時，ASEAN 也有諸多挑戰，諸如貿易和投資環境、產業結構的改變、提高生產效率與提升教育水準等等。不僅如此，經濟增長也導致能源和糧食需求增加，連帶的出生率下降和人口老齡化，為勞動力提供有潛在風險。為了使 ASEAN 在迅速改變的世界經濟達成永續的繁榮，從 2015 年提出「東協連結總體計畫 2025」（Master Plan on ASEAN Connectivity 2025），提出五個戰略領域，分別是 1，永續基礎設施；2，數位創新；3，無縫物流；4，卓越的監管；5，人員流動。

官方所提出的期中檢驗報告，評估影響 ASEAN 連結的 10 個主要趨勢，包括城市化、消費階層的增加、生產力的停滯、技能差距的擴大、ASEAN 會員國的合作框架數量的增加、不平等加劇、對生物多樣性的壓力增加、數位技術的發展和日益增加的地緣戰略關注。美中對抗雖然為 ASEAN 帶來轉單效應，但是 COVID-19 對健康和經濟的影響更大。疫情已經對 ASEAN 經濟體以及區域連結性造成打擊。相較於 COVID-19 之前，經濟成長率下降達到了將近 11 個百分點。⁸

ASEAN 持續檢視總計畫並且進行調整，以提高計畫的效率和參與度。此外，也需要採用不同的指標，以幫助評估總計畫包括社群媒體和網站的分析。評估指標包括多個標準，評估報告的資訊的品質，

⁸ World Economic Outlook, IMF, June 2020

評估範圍的掌握。

再以東協安全共同體藍圖（The ASEAN Political-Security Community Blueprint, APSC Blueprint）為例，其設想將 ASEAN 政治和安全合作提升到更高水準，企圖建立一個以規則為基礎的包容性共同體，期使人民享有人權、基本自由和社會正義、生活在安全和有保障的環境中，並有能力應對全球化下的挑戰。

對此宏偉的藍圖加以檢視與評估，才能確保這些活動符合 ASEAN 的需求和優先順序，又能回應區域和全球格局的變化動態。運用秘書處在 2017、2019、2021 以及 2023 年分別進行檢視與評估，並由 APSC 理事會在東協政治安全共同體協調會議（Coordinating Conference for the ASEAN Political-Security Community, ASCCO）的協助下分別於 2020 年和 2025 年進行中期和期末審查。在審查和評估過程中，東協會員國可以靈活地更新「APSC 藍圖 2025」。檢視和評估的結果由東協秘書長通過 APSC 理事會向東協高峰會報告。

對東協社會文化共同體（ASEAN Socio-Cultural Community, ASCC）藍圖的檢視和評估，是運用現有的監測和評估（Monitoring and Evaluation, M&E）系統，以實施為重點的監測系統和 ASCC 計分卡組成。在目前的 ASCC 計分卡的基礎上，各部門機構重新審視其部門指標，在公認的區域成果管理制度的基礎，強化計分卡及其指標。此外，各個部門機構也進行內部監控以及促使其他利益相關者參與評估。

2009-2015 年從 ASCC 計分卡評估結果獲得經驗、教訓和建議，並納入到結果，用以檢視和評估「ASCC 藍圖」的進度。並將結果在反饋到 M&E 系統，把更高標準的目標、關鍵結果領域（key result areas, KRA）和指標進行彙編。

而對於東協經濟共同體的部分，東協經濟共同體理事會（ASEAN Economic Community Council, AECC）定期審核 AEC 藍圖，但不得超過每三年一次。藉由 2016-2020 年和 2021-2025 年的中期和期末評估，以檢視進度並評估結果與影響，包括實施 AEC 藍圖的成就和挑戰，目的在於強化東協的經濟整合水準。

四、Horizon 2020 Europe

儘管當前歐洲有各種政治、經濟、移民、恐攻等挑戰，但歐盟在經濟市場的整合，實現人員、貨物、資金與服務的自由移動，已經創造出歐洲有史以來最和平、繁榮、富裕的時代，也是一個不爭的事實。

歐盟執委會 (European Commission) 也希望將這種整合帶到創新與科學、學術研究上。大部分會員國都支持歐盟的研究和創新政策，因為這會同時為各國與歐盟都產生巨大的競爭優勢。

「科技研發框架計畫」(Framework Programmes for Research and Technological Development) 就是在此企圖心下建立起來，透過科技研究經費挹注的方法彙集頂尖人才一起合作。⁹事實上，在全球化的推波助瀾之下，歐洲各國面對的公共政策問題早已無法以單獨的力量解決，不論是能源、貿易、食品藥物安全、環境治理還是犯罪、恐怖攻擊等傳統安全議題，在歐洲整合之後，原本個別的國內問題成為歐盟共同的挑戰，透過鼓勵組建跨國性的學術創新研究團隊，提高效率、減少重疊，加速以科技與創新方式解決問題；更重要的是，可以增強歐盟在全世界的創新與科技影響力。此計畫與其他歐盟計畫不同之處在於，面臨各種的歐洲社會因全球化、區域化帶來的許多挑戰，集合各方智慧開發技術和創新解決方案，¹⁰不僅可以創造可能新的商業模式帶來利潤，更可以拓展市場占有率和創造新的就業機會，以解決去工業化後的勞動人口失業問題，進一步為歐盟的競爭力提供關鍵支持。

比利時政治家 Étienne Davignon 擔任歐洲工業事務與能源執委 (the European Commissioner for Industrial Affairs and Energy) 決定將能源、環境以及分子生物學等研究計畫整合成為一個單一具有連貫性的架構計畫，這也直接催生了 1984 年第一期 (每期 6 年) 的研究與科技發展架構計畫 (The Framework Programmes for Research and Technological Development) 的誕生，¹¹至今已經進入第八期的 Horizon 2020 (Horizon 2020, 2014—2020)，成為一個非常龐大的全球規模的研究補助計畫。這個企圖心遠大的計畫主要是由歐盟執委會負責，但下轄各個不同總署有不同階層的綜理。比如，研究和創新總署 (The directorate general for research and innovation, DG RTD) 或電信網路、內容與技術總總署 (The directorate general for communications networks, content and Technology)；此外，負責行政作業的機構則有研究執行署 (The Research Executive Agency, REA)、中小企業執行署 (The Executive Agency for SMEs, EASME)、歐盟研究理事會執行署 (European Research Council Executive Agency, ERCEA) 等機構。

歐盟執委會在 2017 年底對 Horizon 2020 進行期中評估，檢視該

⁹ Peter O'donnell & Ben Deighton, "Europe's Framework Programmes: A Key Element of research policy in Europe," *Horizon Magazine*, (Special Issue, March 2015).

¹⁰ European Commission, *Seventh FP7 Monitoring Report Monitoring Report 2013*, DG Research and Innovation - Evaluation Unit, (Brussels: European Commission, 2015).

¹¹ Peter O'donnell & Ben Deighton, *Ibid*, p. 4.

計畫是否持續履行上述政治目標，以及可能需要進行哪些修改才能更有效地實現這些目標。期中評估有助於確認歐洲的科學技術基礎是否獲得加強，並提供可靠的數據來說明這些成就。此外，需要以一種差異化的方式評估對社會的益處，特別是在應對社會挑戰方面。

Horizon 2020 能夠為歐盟創造附加價值：為了產生這種歐洲附加價值，而且沒有一個歐洲國家能獨自應對氣候變化的挑戰，共同科研計畫勢在必行。歐盟能否通過協調一致的行動取得實質性進展，還有待證明。對歐洲附加價值的評估之外，有效分配資金資源也很有用。歐洲合作項目來自科學和工業界的跨界、跨學科合作，以及由此產生的知識、方法、基礎設施、人才與數據等的交流或聯合使用，最後將這些列為檢視與評估報告的核心。

結語

總結上述對於四個國際組織推動計畫評估方法之分析，各個國際組織對於所屬推動之計畫，視計畫之性質、效益、目標與期程等，均設計不同的評估方法，甚至以東協推動的「東協連結總體計畫 2025」來看，也分由三個共同體執行委員會自行設計符合該目標的評估方法，包括以質性或量化數據，據以評估。

因此，鑒於 APEC 成立多年且陸續有中期計畫在進行推動，許多中期計畫也都設有評估機制，因此，在討論如何評估太子城願景執行成效的議題上，除了可參考外部國際組織的經驗外，也應該檢視 APEC 內部正執行的中期計畫，包括中期計畫的後續延伸性、現行評估方法的可用性等，作為太子城願景執行行動的評估基礎。