

Ewaluacja w Europie

Wyzwania przyszłości

FRANS L. LEEUW
European Evaluation Society

Czwarta konferencja Europejskiego Towarzystwa Ewaluacyjnego odbyła się w Lozannie, pod hasłem „Przybliżanie Ewaluacji Ludziom” (tyt. oryg. *„Taking Evaluation to the People”* – przyp. tłum.). Jednym z powodów, dla których konferencję tę zorganizowano w Szwajcarii, był fakt że pośród innych krajów europejskich, zdecentralizowana Szwajcaria najbardziej chyba, na tak wiele rozmaitych sposobów, podkreśla znaczenie obywateli i społeczeństwa obywatelskiego. Badania szwajcarskich badaczy pokazały, że efektywność i wydajność sektora publicznego ściśle zależą od wagi, jaką dana społeczność przykładła do roli, jaką społeczeństwo obywatelskie odgrywa w procesie podejmowania i implementacji decyzji, jak i od ustaleń instytucjonalnych, które tę rolę kształtują (Frey, 1990; Knoepfel, 2000; Pommerehne, 1991; World Bank, 1997).

Ilość badań ewaluacyjnych prowadzonych w Europie wyraźnie rośnie, i przybierają one formę coraz bardziej zinstytucjonalizowaną,¹ co dobrze odzwierciedlone jest przez proces formowania się społeczeństw ewaluacyjnych. Jest to również udokumentowane w badaniach porównawczych na temat rozwoju ewaluacji przeprowadzonych w 30 krajach (Rist i in., 2001). Podobne wnioski płyną z badania, które wraz z innymi badaczami przeprowadziłem przed kilku laty (Leeuw i in., 1999).

Uwzględniając wszystkie te zjawiska, niniejsza praca ogniskować będzie wokół *czterech wyzwań*, którym sprostać musi zarówno *proces ewaluacji w praktyce* jak i *proces badań ewaluacyjnych* we współczesnej i przyszłej Europie.

Pierwsze wyzwanie: Przybliżanie Ewaluacji Ludziom

Pierwsze wyzwanie bierze się już z samego tytułu konferencji: „Przybliżanie Ewaluacji Ludziom”. Kto będzie czerpać korzyści z rosnącej liczby ewaluacji? Czy jedynie menadżerowie, udziałowcy, akcjonariusze, politycy, normodawcy, audytorzy, przemysł, technokraci i biurokraci? Coleman (1990) analizując postęp, jaki dokonał się w badaniach dotyczących polityki społecznej w Stanach Zjednoczonych w latach 50.-80., zauważył iż podczas gdy pierwotnym celem badań naukowych w zakresie polityki (ang. *policy research* – przyp. tłum.) było poddanie wnikliwej ocenie zachowania podmiotów gospodarczych tak w administracji publicznej, jak i w sektorze prywatnym, ten typ badań zdołał przemienić się w samodzielny „przemysł”. Społeczność ewaluacyjna ma już za sobą ponad 30 lat doświadczenia w zakresie wykorzystywania ewaluacji (Sabatier, Weiss, Bucuvalis, Patton, itd.). Często jednak główny nacisk kładziony był na wykorzystanie rezultatów ewaluacji przez podmioty gospodarcze działające wewnątrz i na zewnątrz sektora publicznego – przez parlamenty, ministerstwa, organizacje pozarządowe, organizacje nastawione oraz nie nastawione na zysk. Zasadniczo mniej uwagi poświęcano temu, w jaki sposób przeciętny obywatel/udziałowiec czy też grupa lub stowarzyszenie zwykłych ludzi mogą wykorzystywać te ewaluacje. Pierwsze wyzwanie polega więc na *zwiększeniu* znaczenia ewaluacji dla ludzi, którzy poddani są regulacjom, dla ludzi, którymi się zarządza, dla ludzi, którzy po prostu wykonują swoją pracę. Czy zapewniamy łatwy dostęp do raportów ewaluacyjnych, czy może ograniczamy ich publikację jedynie dla udziałowców gospodarczych reprezentujących kręgi rządowe lub biznesowe? Czy proaktywnie upowszechniamy dane statystyczne i rozsyłamy raporty, czy może, jedynie na żądanie wypuszczamy je do obiegu w formie swego rodzaju profilu firmy skierowanego do potencjalnych inwestorów?

Powyższe pytania o wzmocnienie roli obywateli powinny znaleźć się wysoko na naszej liście celów. W swoim wystąpieniu na wspomnianej konferencji Patton (2000) zasugerował, że ewaluatorzy i audytorzy w swojej codziennej pracy powinni zadać sobie pytanie o to, do jakiego stopnia ich działalność może umocnić demokrację. Fitz-Gibbon (2000), w swoim wystąpieniu zwróciła uwagę na sposób, w jaki jej instytut udziela szkołom informacji zwrotnej, zapewniając im dostęp do danych statystycznych dotyczących szkolnictwa. W Holandii, Instytut Przeglądu Edukacyjnego (ang. *Education Review Office* – przyp. tłum.) wypuszcza coroczne „karty jakości” zawierające dane na temat szkół pierwszego i drugiego stopnia, do których dostęp za pomocą Internetu mogą mieć zarówno rodzice i dzieci, jak i nauczyciele i inne grupy zrzeszające zwykłych obywateli.

Drugie wyzwanie: Oceniając partnerskie państwo i układy o współpracy

Drugie wyzwanie odnosi się do sposobów w jaki osoby oceniające radzą sobie ze społeczeństwem sieciowym, państwem partnerskim czy „układami o współpracy” pomiędzy organizacjami. Partnerstwo podmiotów publicznych i prywatnych, zarządzanie konsorcyjne, „negocjujące” czy też współuczestniczące, są we współczesnej Europie pojęciami bardzo modnymi. Czy wszystkie te układy są nowymi wynalazkami czy jedynie neokorporacyjną przeróbką dawnych form występującą dziś pod nową nazwą, to się jeszcze okaże. Prawdą jest natomiast ich dzisiejsza wszechobecność. Jakie są więc tego implikacje dla osób dokonujących ewaluacji?

Istnieje wiele odpowiedzi na to pytanie (Gray i in., w druku; Marne i in., 2001). Niektórzy twierdzą, że sytuacja ta nie stwarza żadnych nowych wyzwań: osoby oceniające mogą wykorzystywać wszystkie dostępne im narzędzia do ewaluacji istniejących układów o współpracy. Inni sugerują, że zmieniająca się natura relacji pomiędzy podmiotami gospodarczymi w partnerskim społeczeństwie jest tak dogłębna i złożona, że oceniający i audytorzy potrzebują całkiem nowych metodologii i paradygmatów. Zgodnie z trzecim spojrzeniem, ewaluatorzy nie powinni się martwić, gdyż w państwie partnerskim, ewaluacja nie jest potrzebna, ponieważ otoczeni jesteśmy przez zaufanie, zaangażowanie i reputację. Rzeczywiście, ewaluacja relacji opartych na zaufaniu prowadzi do niezamierzonego efektu ubocznego i może nawet doprowadzić do utraty tego zaufania. Ewaluacja w takim przypadku działa więc jako swoisty „pogromca zaufania”.

Przedstawione perspektywy ujęcia tematu pozwalają co prawda zdefiniować sam problem, lecz podchodzą do niego w sposób nazbyt ograniczony. Nie są tym samym w stanie dostarczyć odpowiedzi na pytania o to jak sprostać tym wyzwaniom. Poniżej przedstawione zostaną więc trzy inne spojrzenia na to zagadnienie.

Po pierwsze, ewaluacja może koncentrować się na próbie odpowiedzi na pytanie czy mechanizmy typowe dla układów o partnerstwie, zarządzania sieciowego czy innej współuczestniczącej formy działalności funkcjonują w należyty sposób i czy nie produkują one niewłaściwych bodźców czy dysfunkcjonalnego zachowania. Wskazane jest również zbadanie

społecznych i ekonomicznych korzyści i kosztów układów wysokiego zaufania, gdyż nie można przecież z góry założyć, że istnieje jakiś dowód, zgodnie z którym zaufanie miałoby zawsze działać w społecznie korzystny sposób. Socjologowie pokazali, że grupy typu mafijnego budują wiele relacji wysokiego zaufania i kapitału społecznego, lecz nikt przecież nie może zaprzeczyć, że takie „układy społeczne” niosą ze sobą negatywne konsekwencje (Flap i in., 1998). Zaufanie może również prowadzić do społecznego wykluczenia tych jednostek, które są wyłączone z relacji wysokiego zaufania.

Ramka 1. Elementy sukcesu występujące w układach o partnerstwie

- Społeczny kapitał sieci
- Cel i wynik dyferencjacji
- Kombinacja organizacji partnerskich
- Wzór interakcji
- Stopień osiągalności zasobów
- Struktura układów o partnerstwie
- Zarządzanie zasobami ludzkimi
- Stabilność partnerstwa
- Stopień formalizacji

Po drugie, w partnerskim państwie ewaluacja jest potrzebna aby wnieść porządek do potencjalnie złożonego (niektórzy mogliby wręcz uznać, że chaotycznego) świata wszystkich układów o partnerstwie. I tak na przykład, partnerstwo podmiotów publicznych i prywatnych, zdaje się być proste na etapie formułowania umowy o porozumieniu, lecz za nim ukryta może być cała sieć zależności pomiędzy stronami występującymi po obydwu stronach umowy. Ewaluacja może być więc wykorzystana do zbadania dla dobra tak parlamentu, jak i całego społeczeństwa, zakresu w jakim ludzie i organizacje zaangażowane w daną relację wywiązują się z ustalonych standardów przyjętych dla celów i środków działania.

Po trzecie, ewaluacja może również oceniać sam układ o współpracy. Koszty transakcyjne wysoce zdecentralizowanej współpracy mogą być wysokie, a zatem empiryczna wiedza na temat tych kosztów i samych układów o partnerstwie może okazać się bardzo przydatna. I choć układy o partnerstwie występują w ramach *organizacji* i pomiędzy nimi, *element ludzki i społeczny jest kluczowy* dla sukcesu wszystkich sieci i układów o współpracy. To przecież ludzie wymieniają się ideami, inwestują w siebie nawzajem i swoje sieci powiązań, „tworzą” szklane bariery, które mają ograniczyć pionową mobilność określonych grup społecznych, tworzą (i niszczą) przyjaźnie wewnątrz i pomiędzy tymi organizacjami, proszą o porady i ich udzielają. Wykorzystanie podejść i instrumentów, które socjologia sieci stworzyła w ciągu ostatnich lat może okazać się dla oceniających bardzo pomocne (Marne i in., w druku). Ramka 1 przedstawia podstawowe czynniki wyjaśniające sukces takich układów.²

Trzecie wyzwanie: Pierwszoplanowość teorii objaśniających i programowych

Definicja teorii objaśniającej wykorzystana w tym miejscu zakłada, że teorię tę stanowi zbiór twierdzeń wyjaśniających (zamierzone i niezamierzone) zachowanie ludzi i organizacji oraz wyniki tych zachowań. Przykładami mogą być socjologiczne teorie determinant kapitału społecznego (Coleman, 1990), teoria zachowań zbiorowych wyjaśniająca dlaczego ludzie dobrowolnie uczestniczą w produkcji dóbr prywatnych a nie publicznych (Olson, 1971), teorie kognitywne dotyczące zmniejszenia dysonansu oraz ekonomiczne teorie o ścieżce zależności i kosztach transakcyjnych w naukach ekonomicznych (Picciotto, 2000).

Nie tak dawno „realni” (ale nie realistyczni) ewaluatorzy nie martwili się ani nie zastanawiali nad tymi teoriami, a większość „realnych” teoretyków nigdy nie myślała nawet o ewaluacji. Lee (2000:127) zauważa, że „znani autorzy w dziedzinie

teorii ewaluacji nie uzyskali swojej formalnej edukacji w programie zaprojektowanym do kształcenia profesjonalnych ewaluatorów”. Shadish i in. (1995) w pracy pt. *„Foundations of Program Evaluation”* opisuje teorie związane z konkretnymi jednostkami (takimi jak Scriven, Rossi czy Cronbach) i zauważa, że nie istnieje jedna właściwa teoria ewaluacji. Debata na temat teorii ewaluacji, dobrze podsumowana przez Lee pod nagłówkiem „Zimna Wojna Ewaluacji”, w dużej mierze ogniskowała wokół sporów pomiędzy pozytywizmem, post-pozytywizmem, konstruktywizmem, oraz innymi –izmami.

Kluczowym elementem teorii objaśniającej jest jej niezależny charakter, niepowiązany w jednym *wydarzeniu*, czy konkretną osobą. Wraz z ewaluacją realistyczną efektywna linia myśli teoretycznej wkroczyła w Europie w sferę (praktycznej) pracy ewaluacyjnej (Pawson i Tilley, 1997).

Przydatne jest również rozważenie teorii właściwych dla takich instrumentów oddziaływania jakimi są: subsydia, kontyngenty, kampanie informacyjne, strategie wyboru, vouchery i przepisy, jak również programy działań (Chen, 1990). Istotnym zagadnieniem jest tutaj model wpływu (Rossi i in., 1999). Przez ponad 50 lat, badacze i ewaluatorzy dużo uwagi poświęcali (społecznym i behawioralnym) przesłankom i mechanizmom, które stały za instrumentami i programami, w celu zrozumienia (i przewidywania) powodów dla których programy i instrumenty odnoszą sukces.³ Generalne biuro obrachunkowe (ang. *General Accounting Office*) w Stanach Zjednoczonych nazwało to „rekonstrukcją modeli leżących u podstaw (proponowanych) programów” i wprowadziło rozróżnienie pomiędzy rekonstrukcją conceptualną i operacyjną. Conceptualna rekonstrukcja dotyczy społecznej i behawioralnej logiki stojącej za programem, podczas gdy rekonstrukcja operacyjna koncentruje się na tym, *dlaczego* można założyć, że *konkretni* aktorzy wykonają *konkretne* działania w celu zapewnienia sukcesu danemu programowi.

We wcześniejszej pracy zdefiniowałem pojęcie teorii polityki (ang. *policy theory*), jako:

system społecznych i behawioralnych zachowań, które leżą u podstaw działań publicznych, które zostały przeformułowane w postaci przesłanek (lub twierdzeń). Przesłanki te odzwierciedlają wierzenia decydentów politycznych co do stopnia zrozumienia, postaw i zachowań tych grup społecznych, których dotyczą konkretne działania... Odnoszą się one również do bardziej strukturalnych czynników, na podstawie których decydenci opierają swoje założenia.
(Leeuw, 1991)

Vedung (1997: 138) mówi na temat „teorii interwencji”, które zawierają „wszystkie (dokonane) empiryczne i normatywne presupozycje i założenia zawarte w programie działania przez jego pierwotnych twórców”. Metodologie służące rekonstrukcji „leżących u podstaw teorii” są dostępne w podręcznikach, jak również w formie inteligentnego oprogramowania, takiego jak *Belvedere* and *Sensemaker* (Leeuw, 2000b).

Istnieje cała grupa podstawowych założeń stanowiących podstawę tych podejść. Po pierwsze, rekonstrukcja i ocena logiki kryjącej się za konkretnym programem (działaniem) jest niezbędna do zidentyfikowania przyszłych możliwości dla tego programu. Im solidniejsze są przesłanki/założenia, na których program lub instrument bazuje, tym większa szansa na powodzenie danego programu. Po drugie, logiki te powinny być zrekonstruowane, ponieważ nie są zwykle w wyraźny sposób przedstawiane przez decydentów politycznych, polityków czy biurokratów. Po trzecie, nie istnieje żaden dowód *a priori*, że założenia te – czy „ulubione teorie” – są słuszne, to jest logicznie spójne i empirycznie poprawne.

Biorąc pod uwagę powszechnie obowiązujący konsensus zakładający, że *praktyka ewaluacji* oraz *teoria objaśniająca i programowa* są powiązane, nasuwa się istotne pytanie dotyczące tego w jaki sposób poprawić jakość teorii leżących u podstaw polityk. Czego można się nauczyć z nauk społecznych i behawioralnych na temat determinant zachowania indywidualnego i grupowego? W jaki sposób przełożyć badania poznawcze, przekraczające granice dzisiejszego stanu wiedzy (ang. *frontier research* – przyp. tłum.) z takich dziedzin jak socjologia, psychologia i ekonomia, do naszych praktyk ewaluacyjnych w sposób bardziej produktywny niż tylko poprzez „wypisywanie” zmiennych? Problem ten został poruszony w ciekawym artykule Tilleya (1999) pt. *„Evaluation and Evidence-(Mis)Led Policy”*, w którym przedstawia w jaki sposób badania ewaluacyjne mogą źle pokierować procesem kształtowania i wdrażania polityk. W jednym z przykładów Tilly przedstawia ewaluatorów mylnie odtwarzających teorie polityki, przez co udzielają oni niewłaściwych porad tak decydentom politycznym, jak i całemu społeczeństwu.

Czwarte wyzwanie: Niezamierzone i niepożądane uboczne skutki ewaluacji

Wraz ze zwiększaniem się zakresu prowadzonych ewaluacji, należy zwrócić baczniejszą uwagę na niezamierzone i niepożądane skutki uboczne takich działań. Jedno z pierwszych spostrzeżeń na ten temat można odnaleźć w badaniach Donalda Campbella przeprowadzonych we wczesnych latach 70. (Shadish i in., 1995). Obecnie Robert Schwarz pracuje nad książką, która ma zgłębić poniższy dylemat.

Bardzo pożądaną jest, aby móc zapewnić menażerom, decydentom politycznym, ustawodawcom i szerokiej opinii publicznej informacji ewaluacyjnej, która byłaby rzetelna, wiarygodna i niezawodna. Informacja, która nie posiadałaby takich cech, ma bowiem niewielką szansę na poprawę przejrzystości, odpowiedzialności i demokratycznej zaręczności. Jednakże mechanizmy służące ocenianiu „jakości” (to jest: rzetelności, wiarygodności i niezawodności) informacji ewaluacyjnej budują niewłaściwy obraz tego, co nazwane zostało społeczeństwem audytu, charakteryzującym się wciąż rosnącą liczbą inspekcji, audytu, ewaluacji i oceny. Społeczeństwo audytu poświęca ogromne środki na działania ubezpieczające, czego bezpośrednią konsekwencją jest wzrost biurokratyzacji.

Również Smith (1995) omawia kilka innych niezamierzonych konsekwencji. Prowadzenie przeglądu czy inspekcji („audytowanie”) może powstrzymać rozwój innowacji i prowadzić do skostnienia: paraliżu organizacyjnego wywołanego przez system pomiaru wyników działalności. Zwraca również uwagę na problem wąskiego patrzenia (ang. *tunnel vision*) (Smith, 1995). „Wąskie patrzenie można zdefiniować jako koncentrowanie się – w ramach prowadzonego pomiaru wyników – na zjawiskach mierzalnych kosztem elementów niemierzalnych.” Innym niezamierzonym skutkiem ubocznym, zdaniem tego samego autora są: „pod-optimalizacja”, którą można zdefiniować jako „stawianie przez menadżerów wąskich celów lokalnych, kosztem ogólnych, całościowych celów organizacji” oraz „obsesja pomiaru”, czyli „położenie nacisku na (jeden) pomiar sukcesu, nie zaś na leżące u jego podstaw cele”.

Fitz-Gibbon (1997: 87-95) nawiązała do opisanego podejścia Smitha w prowadzonych przez siebie badaniach obejmujących 104 dyrektorów szkół podstawowych w Wielkiej Brytanii. Jedno z pytań zadawanych w ramach tego badania miało na celu ustalenie do jakiego stopnia monitoring, inspekcja i audyt przyczyniają się do efektywności szkoły. Odkryła ona, że:

za wyjątkiem skostnienia, każda z tych możliwości (to jest niezamierzone efekty uboczne...) była komentowana przez dyrektorów w otwartych pytaniach zawartych w kwestionariuszach. Nie są to dylematy teoretyczne, lecz rzeczywiste, zauważalne problemy.
(Fitz-Gibbon, 1997: 87)

Light (1993) przeanalizował zalety i wady rosnącego znaczenia Inspektorów Generalnych (IGs, Inspector-Generals) w rządzie federalnym USA i odkrył, że „wraz ze zwiększaniem się liczby Inspektorów, efektywność może się ulegać zmniejszeniu” (Light, 1993: 223). Jest to zgodne z tak zwanym „paradoksem działania” (Meyer i O'Shaughnessy, 1993), w którym równoczesna proliferacja, jak i brak korelacji metod pomiaru wyników działań, które mają kluczowe znaczenie dla samej idei audytu, występują równocześnie. Organizacja, która zmierza do zabezpieczenia swojej działalności poprzez zastosowanie instrumentów pomiaru wyników działalności nie musi być koniecznie organizacją efektywną.

Przeglądając badania w tym zakresie z ostatnich 5-10 lat (Leeuw, 2000a), dochodzę do wniosku, że spokojnie można zakładać, że ewaluacja, nawet jeśli dobrze zamierzona, będzie niosła ze sobą *niezamierzone i niepożądane skutki uboczne*, które mogą zaszkodzić skuteczności i efektywności sektora publicznego, prywatnego i mieszanych. Niektórzy porównują te „niepożądane efekty” ze „stratami ubocznymi” (ang. *collateral damage*). Nie możemy, i nie powinniśmy, zaprzeczać występowaniu tych efektów ubocznych. Moim zdaniem, zasadniczym wyzwaniem jest odkrycie i rozwój takich metod, za pomocą których można by wychwycić niezamierzone efekty uboczne, zapobiegać nim i je tłumaczyć. Teorie objaśniające i behawioralne mogły by być tu pomocne. Teorie te mogą również pomóc w zrozumieniu dlaczego niektóre uwarunkowania instytucjonalne powodują powstawanie bardziej „niepożądanych efektów” niż inne.

Konkluzje

Praca nad opisanymi powyżej czterema wyzwaniami jest niezbędna nie tylko ze względu na utrzymanie ewaluacji „aktywnej i żywej” w ramach tego, co niektórzy nazywają „społeczeństwem audytu”, lecz również – co wielce prawdopodobne – może ona zwiększyć profesjonalną i intelektualną jakość pracy prowadzonej przez (pojedynczych) ewaluatorów. Mam nadzieję, że społeczności ewaluacyjne obecne dzisiaj w Europie będą mogły być pomocne w tym zakresie poprzez łączenie ze sobą ewaluatorów, zarówno z obszarów badań jak i praktyki, w ramach projektów i innych działań. Społeczeństwo będzie wówczas mogło ocenić do jakiego stopnia nasza praca jest użyteczna.

Przypisy końcowe

¹ Zdaniem Piccitto (2000) jest to rosnący rynek jeśli chodzi o „rozwijający się przemysł”.

² Strona internetowa stworzona przez Międzynarodową Sieć dla Analizy Sieci Społecznych (ang. *International Network for Social Network Analysis*, INSNA – przyp. tłum), o adresie: www.heinz.cmu.edu/project/INSA zawiera przydatne informacje na temat sieci społecznych, oraz odpowiednie odwołania źródłowe i adresy stron internetowych.

³ Socjolog Karl Mannheim w opublikowanym w latach 30. dziele pt. „Człowiek i Społeczeństwo w dobie Przebudowy” zawarł apel o artykułowanie założeń, na których opierano to, co wówczas nazywane było „społecznym planowaniem”.

⁴ Odnosi się do tego również Chen (1990).

Bibliografia

- Chen, T. (1990) *Theory-driven Evaluations*. Newbury Park, CA: Sage.
- Coleman, J. S. (1990) *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.
- Fitz-Gibbon, C. T. (1997) *The Value Added National Project: Feasibility Studies for a National System of Value-added Indicators*. Hayes: School Curriculum and Assessment Authority.
- Fitz-Gibbon, C. T. (2000) 'Keynote', Paper presented at the European Evaluation Society Conference, Lausanne, 12–14 October.
- Flap, H. D., B. Bulder and B. G. H. Völker (1998) 'Intra-organizational Networks and Performance: A Review', *Computational and Mathematical Organization Theory* 4: 109–49.
- Frey, B. S. (1990) 'Institutions Matter: The Comparative Analysis of Institutions', *European Economic Review* 34: 443–9.
- Gray, A., B. Jenkins and F. L. Leeuw (eds) (in press) *Sharing Power: Evaluation and Management in the Collaborative State*. New Brunswick, NJ: Comparative Policy Analysis Series, Transaction Books.
- Knoepfel, P. (2000) 'Policy Evaluation Games in a (Very) Federalistic State', Paper presented at the European Evaluation Society Conference, Lausanne, 12–14 October.
- Lee, B. (2000) 'Theories of Evaluation', in R. Stockmann (ed.) *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. Opladen: Leske & Budrich.
- Leeuw, F. L. (1991) 'Policy Theories, Knowledge Utilization, and Evaluation', *Knowledge and Policy* 4: 73–92.
- Leeuw, F. L. (2000a) 'Unintended Side Effects of Auditing: The Relationship between Performance Auditing and Performance Improvement and the Role of Trust', in W.
- Raub and J. Weesie (eds) *The Management of Durable Relations: Theoretical & Empirical Models for Households and Organization*. Amsterdam: Thela Thesis.
- Leeuw, F. L. (2000b) 'Analyzing Program Logic: Methods Available and Problems to be Solved', Working Paper presented at the European Evaluation Society Conference, Lausanne, 12–14 October.
- Leeuw, F. L., J. Toulemonde and A. Brouwers (1999) 'Evaluation Activities in Europe: A Quick Scan of the Market in 1998', *Evaluation* 5(4): 487–96.

- Light, P. C. (1993) *Monitoring Government: Inspectors General and the Search for Accountability*. Washington, DC: The Brookings Institute.
- Mayne, J., T. Wileman and F. L. Leeuw (in press) 'Networks and Partnering Arrangements in the Collaborative State: New Topics for Evaluation and Auditing', in A. Gray, B.
- Jenkins and F. L. Leeuw (eds) *Sharing Power: Evaluation and Management in the Collaborative State*. New Brunswick, NJ: Comparative Policy Analysis Series, Transaction Books.
- Meyer, K. and K. O'Shaugnessy (1993) 'Organizational Design and the Performance Paradox', in R. Swedberg (ed.) *Explorations in Economic Sociology*, pp. 249–79. New York: Russell Sage Foundation.
- Olson, M., Jr (1971) *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.
- Patton, M. Q. (2000) *Keynote*, Paper presented at the European Evaluation Society Conference, Lausanne, 12–14 October.
- Pawson, R. and N. Tilley (1997) *Realistic Evaluation*. Thousand Oaks, CA: Sage.
- Picciotto, R. (2000) *Keynote*, Paper presented at the European Evaluation Society Conference, Lausanne, 12–14 October.
- Pommerehne, W. (1991) 'The Empirical Relevance of Comparative Institutional Analysis', *European Economic Review* 34: 458–69.
- Rist, R. C., J. E. Furubo and R. Sandahl (eds) (2001) *International Atlas of Evaluation*. New Brunswick, NJ and London: Transaction.
- Rossi, P., H. Freeman and M. Lipsey (1999) *Evaluation, a Systematic Approach*. Thousand Oaks, CA: Sage.
- Shadish, W. R., T. Cook and L. C. Leviton (1995) 'Donald T. Campbell: Methodologist of the Experimenting Society', in W. R. Shadish, T. Cook and L. C. Leviton (eds) *Foundations of Program Evaluation: Theories of Practice*, pp. 119–71. Thousand Oaks, CA: Sage.
- Smith, P. (1995) 'On the Unintended Consequences of Publishing Performance Data in the Public Sector', *International Journal of Public Administration* 18: 377–410.
- Tilley, N. (1999) 'Evaluation and Evidence-(Mis)Led Policy', *Evaluation Journal of Australasia* 11(2): 48–64.
- Vedung, E. (1997) *Public Policy and Program Evaluation*. New Brunswick, NJ: Transaction.
- World Bank (1997) *World Bank Development Report 1997: The State in a Changing World*. Washington, DC: World Bank.
- Tłumaczyła: Joanna Orzechowska