Systemic evaluation, impact evaluation and logic models

Ralph Renger

Shandiin Wood

Simon Williamson

Stefanie Krapp

Ralph Renger (top left) is a Professor at the Mel and Enid Zuckerman College of Public Health at the University of Arizona, Tucson, Arizona. Email <renger@u.arizona.edu>

Shandiin Wood (top right) is a Research Evaluation Specialist at the Mel and Enid Zuckerman College of Public Health at the University of Arizona, Tucson, Arizona. Email: <shandiin.wood@gmail.com>

Simon Williamson (bottom left) is a Senior Evaluation Advisor in the Strategy, Advisory and Evaluation Group with the New Zealand Ministry of Foreign Affairs and Trade, Wellington. Email: <crisisresponse@gmail.com>

Stefanie Krapp (bottom right) is a Senior M&E Officer with the Monitoring and Evaluation Unit at Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), Eschborn, Germany. Email: <stefanie.krapp@giz.de>

Evaluators are yet to define systemic evaluation adequately. With respect to impact evaluation, systemic evaluators are critical of current methods such as simple logic models, or logframes, for capturing the contextual complexity in which programs operate. Despite a lack of clarity about what constitutes a systemic evaluation, many evaluators are suggesting it is a better alternative for capturing the true complexity of the context in which a program operates. However, capturing complexity is itself complex, resulting in amorphous and ambiguous approaches and results. This can create problems in relation to evaluation capacity building, usability of results, and evaluator credibility. This article demonstrates how methods such as simple logic models or logic models grounded in situational awareness can be an ally in bridging clients' understanding, so that systemic evaluation is not perceived to be new, but rather something that continues to build on past investments.

What is systemic evaluation?

Systemic evaluation is an approach that shows great promise and continues to attract attention as a model for thinking differently about how to conduct evaluations[1] (Cabrera, Colosi & Lobdell 2008). But while it is accepted and expected practice to begin any article by defining the topic under discussion, the challenge remains that systemic evaluation is not well defined.

Systemic evaluation has its basic tenets in systems theory. To that extent there is some agreement about what constitutes systemic evaluation in that it includes principles such as boundaries, interrelationships and multiple perspectives (Boyd et al. 2007). 'Boundary critique' involves reflecting on and discussing what issues should be included in, or excluded from, an initial definition of the situation (Boyd et al. 2007). 'Interrelationships' 'consider the web's internal and external relationships that constitute

an organization'. These need to be considered in relationship to each other (Boyd et al. 2007, p. 1310). 'Perspectives' involves including as many points of view as possible, including stakeholder evaluation, goal-based evaluation and organisational evaluation.

Many evaluators already consider the methods they employ to include elements of systems theory.[2] For example, obtaining multiple perspectives and considering boundaries are essential elements within empowerment evaluation (Fetterman 2001), utilisation-focused evaluation (Patton 2000), and in theory-driven evaluation (Donaldson 2002; Renger & Titcomb 2002; Rosas 2005).

Given that systems theory principles are already integrated to some extent in current evaluation approaches, evaluators struggle to define what differentiates systemic evaluation from other approaches. This sentiment is echoed by Cabrera, Colosi and Lobdell (2008) who described the process and end result of systemic evaluation as ambiguous and amorphous. Scriven (1991, p. 345) highlighted further the problem of defining what is meant by a 'system' and therefore concluded that 'the approach tends to be more of an orientation than an exact formula'.

More recently, Williams (2010) adds that there is a relationship between systems theory and evaluation principles and between systems and evaluation methods, but that significant gaps exist between them. Although systems theory abounds with methods, the extent to which these methods are useful within systemic evaluation is still being determined (Williams 2010).

Furthermore, in the authors' opinion, although some theorists may be clearer about what constitutes systemic evaluation, the broader evaluation community remains confused/uncertain. For example, at a recent international conference on systemic approaches in evaluation, funded by GIZ (2011), practitioners worldwide routinely confused systemic evaluation with systematic approaches, qualitative methods or evaluating a system. This caused some participants to surmise that system theory is a way of thinking that should be considered in any evaluation, but that it does not constitute a separate evaluation approach. In addition, at the 2011 Australasian Evaluation Society International Conference in Sydney similar sentiments were expressed by participants in a discussion about their understanding of systems theory (Renger, Williamson & Wood 2011).

Despite the absence of a clear definition, systemic evaluation is characterised by its underlying constructivist perspective where the 'truth' is not an objective, knowable thing but is individually and collectively *constructed* by the people who experience it. Systemic evaluation rejects the tenets of logical positivism and criticises its focus on objectives, description and judgement. A constructivist perspective assumes that there are many different ways of interpreting or understanding the intent or impact of a policy or program. It is expected that evaluators empower the stakeholders to shape and control the evaluation activities in their preferred ways. Therefore, stakeholders need to play a key role in determining the evaluation questions, variables and interpretive criteria. Furthermore, evaluators should steer evaluations in a way that empowers involved or affected disenfranchised people towards democratic participation (Guba & Lincoln 1989; Stufflebeam & Shinkfield 2007).

Cabrera, Colosi and Lobdell (2008) proposed that systemic evaluation is differentiated by its use of multiple methods, which have been referred to as 'methodological pluralism'. This is exemplified by Flood's (1999) synthesis of systemic evaluation, which used both formative and summative methods and Boyd et al. (2007) who used a combination of stakeholder, goal-based and organisational evaluation. To illustrate the range of definitions used to define systems evaluation further, some researchers at the 2011 American Evaluation Association Annual Conference defined systems evaluation as 'community-based outputs attributable to coalition work' (Yang 2011).

Despite not being able to agree on what constitutes a systemic evaluation, or being able to offer an operational definition, it does seem reasonable to posit the principles of systems theory that could be applied to meet different evaluation purposes. For example, the consideration of boundaries could be critical in deciding which data should be collected for purposes of general accountability—what Mark, Henry and Julnes (2000) refer to as 'oversight and compliance'. The principle of multiple perspectives would seem to be critically important when conducting a process evaluation. To determine how well a program has been implemented and how to improve delivery requires perspectives of those providing and receiving services—and perhaps even a neutral third observer.

Finally, significant literature is currently focused on the benefit of systemic evaluation in improving impact evaluation, or determining the merit and worth of a program (Mark, Henry & Julnes 2000). This article goes on to focus specifically on the arguments made by systemic evaluators regarding limitations of the current approaches to evaluating impact, and the merit of these arguments. It then offers an alternative point of view about how 'traditional' and systemic evaluation approaches to evaluating program impact and outcome might be bridged.

Systemic evaluation: criticisms of current approaches to evaluating impact

One of the key tools used in impact evaluation is the logic model. Logic models variously describe the linkages between the program theory, the inputs, activities and outcomes (Leeuw 2003). This type of logic model has come under heavy criticism by systemic evaluators, who point out its oversimplification of the context in which programs actually operate[3] (Hummelbrunner 2010a, 2010b; Lithman 2010, Rogers 2010; Williams 2010). It is also argued that the traditional 'box and table' logic model is too rigid to capture the contextual complexity underlying a program or, as Rogers (2010) refers to it, 'the truth'. Systemic evaluators contend that there is a direct relationship between the extent to which the truth is captured and interpretation of results. For example, proponents of the logical framework argue that its beauty lies in its simplicity. However, Hummelbrunner (2010a) argues that the logic model's depiction of a *single causal strand* can create confusion when attempting to interpret the results, because it is a misrepresentation of the truth or is at least so limiting as to distort the truth. That is, the findings often do not make sense (and therefore are unusable) because the program is not operating in isolation, as the logic model would suggest (Hummelbrunner 2010b).

From a systemic evaluation perspective the less complex the logic model, the less it is likely to capture reality. In an environment of limitless dependent and interdependent variables, it is important to identify as many of the interdependent variables as possible, regardless of whether they result directly from the proposed program or activity.

Other issues pertaining to logic models have been put forward by Rogers (2010) who noted limitations related to six dimensions, including focus, governance, consistency, 'necessariness', sufficiency, and change trajectory. For example, the focus of logic models tends to be very narrow and this does not lend itself to: the consideration of different objectives valued by different stakeholders; multiple competing imperatives; or to objectives at multiple levels of a system. Williams (2010) has also criticised logic models for their inability to capture the three key elements of a systems approach noted earlier: dynamic interrelationships, multiple perspectives, and the consideration of boundaries.

On the surface, criticisms levied against logic models by systemic evaluators seem well founded and justified. However, not all logic models are derived in the same way. There are at least two different ways in which logic models are derived and these approaches differ according to the extent to which systems theory is being incorporated.

On the one hand, there are what might be called intervention-driven logic models; that is, a commitment has already been made to an intervention (activity or strategy) and then other elements of the logic model (e.g. programmatic assumptions, outcomes) are aligned with the intervention. Such a process of completing a logic model in reverse or by retrofitting the logic (Brouselle 2010) serves primarily to substantiate and justify the intervention or, more sceptically, as leverage to ensure control over a program (Hummelbrunner 2010a). Many evaluators refer to logical frameworks, or logframes, as simple logic models (Rogers 2008; Rosenberg & Posner 1979). They are characterised by a single causal strand of logic and are most appropriate for use with a single organisation (though it should be noted that single organisations can also be highly complex) (Rogers 2008). Simple logic models are most often intervention driven (Rogers 2010). It is assumed that the aforementioned criticisms of logic models raised by systemic evaluators, as related to impact evaluation, are directed at simple logic models.

But even simple logic models can be utilitarian. For example, Wholey (2003) noted that simple logic models might be appropriate under some conditions, such as result-oriented management where goals are clearly defined and measurement of progress towards achievement of them is reliable.

Simple logic models are also useful for illustrating the basics of impact evaluation, including defining the program theory, linking strategies to the programmatic assumptions, and selecting outcomes aligned with what is being targeted for change. No matter how complex, emergent and changing the intervention, how many more interrelationships are demonstrated, how many strategies there are, or how complex the measurement, the basic goal in impact evaluation is to ensure that programmatic assumptions, strategies and measurement are linked and always remain the same. To that extent, simple logic models provide an important foundation for understanding the principles of impact evaluation.

There are many types of logic models (e.g. Chen 1990; Chen & Rossi 1983; Kellogg Foundation 2004; McLaughlin & Jordan 1999; Renger & Titcomb 2002; Rogers 2008; Rosas 2005) and not all logic models are oversimplified and rigid as with the simple logic model (Rogers 2008). Several begin with situational analysis trying to capture the context in which a program operates and incorporate several elements of systems theory (Cole 1999; Renger & Titcomb 2002; Rosas 2005). For example, in the first step of their approach to logic modeling Renger and Titcomb (2002) employ a form of root cause analysis to understand the complexity of the conditions and their interrelatedness. Perspectives of multiple stakeholders are deliberately sought in depicting

visually the antecedent conditions of a problem. Step 2 of their methods involves targeting strategies. This step clearly must consider boundaries and motivations of decision-makers selected to prioritise targeted antecedent conditions.

The challenge facing systemic evaluation

The challenge for systemic evaluators is the pursuit of truth (i.e. accurate depiction of the context in which a program operates) and this can result in such complexity that it is often too difficult to understand, explain and use. Increased complexity can lead to more measurement (but not necessarily more quality), increased evaluation cost, and negative impacts on the lives of employees (Goodyear & Dahler-Larsen 2010; Zurcher 2011).

As evaluators employ ever more sophisticated methods and designs to get closer to the 'truth', they must consider several important questions. Do the clients possess the necessary level of sophistication to understand the truth and all the complexity that entails?—because the failure of a client to understand the evaluation can impact significantly on the building of trust, evaluator credibility and evaluation capacity building. Then are the clients receptive to the truth or to findings that may challenge their existing perception of the truth? Are the results and recommendations usable? One of the common errors listed under the 'Feasibility Standards' is the failure to weigh accuracy versus practicality (Joint Committee on Standards for Educational Evaluation 1994). In general, the more accurate and precise we try to be and the more we try to convey the true complexity of a program, the less usable the information (Rogers 2008). This concern was noted by Rogers (2008) who cited researchers such as Pinnegar (2006) who argued that the objectives of complex programs might be too difficult to explain in tangible/readily accessible terms.

Another challenge relates to the credibility of our profession. This manifests itself in at least two ways. First, there are examples of where systemic approaches have failed (Zürcher 2011)—not failed in the sense that the results of the evaluation were negative, but that the evaluation plan promised to the client could not be delivered. It is the authors' opinion that presenting clients with something that is not well understood and that potentially results in a failure, undermines an evaluator's credibility. Even when the context for systemic evaluation seems warranted, caution should be exercised by evaluators to ensure that they do not overextend their skill set and that they do not make promises they cannot keep. It is important to note, however, that the authors' position is not shared by some European colleagues (Renger 2011).

The second credibility issue relates to the evaluation profession as a whole. Significant investments have been made by all levels of government and foundations in investing in logic models (Fielden et al. 2007; McLaughlin & Jordan 1999; Millar, Simeoneb & Carnevalec 2001). These investments cannot be ignored. As was noted at the recent GIZ (2011) International Conference, there is a risk that systemic evaluation could be viewed as nothing more than the current fad. Such a perception could have significant consequences for the credibility of the evaluation profession. For instance, when logic models were first introduced in the United States, the initial reaction by many was that they were just the 'flavour of the month' or a hoop that needed to be jumped through, and that something new would replace the approach later.[4] To introduce systemic evaluation as a new approach in the same way would only reinforce this perception. The pushback may make it difficult for evaluators to gain necessary buy-in, which could be argued could be even more critical when implementing an evaluation plan that is potentially emergent, as opposed to static.

Zürcher (2011) noted that systemic evaluation was tried in the early 1980s, but with limited success, causing the client to become disenchanted with the process and opting for a simpler, more understandable evaluation, namely to return to the logical framework. So, the challenge for evaluators wanting to use systemic evaluation is to ensure it is not perceived as a change in direction, but rather to demonstrate how it advances evaluation by building on past investments.

Seeing logic models as an ally not a foe

If systemic evaluation is going to be accepted or integrated successfully into an agency, then it requires organisational staff to understand it (Preskill & Boyle 2008). Preskill and Boyle go on to note that the strategies used to transfer learning depend on the level of the staff members' experience in evaluation. Some may possess very little experience and so may require a slower, building-block approach to learning.

Similarly, although evaluation capacity and interest in conducting evaluations is growing worldwide[5], many agency staff members are novice evaluators. This is evident from the presence of specialisations within evaluation, such as Fetterman's empowerment evaluation (2001) that flourished in response to a specific need. However, it seems reasonable to suggest that those with little experience need to learn the basics of evaluation before being introduced to more complex ideas and approaches associated with the discipline.

One way forward is for advocates of systemic evaluation to realise logic models are an ally rather than a foe. The approach thus far has been to criticise the logic model as a method for

evaluating impact. However, methodology is not a zero sum game and there is agreement that the evaluation methods employed should fit the situation (Rossi, Lipsey & Freeman 2004). Thus, in some impact evaluations the simple logic model is more than adequate to meet client needs, while in other situations systemic evaluation may be more appropriate. We do not need to denigrate one method to promote another.

In addition to being appropriate under some circumstances (Wholey 2003), an important additional benefit of simple logic models is that they can help teach how to build a program theory and how to use the evolving program theory to align activities and measurement in order to develop the program logic further. Such a foundation could facilitate the understanding and adoption of more sophisticated and abstract approaches such as systemic evaluation. Boyd et al. (2007) support this position, noting that 'people with minimal knowledge to experience evaluation may be better served by an introduction to the basic concepts before trying to use it' (pp. 1317).

For clients with exposure to logic models grounded in situational analysis, the relationship between their current investments to systemic evaluation could be made to seem even more natural and seamless. For example, Renger and Titcomb's three-step Antecedent Target Measurement (ATM) approach (2002) has similarities with systemic evaluation approaches, such as Adaptive Action (AA) (Holliday 2010; Tytel 2010). AA consists of an ongoing iterative evaluation cycle that answers the questions: What? So what? and Now what? The questions asked under 'What?' are consistent with those asked in Step 1 of the ATM approach. The major difference is that AA is an ongoing, emergent process, while ATM is defined before the program commences, is static, and is not sensitive to capturing unintended outcomes (Morell 2010; Renger & Hurley 2006). Step 2 of the ATM process, Targeting, must consider issues of boundaries carefully when selecting the stakeholders who will decide how resources are to be allocated for targeting and measuring antecedent conditions.

In summary, logic models, grounded in situational analysis—such as outcome mapping (sometimes used in aid and development evaluations), concept mapping (Rosas 2005), the Aetiologic Theory Structuring Guide (ATSG) (Cole 1999) and the ATM approach (Renger & Titcomb 2002)—all contain elements of systems theory. These can be used to help clients understand the concepts and to realise systemic evaluation is not a new way to undertake impact evaluation, but rather is an extension of past investments and understanding.

Summary

Systemic evaluation is in its infancy and evaluators are still struggling with defining its parameters and methods. Introducing a complex method to clients is a formidable challenge in itself, but the wisdom of doing so when the approach is not yet well defined by evaluators can do more harm than good. Introducing a concept more sophisticated than the maturity of the user can undermine efforts at capacity building, create distrust, and affect evaluator credibility. We must walk before we run.

Simple logic models and logic models built on situational analysis contain some of the elements of systems theory underpinning systemic evaluation. When systemic evaluation is deemed the most appropriate approach for a particular situation, evaluators can build on the principles of simple logic models and many of the elements of logic models grounded in situational awareness can be employed to demonstrate how the suggested approach is not new, but rather is an evolution of past investments leading to a more comprehensive evaluation.

Striving to capture the truth is a noble goal and should be encouraged. However, in some instances there is no need to capture additional complexity and under these conditions simple logic models are appropriate (Rogers 2008; Wholey 2003). Knowing the conditions under which systemic evaluation could prove beneficial is important. Knowing how to convey the complexity is even more important. Given the current stage of genesis, careful thought must be given to the consequences of introducing something such as systemic evaluation, to clients when it is not fully understood by our own profession. Building on existing knowledge and investments when possible would seem prudent.

Notes

1. This was noted at the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)International Conference on Systemic Approaches in Evaluation, Eschborn, Germany, 25–26 January.

2. See note 1 above.

3. See note 1 above.

4. This view was formed as a result of discussions with participants at the BHPr All Grantee Conference held in Washington, DC on 3 June 2005.

5. The OECD website has numerous articles supporting our contention that evaluation capacity building is growing worldwide. For further information refer to <http://www.oecd.org>.

References

Boyd, A, Geerling, T, Gregory, WJ, Kagan, C, Midgley, G, Murray, P & Walsh, MP 2007, 'Systemic evaluation: a participative, multi-method approach', *Journal of the Operational Research Society*, vol. 58, no. 10, pp. 1306–1320, doi: 10.1057/palgrave.jors.2602281.

Brouselle, A 2010, 'Assessing the validity of program theory: logic analysis', paper presented at the 9th Annual European Evaluation Society International Conference, Prague, 6–8 October.

Cabrera, D, Colosi, L & Lobdell, C 2008, 'Systems thinking', *Evaluation and Program Planning*, vol. 31, no. 3, pp. 299–310.

Chen, HT 1990, *Theory-driven evaluations*, Sage, Newbury Park, California.

Chen, H & Rossi, PH 1983, 'Evaluating with sense: the theory-driven approach', *Evaluation Review*, vol. 7, no. 3, pp. 283–302.

Cole, GE 1999, 'Advancing the development and application of theory-based evaluation in the practice of public health', *American Journal of Evaluation*, vol. 20, no. 3, pp. 453–470.

Donaldson, SI 2002, 'Theory-driven program evaluation in the new millennium', in SI Donaldson & M Scriven (eds), *Evaluating social programs and problems*, Lawrence Erlbaum Mahwah, New Jersey.

Fielden, SJ, Rusch, ML, Masinda, MT, Sands, J, Frankish, J & Evoy, B 2007, 'Key considerations for logic model development in research partnerships: a Canadian case study', *Evaluation and Program Planning* vol. 30, no. 1, pp. 115–124.

Fetterman, DM 2001, *Foundations of empowerment evaluation*, Sage, Thousand Oaks, California.

Flood, RL 1999, *Rethinking fifth discipline: learning within the unknowable*, Routledge, London.

Goodyear, L & Dahler-Larsen, P 2010, 'Evaluation systems in the public interest?', paper presented at the 9th Annual European Evaluation Society International Conference, Prague, 6–8 October.

Guba, EG, Lincoln, YS 1989, *Fourth generation evaluation*, Sage, Newbury Park, California.

Holliday, R 2010, 'A case study on adaptive action in education', paper presented at the 24th Annual Conference of the American Evaluation Association, San Antonio, Texas, 10–13 November.

Hummelbrunner, R 2010a, 'Beyond logframe: critique, variations and alternatives', in N Fujita (ed.), *Beyond logframe: using systems concepts in evaluation*, Foundation for Advanced Studies on International Development, Tokyo.

Hummelbrunner, R 2010b, 'Changing institutional approaches to using logic models', paper presented at the 24th Annual Conference of the American Evaluation Association, San Antonio, Texas, 10–13 November.

Joint Committee on Standards for Educational Evaluation 1994, *The program evaluation standards: how to assess evaluations of educational programs*, 2nd edn, Sage, Thousand Oaks, California.

Kellogg Foundation 2004, *Logic model development guide*, WK Kellogg Foundation, Battle Creek, Michigan, viewed 1 December 2011, <http://www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx>.

Leeuw, FL 2003, 'Reconstructing program theories: methods available and problems to be solved', *American Journal of Evaluation*, vol. 24, no. 1, pp. 5–20.

Lithman, E 2010, 'Evaluability assessment 2.0—or: on a reflexive approach to evaluation system in the knowledge society', paper presented at the 9th Annual European Evaluation Society International Conference, Prague, 6–8 October.

Mark, MM, Henry, GT & Julnes, G 2000, *Evaluation: an integrated framework for understanding, guiding, and improving policies and programs*, Jossey-Bass, San Francisco.

McLaughlin, JA & Jordan, GB 1999, 'Logic models: a tool for telling your programs performance story', *Evaluation and Program Planning*, vol. 11, no. 2, pp. 54–61.

Millar, A, Simeoneb, RS & Carnevalec, JT 2001, 'Logic models: a systems tool for performance management', *Evaluation and Program Planning*, vol. 24, no. 4, pp. 73–81.

Morell, J 2010, *Unintended consequences: Jonny Morell on logic models and unintended consequences*, AEA365, a Tip a Day for Evaluators, viewed 1 December 2011, <http://aea365.org/blog/?tag=unintended-consequences>.

Patton, MQ 2000, 'Utilization-focused evaluation', in DL Stufflebeam, GF Madaus & T Kellaghan (eds), *Evaluation models: viewpoints on educational and human services evaluation*, 2nd edn, Kluwer, Norwell, Massachussets.

Pinnegar, S 2006, *Are complex programs the best response to complex policy issues?* City Futures Research Centre issues paper, University of New South Wales, Kensington, NSW.

Preskill, H & Boyle, S 2008, 'A multidisciplinary model of evaluation capacity building', *American Journal of Evaluation*, vol. 29, no. 4, pp. 443–459.

Renger, R 2011, 'Systemic evaluation: much too soon?', paper presented at the GIZ International Conference on Systemic Approaches in Evaluation, Eschborn, Germany, 25–26 January.

Renger, R & Hurley, C 2006, 'From theory to practice: lessons learned in the application of the ATM approach to developing logic models', *Evaluation and Program Planning*, vol. 29, no. 2, pp. 106–119.

Renger, R & Titcomb, A 2002, 'A three-step approach to teaching logic models', *American Journal of Evaluation*, vol. 23, no. 4, pp. 493–503.

Renger, R, Williamson, S & Wood, S 2011, 'Systemic evaluation: using logic models to bridge capacity building gaps', roundtable presentation at the Australasian Evaluation Society 2011 International Conference, Sydney, 31 August – 2 September.

Rogers, PJ 2008, 'Using programme theory to evaluate complicated and complex aspects of interventions', *Evaluation*, vol. 14, no. 1, pp. 29–48, doi: 10.1177/1356389007084674.

Rogers, P 2010, 'Representing simple, complicated, and complex aspects in logic models for evaluation quality', paper presented at the 24th Annual Conference of the American Evaluation Association, San Antonio, Texas, 10–13 November.

Rosas, SR 2005, 'Concept mapping as technique to program theory development: an illustration using family support program', *American Journal of Evaluation*, vol. 26, no. 3, pp. 389–401.

Rosenberg, L & Posner, L 1979, *The logical framework: a manager's guide to a scientific approach to design and evaluation*, Practical Concepts Incorporated, Washington, DC.

Rossi, PH, Lipsey, MW & Freeman, HE 2004, *Evaluation: a systematic approach*, 7th edn, Sage, Thousand Oaks, California.

Scriven, M 1991, *Evaluation thesaurus*, Sage, Newbury Park, California.

Stufflebeam, DL, Shinkfield, AJ 2007, *Evaluation theory, models, and applications*, Jossey-Bass, San Francisco.

Tytel, M 2010, 'Translating lessons learned into lessons applied', paper presented at the 24th Annual Conference of the American Evaluation Association, San Antonio, Texas, 10–13 November.

Wholey, JS 2003, 'Using policy evaluation to support decision making and improve program performance', *Japanese Journal of Evaluation Studies*, vol. 3, no. 2, pp. 6–21.

Williams, B 2010, 'Fitting the key to the lock: matching systems methods to evaluation questions', paper presented at the 24th Annual Conference of the American Evaluation Association, San Antonio, Texas, 10–13 November.

Yang, E 2011, 'National efforts to build coalition and evaluator capacity to track community and systems change', American Evaluation Association Annual Conference, Anaheim, California, 2–5 November.

Zürcher, D 2011, 'Why the (development) world is not a system', paper presented at the GIZ International Conference on Systemic Approaches in Evaluation, Eschborn, Germany, 25–26 January.