REFEREED ARTICLE
Evaluation Journal of Australasia, Vol. 12, No. 2, 2012, pp. 4–14

Using root cause analysis for evaluating program improvement

A common evaluation purpose is to determine whether a policy or program was implemented as intended: referred to as formative evaluation, process evaluation, or evaluating program improvement. A well-designed formative evaluation is important in: detecting program drift; providing timely feedback to program staff to make cost-saving mid-course corrections; reassuring the sponsor that quality assurance measures are implemented to protect investments; and interpreting impact/outcome evaluation. A formative evaluation should not just gather data on deviations from an anticipated course of action, but provide recommendations for improvement. Current methods for program improvement vary in their ability to solicit targeted recommendations. Root cause analysis (RCA) is a well-established, robust methodology used in a variety of disciplines. RCA has been primarily used by evaluators operating from a theory-driven orientation to evaluate the merit and worth of a program or policy. Surprisingly, a review of the literature suggests that RCA's utility as a program improvement tool has remained largely unrecognised in evaluation. This article illustrates the application of RCA in evaluating program improvement. The conditions under which RCA might be preferred over other formative evaluation methods are discussed.

Rebekah Coşkun

Abigail Akande

Ralph Renger

Rebekah Coşkun (top left) is a DrPH student at the Mel and Enid Zuckerman College of Public Health at the University of Arizona, Tucson, Arizona.
Email <bekahc@email.arizona.edu>

Abigail Akande (top right) is a PhD Candidate in Rehabilitation in the College of Education at the University of Arizona, Tucson, Arizona.
Email: <aakande@email.arizona.edu>

Ralph Renger (bottom) 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>

Introduction

There are many purposes of program evaluation, including evaluating oversight and compliance, merit and worth, and program improvement (Mark, Henry & Julnes 2000). The focus of this article is on the methods used for evaluating program improvement. Formative evaluation, process evaluation, quality assurance, and program improvement are all synonymous terms[1] in the evaluation literature for determining the extent to which a program or policy was delivered with fidelity.

Assessing program fidelity is important for two main reasons. First, results of an outcome evaluation can only be clearly interpreted if it is first established that the program was delivered with fidelity. If the program was not delivered as it was originally intended, then it is impossible to determine whether the failure to observe changes in outcomes was due to the design of the intervention or simply because the program was not executed correctly (Chen 1990; Mills & Ragan 2000). Second, assessing fidelity helps to detect program drift (Bond 1991). Detecting drift early on can result in significant cost savings and/or the identification of alternative implementation strategies (Reijers & Mansar 2005).

Program improvements can be: minor changes, such as the editing of an intake form or other paperwork; moderate changes, such as training and staff development

needs; or broad changes, such as system modifications or changes in policy. One of the greatest benefits of a formative evaluation is the subsequent enhancement of program sustainability, or its ability to be replicated and/or modelled in other settings (Speelman et al. 2007).

Given the importance of formative evaluation in our field, it is advantageous for the evaluator to have a variety of methods at his or her disposal to find the best possible fit of method for the context. Formative evaluation methods include, but are not limited to, interviews, focus groups, observations, and survey questionnaires. One method that has received surprisingly little attention in the evaluation literature as a tool for evaluating program improvement is root cause analysis (RCA). It is unclear why RCA has been overlooked as a tool for evaluating program improvement. The purpose of this article is threefold: first, to review some of the more frequently used formative evaluation methods, highlighting their strengths and weaknesses; second, to show how the RCA methodology can be adapted for the purpose of evaluating program improvement; and third, to discuss how the RCA methodology compares to other formative evaluation methods.

A review of four formative evaluation methods

Interviews

Interviews are used to obtain detailed responses from individuals (Ash & Guappone 2007; Creswell 2009; Glesne 2011). Interview structures vary and are found on a continuum including open-ended, semi-structured and structured interviews (Ash & Guappone 2007; May 1997). Group and individual interviews may be conducted with clients and/or program or organisational staff (Patton 2003).

Under time constraints group interviews may be chosen (Gaskell 2000), but individual interviews offer the undivided attention of the evaluator as well as account for status and power differentials that make speaking in a group uncomfortable (Lewis 2003). Also, individual interviews offer flexibility in scheduling time and location (Lewis 2003). Interviews may last an hour or longer depending on the setting and structure of the interview (Ash & Guappone 2007; Gaskell 2000).

Individual interviews are recommended when the nature of the interview involves sensitive issues that someone might not feel comfortable discussing in a group setting and may hold back in their response (Gaskell 2000; Lewis 2003). Interviews are often voice recorded (McDavid & Hawthorn 2006; Posavac 2011) and must be transcribed and coded at a later time (Ash & Guappone 2007; McDavid & Hawthorn 2006). Evaluators must prepare adequately for an interview and explain carefully to the interviewee(s) what will occur; building rapport with the interviewee is crucial and can often be done in the beginning by asking simple questions to allow the individual to feel comfortable (Gaskell 2000; Posavac 2011).

The evaluator may work with the interviewee to facilitate and prompt discussion as well as to ask follow-up questions (Ash & Guappone 2007; Posavac 2011). Yes or no questions must be avoided, and questions must be designed to prompt the interviewee to reveal useful information (Posavac 2011). An evaluator must be trained to detect and use verbal and non-verbal signals in interviews (Posavac 2011), and as such, the evaluator must possess aptitude in observation (Patton 1987). Additionally, the evaluator bears the responsibility of conveying the tone, emotions and perceptions in textual form revealed during the interview (Gaskell 2000). He or she must also possess advanced note-taking skills and the ability to appropriately pace the interview (McDavid & Hawthorn 2006). Evaluators must use the common language of the interviewee, be cautious not to infer too much from what is said (Gaskell 2000), and not bias the responses of the interviewee (McDavid & Hawthorn 2006). An evaluator employing interviewing techniques must be extremely sensitive to their surroundings and possess '… concentration, interpersonal understanding, insight, mental acuity, and discipline' (Patton 1987, p. 108).

Interview data has limitations, as the interviewee is usually recalling past information (Gaskell 2000). Scriven (1991) argues if interviews are designed and implemented well, then they can yield valuable information; but he cautions that many times interviews are not conducted correctly. Open-ended interview questions provide an evaluator with detailed data based on an individual's '… experiences, perceptions, opinions, feelings, and knowledge' (Patton 2003, p. 2) and often relies on direct quotations to reflect responses. Although, the breadth of the data accumulated from successfully completed, open-ended interviews cannot be denied, this method is time intensive. The evaluator must also sift through a large volume of data, and a balance must be met between the depth and breadth of interviews (Patton 1987).

Focus groups

Focus groups are a type of semi-structured group interview. They are led by an evaluator or moderator who facilitates discussion around a specific topical area and thematic questions. One benefit of this method is the opportunity for individuals to interact with one another in discussing the various themes, often resulting in richer responses (Ash & Guappone 2007; Gaskell 2000; Patton 2003; Posavac 2011). Focus groups are more informal than other methods (Posavac 2011). Open-ended questions are usually asked by the moderator who must balance staying in the background with keeping participants from straying off the issue (Posavac 2011). Also, they must remain as neutral and objective as possible (Lewis 2003). Focus groups usually include seven to 10 individuals (Posavac 2011) and are held for an hour or slightly longer.

Compared to individual interviews, focus groups are more efficient, permitting more data to be collected within a short period of time (Ash & Guappone

2007; Gaskell 2000). Participants of focus groups should be more homogeneous, and commonalities among individuals should exist, such as educational attainment, age and socioeconomic status, in order to promote an open and comfortable discussion (Lewis 2003; Patton 1987; Posavac 2011). Challenges arise in the information gained from focus groups when the participants do not feel comfortable and hold back from sharing their perspectives and opinions (Posavac 2011). Group dynamics need to be taken into consideration, as this setting may not be appropriate for the discussion of highly sensitive issues (Gaskell 2000). Focus groups do not allow for the same type of personalised attention that individual interviews allow and it may be difficult to recruit participants who are very busy; as this method requires the scheduling of a common time and location that is convenient for everyone (Gaskell 2000; Lewis 2003).

Observation

When using observation, the evaluator is immersed in the organisation, department or program for short or long periods (Ash & Guappone 2007). Non-participant observation, which observes behaviors without the need for interacting with participants, is best used in more public spaces, and participant observation in more private settings (Posavac 2011). Participant observation allows for a comprehensive first-hand look at the different processes occurring (such as staff behaviors and other day-to-day activities) as opposed to being told what happened (Patton 1987). At times, observations can be enhanced with the use of checklists (Scriven 1991). This method results in rich descriptive data and allows the evaluator to collect information that may be otherwise missed (Ash & Guappone 2007; Patton 2003), as the evaluator is able to put the observations into context and see things that individuals may omit during interviews or open-ended surveys (Patton 1987).

While observation methods can yield high-quality and detailed descriptive data, the data can be very time consuming and resource intensive to analyse. In addition, the evaluator must be careful to realise his or her own biases in the interpretation of the data (Ash & Guappone 2007; Patton 1987). Scriven (1991) warns that a grey area exists between interpretation and observation, and as such, those who employ observation should be trained carefully not to blur the lines between describing and inferring. Further, this method is not for novice evaluators (Scriven 1991). The presence of an observer can alter the ways in which people act (Patton 1987; Posavac 2011), and gaining access to conduct observations may prove difficult (May 1997). Patton (1987) explains that descriptive data obtained from fieldwork observations '… must be factual, accurate, and thorough without being cluttered by irrelevant minutiae and trivia' (p. 12). Because of the challenging nature of observational fieldwork, the evaluator must be skillfully trained to be meticulous and precise as the evaluator/observer '… is the instrument' (Patton 1987, p. 12).

Survey questionnaires

Survey questionnaires allow for a more systematic way of collecting information (McDavid & Hawthorn 2006). They can include closed (Ash & Guappone 2007) or open-ended questions and are used to elicit a wide variety of responses (McDavid & Hawthorn 2006). Questionnaire formats are flexible (May 1997) and can be completed online, in person, by mail or by phone (Centers for Disease Control and Prevention 2008). Questionnaires can be completed by a large number of respondents and are often more cost effective than other methods (Patton 1987). However, some authors have noted that questionnaires can have significant costs associated with them, which are discussed below (McDavid & Hawthorn 2006).

Questionnaires can be completed after implementation or they can be administered continually. They are often used to gauge client satisfaction and perceptions (Carey & Seibert 1993) and can serve as a checklist of what and how something was completed (Saunders, Evans & Joshi 2005). Potential drawbacks of questionnaires include issues with length, low response rates, and trouble in having adequate response choices (Scriven 1991). Good surveys are resource intensive as time and money are spent creating the instrument, and special attention must be placed on how questions are worded (May 1997; McDavid & Hawthorn 2006). A good questionnaire should be piloted (May 1997; McDavid & Hawthorn 2006; Scriven 1991), which consumes precious time. Questionnaire responses need to be coded as well (May 1997). Also, surveys usually ask respondents to recall something from the past (McDavid & Hawthorn 2006), which can limit the quality of information received.

How participants are selected and the number of individuals needed to complete any of the formative evaluation methods reviewed varies depending on time, resources and context (Gaskell 2000; Patton 1987). Individuals selected can include a variety of key stakeholders, such as program staff, personnel and clients (Saunders, Evans & Joshi 2005).

All formative methods can be used to elicit a variety of responses (Gaskell 2000) about individual or group perceptions, experiences and understandings of a program.

Root cause analysis in evaluation

Root cause analysis (RCA) is a well-established, robust methodology used in a variety of disciplines (Adams et al. 1999; Aladwani 2001; De Grave, Boshuizen & Schmidt 1996; Hollnagel 1999; Liang et al. 2006; Rasmussen 1997; Senders 2004). Within the field of evaluation, RCA has been championed as a rigorous and user-friendly process (Morell 2000). Interestingly, a review of the evaluation literature suggests that the power of RCA has been recognised primarily in evaluating program merit and worth and not necessarily for the purposes of evaluating program improvement. For example, RCA principles

are frequently used in theory-driven evaluation (Donaldson 2003; Renger & Titcomb 2002). More specifically, adaptations of RCA can be found in theory-driven evaluation methods such as concept mapping (Caracelli & Riggin 1994; Rosas 2005; Shern, Trochim & LaComb 1995; Trochim 1989; Yampolskaya et al. 2004); the Aetiologic Theory Structuring Guide (Cole 1999); and the antecedent, targeting and measurement approach (Renger & Titcomb 2002). These approaches use variations of RCA to identify and make programmatic assumptions necessary for evaluating merit and worth explicit.

For the most part, the utility of RCA for evaluating program improvement in the evaluation literature has been overlooked. One exception to this is the work of Shern, Trochim and LaComb (1995) who extended the application of a specific type of RCA, concept mapping, to assess the fidelity of a psychiatric rehabilitation program to the original program after which it is was modelled. The focus of the study was to demonstrate the use of concept mapping in the assessment of model transfer fidelity. In other words, the authors used RCA to identify inconsistencies between the actual and model program, which could then be used to offer recommendations for program improvement. Shern, Trochim and LaComb (1995) did not use RCA to evaluate day-to-day program implementation.

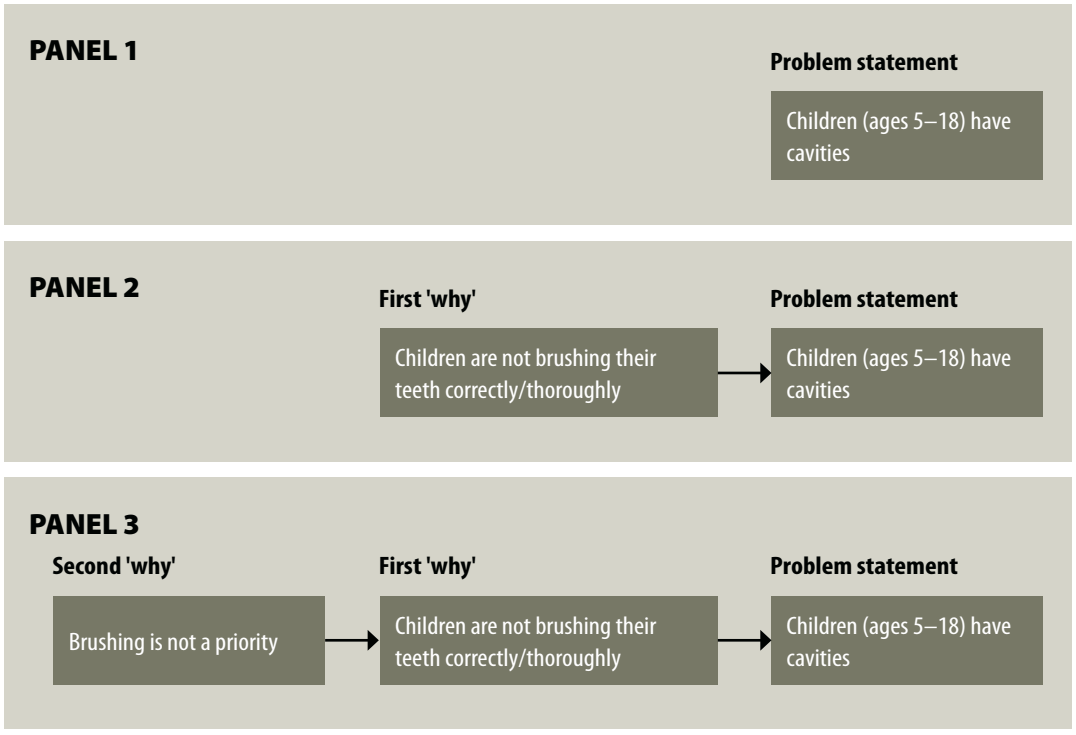
RCA methodology

RCA is a structured approach to recognising the real cause of a problem and identifying actions resulting in a permanent solution to that problem (Duffy, Moran & Riley 2012). The principles of RCA underpin many quality control methods such as failure analysis (Liang et al. 2006), problem investigations (De Grave, Boshuizen & Schmidt 1996), continuous improvement process (Adams et al. 1999), accident analysis (Hollnagel 1999), change management (Aladwani 2001), failure mode and effects analysis (Senders 2004), risk management (Rasmussen 1997), and so forth. Although slightly altered and referred to somewhat differently across various fields, the essence of RCA remains the same: program improvement.

Before RCA can begin, a problem must be identified and defined. Since differing definitions of a problem can lead to the identification of different root causes, it is important that the problem be operationally defined and stated in as concrete of terms as possible. In the example from Figure 1, the population is defined as children specifically between the ages of 5 and 18. However, it would be important to clarify whether this includes both males and females, which ethnicities, and to ensure there is a common understanding of the term cavity.

RCA begins by placing the problem statement on the right-hand side of a whiteboard as shown in Figure 1, panel 1. Therefore, it is critical that the RCA

FIGURE 1: A SAMPLE RCA PROBLEM STATEMENT AND FIRST AND SECOND 'WHY' QUESTION RESPONSES



process begin by establishing an operational definition of the problem.

Subject matter experts (SMEs) are then recruited to assist with the RCA. SMEs are individuals who have substantive content expertise. The number of SMEs needed for RCA depends on the purpose and size of the program to be evaluated, but Renger and Hurley (2006) have noted that the point at which little new information is gained (i.e. redundancy is approached) occurs at around eight to 10 SMEs. SMEs can be program staff, recognised content experts, and although often neglected, representation of the target population that the program is intended to serve.

RCA engages SMEs in an organised way to help understand and make explicit the reasons why a problem exists. RCA with SMEs can be conducted individually or in groups. The advantages and disadvantages of conducting RCA as a group or individually are the same as those noted for conducting focus groups or individual or group interviews (Gaskell 2000).

With the problem statement clarified, the SME is then asked the first 'why' question. For example, 'Why do children have cavities?' The SME's response is placed to the immediate left of the problem statement, as shown in Figure 1, panel 2. Note that the RCA process works from right to left. However, the logic of the evolving thread is checked from left to right using 'if-then' statements. For example, 'IF children are not brushing their teeth correctly/thoroughly, THEN children have cavities'. The 'if-then' technique for checking logic is used by several researchers (Chien, Wang & Chen 2007; Doggett 2005; McLaughlin & Jordan 1999; Renger & Titcomb 2002; Venkatasubramanian et al. 2003).

SMEs are then asked the second 'why' question. This response is placed to the left of the initial antecedent condition, as shown in Figure 1, panel 3. Once again, the evolving thread of logic is verified using 'if-then' statements by moving from left to right. For example, 'IF brushing is not a priority, THEN children are not brushing their teeth correctly'.

The process of asking SMEs 'why' questions is repeated until a root cause for the problem is identified. Ohno (1988) noted that it is seldom necessary to ask 'why' more than five times along a single thread to uncover a root cause. Deciding where to stop a thread is more of an art and based on facilitator experience (Renger & Hurley 2006). However, continuing to probe with 'why' questions is unnecessary if an identified root cause is unlikely to change, such as poverty or a genetic predisposition. Another factor in deciding how far to query is whether the root cause being identified falls under the agency's mission. For our example, it might be out of the realm of practice, interest or expertise for a group of dentists to provide nutritional counselling and interventions to a group of children with cavities in a rural community of a developing country, although nutritional factors are a root cause of cavities.

Once a thread of logic is complete and the root cause is identified, SMEs are asked to revisit the problem statement and identify another first-level 'why'. The process of asking 'why' is repeated again until a root cause is identified. For many problems, the process results in the identification of several root causes. An example of a partially completed RCA diagram is shown in Figure 2.

Adapting RCA for the purpose of evaluating program improvement

The example illustrated in Figures 1 and 2 is typical of how RCA is used in theory-driven evaluation to identify programmatic assumptions necessary for evaluating merit and worth. However, when evaluating program improvement the focus is on establishing whether a program was implemented with fidelity (Holzer et al. 2006; Mark, Henry & Julnes 2000).

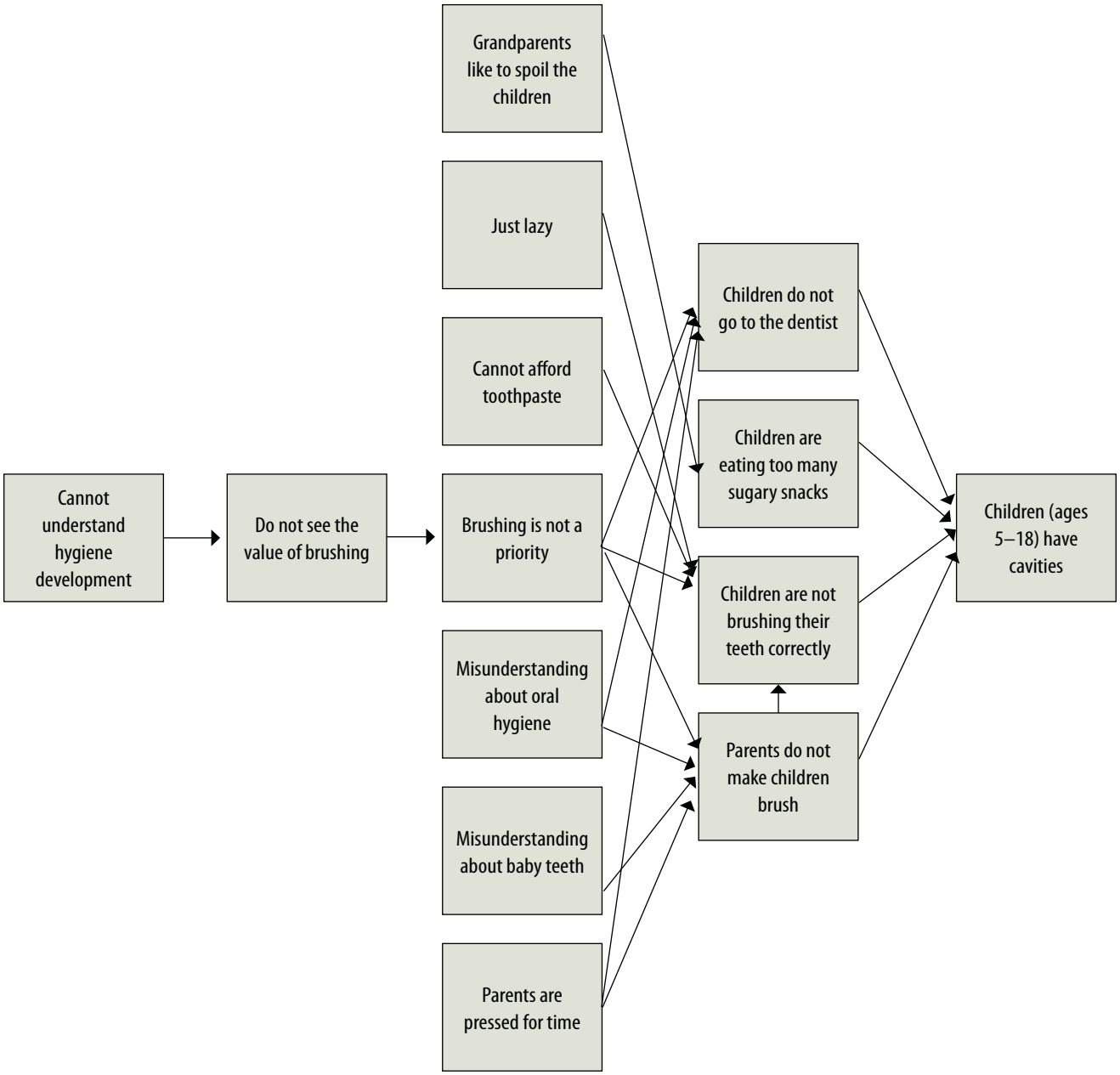
When RCA is used for program improvement, the problem statement will always centre on the success of implementation. For example, a problem statement such as, 'The program was not implemented as intended' is fairly generic and can apply to almost every program. It is still important to ensure that there is a common understanding of the problem statement, but in the authors' experience, problem statements for the purpose of program improvement are easier for SMEs to understand and achieve consensus as compared to problem statements used to derive programmatic assumptions when evaluating merit and worth.

In the example in Figure 3 (page 10), the problem statement is: 'The university's academic integrity workshop was not implemented as intended'. In this case the SMEs would be those with direct knowledge about the Academic Integrity Workshop and its implementation, such as: workshop instructors; the instructors' supervisors; and individuals who serve on the university's Academic Integrity Council, which might include faculty members, deans, and/or other administrators; or any other individuals who have direct knowledge and input about how the workshop is constructed or run.

As before, the RCA process begins by asking the SME the first 'why' question. For example, 'Why was the Academic Integrity Workshop not implemented as intended?' The logic of the evolving thread is checked from left to right using 'if-then' statements. The process of asking 'why' questions continues until a root cause is identified. The problem statement is then revisited and another first 'why' question is asked. In short, the RCA methodology does not change.

In the example, the RCA process identified several root causes. With the root causes made explicit, it is now possible for the department to use this information to make decisions about how to improve program delivery. One of the first decisions is to determine which factors are within the department's direct and immediate control to change (Huntington & Renger 2003). In the example, the root cause 'people involved with the workshop frequently change' is arguably beyond the sphere of the department's influence. Turnover may be a common occurrence, especially when the workshop instructors

FIGURE 2: AN EXAMPLE OF A PARTIALLY COMPLETED RCA DIAGRAM ONCE A ROOT CAUSE IS IDENTIFIED FOR THE PURPOSE OF MERIT AND WORTH



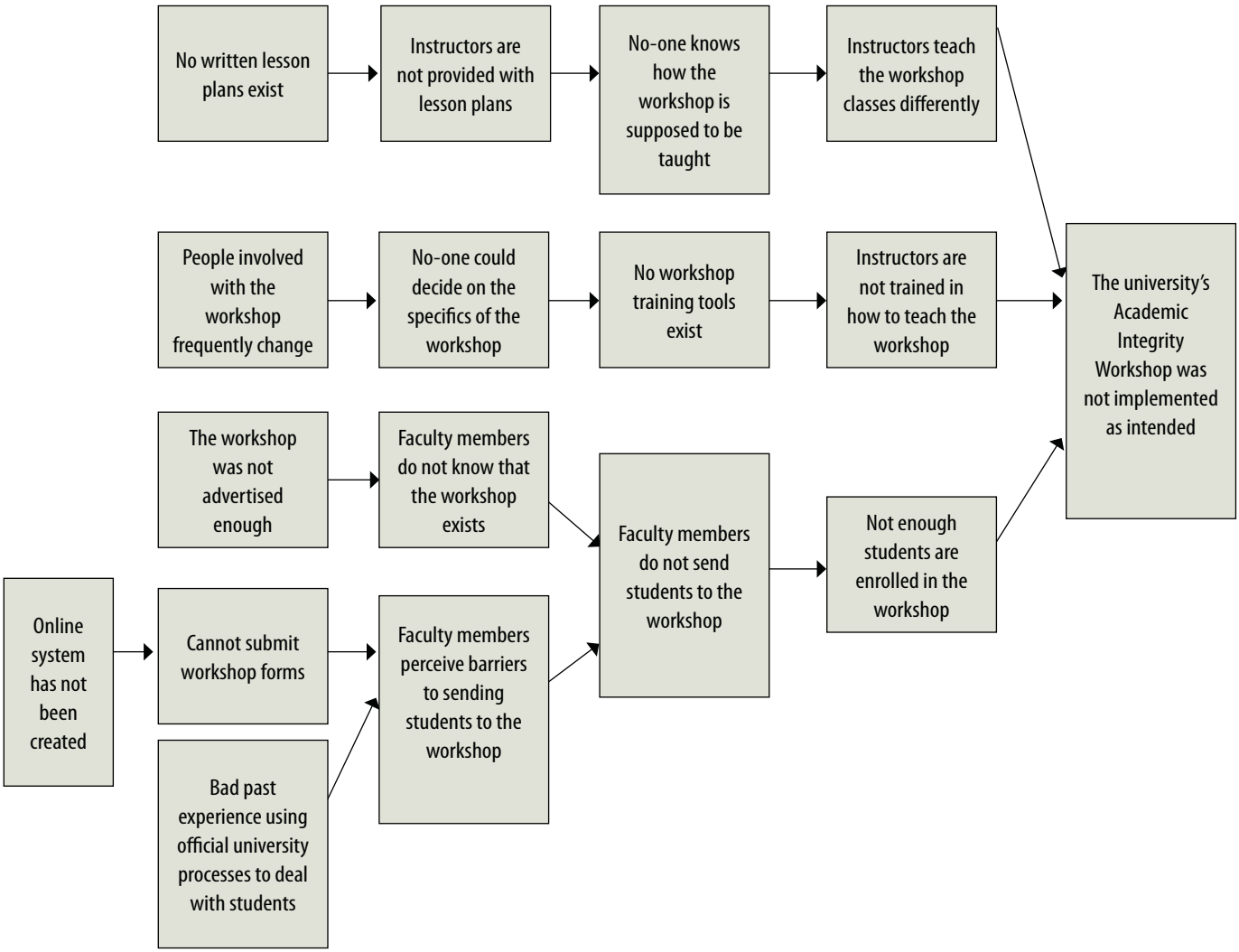
are graduate students. On the other hand, it may be that the department could potentially improve the availability of written lesson plans, improve on advertising the workshop, or resources permitting, create an online registration system.

Strengths of the RCA methodology for program improvement

The importance of understanding the root causes before developing strategies to resolve program implementation problems cannot be overstated. Failing to first identify a problem's root causes can lead to activity traps, where the tasks are carried out correctly even though the tasks themselves are unrelated to solving the issue(s) at hand (Renger & Titcomb 2002). The consequences of activity traps on evaluation are immediate. The likelihood of a misdirected corrective action having its intended impact will be small, regardless of how well it is implemented (Anjard 1998; Cale & Harris 2009; Renger & Titcomb 2002; Spath 2003).

One of the major advantages of using RCA over other formative evaluation methods is that it provides specific information from which corrective actions

FIGURE 3: AN EXAMPLE OF A COMPLETED RCA DIAGRAM IDENTIFYING THE ROOT CAUSES FOR THE PURPOSE OF PROGRAM IMPROVEMENT



can be easily derived. RCA visually presents the causal linkages, or antecedent conditions, that lead up to the problem and removes the ambiguity associated with corrective actions. Other methods, such as interviewing or focus groups, may provide useful information but do not necessarily identify specific items to be corrected or how best to correct them. The same is true for observations and survey questionnaires.

A second advantage of RCA is that issues are placed in context. For example, when using an open-ended exit survey questionnaire that asks SMEs to identify program weaknesses, the result is typically a list of issues such as:

- faculty members do not know that the workshop exists
- faculty members do not send students to the workshop
- faculty members perceive barriers to sending students to the workshop

- not enough students are enrolled in the workshop
- cannot submit workshop forms online.

Evaluation data presented this way makes it difficult to identify where best to begin with respect to making programmatic decisions to improve delivery. In other words, when each listed weakness is viewed in isolation, it is difficult to understand what might be the most efficient and effective corrective actions. A major benefit of RCA is its depiction of the relationships between antecedent conditions and root causes. It is the root cause for which corrective actions are sought. A corrective action targeted at a root cause should be more efficient because changing it will lead to a change in several other 'downstream' issues. For example, in Figure 3 one identified root cause is 'no written lesson plans exist'. It is more efficient for the department to focus on developing and recording workshop lesson plans because it should lead to changes in other downstream issues such as 'no-one knows how the workshop is supposed to be taught' and/or 'instructors teach the workshop classes differently'. If the department focused on the downstream issues as opposed to the root cause, then it is less likely that it will achieve the desired improvements in program delivery.

The quality of program improvement recommendations is directly related to the specificity of the evaluation data. Often, specificity is derived from source documentation (Renger 2010), which details the program implementation steps. From this, documentation checklists and tools can be developed a priori to evaluate the extent to which the program was implemented with fidelity. RCA may be especially useful when (a) there is no source documentation from which to develop a priori evaluation tools and/or (b) no program improvement tools have been developed and the implementation of the program has commenced or is finished. RCA by its nature does not depend on source documentation and is typically used once a bottleneck or problem has occurred, thus negating the need for a priori data collection tools.

Comparing RCA to other formative evaluation methods

The RCA approach for program improvement is preferred over interviews, focus groups, observation and open-ended surveys or questionnaires focusing on client satisfaction because it: (1) provides structure to the process evaluation while still permitting the SME to describe the situation using their own terms; (2) assists in the prioritisation process needed for strategy development; and (3) leads to new ideas, as opposed to simply fixing an existing system, which has been its main purpose in other industries. Table 1 compares the strengths and weaknesses of the four reviewed formative evaluation methods and RCA.

While other formative methods may yield larger volumes of data, RCA only collects data specific to program implementation. Because of its focus on the true root cause(s), targeted corrective actions can be identified. Other methods may never discover the actual root causes of the problem or show the relationships between them. Additionally, RCA requires only eight to 10 interviews/maps and summaries, and can be finished quickly. The whole process requires less time and resources than other methods. Patterns in RCA interviews can often be identified after just three to four interviews have been completed. The RCA data is presented visually in a map, which aids staff and personnel in understanding the issues and leads to them quickly identify corrective actions.

A potential issue to consider when employing RCA is that it highlights problems in implementation. Often implementation involves human beings, so the potential of RCA to be equated to assigning blame is high (Renger, Davis, & Granillo, in press). In these cases, it is of the utmost importance to make certain that the staff involved in the process do not feel blamed or as if their jobs are in jeopardy. It is essential that the root causes delineated as a result of RCA do not focus on 'who' issues, but rather 'what' and 'why' issues (Williams 2001). This less-threatening approach creates a more positive environment that encourages problems to be detected earlier, thereby increasing the likelihood that a program will meet its goal and utilise resources more efficiently. With the use of other formative evaluation methods, the likelihood of the 'blame issue' may not be as high, especially with questionnaires due to the anonymity of respondents. Because of the smaller sample sizes utilised in RCA, it may be more likely to pinpoint a person to blame as well as the individual who blames them. With that said, individual and group interviews as well as focus groups may also have smaller sample sizes and have to deal with issues of blame. With all formative evaluation methods, the evaluator must take precaution to ensure anonymity as much as possible and be aware of the potential for assigning blame.

Conclusion

RCA seems to have gone unrecognised or has been undocumented in the evaluation literature as a robust methodology for evaluating program improvement. As we have demonstrated, RCA's typical use as a method for generating program theory in evaluating merit and worth can be expanded to identify corrective actions for program improvement. RCA is a useful, easy-to-understand and targeted approach to collect program improvement data. The principal benefit of RCA over other formative evaluation methods is in the depiction of the relationship between issues and the identification of root causes. It is this fact that allows for targeted recommendations and corrective actions that are more likely to lead to improvements in program delivery.

TABLE 1: A COMPARISON OF THE STRENGTHS AND WEAKNESSES OF THE FOUR REVIEWED FORMATIVE EVALUATION METHODS AND RCA

Method	Strengths	Weaknesses
Individual interviewing	Detailed responses Opportunity for follow-up questions Rapport building Allow for sensitive topics Flexible scheduling Personalised attention	Respondent recall bias More formal
Focus group	Participant interaction Detailed responses Resource conservative Rapport building Less formal	Respondent recall bias Group dynamics: participant inhibition, participant dominance, participant influence/sway More prone to going off topic
Observation	'First-hand' look Descriptive data Gathering of pertinent data that might have otherwise been missed or not discussed	Resource intensive Observer bias Observer can be seen as intrusive Requires an experienced researcher Ethical dilemma: observation of confidential information 'Grey area' of interpretation
Survey questionnaire	Flexible formats Potentially large sample size Resource conservative Less formal Standardised data collection	Respondent recall bias Low response rates Limited response choices and detail Typically too long Resource intensive
Root cause analysis	Provides information specific to the problem, leading to strategy development Issues placed in context Highly structured collection of data relevant to the problem Detailed responses Opportunity for follow-up questions Rapport building Allow for sensitive topics Flexible scheduling Personalised attention Resource conservative Less formal Visually represents the causal linkages	May present opportunities for assigning blame Requires an experienced facilitator May not be as comprehensive Best used after fidelity issues present

Note

1 These terms are used interchangeably throughout this article.

References

Adams, M, Componation, P, Czarnecki, H & Schroer, BJ 1999, 'Simulation as a tool for continuous process improvement', in PA Farrington, HB Nembhard, DT Sturrock & GW Evans (eds), *Proceedings of the 1999 Winter Simulation Conference*, ACM Press, Piscataway, New Jersey.

Aladwani, AM 2001, 'Change management strategies for successful ERP implementation', *Business Process Management Journal*, vol. 7, no. 3, pp. 266–275.

Anjard, RP 1998, 'Total quality management: key concepts', *Work Study*, vol. 47, no. 7, pp. 238–247.

Ash, JS & Guappone, KP 2007, 'Qualitative evaluation of health information exchange efforts', *Journal of Biomedical Information*, vol. 40, no. 6, suppl., pp. S33–39.

Bond, GR 1991, 'Variations in an assertive outreach model', *New Directions for Mental Health Services*, no. 52, pp. 65–80.

Cale, L & Harris, J 2009, 'Fitness testing in physical education: a misdirected effort in promoting healthy lifestyles and physical activity?', *Physical Education & Sport Pedagogy*, vol. 14, no. 1, pp. 89–108.

Caracelli, VJ & Riggin, LJC 1994, 'Mixed-method evaluation: developing quality criteria through concept mapping', *American Journal of Evaluation*, vol. 15, no. 2, pp. 139–152.

Carey, RG & Seibert, JH 1993, 'A patient survey system to measure quality improvement: questionnaire reliability and validity', *Medical Care*, vol. 31, no. 9, pp. 834–845.

Centers for Disease Control and Prevention 2008, *Data collection methods for program evaluation: questionnaires,* Evaluation briefs, no. 14, Centers for Disease Control and Prevention, Atlanta, Georgia, viewed 3 August 2012, <http://www.cdc.gov/healthyyouth/evaluation/pdf/brief14.pdf>.

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

Chien, C, Wang, W & Cheng, J 2007, 'Data mining for yield enhancement in semiconductor manufacturing and an empirical study', *Expert Systems with Applications*, vol. 33, no. 1, pp. 192-198.

Cole, G 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.

Creswell, JW 2009, *Research design: qualitative, quantitative, and mixed method approaches*, 3rd edn, Sage, Thousand Oaks, California.

De Grave, W, Boshuizen, H & Schmidt, H 1996, 'Problem-based learning: cognitive and metacognitive processes during problem analysis', *Instructional Science*, vol. 24, no. 5, pp. 321–341.

Doggett, AM 2005, 'Root cause analysis: a framework for tool selection', *Quality Management Journal*, vol. 12, no. 4, pp. 34–45.

Donaldson, SI 2003, 'Theory-driven program evaluation in the new millennium', in SI Donaldson & M Scriven (eds), *Evaluating social programs and problems: visions for the new millennium*, Lawrence Erlbaum, Mahwah, New Jersey.

Duffy, G, Moran, J & Riley, W 2012, *Solve the real problem using root cause analysis*, Public Health Foundation, Washington, DC, viewed 5 May 2012, <http://www.phf.org/resourcestools/Pages/Solve_The_Real_Problem_Using_RCA.aspx>.

Gaskell, G 2000, 'Individual and group interviewing', in MW Bauer & G Gaskell (eds), *Qualitative researching with text, image and sound: a practical handbook*, Sage, London.

Glesne, C 2011, *Becoming qualitative researchers: an introduction*, Pearson Education, Boston.

Hollnagel, E 1999, *Accident analysis and barrier functions*, Department of Information Technology, Uppsala University, Sweden, viewed 3 April 2012, <http://www.it.uu.se/research/project/train/papers/AccidentAnalysis.pdf>.

Holzer, PJ, Higgins, JR, Bromfield, LM, Richardson, N & Higgins, DJ 2006 'The effectiveness of parent education and home visiting child maltreatment prevention programs', *Child Abuse Prevention Issues*, no. 24, pp. 1–24, viewed 2 August 2012, <http://www.aifs.gov.au/nch/pubs/issues/issues24/issues24.html>.

Huntington, C & Renger, R 2003, *NAO's response to the OMB–PART report: April–June*, NAO News of the National AHEC Organization, Oak Creek, Wisconsin.

Lewis, J 2003, 'Design issues', in J Ritchie & J Lewis (eds), *Qualitative research practice: a guide for social science students and researchers*, Sage, London.

Liang, Y, Zhang, Y, Jette, M, Sivasubramanium, A & Sahoo, R 2006, 'BlueGene/L failure analysis and prediction models', *Proceedings of the 2006 international conference on dependable systems and networks*, vol 3, Philadelphia, Pennsylvania, pp. 425–434.

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

May, T 1997, *Social research: issues, methods and process*, 2nd edn, Open University Press, Buckingham, UK.

McDavid, JC & Hawthorn, LRL 2006, *Program evaluation and performance measurement: an introduction to practice*, Sage, Thousand Oaks, California.

McLaughlin, JA & Jordan, GB 1999, 'Logic models: a tool for telling your program's performance story', *Evaluation and Program Planning*, vol. 22, no. 1, pp. 65–72.

Mills, SC & Ragan, TJ 2000, 'A tool for analyzing implementation fidelity of an integrated learning system', *Educational Technology Research and Development*, vol. 48, no. 4, pp. 21–41.

Morell, J 2000, 'Internal evaluation: a synthesis of traditional methods and industrial engineering', *American Journal of Evaluation*, vol. 21, no. 1, pp. 41–52.

Ohno, T 1988, *Toyota production system*, Productivity Press, Portland, Oregon.

Patton, MQ 1987, *How to use qualitative methods in evaluation*, Sage, Newbury Park, California.

Patton, MQ 2003, *Qualitative evaluation checklist*, Evaluation Checklists Project, viewed 3 August 2012, <http://dmeforpeace.org/sites/default/files/Patton_Qualitative%20Evaluation%20Checklist.pdf>.

Posavac, EJ 2011, *Program evaluation: methods and case studies*, 8th edn, Prentice Hall, Boston.

Rasmussen, J 1997, 'Risk management in a dynamic society: a modeling problem', *Safety Science*, vol. 27, no. 2/3, pp. 183–213.

Reijers, HA & Mansar, SL 2005, 'Best practices in business process design: an overview and qualitative evaluation of successful redesign heuristics', *The International Journal of Management Science*, vol. 33, no. 4, pp. 283–306.

Renger, R 2010, 'Constructing and verifying program theory using source documentation', *The Canadian Journal of Program Evaluation*, vol. 25, no. 1, pp. 51–67.

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, Davis, M & Granillo, B in press, 'Using root cause analysis (RCA) to facilitate corrective actions, after action reports (AARs), and improvement plans', *Journal of Emergency Management*.

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.

Saunders, RP, Evans, MH & Joshi, P 2005, 'Developing a process-evaluation plan for assessing health promotion program implementation: a how-to guide', *Health Promotion Practice,* vol. 6, no. 2, pp. 134–147.

Scriven, M 1991, *Evaluation thesaurus*, 4th edn, Sage, Newbury Park, California.

Senders, JW 2004, 'FMEA and RCA: the mantras of modern risk management', *Quality and Safety in Healthcare,* vol. 13, no. 4, pp. 249–250.

Shern, DL, Trochim, WMK & LaComb, CA 1995, 'The use of concept mapping for assessing fidelity of model transfer: an example for psychiatric rehabilitation', *Evaluation and Program Planning,* vol. 18, no. 2, pp. 143–153.

Spath, P 2003, 'Don't get caught in the activity trap', *Hospital Peer Review*, vol. 28, no. 9, pp. 130–132.

Speelman, EN, Lopez-Ridaura, S, Colomer, NA, Astier, M & Masera, OR 2007, 'Ten years of sustainability evaluation using the MESMIS framework: lessons learned from its application in 28 Latin American case studies', *International Journal of Sustainable Development and World Ecology*, vol. 14, no. 4, pp. 345–361.

Trochim, W 1989, 'An introduction to concept mapping for planning and evaluation', *A Special Issue of Evaluation and Program Planning*, vol. 12, no. 1, pp. 1–16.

Venkatasubramanian, V, Rengaswamy, R, Kavuri, SN & Yin, K 2003, 'A review of process fault detection and diagnosis part III: process history based methods', *Computers and Chemical Engineering*, vol. 27, no. 3, pp. 327–346.

Williams, PM 2001, 'Techniques for root cause analysis', paper presented at the pathology fall symposium, disaster and emergency management: knowledge gained, experience applied, Baylor University Medical Center, 2 November, *Baylor University Medical Center Proceedings*, vol. 14, pp. 154–157.

Yampolskaya, S, Nesman, TM, Hernandez, M & Koch, D 2004, 'Using concept mapping to develop a logic model and articulate a program theory: a case example', *American Journal of Evaluation,* vol. 25, no. 2, pp. 191–207.