
A New Realistic Evaluation Analysis Method: Linked Coding of Context, Mechanism, and Outcome Relationships

American Journal of Evaluation
33(3) 339-349
© The Author(s) 2012
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1098214012440030
<http://aje.sagepub.com>



Suzanne F. Jackson¹ and Gillian Kolla¹

Abstract

In attempting to use a realistic evaluation approach to explore the role of Community Parents in early parenting programs in Toronto, a novel technique was developed to analyze the links between contexts (C), mechanisms (M) and outcomes (O) directly from experienced practitioner interviews. Rather than coding the interviews into themes in terms of context, intervention elements (mechanisms) and outcomes separately and which could be assembled into CMO configurations by the analyst, they were coded as linked dyads and triads directly from the practitioner narratives. Out of all of the linked codes entered, there were a maximum of three with the same combination, presenting challenges for typical qualitative data analysis. This article examines a novel technique that was developed in an attempt to expand this method beyond the circumstances described in the realistic evaluation literature to date. The bulk of the article focuses on the linked coding and analysis procedures, the challenges faced, and the original solutions that were developed to analyze the CMO relations and generate the mid-range theories necessary to move to the next stage of a realist evaluation approach. The features that distinguish this linked coding method from other methods (e.g. Qualitative Comparative Analysis), the major benefits and drawbacks, the utility of the approach within evaluation practice, and its application to realist synthesis and research are discussed.

Keywords

realistic, qualitative, analysis, evaluation

The Issue

The Realistic Evaluation method described by Pawson and Tilley (1997) requires the generation of mid-range context–mechanism–outcome (CMO) hypotheses about program functioning in different contexts. In the interests of developing these theories in an area with very little preexisting theory

¹ Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Suzanne F. Jackson, Dalla Lana School of Public Health, University of Toronto, 155 College St. Room 516, Toronto, ON M5T 3M7, Canada

Email: suzanne.jackson@utoronto.ca

Table 1. Summary of Major Ways That Context-Mechanism-Outcome Configurations or Testable Hypotheses can be Generated as Reported in the Literature Reviewed

Domain	Recommended Process
Review and synthesis of literature	(a) Realistic synthesis of literature (Pawson & Tilley, 1997; Greenhalgh et al., 2007) (b) Selection of key relevant theories from published literature, including applying relevant theories from another field of study (Blamey & Mackenzie, 2007; Evans & Killoran, 2000; Kennedy et al., 2005; Leone, 2008; Pawson & Tilley, 2001; Timmins & Miller, 2007; Tolson et al., 2007)
Talking to staff or stakeholders	(a) Interviews or focus groups with staff about what works and why (Priest & Waters, 2007) (b) Co-creative dialogue about CMOs between staff and evaluator (Pawson & Tilley, 1997) (c) Delphi method with interviews (Fisher & Downes, 2008) (d) Reflexivity and critical discourse (Hurley et al., 2010)
Empirical research	(a) Quasi-systematic review of literature in interaction with field observations (Blaise & Kegels, 2004) (b) Process-oriented interviews with RCT to provide outcomes (Byng et al., 2005)

and published evidence, the literature typically describes a qualitative interview process with the plausible hypotheses derived logically from the data. Pawson and Tilley describe an interactive interview or dialogue as a method to understand context, mechanism, and outcome relations where the interviewer and the interviewee co-develop these hypotheses (Pawson & Tilley, 1997). If many practitioners are involved, is there another way that these relations can be uncovered using interviews with a group of practitioners?

Several methodological problems are posed in trying to do this. One is the need to ask the right questions to collect information about the contexts, and the program activities and characteristics that could become mechanisms. A second is how to recognize the CMO connections that are made implicitly or explicitly in the data. A third major problem is how to analyze connected strings of data in a way that is different from grouping codes under common themes. This article will examine a novel technique that was developed in an attempt to expand this method beyond the circumstances described in the realistic evaluation literature to date. We begin by examining the realist evaluation literature related to coding, and then describe the case and data collection methods. The bulk of the article focuses on the analysis methods, the challenges we faced, and the solutions that were developed. We end by describing the strengths and limitations of the linked coding approach and suggest implications for evaluation practice and future research.

Background

The first step of the realistic evaluation process described by Pawson and Tilley (1997) is to identify a set of plausible midrange theories that may explain which intervention components work for whom and in what circumstances. The intent is to develop a set of possible relationships between the context or circumstances (C), the intervention components or mechanisms (M), and the outcomes or results (O). The second step in the process is to then test and refine these theories during program implementation. From the literature on realistic evaluation, several different methods have been used to generate the initial CMO configurations that form the first step of a theory-driven evaluation (Table 1). It is possible to derive theories from the literature if there are preexisting theories or existing evidence on the effectiveness and implementation details (Greenhalgh, Kristjansson, & Robinson, 2007; Pawson & Tilley, 2001). If published evidence is not readily available and/or the intervention is in an experimental or pilot phase, techniques rely on consultations with experienced staff (Priest & Waters, 2007). Some other empirical research

methods have been employed which use field observations (Blaise & Kegels, 2004) or randomized control trial results (Byng, Norman, & Redfern, 2005).

In all of these cases, it is important to understand the connections between the contexts, mechanisms, and outcomes. We are particularly interested in the cases where empirical information is gathered about an intervention in order to develop (Step 1) or test (Step 2) mid-range CMO configurations. In order to maximize the information gathered and to ensure that details about all three aspects are captured, information is usually gathered using qualitative approaches, such as interviews or focus groups. Semi-structured interviews (Kennedy, Rogers, & Gately, 2005), a reflexive critical discourse (Hurley, Kelly, Mears, Morrison, & Wiseman, 2010), or a co-created dialogue (as recommended by Pawson & Tilley, 1997) have been used. Fisher and Downes (2008) used a Delphi consultation process with a wide group of stakeholders to generate and validate a set of initial scenarios representing mid-range theories. Usually, the results of the interviews are presented as a table listing all of the contexts, mechanisms, and outcomes in columns but the explicit connections are not always clear (Evans & Killoran, 2000; Kennedy et al., 2005). The connections are derived logically by the evaluation team and initial plausible hypotheses or CMO configurations are presented. In all of the articles reviewed, there was no indication that direct links between Cs, Ms, and Os were recorded from the interviews. The information was always reported in aggregated form in tables or final statements. The only exception was Byng, Norman, and Redfern (2005) who developed an elaborate procedure for analyzing their data showing direct links between Cs, Ms, and Os.

Given the emphasis that realistic evaluation places on the links between contexts, mechanisms, and outcomes, it seemed to us that it would be important to gather information from those most experienced in the program operations about what they know explicitly and implicitly in terms of CMO links. Other studies have used professional experts with education and conceptual abilities to co-create hypotheses (Blamey & Mackenzie, 2007). Even then, the evaluator or the authors developed the mini-theories based on the connections they saw in the data. Qualitative methods such as open coding and thematic analysis have been used (Tolson, McIntosh, Loftus, & Cormie, 2007). We wanted to uncover the links as understood directly from the program implementers. When they talk about their program, what are the links they make between the features of the context, the program elements and the outcomes they observe? Other methods, such as Qualitative Comparative Analysis (QCA), are also concerned with “causal complexity,” or the recognition that the constellation of multiple contexts and mechanisms must be taken into account when evaluating the factors that lead to a certain, specified outcome or set outcomes (Befani, Ledermann, & Sager, 2007; Ragin, 1987; Sager & Anderreggen, 2012). A synthesis of QCA and realist evaluation have been developed which provides an interesting method for reducing and analyzing data by summarizing it into a truth table, followed by minimization based on Boolean algebra (Befani et al, 2007; Rihoux, 2006; Sager & Anderreggen, 2012). However, our team was interested in the steps prior to data reduction, and in exploring ways to record and analyze the dyads (CM, MO, CO) and triads (CMOs) that emerge directly from interviews with practitioners. We could find no literature on “linked coding” and working with dyads and triads of data within a qualitative analysis. Thus, we believe the following details about the methods we employed to be an original contribution to the field.

The Case and Data Collection Method

Program coordinators in two local children’s mental health agencies were interested in examining the unique role of Community Parents (CPs) in early parenting programs (EPPs). CPs are parents with first-hand experience similar to the experience of the participants in the program, who are not required to have formal qualifications in health care or early childhood education, but instead rely on their lived experience as community members and parents. The program coordinators were interested in examining the contribution of CPs as the local public health department was piloting a new program model within high-risk communities that did not include CPs. These coordinators believed very strongly in the

importance of the role of CP due to their experience over the previous 10 years in developing and implementing EPPs in a large urban center that were aimed at people living in new immigrant communities who were experiencing risk conditions such as low income, unemployment or underemployment, lack of knowledge of community services, and few social supports. As Pawson and Tilley argue, in the course of program delivery, practitioners create hypotheses about what works for whom in what circumstances (1997). We felt it was an ideal opportunity to uncover the explicit and implicit understandings of practitioners working in EPPs about the role of CPs. This study was reviewed by the Health Sciences Research Ethics Board at the authors' university.

Eleven interviews were conducted with CPs, early childhood educators working in EPPs in coordination roles, and public health nurses who both worked with and did not work with CPs. Participants were chosen based on their familiarity with EPPs and their depth of experience, as they were ideally placed to provide insights and rich information about the functioning of these programs in a variety of contexts. Each semi-structured interview lasted approximately 30 min and explored the practitioner's role, the characteristics of the populations they worked with, markers of success, what contributed to success, the role and contribution of the community parent, and when they thought it was necessary to include a community parent in the program and why. The interviews were transcribed and entered into NVivo for analysis.

The Analysis Methods

Initially, the interviews were coded in terms of statements related to contexts (C), mechanisms (M), and outcomes (O) with discrete codes, following a typical qualitative thematic analysis process (Patton, 1990). Afterward, the codes were grouped and hierarchies were created. Attempts were made to conduct data reduction by following the method described by Byng et al. (2005). In this method, attention is focused on an outcome of interest, and analysis and reduction proceed by identifying the mechanism most associated with this outcome to form a MO dyad. Data are then examined to determine which contexts are associated with MO dyads. When a context is identified, it is incorporated into the MO dyads to form a "conjectured Context–Mechanism–Outcome configuration" or "CCMOC" that functions as a hypothesis which can be tested by examining positive and negative cases of the phenomena under study (Byng et al., 2005). As with much of the realist evaluation literature, the process of analytic induction described by Byng relies heavily on the existing literature and theory to identify the discrete contexts, mechanisms and outcomes under investigation. Additionally, this process relies on the associations and connections made by the analysts for the development of the CMO configurations.

Given the lack of an established literature on our topic, our team wondered if there was a method that could be used to generate the CMO connections empirically from the data. We began to question whether, by sticking very closely to the descriptive accounts of the interviewees, it would be possible to generate midrange CMO configurations directly from the narratives of the practitioners. In order to do this, all the interviews were completely recoded. This time, the focus was on looking for CMO connections, or strings, within the narratives of the interviewees. So in addition to coding individual units (a discrete C, M or O), we focused on identifying dyads, triads or more complex strings in the narratives. It was a complete surprise to the authors to find that almost every sentence of each interview was a CM, MO, CMO, CMMO, or some variation. No matter what question was asked, the respondents spoke naturally in terms of context, mechanism and outcome connections. We realized that we were exploring a situation where using the respondents' narratives to generate these connected or "linked" codes was more suited to examining our data. By coding these linkages directly, we realized that it would allow us to explore the different connections or constellations of specific contexts and outcomes that practitioners themselves identified, allowing us to then examine whether these differed from what existed in the formal literature. Two example sentences were:

They [parents] know how to communicate, how to act if the child is crying. We are giving them tools and these tools are used, that they didn't have when they came to the program.

A lot of our moms are first time moms, we do have people coming back a second time, but I think it benefits from having the community parent there. It is that more comfortable person to talk to . . . they speak your language or come from the country you are from.

In terms of CMO connections, the sentences were characterized as follows:

They know how to communicate, how to act if the child is crying. We are giving them tools and these tools are used, that they didn't have when they came to the program. [outcome—mechanism—context]

A lot of our moms are first time moms, we do have people coming back a second time, but I think it benefits from having the community parent there. It is that more comfortable person to talk to . . . they speak your language or come from the country you are from. [context—outcome—mechanism]

The criterion used to determine if the phrase was a context was that it described something that existed prior to the introduction of the program. This could be characteristics of the location or community, or the personality or knowledge of the parent. The mechanisms were activities or actions taken in the program or by the community parent or public health nurse. A phrase was coded as an outcome if it described a result of the program or intervention. One author coded all interviews first and the second author reviewed and checked the coding before entering it into NVivo. While our team chose to use NVivo to assist with coding, any program that allows for the coding of qualitative material should be suitable for this task. Differences of opinion were discussed and resolved.

The contexts, mechanisms and outcomes were given subcodes as in a typical qualitative analysis. The study generated 46 context codes, 58 mechanism codes and 24 outcomes codes. The above sentences were coded as follows:

They know how to communicate, how to act if the child is crying [Outcome 9—positive interactions between parent and child] . . . We are giving them tools and these tools are used, [Mechanism 39—teaching tools and skills re bonding and attachment] that they didn't have when they came to the program." [Context 22—Parent's level of knowledge of child interaction]

A lot of our moms are first time moms, [Context 13—participants are first time parents] we do have people coming back a second time, [Outcome 7—frequent attendance] but I think it benefits from having the community parent there. [Outcome 12—General success of program] It is that more comfortable person to talk to . . . [Mechanism 20.2—CP very approachable] they speak your language or come from the country you are from. [Mechanism 15.3—CP shares cultural background]

Specifically, the first sentence was coded as a C22–M39–O9 string and the second selection was coded as a C13–M15–M20–O7–O12 string. As can be seen from the coding above, the CMO coding configurations were drawn directly from the narrative of the interviewees. This was done in order to stay as true as possible to the links made by the interviewees themselves, and to examine whether a structured analysis of how the programmatic context and the mechanisms within program delivery-influence outcomes could be conducted *without* interfering with the natural story-telling process.

The Analysis Challenge and Solution

The major challenge using this form of analysis was that most of the linked strings of codes were unique. There were a few dyads that came up more than once and a maximum of three CMO strings in the same combination. Of the hundreds of linked codes, no pattern emerged. Although the initial challenges around coding as dyads and triads had been solved, the analysis itself posed a problem.

Table 2. Example of a Page of Linked Codes for the Outcome Code—Frequent Attendance

O7—Frequent Attendance
 M9-M42.3—O7—O11
 M38—M44—O7—O8
 M35.5—O7
 M18—O7
 C46—M6—M20.2—M20.3—O7—O12
 C46—M4.3—O7
 C25.6—M20—O7—M1.1—O12
 C25.6—M20—O7
 C21—M42—O7—O10
 C1—M37—O7

Starting with the grouping with the least codes (the Outcomes), the authors found an “edge” for the analysis by dividing the mechanisms into those which were unique to CPs and those which were not. This created two groupings of mechanisms related to each outcome. (Please note that in other cases using this technique, there could be three or more groupings of mechanisms). For example, a typical output page for one outcome (frequent attendance) is shown in Table 2. The mechanisms linked to frequent attendance that were specific to CPs were:

M1.1—CP refers to other parenting programs
 M4.3—CP makes personal connections with participants and “becomes like a friend”
 M6—CP does follow-up call at home
 M9—CPs do outreach in the community
 M18—CP is a person the participants can identify with
 M20—CPs nonjudgmental attitude to parents

This is in contrast to the list of mechanisms linked to frequent attendance that were not related to CP.

M35.5—Room is cold
 M37—Parents participate in EPP program
 M38—EPP group provides space for socialization for parents
 M42—Different topic in EPP program every week

Once this step of linking the two groups of mechanisms with each outcome was accomplished, the analysis could then focus on their relationships to the contexts that had been identified, and look for emerging patterns. As it turned out, the contexts connected to the two different groupings of mechanisms separated neatly in almost every case. This allowed for the development of mid-range hypotheses, empirically linked to the data. In our example, the CP mechanisms were linked to two contexts:

C25.6—Participants are negative about value of participating in EPP
 C46—Participants are new parents

The non-CP mechanisms were linked to two different contexts:

C1—Neighborhood Location (with many new immigrants)
 C21—New parents with the presence of two or three risk factors

The next step in this analysis was to repeat the above process for each of the 24 outcomes and look for patterns (see Figure 1 for a presentation of the procedures). Looking at the patterns for the MO

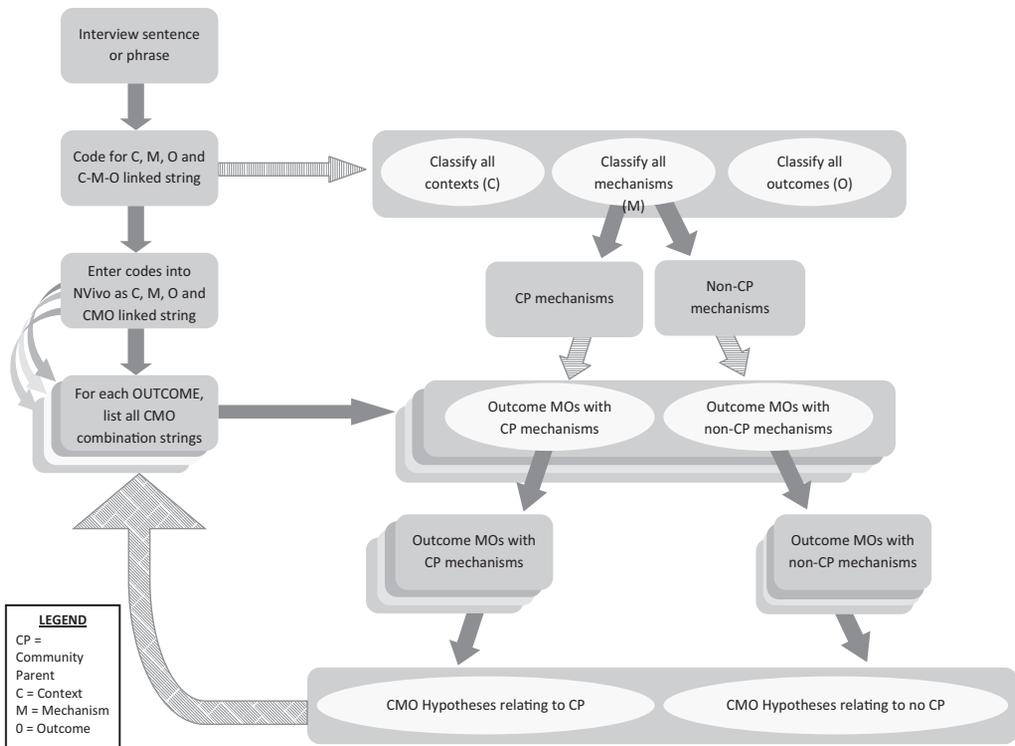


Figure 1. Linked code analysis procedures.

and CM connections for each outcome, the significance for generating mid-range hypotheses emerged. Overall, using this analysis procedure, 14 mid-range hypotheses about the role of CP in EPPs in connection to different contexts were created. The two hypotheses related to the above example about frequent attendance are:

- A. Need for CP—In order to get new parents and those who do not see the need to attend such a program to attend EPPs frequently, CP need to do outreach, follow-up calls, be perceived to be a peer, and provide one-on-one support/confidential listening with a nonjudgmental attitude.
- B. No need for CP—In order to get frequent attendance at an EPP in a community with high risk new parents, it is important that different topics are covered each week, and the space is comfortable and enables participation and socializing.

In this way, the impact of the CP was distinguished from other factors that affect EPP outcomes in the interview data and could be highlighted. Our team was satisfied with the results achieved using this form of linked coding as they could be directly connected to the statements and experiences that were expressed in the narratives of the 11 experienced practitioners involved in early parenting programming. There are two ways that the hypotheses generated were significant: (a) the 14 hypotheses represented 14 different CMO links that the interviewees made in the interviews; (b) they represented mechanisms that specifically included the CP and mechanisms that did not. At this point, it is possible to look for common contexts, mechanisms, and outcomes across the hypotheses and generate one or two overall mid-range hypotheses for testing in the next step, either empirically through an evaluation process or through a review of the literature. At this point, when deriving the

overall midrange hypotheses, one could either maintain the linked codes within the subsequent analysis, using a method such as thematic analysis (Braun & Clarke, 2006) or depart from the linked coding and proceed to use another method of summarizing or synthesizing the data, such as QCA (Befani et al., 2007; Rihoux, 2006).

Strengths and Limitations

This technique of linked coding to derive mid-range realistic evaluation theories from multiple practitioners worked well for a small number of semistructured interviews. A large number of code strings were generated from a relatively small amount of data and the focus of interest was only one aspect of a multi-strategy EPP. It is unclear whether this approach would be applicable in situations where there are large amounts of data to analyze and when there are multiple program elements under review. In these cases, a method such as QCA may be preferable, as adaptations to QCA have been proposed for situations where there are a small, medium, and larger number of cases to be analyzed (Rihoux, 2006).

One of the major drawbacks of the method is that it is time-consuming and generates many hypotheses that require further grouping in order to generate a couple of functional and testable hypotheses. We propose that its value is in the initial stage of generating a theory when literature is sparse or in generating the initial review question for a realist literature review. It may be too cumbersome for application in later stages of evaluation.

Although the focus of this article is not on testing all possible applications of this method, we did explore its possible applicability to document analysis and a realist literature synthesis. We found that fragments of the articles or documents could be coded rather than the article as a whole, which is consistent with Pawson's advice on literature synthesis (Pawson, 2006, p. 88). However, with only one to five CMO strings identified per article or document fragment, it did not seem worthwhile to go through the procedures we outline in this article. We speculate that our linked coding method may be of some assistance in the literature synthesis process at the beginning stage when program theories are being identified and a formal subset of the hypotheses to be tested or examined further during the synthesis are being proposed. Coding the fragments of a few hundred primary articles into CMOs as linked strings may help track the fragments and distinguish different hypotheses. When reviewing a small number of case notes that were somewhat scattered and did not form a coherent "story," the linked coding process had potential to bring some coherence to the fragments of information present.

As a program evaluation technique, this method would likely have limited applicability for busy practitioners because the coding, data entry, and analysis are time-consuming. However, this method could be very effective during the experimental or pilot phase of programs, where testable hypotheses can be generated in preparation for further testing either empirically through an evaluation design or to begin a realist synthesis of the literature. Other positive aspects of this coding process are that it can generate CMO hypotheses that are closely aligned to practitioner views at an early stage of the realist evaluation process which can be useful when there is little existing literature identifying these theories. It is also possible to track the CMO connections made in practitioner narratives with their associated nuances to the mid-range practitioner hypotheses generated across interviews and the final overall hypothesis that may serve to move on to the testing stage. It is possible to link practitioner quotes to the hypothesis to illustrate the conclusions which is consistent with common practice in reporting of qualitative research. This would be familiar and potentially useful for qualitative evaluators and researchers.

Implications for Evaluation Practice

Based on this experience, the main value of the method described above is in being able to empirically generate mid-range theories derived directly from practitioner descriptions of their experience. Practitioners easily generated CMO connections in every sentence when interviewed about the role of CP. While time-consuming in the analysis stage, our experience suggests that this method is useful for uncovering the CMO connections made by practitioners that may be missed within more traditional qualitative analysis techniques. Allowing practitioners to tell the story of their experience of program implementation is a more natural process than attempting to direct participants to focus on the CMO connections consciously. Narratives about program implementation experience can then be analyzed for the implicit (and sometimes explicit) CMO connections that emerge organically from practitioners. This analysis holds promise as a method for discerning the key elements to be monitored in future program implementations, as well as in generating an overall midrange theory to be tested in subsequent evaluations about how the program works, for whom and in what circumstances.

It is possible that the linked coding method could be used to simplify the analysis of data from multiple sources. CMO links could be made in literature fragments, document fragments, and practitioner narratives and then grouped into a few hypotheses. These then become the material (with an empirical base) for developing one or two complex, mid-range, multi-chain hypotheses for testing or to guide a literature synthesis.

Additionally, this method allows for outcomes to be specifically connected to contexts. This is particularly useful feature when attempting to discern the mechanisms that are functioning as mediating factors and are producing consistent outcomes across a wide range of contexts. This can allow for the easy identification of the context–mechanism dyads that are necessary pairings in order to achieve certain outcomes. One value in the method described in this article is that it takes Pawson and Tilley’s key observation that the relationships between factors influencing program success are not fixed. They are, in fact, contingent on the arrangement of specific mechanisms together, within particular contexts, to produce certain outcomes (Pawson & Tilley, 1997). The key is in identifying the combination of context, mechanism, and outcome constellations that produce desired outcomes, and then testing these constellations through further evaluation. The strategy described in this article provides a method to identify these constituent and interconnected elements in a program from the narratives of the practitioners themselves, taking advantage of their first-hand knowledge and experience.

There is a growing emphasis on unimpeded “storytelling” as a way to encourage fuller descriptions of program experiences by practitioners and participants (Frank, 2010). The technique we describe in this article provides a novel method for conducting a structured analysis of how the programmatic context and the mechanisms within program delivery influence outcomes *without* interfering with the natural storytelling process. This technique may also have particular benefits for practitioners and participants who have varying amounts of formal education, or experience with formal evaluation and research process, as it allows them to reflect fully on the range of their experiences without requiring them to understand the conceptual nuances of words like context, mechanism, and outcomes. Concrete information on context, mechanism, and outcomes can then be gleaned within the analysis process without having to engage in structured questioning of participants during data collection.

Implications for Future Research

In the principles for realistic evaluation set out by Pawson and Tilley (2001), they state “Be unafraid to ask big questions of small interventions and to use small interventions to test big theories.” With this in mind, we used a small study of 11 interviews to develop this method. In the future, we will be

investigating the use of this technique of linked coding in realistic evaluations of other programs, combining interview data with document analysis and other data, which report outcomes and contextual features and describe implementation processes. We will also test mid-range hypotheses developed from this linked coding approach. In such cases, interviews with practitioners could be used to generate information about the perceived CMO links, and data collected about the Cs, Ms, and Os using other means could be tested and potentially corroborated.

A second future research initiative would be to use linked coding to analyze the correlation of empirical field experience with evidence from the literature. This approach would be useful in situations where a solid evidence base exists in the literature. However, even when evidence is available, it often does not explicitly describe the details of the program implementation, which can lead to difficulties in using this evidence to construct testable theories. In those situations, it is helpful to combine literature searching with empirical field observations (Blaise & Kegels, 2004). Most other authors choose to fill the gaps by talking directly to program implementation staff, using a variety of methods. Our linked coding method has the potential to facilitate this process, as it could be used in coding both the literature and the staff interviews into CMO strings. Codes from both sources could then be analyzed together to generate the mid-range hypotheses necessary for realistic evaluations.

Acknowledgments

This work was conducted as part of EPP evaluations with Julie Burdon, Building Brighter Futures Coordinator, Hincks-Dellcrest and Dawn Wellington, Coordinator of Better Beginnings Now, Adventure Place funded by the United Way. The interviews were conducted by Suzannah Bennett as part of her Masters of Health Sciences degree and the coding and data entry were done by Gillian Kolla, in partial fulfillment of her Master of Public Health degree, both under supervision of Dr. Suzanne Jackson at the University of Toronto. Funding for this work was received from United Way as part of their "Success By Six" fund.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: United Way.

References

- Befani, B., Ledermann, S., & Sager, F. (2007). Realistic evaluation and QCA. *Evaluation, 13*, 171–192.
- Blaise, P., & Kegels, G. (2004). A realistic approach to the evaluation of the quality management movement in health care systems: A comparison between European and African contexts based on Mintzberg's organizational models international journal of health planning and management. *International Journal of Health Planning and Management, 19*, 337–364.
- Blamey, A., & Mackenzie, M. (2007). Theories of change and realistic evaluation: Peas in a pod or apples and oranges? *Evaluation 13*, 439–455.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101.
- Byng, R., Norman, I., & Redfern, S. (2005). Using realistic evaluation to evaluate a practice-level intervention to improve primary healthcare for patients with long-term mental illness. *Evaluation: International Journal of Theory, Research and Practice, 11*, 69–93.
- Evans, D., & Killoran, A. (2000). Tackling health inequalities through partnership working: Learning from a realistic evaluation. *Critical Public Health, 10*, 125–140.

- Fisher, C., & Downes, B. (2008). Performance measurement and metric manipulation in the public sector. *Business Ethics: A European Review*, *17*, 245–258.
- Frank, A. (2010). *Letting stories breathe: a socio-narratology*. Chicago: University of Chicago Press.
- Greenhalgh, T., Kristjansson, E., & Robinson, V. (2007). Realist review to understand the efficacy of school feeding programmes. *British Medical Journal*, *27*, 335, 858–861.
- Hurley, J., Kelly, M., Mears, A., Morrison, C., & Wiseman, F. (2010). Developing unplanned care teams in Scotland: Evaluating outcomes. *Journal of Psychiatric Nursing and Mental Health Nursing*, *17*, 348–354.
- Kennedy, A., Rogers, A., & Gately, C. (2005). Assessing the introduction of the expert patients programme into the NHS: A realistic evaluation of recruitment to a national lay-led self-care initiative. *Primary Health Care Research and Development*, *6*, 137–148.
- Leone, L. (2008). Realistic evaluation of an illicit drug deterrence programme: Analysis of a case study. *Evaluation*, *14*, 9–28.
- Patton, M. Q. (1990) *Qualitative evaluation and research methods* (2nd ed.). Thousand Oaks, CA: SAGE Publications, Inc.
- Pawson, R. (2006). *Evidence-based policy: A realist perspective*. London, England: Sage.
- Pawson, R., & Tilley, N. (1997) *Realistic evaluation*. London, England: Sage.
- Pawson, R., & Tilley, N. (2001). Realistic evaluation bloodlines. *American Journal of Evaluation*, *22*, 317–324.
- Priest, N., & Waters, E. (2007). ‘Motor Magic’: Evaluation of a community capacity-building approach to supporting the development of preschool children (Part 2). *Australian Occupational Therapy Journal*, *54*, 140–148.
- Rihoux, B. (2006). Qualitative comparative analysis (QCA) and related systematic comparative methods. *International Sociology*, *21*, 679–706.
- Sager, F., & Andereggen, C. (2012). Dealing with complex causality in realist synthesis: The promise of qualitative comparative analysis. *American Journal of Evaluation*, *33*, 60–78.
- Timmins, P., & Miller, C. (2007). Making evaluations realistic: the challenge of complexity. *Support for Learning*, *22*, 9–16.
- Tolson, D., McIntosh, J., Loftus, L., & Cormie, P. (2007). Developing a managed clinical network in palliative care: a realistic evaluation. *International Journal of Nursing Studies*, *44*, 183–195.