A Qualitative Systems Approach to Studying Innovation Implementation in an
Inter-Organizational Smoking Cessation Network

by

Jennifer Lynn Terpstra

B.A., The University of Lethbridge, 2000
M.P.H., San Diego State University, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Interdisciplinary Studies)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2011

Abstract

The overarching purpose of this study was to explore the implementation of an innovation in the North American quitline network using qualitative data and a systems approach. The innovation chosen to explore in more depth was evaluating effectiveness of the tobacco cessation quitlines. The three research questions guiding the study were 1) What are the factors influencing implementation of the innovation, 2) How do system structure and dynamics impact implementation of the innovation, and 3) What strategies can be used to achieve successful implementation of the innovation. To answer the research questions, 19 semi-structured interviews were conducted with decision-makers in the quitline network. The interview data were analyzed using a thematic analysis technique and a systems change framework from the literature. The findings suggest that there were a broad range of factors influencing implementation of the innovation at multiple levels of the system. The findings also provide insights into how differences in quitline structure and system dynamics influenced implementation of the innovation. Lastly, these findings were used to identify potential strategies and provide recommendations to improve future implementation efforts in the quitline network.

Preface

This study was approved by the Human Ethics Board at the University of British Columbia. The ethics certificate number provided by the Human Ethics Board for this study is H09-01651.

Table of Contents

List of Tables

List of Figures

Acknowledgements

There are many people that have helped me to reach this point, and I thank them all. In particular, I would like to thank my committee members for their guidance over the last five years. Thanks to Allan for introducing me to the world of systems thinking, for always treating me with the respect of a colleague, and for encouraging me to think big. Thanks to Chris for her encouragement and mentorship in evaluation. Thanks to Scott for inviting me to join the KIQNIC team and for providing the financial support for this project. And thanks to Wendy for her diligent editing, guidance through the world of qualitative research, and for teaching me to recognize and question the underlying assumptions of my research. I would also like to thank my family and friends for their ongoing support throughout my academic career. I would especially like to thank my mother for always believing in me. Lastly and most importantly, I'd like to thank my husband for being the most amazing partner throughout the entire process, without whose unconditional support and encouragement this would not have been possible.

To my husband, Chad.

Chapter 1. Introduction

Chapter Overview

In recent years, the need to close the gap between research and practice in public health has become increasingly apparent because all too often, critical evidence produced by research fails to be implemented into public health practice and lessons learned from current practice are not always incorporated into new public health research (Davies, Nutley, & Walter, 2005). Despite efforts and resources dedicated to knowledge translation (KT) in public health, this gap "remains a substantial obstacle to improving the quality of health care" (AHRQ, 2001, p.1). Over the last five decades our understanding and conceptualization of the KT process has evolved from a linear, reductionist type approach to a systems approach (Best, Hiatt, & Norma, 2008). A systems approach contrasts with the linear approach in that it recognizes the complexity of both the innovation and the system and considers the network of interdependencies influencing the KT process. A systems approach also acknowledges the context-innovation interaction as a key element to successful KT.

A systems approach provides an alternative paradigm that requires different models and methodologies but can in return produce unique findings and insights to understanding implementation problems. A reductionist approach to implementation attempts to reduce the phenomenon to the smallest parts possible, studying it at a subsystem level. The reductionist approach is a mechanistic view that assumes that the individual parts of the system can be studied separately to understand and predict the properties and behaviours of the whole.

In contrast, a systems approach assumes that properties can emerge at the macro level that cannot be identified or explained at the subsystem level, or from the sum of the parts. Emergence is a key distinction between systems science and reductionism. An example that illustrates this distinction is a human being. One could study the subparts (e.g., heart, brain, muscles) but this would not be the same as looking at the whole person. Even in systems where it is feasible to study every subpart, this would still not provide an understanding of the emergent properties and behaviours at the macro level.

Systems science is a broad area of study and there are many streams within the systems field. Complexity science is a relatively new area of study within systems science that provides specific principles to help guide a systems approach to studying social phenomenon (Flood, 2010). The purpose of this study was to contribute to studying implementation phenomena by adding a qualitative systems approach to a larger positivist quantitative study. This study applies a systems approach and specifically uses complexity science as the underlying theoretical tool.

Study Rationale

My dissertation project was developed based on a larger study funded by the National Institutes of Health (NIH) titled Knowledge Integration in Quitlines: Networks that Improve Cessation (KIQNIC). A primary goal of KIQNIC is to assess how decision-making is a moderator for network characteristics and implementation outcomes in quitlines (QLs). The implementation outcome variable is measured as a summative score of 23 innovations in the QLs. Further details regarding the KIQNIC study and the innovations are provided in chapter three (context). A QL is a telephone-based cessation

service that has been shown to be effective in helping people who want to quit using tobacco (Zhu et al., 2002). QLs offer telephone support primarily through counselling, information, and self-help materials. The number of states and provinces in North America offering QL services for smokers and other tobacco-users has increased exponentially in the last decade. Currently, there are QLs available in all ten provinces in Canada and all 50 states, plus the District of Columbia and Puerto Rico in the United States (U.S.).

KIQNIC researchers have been collecting quantitative data from the QLs through an annual survey conducted over three years. My dissertation project was conducted between the first and second wave of survey data collection, July 2009 and June 2010, respectively. I participated as a collaborator on the KIQNIC project for approximately one year, assisting with various aspects of the research and became interested in the implementation part of the project. Specifically, I became interested in how different methodological approaches to studying the phenomenon could confirm or yield different findings and insights.

Although KIQNIC is using social network analysis (SNA), which is a systems approach, it is based primarily on a positivist paradigm and is collecting only quantitative survey data. For example, one of the objectives of the study is to identify mediator variables for implementation that are reproducible across innovations. The value of this type of research approach is contested by some researchers who argue that because of the "complexity of the innovation, the dynamic and contingent nature of the implementation process, and the shifting environmental context (political, economic, technological), complex service level innovations are inherently unpredictable and that the search for

reproducible "effect sizes," "mediators," and "moderators" is likely to prove fruitless" (Greenhalgh et al., 2008, p.2). An alternative approach is to observe, reflect, and describe the phenomenon using theory, as opposed to trying to identify specific determinants of implementation and predict the system (Greenhalgh et al., 2010).

Implementation of innovations is an iterative process with complex interdependencies in the social system and it has been argued that understanding it requires both qualitative and quantitative methods (Greenhalgh et al., 2010). Qualitative data provide a different perspective on the problem and can complement quantitative data. According to Easterby-Smith, Lyles and Tsang (2008), qualitative methods are better when describing how things change over time and for investigating processes of knowledge transfer, whereas quantitative studies have been better at capturing what is happening at a single point in time.

Qualitative research has many definitions, but a key difference between quantitative and qualitative research is that the latter is naturalistic, meaning the researcher studies the phenomenon in its natural setting and is part of the world by interpreting and making sense of the phenomenon based on the meanings provided by social actors (Denzin & Lincoln, 2005). Denzin and Lincoln (2005, p.10) also suggest that qualitative research emphasizes social experiences and meanings, while quantitative emphasizes "measurement and causal relationships between variables, not processes." Furthermore, qualitative research is well suited to hypothesis generation, whereas quantitative research is better suited to hypothesis testing (W.K. Kellogg Foundation, 2007).

Given the characteristics of the system and the implementation process described above, a systems approach, particularly complexity science, provides a solid theoretical grounding for this study. A system refers to "a set of elements interrelated among themselves and within the environment" (National Cancer Institute, 2007, p. 14). The QL system includes the QL organizations and the actors, as well as less obvious system elements such as policies, culture, and incentives. A more detailed description of the QL system is provided in chapter three. The QL network is a complex adaptive system, meaning that there are a multitude of interconnected parts that are constantly interacting and adapting over time (Holland, 1992). The innovations in the QLs are also high in complexity and successful implementation is often contingent upon simultaneous changes in various system parts. The fluidity and complexity of the system and innovations makes a strictly positivist approach inappropriate because it is not possible to predict the system outcomes or to reproduce the results. The specific theoretical approach driving this dissertation study is complexity science, which integrates elements of interpretivist and positivist approaches and will be described in greater detail in chapter two (the literature review).

Study Purpose & Research Questions

The overarching purpose of my study was to build upon the positivist quantitative KIQNIC study by exploring the implementation of an innovation in the QLs using qualitative data and a systems approach based on complexity science. Although I will not be integrating the qualitative and quantitative data in my study, the qualitative data could be combined with the KIQNIC findings at a later date to move towards a mixed-methods

integrative approach. A single innovation was identified from the list of innovations included in the implementation section of the KIQNIC survey (Appendix A). The innovation "to conduct an evaluation of the effectiveness of the QL," was selected because of a high level of interest in the innovation and the potential for the findings to have practical implications for the QLs. Greater detail regarding the innovation selection process is provided in chapter three. The implementation of this innovation was explored using qualitative data collected via nineteen interviews with decision-makers in the QL community. This study was guided by the following research questions:

Question 1: What are the factors influencing implementation of the innovation?

Question 2: How do system structure and dynamics impact implementation of the innovation?

Question 3: What strategies can be used to achieve successful implementation of the innovation?

Dissertation Outline

This dissertation consists of nine chapters that together present a comprehensive overview of the literature, the study, and the subsequent findings. Chapter one provides a brief overview of the dissertation, including the research questions and the impetus behind it. Chapter two provides a review of the literature presented in a flow of logic leading from KT and evidence-based practices, through implementation, and concluding with an argument for the system approach. I also provide an overview of key complex system principles, which is the theoretical lens applied to my study. Chapter three provides the context of the study including an overview of tobacco related health

mortality and morbidity, a description of the North American QLs, the KIQNIC project, and the innovation studied in this project. The fourth chapter presents my research methods and provides a detailed account of the process of the study including my researcher role and perspective, the study design and methods, ethics, and key decisions made throughout the study. Chapters five, six and seven cover the findings of the study from the thematic analysis. Specifically, in chapter five, I present the normative elements, chapter six is the system resources, and chapter seven is the system regulations and operations. The eighth chapter is the discussion where I reflect on the findings in the context of the literature review and address how the results answer the three research questions. Chapter nine is the final conclusion chapter where I describe the theoretical, methodological, and practical contributions of the study, present lessons learned as well as recommendations for practice and research, and strengths and limitations of the study.

Chapter 2. Literature Review

Chapter Overview

In this chapter, I present the relevant areas of the literature that formed the basis for the theoretical grounding and rationale for my study. I begin by describing the connection between evidence-based practice (EBP) and implementation, followed by an overview of the implementation literature including definitions of the concept, challenges of the field, and the shift from a focus on linear reductionist models to a systems approach to implementation. I conclude the chapter by describing the systems change field including systems thinking followed by an overview of complex systems principles and their relevance to this implementation study. It is important to understand key complex systems principles, as this was the theoretical approach used to frame the study and analyze the results.

Knowledge Translation & Evidence-Based Practices

Evidence-based practice is a priority topic in tobacco control, as well as other areas of public health (AHRQ, 2001). The general consensus is that the utilization of evidence in practice will increase effectiveness and quality of public health practice. As such, significant research efforts have been directed towards understanding how to translate evidence from research to practice (Greenhalgh, Glenn, Bate, Macfarlane, & Kyriakidou, 2005). The last five decades have witnessed an evolution in the KT field whereby linear models have been replaced by relationship models and most recently by systems models (Best et al., 2008). A landmark study conducted by Rogers (1995) developed the diffusion of innovations theory, which has been a foundation of knowledge

for future work exploring the translation of evidence to practice. However, Roger's diffusion of innovation theory and much of the early work in the KT field has approached the problem from a linear lens (Best et al., 2008). With this approach, KT is conceptualized as a one-directional process where evidence is produced by researchers and passed onto practitioners to implement into practice. This approach does not account for the potential context-innovation interaction or the dynamic complexity of the innovation and the system.

Recent definitions of KT reflect this evolved understanding of it as "a process that takes place within a complex system of interactions between researchers and knowledge users" (CIHR, 2005, p. 1). The Canadian Institutes of Health (CIHR) define KT as "a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically-sound application of knowledge to improve the health of Canadians, provide more effective health services and products and strengthen the health care system"(CIHR, 2005, p. 1). CIHR also uses the term Knowledge-to-Action (KTA) and has adopted a specific action cycle or process. The addition of the 'action' part of KT is reflective of the increased understanding that evidence is not automatically taken up in practice once it has been 'translated' or disseminated to practice settings (Graham et al., 2006). "The action part of the process can be thought of as a cycle leading to implementation or application" (CIHR, 2005, p. 1). This term emphasizes the need for active efforts for implementation once an innovation has been diffused and disseminated.

Implementation

The recognition of the need for active efforts to implement innovations after successful dissemination and diffusion has spurred the field of implementation sciences. Implementation science is the "investigation of methods, interventions (strategies), and variables to influence adoption of evidence-based healthcare practices by individuals and organizations to improve clinical and operational decision making, and includes testing the effectiveness of interventions to promote and sustain use of evidence-based healthcare practices" (Titler, Everett, & Adams, 2007, p. S53).

Until recently, there has been a dearth of implementation literature to draw from and the little available, which overlaps with change management and organizational development, is complex in that there are no clear solutions provided to achieve implementation success (Greenhalgh et al., 2005). However, the implementation field has taken root within the last decade, as is evident by the creation of the open access journal specifically for implementation research (Eccles & Mittman, 2006). Despite the advent of the implementation journal and increased funding in this area, there are still many challenges to this field. A primary one is that it lacks a common language and is dispersed across numerous different disciplines. For example, the management and policy literatures refer to innovation implementation and organizational change, the educational literature refers to transformative change, and the health research literature refers to research utilization, implementation, knowledge exchange, and knowledge translation often as synonymous concepts.

Furthermore, implementation as a term can refer to both an outcome and a process. Implementation has been defined as "the early usage activities that often follow

the adoption decision" (Meyers, Sivakumar, & Nakata, 1999, p. 295). It has also been defined as a process by Timmreck (1997, p. 328) that encompasses "the act of converting planning, goals, and objectives into action through administrative structure, management activities, policies, procedures, and regulations, and organizational actions of new programs." This definition describes the necessary activities that take place in order to successfully implement an innovation. Institutionalization and sustainability are terms that denote a high level of implementation in which an innovation has been integrated into the organizational functioning or routine use (Stetler, Ritchie, Rycroft-Malone, Schultz & Charns, 2009). Although implementation and institutionalization are technically different, they are often used interchangeably in the literature.

Implementation of innovations requires organizational change to occur and as such overlaps significantly with the organizational change and management literatures (Moss, 1983; Shortell, 1981). In the health sciences, the term 'practice' is used, whereas the management literature more commonly uses the term 'innovation'. Innovations can be products with distinct boundaries, but they can also be less tangible entities such as practices, policies, or processes that are new to an organization. The KIQNIC project uses the term 'practice,' although the list of practices could also be referred to as innovations as they are a mix of practices, policies, and processes that are relatively new to the QLs. In fact, the majority of the relevant literature would use the term 'innovations.' Evaluating effectiveness of the QLs could be considered an innovation because it is a relatively new practice in the QL system. Furthermore, there is a new 'policy' in the QL system for all QLs to be conducting evaluation and using a standardized evaluation framework. The practice of evaluating effectiveness in the QL system has not yet been

successfully institutionalized. This means that evaluating effectiveness as a practice has not yet been successfully taken up into routine practice and become part of regular daily functioning in the QLs. Therefore the terms practice, policy, and innovation are all appropriate concepts for describing evaluating effectiveness in the QLs. From this point forward, I will primarily use the term innovation in reference to the practice of evaluating effectiveness in the QLs.

Attempts at innovation implementation are often unsuccessful and the reasons why some innovations are successfully institutionalized and others are not is unclear (Repenning, 2002). As such, understanding determinants of successful implementation has become a topic of significant interest in the innovation and implementation literature.

Approaches to Implementation

There exist two distinct methodological approaches for studying organizational change and implementation in the literature: 1) the traditional variance approach (e.g., predictive linear models), and 2) a process narrative approach (e.g., ethnographies) (Van de Ven & Poole, 2005). The majority of organizational change and implementation studies use the traditional variance method with a few applying the process narrative. Even fewer apply a mix of approaches, despite studies that have demonstrated the advantage of an integrative approach (Saberwhal & Robey, 1995; Poole & Van de Ven, 1989).

The different scientific approaches (i.e., variance versus process) are reflected in the different types of implementation models in the literature. According to Marble (2000), there are two types of implementation models, the positivist and interpretivist

models. A positivist position "assumes an external and knowable reality that can be objectively measured, an impartial researcher, and the possibility of producing generalizable statements about the behaviours of the natural and social world" (Greenhalgh, Potts, Wong, Bark, & Swinglehurst, 2009, p. 734). In contrast, an interpretivist position "assumes a socially constructed reality that is never objectively or unproblematically knowable and a researcher whose identity and values are inevitably implicated in the research process" (Greenhalgh et al., 2009, p. 734). The scientific approach used to study the phenomenon will determine the potential findings that can be obtained.

Many of the implementation models in the literature fit into Marble's positivist school (Greenhalgh et al., 2005). Greenhalgh et al. (2005. p.178) suggest that the positivist models have "in common the notion that the implementation process occurs as a sequence of stages that can be planned and controlled, and that planning, controlling and evaluating against predefined success criteria is the key to implementation." An example of a staged and controlled approach to implementation is provided by Graham et al. (2006) in their implementation cycle model. This model does provide some consideration of the need for iteration in the implementation process, as is evident by the fact that the model is circular and not linear. However, it is still a specific, controlled, step-by-step process for implementation.

Implementation research from a positivist approach also applies a mechanistic approach, assuming that it can be dissected into individual determinants that can be studied independently and assessed for their impact on implementation (e.g., adopter skills and type of evidence). This approach assumes that an ideal combination of these

factors can be identified and generalized to different settings and innovations. The problem with the examination of only specific factors or determinants of change (i.e., a reductionist approach) is that it ignores the interdependencies of the factors/agents and the dynamic complexities of the phenomenon that arises from the interactions (Sterman, 2000). Some researchers suggest that the rationale approach to implementing innovations in health services lies at the root of many of the failed attempts to introduce new innovations because it neglects the complexity of the phenomenon (Fonesca, 2001; Plsek & Greenhalgh, 2001). In fact, studies that have attempted to develop a universal formula for successful implementation have been largely inconsistent. For example, one study may find that receptive culture for change is a necessary determinant, whereas another study may find it has no significant influence on the implementation outcomes. The inconsistency in determinants is often attributed to the differences in contextual conditions across settings (Kitson et al., 1998). Another problem with many of the implementation studies using the traditional variance approach is that they assume that the factors/determinants are fixed, when in fact they are dynamic (Bucknall, 2007). For example, Bucknall (2007) describes how models for decision-making in research utilization assume that the environment is static, when in fact it is constantly changing. In Bucknall's study, the behaviours of clinicians varied across time depending on the characteristics of the context at particular points in time.

In contrast, an interpretivist approach to implementation assumes that implementation is not a staged and controlled process, but rather occurs as a result of "social interaction, exchange of ideas, and mutual sense-making" (Greenhalgh et al., 2005, p. 177). In this approach, context and social processes are considered to be central

to knowledge production and utilization (Dopson & Fitzgerald, 2005). Determinants of successful implementation are highly contextual and interact in complex ways, which is why it is not possible to generalize results from one implementation study to another although patterns in the data may be similar if the contexts under examination are also similar (Plsek, 2003). Plsek (2003) suggests that it is in large part because of context interaction that there can be no universal formula for successful implementation. Furthermore, because of the interdependencies in the system it is not possible to study the problem in a 'strict' mechanistic way. Instead, it is ideal to consider the whole system including both mechanistic components and the dynamic context interaction. The acceptance of a whole system, or a systems change approach to implementation, is growing in the implementation field and researchers are beginning to explicitly state the need for a systems change approach to implementation (Kitson, 2009). Aligned with this approach, the evaluation innovation in the QL system requires systems change in order for it to be fully implemented and institutionalized to the point of sustainability, which has yet to be achieved. The following section describes the meaning of systems change and why systems change is needed for full implementation and institutionalization of the evaluation innovation.

Systems Change

Implementation and institutionalization of an innovation is essentially a process of change. The difficulty in achieving change is directly related to the dynamic complexity of both the innovation and the system (Greenhalgh et al., 2005). Foster-Fishman, Nowell, and Yang (2007) suggest that a systems change approach is needed when a problem is

deeply embedded in a system's dominant norms and other system structures and parts.

Systems change refers to a fundamental shift in the nature of the system and substantial

changes to the structural, relational makeup of a system (Hirsch, Levine, & Miller, 2007).

Systems change also requires consideration of contextual factors as an active component

in the process (Suppovitz & Snyder, 2005; Netting, O'Connor & Fauri, 2007). There are

numerous different contexts and the systems change agents must identify the most

important contexts to consider. For example, the social, cultural, and political contexts

are all important to consider when planning systems change (Kreger, Brindis, Manuel, &

Sassoubre, 2007). The need to consider multiple contexts as active agents is yet another

reason why it is not possible to develop one universal model for implementation or

systems change.

It is also necessary to consider the principles underlying the system in order to

achieve it. To assist with this process, Foster-Fishman et al. (2007) developed a

framework for assessing and creating systems change based on organizational change and

systems thinking literature. According to them (2007, p. 201), systems change requires

three considerations: "1) understanding different perspectives concerning the problem

situation; 2) locating root causes to systemic problems by identifying system parts and

their patterns of interdependency that explain the status quo; and 3) using this

information to identify leverage points that will cultivate second-order change." The

authors also note that systems change requires changes to deep structures of the system,

such as normative elements (e.g., attitudes, values, expectations) as well as other system

elements such as system resources (e.g., human, social, economic capital), regulations

(e.g., policies and procedures) and operations (e.g., decision-making structures), which

are the root causes of the system problems. The goal when using this framework is to look at these system parts across levels, niches, organizations, and actors to determine differences between system parts or interactions that create patterns in the system. The patterns can be used to identify leverage points that can shift the system towards the desired state.

Despite Foster-Fishman's framework for assessing and creating systems change, achieving it remains difficult. One of the main reasons is a paucity of literature to guide such efforts. In fact, Greenhalgh et al. (2005) identified only one large-scale program (Riley, Taylor, & Elliott, 2001) that was designed around a whole systems approach in their comprehensive review of the literature. Thus, it is not surprising that scientists trained in the linear, reductionist approach have difficulty moving toward a whole systems approach to implementation because it requires a different type of thinking. This is clear in that a traditional change approach focuses heavily on specific actors or parts in the system; whereas, systems change requires consideration of the patterns in the system, which are created by the interactions between actors and system parts (Olson, Eoyang, Beckhard, & Vaill, 2001).

Qualitative data have been used in recent systems change studies and have demonstrated 'added value' to a strictly quantitative approach. Qualitative data can be used to answer research questions pertaining to what the innovation meant to the stakeholders, as well as the social and technical challenges involved (Greenhalgh et al., 2010). In a recent qualitative study by Greenhalgh et al. (2009), the authors showed that there is no simple recipe for systems change because contexts are complex and rapidly changing.

In another recent systems change study, Greenhalgh et al. (2010) used a mix of qualitative and quantitative methods to show that challenges to implementing a technology innovation were complex interdependencies of both social and technical nature. Together these studies provide evidence for the innovation-context interaction and the inherent complexity of the implementation process. The researchers suggest that for numerous reasons, it was not ideal to use a strictly positivist approach that involved pre-post comparisons or identification of linear causal relationships. For example, there was a dynamic local context and wider policy environment that was influencing the systems change process (Greenhalgh et al., 2009). Due to the dynamic complexity of both the system and the innovation, a mixed-methods integrative approach would be ideal for studying the systems change process. Similarly, the QL system and evaluation innovation are also high in dynamic complexity. Therefore, adding a qualitative component that complements the existing quantitative data being collected by KIQNIC allows for movement toward the ideal approach for studying systems change.

Complexity of Evaluation Innovations

Successful implementation is a result of both the innovation and system characteristics (Greenhalgh et al., 2005). Thus far, I have focused on the importance of examining system characteristics for successful implementation. However, it is also important to consider the characteristics of the innovation in terms of implementation because not all innovations are equal. Specifically, "the more complex the innovation, the more iterative, complex and multidirectional will be the implementation process" (Greenhalgh et al., 2005, p. 175).

Institutionalization becomes increasingly more difficult when the innovation lacks clearly defined boundaries and when the implementation target is complex. For example, implementing a hand-washing protocol for physicians in a hospital is less complex than implementing an evaluation policy in a diverse network of organizations with decentralized decision-making. In the hand-washing example, the innovation can be easily defined and agreed upon by the different stakeholders and there is likely only one governing body with the authority to make the decision to implement the policy. In the evaluation example, the implementation process is more difficult because there are many stakeholders involved with different opinions on the policy. Although both examples require consideration of the context and barriers to implementation, the evaluation policy example often requires systems change in order to achieve successful institutionalization.

Evaluation innovations inherently have a high degree of dynamic complexity. In order for evaluation results to be used, it is necessary to involve members in the evaluation process and to include a process of reflection and adaptation to ensure that the evaluation is relevant and useful (Patton, 2002; Patton, 2008). As a result, institutionalizing evaluation innovations usually requires substantial changes to the system itself. Also, successful institutionalization often requires changes to multiple parts of the system. Because of the high degree of dynamic complexity of both the evaluation innovation and the QL system, institutionalization of the evaluation innovation requires systems change, which in turn requires systems thinking.

Systems Thinking

Systems are defined as bounded entities with interdependent parts, where the whole is greater than the sum of the parts (Stacey, 1996). In the context of systems change, a system refers to, "a set of actors, activities, and settings that are directly or indirectly perceived to have influence in or be affected by a given problem situation" (Foster-Fishman et al., 2007, p. 198). Innovation implementation is viewed very differently from a systems thinking approach relative to a reductionist approach. From a systems approach, innovation is seen as an emergent phenomenon, resulting from underlying patterns of interactions between the actors and system parts (Fonesca, 2001). The systems approach recognizes the dynamic complexity and interdependency of both the innovation and the system (Fonesca, 2001).

To further understand the meaning of a systems approach to studying implementation, consider the example of the hand-washing protocol described in a previous section. Reductionism usually assumes that the best approach is to simplify the phenomenon by reducing down to the smallest components possible, to study it at the subsystem level, and abstract the independent parts from the rest of the system. The assumption in this approach is that the individual parts can be studied independently in order to understand the whole. As such, a reductionist approach to implementing a hand-washing protocol for physicians in a hospital may include an intervention (e.g., presentation) to increase the physicians' knowledge of, and attitudes towards, disease spread through physician-patient contact.

In contrast, the systems approach is more ecological in nature as it would view the physician as embedded in the organization, which is embedded in the larger socio-political system, and all of these levels are constantly interacting. A systems approach would recognize the interdependencies in the system and consider those that are influencing the physicians' behaviours. A systems approach would consider the political and institutional context, the relationships between different actors, and also the physical environment of the hospital. An intervention from a systems approach would not just address the physicians' attitudes and knowledge. It may consider physician knowledge as one aspect in addition to factors at the organizational and socio-political level (e.g., organization culture, physician training, hospital scheduling policies, incentives, physical layout of hospital, nurse-physician relationships, etc.). An intervention from a systems approach would likely intervene at multiple parts and levels of the system. For example, interventions might include the physician presentation, placement of antibacterial lotion above each patient's bed, and implementation of a reward policy that incentivizes physicians to wash their hands between each patient visit. In order to create a hospital culture that emphasizes hand washing the intervention might also target other stakeholders in the hospital such as the administrators and nurses. Another aspect of a systems approach is consideration of unintended consequences. Because of the interconnectivity of system parts, changes in one part of a system will result in changes in other parts of the system. For example, providing a financial incentive to physicians for hand washing could result in resentment from other hospital medical staff and a culture in the hospital that is driven by monetary gains over patient welfare.

Systems thinking aims to capture these interdependencies between different levels and parts of the system to better understand how an innovation can be implemented. Systems thinking also recognizes that there are different types of systems, such as simple systems, open/closed systems, and complex systems; each with different principles guiding the system's behaviour. I will describe complex systems in more detail below as it is this type of system that best characterizes the QL system.

Complexity Science as a Theoretical Approach for Systems Change

Complexity science is a branch of systems science that looks specifically at the behaviour of complex systems (Zimmerman, 2001). From this approach, a system is viewed as a living organism and as such is seen to have behaviours. A complex system is a network of interdependencies that is constantly adapting, learning, and changing over time (Cilliers, 1998). Complex systems operate based on unique principles and characteristics that guide the system's behaviour. The majority of implementation problems in the health area are located within complex systems (Plsek, 2001). As such, it is important to understand complexity principles in order to change the system.

Complexity science can be used as a theoretical grounding and conceptual framework for guiding implementation and systems change case studies (Anderson, Crabtree, Steele, & McDaniel, 2005). Complexity science is viewed as a new type of science that incorporates elements of both positivism and interpretivism (Vogel, 2009). For example, similar to interpretivism, it emphasizes context, integration of information across different perspectives, and does not aim for reproducibility or predictability of specific outcomes. However, similar to the positivist approach, it assumes that there are

22

identifiable causal relationships that create patterns in the system, albeit the relationships are viewed as nonlinear. As previously mentioned, an integrative approach has been shown to have an advantage over a strictly interpretivist or positivist approach in implementation and organizational change studies (Saberwhal & Robey, 1995; Poole & Van de Ven, 1989).

Although the current study is qualitative and interpretivist, it is intended to complement the KIQNIC study which uses a positivist quantitative approach. By using complexity science as a theoretical approach, the results of this study can be incorporated into the KIQNIC study findings at a later date to create an integrative mix-methods study of systems change and implementation. Furthermore, the use of theory in implementation research has been inadequate and there is a need for more theory (Cummings et al., 2007; Grimshaw et al., 2004). Complexity science has been shown to be useful for studying the implementation of evidence-based practices and systems change (Murphy-Smith, 2004) and is therefore the theoretical approach used to frame this study. The Foster-Fishman (2007) framework was developed based on the organizational change and systems thinking literature specifically to guide systems change efforts.

Principles of Complex Systems

As previously mentioned, complexity science makes the application of systems thinking easier (Flood, 2010). Complexity science provides specific principles to help understand the behaviours of a social system. The purpose of the following section is to describe key complex systems principles as well as how they are related to systems change for innovation implementation. The goal of this study is not to test complexity

science or to assess the complexity of the QL system. The impetus for describing the complexity principles is to be explicit about the assumptions underlying the research paradigm with respect to the system's behaviour. It should be noted that a combination of both interpretivist (e.g., change via acting on relationships) and positivist (e.g., causal feedback loops) approaches are demonstrated throughout the description of complexity principles. The assumptions underlying the research paradigm of this study include that the QL system is robust, has a multitude of dynamic interconnections between parts, and has nonlinear causal relationships that create feedback loops in the system. These underlying assumptions of the QL system's behaviour are described in greater detail below.

Robustness

One characteristic of complex systems is that they are thought to be robust and resistant to change (Carlson & Doyle, 2002). Robustness refers to "the maintenance of some desired system characteristics despite fluctuations in the behaviours of its component parts or its environment" (Carlson & Doyle, 2002, p. 2539). Perturbations in complex systems may cause upset initially, but the system will quickly re-organize back to its initial state of equilibrium. This characteristic is important in the context of systems change because it helps to explain why many change efforts have little effect or success. Complex systems have an internal structure and patterns that maintain the status quo or equilibrium (Carlson & Doyle, 2002). In order to achieve systems change, it is important to understand all aspects of the system that are contributing to maintaining the status quo such as structures, relationships, and perspectives (Behrens & Foster-Fishman, 2007).

Interconnectedness

Part of what makes complex systems so robust and resistant to change is the interconnectedness of the system (The National Academies Keck Futures Initiatives, 2009). Complex systems have inter-linkages between components that are dynamic so that change in one system component affects other components. There are a couple of pertinent implications for this principle of interconnectedness, one is related to robustness and the other to nonlinearity.

In terms of robustness, if a change intervention is directed to one component of the system, it may change that one component momentarily. However, because of the inter-linkages in the system that one component is being 'pushed' on by many other system components. If change has only been made in that one component and the rest of the system is still in the initial state, then the rest of the system will 'push' the changed component back to its initial state. Take, for example, an attempt to shift a healthcare organization from a focus on treating disease to a focus on health promotion and disease prevention. One intervention that has been advocated for is having health practitioners (e.g., primary care physicians) provide counseling to patients on positive lifestyle modification (Egede & Zheng, 2002). Although the intervention may succeed in educating the physicians and changing their values so that they will want to counsel the patients on healthy lifestyle changes, this single intervention is unlikely to be successful if done alone. The problem is that other components of the system are not congruent with this change. That is, if lifestyle counseling is not on the reimbursement schedule, then physicians will not likely do it even if they believe it is important (Sesselberg, Klein,

O'Connor, & Johnson, 2010). Another example of push back could be if a pharmaceutical company has been visiting physicians marketing a new blood pressure drug and providing free samples, then the physicians may be more likely to give the drug to a patient with high blood pressure (Vancelik, Beyhun, Acemoglu, & Calikoglu, 2007). The problem is further compounded by the fact that physicians perceive a major barrier to be the patients' unwillingness to change their lifestyle in order to reduce risk factors (Jallinoja et al., 2007). The perceived patient's unwillingness then interacts with pharmaceutical marketing to reinforce the physician's choice to treat the patient with a drug instead of lifestyle counseling. This example illustrates why targeting change in one component of the system is unlikely to achieve success.

In order to achieve change in a complex system, it is necessary to have parallel mechanisms for change across different parts and aspects of the system. It is critical that any systems change planning and evaluation efforts consider the coherence and alignment of system components (Suppovitz & Snyder, 2005). The solutions, or change interventions, must be interdependent in the same way that the system is interdependent (Janzen, Nelson, Hausfather, & Ochocka, 2007). Incongruence between parts of the change efforts, or between different system components, will increase the likelihood that the change efforts will fail. For example, if decision-makers in a health system are advocating a shift towards health promotion, but they do not make healthy lifestyle counseling a billable treatment for physicians, then this incongruence between change goals and current policies will cause resistance to change. Similarly, if physicians are receiving compensation or "kick-backs" from pharmaceutical companies, then this will

also create resistance to changing the system towards the desired shift to health promotion.

Nonlinearity

The other important implication of interconnectedness of the system is the principle of nonlinearity. Complex systems are nonlinear systems, which in technical terms means that the input is not necessarily equal to the output (Willy, Neugebauer, & Gerngroß, 2003). In this case, a very small input (i.e., an intervention) could result in a massive output (i.e., change), and conversely, a very large input could result in little to no output (Eoyang, 1998).

The key characteristic that allows for this phenomenon to occur is that variables can be both a cause and an effect of specific phenomena in a nonlinear system. In a linear system, it is assumed that a variable is either a cause or an effect; it is never both simultaneously. In a nonlinear system, cause and effect relationships are distal, not so obvious or easy to identify and include variables that are a cause and an effect simultaneously. An example of a linear relationship is the relationship between the gene for Huntington's disease and an individual getting the disease. In this example, the gene is the cause and the disease is the effect, and the gene cannot also be the effect in the relationship. However, linear relationships such as this are rare in implementation and social systems. More common are nonlinear reciprocal relationships. For example, the relationship between expectations and perceptions is nonlinear because expectation has a causal effect on perceptions, and perceptions in turn can have a causal effect on

expectations. In this way, both variables in the model are both a cause and an effect in the relationship. It is because of this nonlinearity that outcomes can be greatly amplified.

Nonlinearity is an important concept for systems change for several reasons. For one, nonlinearity can cause unintended consequences in the change efforts. Change in the targeted component can cause a chain reaction of changes that can ultimately come back around and affect the original targeted component. The unintended consequences can result in amplifying or dampening of the desired outcome in the targeted component and it can also create completely different changes in the system. It is not possible to predict exact outcomes in complex nonlinear systems because there are too many extraneous variables that cannot be controlled for (Eoyang, 1998). Instead, it is important to identify the interdependencies in the system and think through the potential non-linear relationships when trying to change the system.

Feedback Loops

Another reason why interconnectedness and nonlinearity are important for systems change is because these principles create feedback loops which are at the heart of systems thinking and are critical for systems change. They are a result of nonlinearity in the system, which allows a variable to be both a cause and an effect at the same time (Sterman, 2000). According to system dynamics theory, all complex systems are made up of two kinds of feedback loops: positive (i.e., self-reinforcing) and negative (i.e., self-correcting) (Sterman, 2000). The terms 'positive' and 'negative' are not value laden in this context, they only refer to amplifying (positive) or dampening (negative) an initial condition. An example of a negative feedback loop is a thermostat that corrects the

temperature by changing the initial condition, such as the heat coming from the heater. An example of a positive feedback loop is the broken windows paradigm (i.e., when a neighbourhood has abandoned buildings with broken windows it encourages the vandalism of other buildings) (Foster-Fishman et al., 2007). In this case, the broken windows can amplify the initial condition of vandalism. Also, variables can be part of multiple feedback loops, which increases the dynamic complexity of the system (Hirsch et al., 2007). Foster-Fishman et al. (2007) instruct systems change agents using their framework to identify key feedback loops influencing the systems change.

System thinking assumes that the system's behaviour is a result of the underlying feedback mechanisms and that it is necessary to understand both the behaviours and the feedback mechanisms (Hirsch et al., 2007). A complex system adapts and changes over time because it learns and all learning depends on feedback loops. One strategy for systems change is to consider existing feedback loops in the system, as well as develop new feedback loops that help achieve the desired outcome.

Self-Organizing

The object of change in a complex systems approach is on influencing the interactions and exchanges in the system in order to alter the path of self-organizing (W.K. Kellogg Foundation, 2007). Self-organization is defined as the process "whereby new emergent structures, patterns and properties arise without being externally imposed on the system" (Zimmerman, 2001, p. 270). Although the system is complex, there are patterns of interaction that can provide cues for interventions. Patterns in the system emerge over time as a result of these interactions and adaptations within the system. In

addition, agents in a system are constantly learning and adapting as a result of interactions with each other and the system. It is important to understand the self-organizing patterns of the system in order to change the direction of the system.

Patterns for innovation in networks are discernable and these patterns have implications for service and policy decision-making (Kash & Rycoft, 2000). It is important to study and understand the underlying dynamics and structure in interorganizational networks leading to innovation (Gay & Dousset, 2005). A complex systems approach and identification of self-organizing patterns for innovation in networks is useful to decision-makers as it can be used to inform policy decisions (Frenken, 2000).

Leverage Points

Leverage points are the "places in a complex system where a small shift in one thing can produce big changes in everything" (Meadows, 1999, p. 1). As such, leverage points are considered a strategy for achieving implementation and other change initiatives in complex systems. Leverage points are possible because of interdependencies/interactions in complex systems. As a result of the interdependencies, change in one part of the system can create change in other interconnected parts of the system. Numerous possible leverage points exist in a system, with different degrees of potential impact on system change. Leverage points can be counterintuitive and it can be difficult to identify the most powerful and correct leverage points for achieving systems change (Meadows, 1999). Meadows acknowledges that there are no definitive rules that can be generalized to all complex systems, but provides a suggestive list of leverage

points to serve as a benchmark. The leverage points, in order from most to least likely to create change are: 1) power to transcend paradigms, 2) mindset out of which the system arises, 3) goals of the system, 4) organize system structure, 5) rules of the system, 6) information flows, 7) positive feedback, 8) negative feedback, 9) length of delays, 10) physical structure, 11) size of system stabilizers, and 12) constants and parameters (Meadows, 1999). Meadows' twelve leverage point levels have been further modified by Malhi et al. (2009) who collapsed them into five leverage/intervention levels for systems change in food policy. For example, one of the leverage points under Meadow's goal category proposed by Malhi et al. (2009) to change the food system is: agricultural policy that maximizes positive health outcomes and minimizes negative health impacts. Another of the leverage points proposed by Malhi et al. (2009) under Meadow's structure category is: public education on consumption of an environmentally sustainable diet. Malhi et al. (2009) propose that these leverage points, along with nineteen other leverage points developed based on their five intervention leverage point levels, will help shift the food system to be more healthy, green, fair and affordable.

Foster-Fishman et al. (2007) propose two types of leverage points: 1) those that shift fundamental parts to be consistent with the desired change, and 2) those that strengthen system parts that are already consistent with the desired change. The authors suggest that leverage points can be parts of the system (e.g., system elements) or patterns and interactions in the system. Understanding leverage points for systems change also requires consideration of the root of the problem or the moral positions at the heart of the system (Kreger et al., 2007). For example, systems change requires modification of the 'deep structures' of the system such as normative elements (e.g., attitudes, values,

expectations), which are often the root causes of system problems (Foster-Fishman et al., 2007). Social processes can also serve as an important lever for change (Tseng & Seidman, 2007).

Summary

The fields of KT and implementation sciences have evolved significantly over the last five decades from linear models to systems models (Best et al., 2008). The system models address the inherent complexity of both the system and the innovation. Unfortunately, there is a dearth of literature available to guide implementation from a systems approach. The majority of the implementation literature applies a positivist quantitative approach that cannot adequately consider the dynamic complexity of both the innovation and the system. A mixed-methods approach that integrates both interpretivist and positivist thinking has been shown to be advantages for studying systems change. The KIQNIC study applies a positivist quantitative approach to studying implementation in the QLs. This approach has many benefits and provides valuable information for studying the problem. For example, it will identify network connections between organizations and causal relationships in the system. However, a qualitative interpretivist approach provides additional insights and information for understanding the problem. The goal of this present study is to demonstrate the value of adding a qualitative interpretivist perspective to studying the implementation of an evaluation innovation in the QLs. Although the integration of the quantitative and qualitative data is outside of the scope of this project, the hope is that at a later date the KIQNIC project can integrate both approaches to create a more comprehensive understanding of the implementation

phenomenon. Complexity science provides a theoretical framework that integrates

interpretivist and positivist concepts and will thereby allow for a merging of the KIQNIC

study findings with the qualitative findings from this study at a later date. The next

chapter provides the context for this study including a description of the KIQNIC study

and the QL system, as well as the evaluation innovation selected to study in more depth.

Chapter 3. Context

Chapter Overview

The following chapter describes the context for this study. The first section provides an overview of the tobacco context. The following section describes the North American quitlines including the structure of the network, the QL system and the key stakeholders discussed throughout this paper. The final sections in this chapter include a description of the KIQNIC project and an overview of innovation implementation. A detailed description of the evaluation innovation studied in this project is included in the final section.

Tobacco Use & Cessation

Despite significant efforts, tobacco related mortality and morbidity continues to be a daunting public health problem. Tobacco use is the leading preventable cause of mortality and morbidity in North America (CDC, 2005) and is responsible for 400,000 deaths per year in the U.S. (CDC, 2009). Lung cancer is the leading cause of cancer-related deaths among both men and women (WHO, 2009) and smoking causes almost 90% of all lung cancer (Wingo et al., 1999). In 2008, an estimated 20.6% (46.0 million) of the United State's adult population (≥18years) self-reported as current cigarette smokers (CDC, 2009). In 2009, approximately 17% (4.8 million) of the Canadian population aged 15 and older self-reported as current smokers (Health Canada, 2010a). In addition, Canada spends over $3.5 billion to provide direct medical care to tobacco users and over $15 billion when indirect costs are factored in (e.g., worker absenteeism) (Health Canada, 2010b). Tobacco-related illnesses kill about 47,581 Canadians each year

(Makomaski-Illing & Kaiserman, 2004) and more than 1,000 Canadians die each year from heart disease and cancer caused by second-hand smoke (Health Canada, 2008).

The majority of smokers recognize the harmful effects of tobacco use and a large percentage attempt to quit each year (U.S. DHHS, 2004). In 2008, approximately 45.3% (20.8 million) of the adult cigarette smokers in the U.S. had attempted to quit smoking within the twelve months prior to the survey used to collect these tobacco statistics (CDC, 2009). Unfortunately, only about six percent are actually successful at quitting for more than one month on a given attempt (U.S. DHHS, 2004). Quitting tobacco use is difficult, because most people who use tobacco regularly are addicted to the nicotine (Benowitz, 2009). Smoking addiction is a complex problem involving a combination of pharmacological and behavioural factors. For these reasons, significant efforts and resources have been directed into tobacco cessation research and initiatives to better understand and help individuals quit. Quitlines (QLs) are one of the primary tobacco cessation services offered throughout North America.

North American QLs

A QL is a telephone-based cessation service that helps people who want to quit using tobacco. QLs offer telephone support primarily through counselling, information, and self-help materials. Some QLs offer additional services such as medications, online cessation information and programs, and referrals to community-based cessation programs. The first QL started operation in California in 1995 and it grew out of a clinical research trial that demonstrated the effectiveness of phone counselling for tobacco cessation (Zhu et al., 2002). The number of states and provinces in North

America offering QL services for smokers and other tobacco-users has increased exponentially in the last decade.

Currently, there are QLs available in all ten provinces in Canada and all 50 states, plus the District of Columbia, and Puerto Rico in the United States. In addition to these 62 QLs, there are also 22 QLs in Europe, eight QLs in Australia, and one in Mexico. QLs represent a unique opportunity to reduce tobacco use in North America and globally. In 2008, forty-seven of the fifty-two U.S. QLs received a total of 409,902 incoming calls from tobacco users (median = 4,847 calls per QL). That same year, 18,125 incoming calls from tobacco users were received by nine of the Canadian QLs (median = 591 calls per QL) (North American Quitline Consortium, 2009). Given the significant amount of resources being directed to QLs, as well as the number of people reaching out to them for assistance with quitting, it is imperative that they be effective and efficient. In order to increase effectiveness and efficiency, the QLs must be able to disseminate and implement research-based and practice-based innovations throughout the network.

Structure & Characteristics of the QL Network

As described above, there are ten QLs in Canada, 52 in the U.S., and one in Mexico, that together make-up the North American QL network. However, for the purpose of this study, the North American QL network will refer to only the Canadian and American QLs, as the Mexican QL was not included in either this study or the larger KIQNIC study. Also, it is important to note that although both American and Canadian QLs were included, the primary focus is on the QLs in the U.S. There are a couple of reasons for the dominant focus on QLs from this particular country. First, there are far

fewer Canadian QLs in the network and in general they receive far fewer calls than the American QLs do. The second reason is that although there was equal opportunity for decision-makers from both countries to be recruited into this study, only two Canadians were successfully recruited. As such, the study is focused predominantly on the American QLs with significantly less information and findings provided on the Canadian QLs.

The structure of the QLs in these two countries is similar, as for the most part they are composed of two primary entities, the funder organization (e.g., state health department) and the service provider organization. There is, however, some variation in this structure and figure one illustrates some of the different QL structures. In this figure, model A represents a service provider dedicated to a single QL and model B represents two QLs with the same service provider. In addition to the service providers and funder organizations, other entities within the QL community include the North American Quitline Consortium (NAQC), external evaluation contractors, and the Center for Disease Control (CDC) in the U.S. Model C in the figure represents a QL that has an external organization as an evaluation contractor. Lastly, model D represents a QL where there are two service providers and one funder organization. This is a sample of some of the possible variations in QL structure but not all. Most of the American QLs receive their funding from one or both of two possible funding sources, state funding and CDC. The Canadian QL funders are different for each province and include: the provincial ministry of health, the Canadian Cancer Society, Heart and Stroke Foundation of Canada, Health Canada, and Alberta Health Services. Below is a description of the different organizations and actors that are part of the QL community that I will refer to throughout this manuscript.

Figure 1. Different Quitline Structure Possibilities
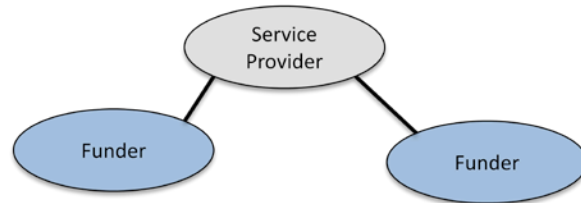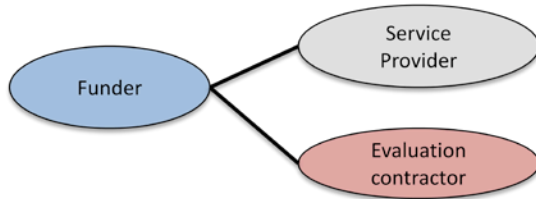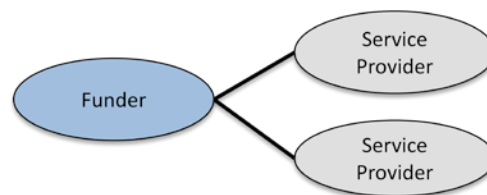
Decision-Makers

The KIQNIC study, which is described in greater detail below, defined decision-makers as, "any individual at a QL funder, service provider or coordinating organization who is involved in decision-making about the implementation of QL practices." According to the KIQNIC project, there are 276 decision-makers in the QL system. The decision-makers include a variety of positions in the QLs including managers, directors, and coordinators. The decision-makers vary significantly in terms of their educational backgrounds and experiences. For example, some of the decision-makers are PhD level researchers with positions in universities and others are administrators with no research background.

Service Providers

The service provider organization is the entity that is responsible for providing the telephone counselling services. The service provider characteristics and additional

responsibilities vary significantly across providers. There are private for profit service providers and public not for profit service providers. Three of the public service providers are located in universities, for example the first QL was established with the University of California - San Diego (UCSD) as the service provider. The university service providers, as well as some of the other service providers are contracted to service only one QL. For instance, UCSD provides telephone counselling services only to the California QL. In contrast, there are larger service providers that contract with multiple QL funders.

Until recently, the two largest QL service providers in the U.S. were Free and Clear, a for-profit company, and the American Cancer Society (ACS), a not for profit organization. Together, these two service providers held the majority of the QL contracts in the states and were in competition with each other. In fall of 2010, these two service providers formed a collaborative wherein Free and Clear assumed all of ACS's contracts. In Canada, all of the QLs with the exception of Alberta and British Columbia are operated by the Canadian Cancer Society. The QL in Alberta is operated and funded by Alberta health services and the QL in British Columbia is operated by Sykes, a for profit tobacco cessation company.

The QLs also have different protocols for the provision of telephone counselling. Although they all provide telephone counselling, which is viewed as an evidence-based practice (innovation), there are significant differences in the details of how these services are provided. For example, some QLs may have all Master degree level counsellors conducting the calls with QL clients and others may have non-college graduates working as operators in a call centre trained in a counselling protocol.

State Funders

The majority of QLs in the U.S. receive some state funding, usually from either Tobacco Master Settlement Agreement (MSA) funds, or from tobacco excise taxes. The state funding is administered through the state health departments and these are the organizations referred to by the term "funder organization" in the American QL partnerships. The state funders have contracts with service providers, sometimes the contracts are open for competitive bidding and sometimes they are not. For example, the UCSD service provider does not bid for their state-funding contract because they have an inter-agency agreement. The role of the state funder also varies across the QLs, with some being more involved in decision-making than others. The state funders usually are responsible for decision-making with respect to evaluation and many of the state funders contract with external organizations to evaluate their QL.

Canadian QL Funders

As previously mentioned, there is more variation in funders for the Canadian QLs. Alberta is the only QL that is funded and operated by its provincial Health Services. The QLs in British Columbia, Ontario, and Quebec are funded by various departments within the Ministry of Health. The Saskatchewan QL has two funders, the Canadian Cancer Society and the Heart and Stroke Foundation of Saskatchewan. Health Canada funds the QLs in Manitoba, Newfoundland, and New Brunswick. The Prince Edward Island (PEI) QL is funded by the Canadian Cancer Society, PEI division.

Evaluation Contractors

In the U.S., some of the QLs conduct evaluation in-house and others contract with an external evaluation organization. The reasons for contracting with an external evaluation entity are either: 1) not having the capacity and resources to evaluate in-house or 2) having a mandate from the state funder to use a specified external contractor. In many cases, the QL is assigned a third party evaluation contractor by the state. The QLs that are not assigned an evaluation contractor by the state may or may not conduct their own evaluation. Some of the QLs choose to hire external entities to conduct their evaluation or to do the database management because they do not have the internal capacity to conduct it themselves. In some cases, a QL funder will contract with the service provider to do both the service and the evaluation of the QL. There are however, examples of QLs that conduct all aspects of the evaluation in-house. Examples of the evaluation contractors are private consulting companies and evaluation units in universities.

In Canada, there is one primary evaluation entity that conducts evaluation for the majority of the Canadian QLs and this entity is located within the University of Waterloo. There is no similar primary evaluation entity in the U.S. with most of the QLs having different evaluation contractors. However, funding for evaluation of the QLs in Canada has been very unstable over the years due to changes in funding policies at the national and provincial level. At the time of conducting this study, government funding was not being provided to evaluate the QLs with the exception of Ontario. It was also unclear whether any of the other Canadian QLs were still collecting evaluation data at the time of this study, given the lack of funding for it.

Centers for Disease Control and Prevention

The Centers for Disease Control and Prevention (CDC) in the U.S. plays multiple roles in the QL community, including evidence source, funder and NAQC partner. One of CDC's primary responsibilities in general is the translation of evidence to practice and they also play this role for the QLs. In recent years, the CDC has taken an increasing role as a funder for the American QLs and as such has increasing power over decision-making. At the time of conducting this study (2010), the federal government in the U.S. gave $45,000,000 in funding to the QLs as part of the American Recovery and Reinvestment Act (ARRA). The CDC was responsible for administering this substantial amount of funding and under the instructions of the federal government, included strict regulations for accountability reporting in the QL contracts for the funding. As part of the ARRA funding mandates, the QLs will be required to provide the CDC with their data from the Minimal Data Standards (MDS), which will be explained in greater detail below in the 'Overview of the Innovation' section. CDC will be collecting the MDS data from the American QLs, entering it into a database, and making the data public. This will be the first time in QL history that evaluation data of any kind will be collected and put into an aggregate form (i.e., all QLs together). The other role that CDC plays is as a support to NAQC, they have a strong partnership and an explicit contract of support.

NAQC

In 2004 the North American QL Consortium (NAQC) was established with the primary role to promote evidence-based services across North American QLs. NAQC is a

non-profit organization, which receives funding from a variety of organizations including CDC in the U.S. and Robert Wood Johnson Foundation, to help ensure that NAQC can continue to support the QLs in North America. NAQC consists of a team of six staff members (including a director of research), a board of directors, and an advisory council. NACQ provides leadership and works to bring together diverse partners such as state and provincial QL administrators, QL service providers, researchers and national organizations in the U.S. and Canada. NAQC provides a forum for shared learning in hopes of improving the operations and effectiveness of the QLs. The QLs must pay for memberships to NAQC and the memberships provide a variety of benefits including access to forums and professional development activities. Not all members of the QL community are members of NAQC and the QLs' ability to provide memberships to all individuals in a QL vary. For example, not all QLs can afford to provide memberships to their service providers and evaluation contractors, so in some cases evaluation contractors are not part of NAQC and therefore do not have access to NAQC forums and member resources.

QL Funding

There is no standardization of funding across the QLs and there is significant variation in both the source and amount of funding. In 2009 the QL budgets ranged from $77,218 for the smallest budget to $17,869,238 for the largest. The range for the Canadian QL budgets is within the range just described but has not been reported publicly and therefore is not reported in this study. The reason why it has not been reported publicly is that because there are so few QLs it would be too easy to connect the

various budgets to the respective QLs. Similar to the American QLs, there is also a significant variation in budgets across the Canadian QLs.

The following information is specifically related to the American QLs. As previously mentioned, some of these QLs receive their funding through tobacco Master Settlement Agreement (MSA) dollars and others through tobacco taxes. Also, the way that tobacco settlement dollars were allocated and secured differs across states and subsequently impacts the QL funding in terms of stability and amount received. For example, one state had chosen to securitize payments of the MSA. The governor of the state and the legislature had made the decision to sell the MSA payments to a securitization company, who gave the state a lump sum, in exchange for the annual payments. Another state had put the MSA funds into an interest bearing account, which had then been used to supplement the budget during budget cuts. There is also significant variation in the degree of stability of the funding across QLs, which is partly attributed to the source of the funding and partly to the political context of the state. The different funds across states also come with different reporting requirements, as well as different regulations about what the funds can be used for. These huge variations in funding result in very different QL contexts that inevitably impact the implementation of evaluation innovations in the QLs.

The QL System

The overarching goal of this study is to explore implementation of the evaluation innovation in the QLs using a qualitative systems approach. As such, it is important to understand the complexity of the QL system, in order to think through how and why a

systems approach is necessary. It is impossible to describe every facet of the QL system because there are simply far too many components to include. However, in this section I describe some of the main components that will be relevant to the later findings presented that aid in understanding systems thinking in the QLs. Specifically, I will present some of the key system components that will be included in this study, including system levels, actors, organizations, and niches (Foster-Fishman, 2007; Janzen, 2007). It should be noted that there is no single correct way to describe the system as there are many alternative ways to arrange the components. The objective here is to illustrate how my approach is using a systems perspective by exploring different levels and parts of the system.

The system can be viewed as having five different levels: 1) federal, 2) state/provincial, 3) QL network, 4) individual QLs, and 5) individual organizations (see figure 2). The federal level is primarily involved in funding of the QLs and includes entities such as the CDC. The primary entities at the state/provincial level are the state health departments, which represent the funder organization in the QL partnership. The QL network level represents all of the QLs together, whereas the QL level represents individual QLs which usually consist of a service provider and a funder. The organizational level primarily includes the funders, service providers, and third party evaluation contractors. And lastly, the actors are the individuals within the different organizations.

Figure 2. Description of the QL System
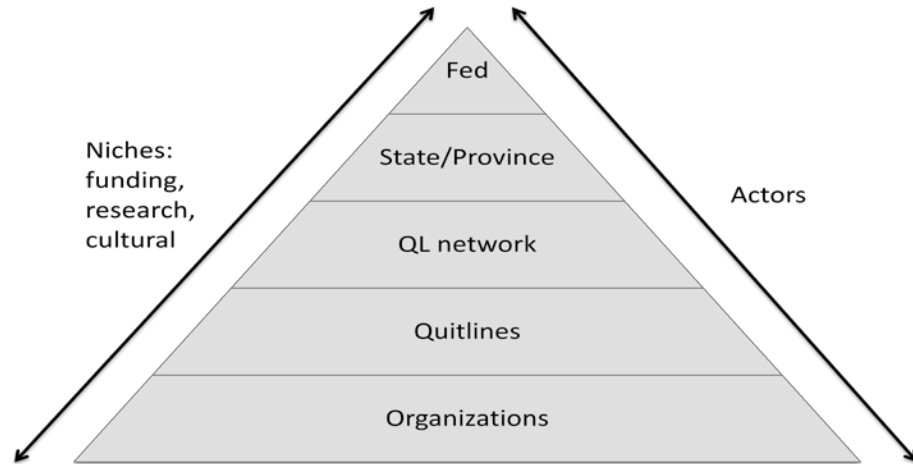
The actors can cross system levels and niches, although they usually do not cross between different organizations. Some of the key actors include: researchers, project officers (CDC), tobacco control managers (state) evaluators (contractor), directors, managers, and other QL staff. The different actors have different perspectives of the evaluation innovation that are influenced by their location in the system.

Niches generally have less tangible boundaries and cross multiple system levels. Examples of niches include funding, research, culture, and incentives. Some of the QL's service providers are part of a university and as a result are part of the research niche. The research niche has a pressure to evaluate and publish whereas QLs outside of the research niche do not necessarily have this pressure. The pressure and culture of the research niche creates a difference in priorities for the different QLs. These niches in the system can also be referred to as different system parts and are another element to consider in a systems approach to implementation.

Innovations in the QLs

There has been a plethora of tobacco related research conducted over the last few decades, much of it specifically looking at effective practices for promoting and supporting tobacco cessation. For example, research has demonstrated that smoking cessation rates are significantly improved if behavioural therapy and pharmacotherapy are used in conjunction, as opposed to either applied independently (Hughes, 1995). The evidence-based practices come from various sources including CDC, the Public Health Institute, and the research literature. The QLs also develop practice-based innovations that are shared between QLs. The practice-based innovations are often service norms or organizational policies that improve their services. For example, one of the innovations on the KIQNIC list is a faxed-based referral, which is a practice-based innovation or service norm, that has spread throughout the QL network as a suggested practice to improve reach in QLs (appendix A). Both the research-based and the practice-based innovations on the KIQNIC list have the potential to improve the QLs effectiveness. However, these innovations are not easily diffused and implemented throughout the network. There are many reasons why dissemination and implementation of innovations is difficult, many of which pertain to the network characteristics and structure.

Implementing Evidence-Based Innovations

The implementation of evidence-based innovations is difficult in most settings, but the difficulty is compounded in complex systems such as the QLs. The QL structure and characteristics add several dimensions of complexity to efforts to promote and

implement evidence-based innovations. For example, the QLs have a decentralized decision-making structure, which means that there is no single agent responsible for making and enforcing decisions on what innovations to implement. Each QL is a semi-autonomous entity, with a unique perspective and context, and each QL experiences different barriers to implementation, making it impossible to apply one standardized intervention for implementation. Furthermore, the diverse QL contexts create a tension between fidelity versus adaptation, where QLs must balance maintaining fidelity, with adapting innovations to be appropriate for a QL context. These factors are part of the reason why successful dissemination and implementation of innovations is challenging in the QL community. In response to the challenge, and also the potential benefit to achieving the implementation of evidence-based innovations, a research grant was funded to explore this issue in more depth.

Knowledge Integration in QLs: Networks that Improve Cessation (KIQNIC)

KIQNIC is a large research grant funded by the National Institutes of Health (NIH) in the U.S. A primary goal of KIQNIC is to assess how decision-making in QL organizations is a moderator for network characteristics and implementation outcomes. My dissertation project was developed based on my work with the KIQNIC project. I was invited to join the research team in the second year of the grant, during the instrument development phase, in order to assist with the implementation measurement piece of the study. The principal investigator of the grant is located at the University of Arizona, but the KIQNIC research team consists of researchers and NAQC members located

throughout the U.S. and Canada. One of the KIQNIC team members is the Research Director of NAQC and this individual plays a key mediator role between the KIQNIC study and the QL community. The Research Director was also my primary resource in developing my dissertation topic and research questions. I initially had numerous discussions with this individual about different research topics and their potential value to the QL community. The Research Director also provided guidance in determining which innovation to select from the KIQNIC survey to explore in more depth for this project. More details describing the innovation and the reason for choosing it are described in the following section. Another group of actors in the KIQNIC project is the workgroup, which consists of decision-makers in the QL community. The role of the workgroup is to provide guidance and feedback to the KIQNIC research team. To date the workgroup has primarily been involved in reviewing data collection instruments and providing feedback on preliminary results. The workgroup was also heavily involved in developing the list of 23 innovations (Appendix A).

KIQNIC is collecting quantitative data from the QL decision-makers through an annual survey conducted over three years. My dissertation project was conducted between the first and second wave of survey data collection, July 2009 and June 2010, respectively. There are three primary constructs to the KIQNIC survey, social networking, decision-making, and implementation of innovations. The social networking construct measures connections between different organizations in the network. The decision-making construct measures how decisions to adopt innovations are made in the QLs and what factors are considered when making decisions to adopt innovations. The

decision-making construct was developed using the theory of planned behaviour, which estimates an individual's intention to perform a particular behaviour (Ajzen, 1985).

The outcome measure for KIQNIC is the implementation of innovations measured using a summative score (range 0-23) of the 23 innovations (Appendix A). The list of innovations was developed by the KIQNIC workgroup and it includes both evidence-based (i.e., best practices) innovations from the literature, as well as practice-based innovations, which are referred to as 'service norms'. At the request of the KIQNIC workgroup, the innovations are referred to as 'new practices (innovations),' instead of, 'best practices (innovations)' by the KIQNIC research team. For the implementation section of the online survey, respondents were asked to report their QL's level of implementation for each of the 23 innovations on the list. Respondents were led through a skip pattern question that determined which stage of implementation the respondents' QL was in for each innovation (Appendix B).

Results from the baseline KIQNIC survey demonstrated significant inconsistency in responses to the implementation section for respondents from the same organizations. For example, four different respondents from the same organization selected four different options for implementation level of the same innovation (e.g., aware, decided not to implement; aware, in discussion; and fully implemented). The inconsistency in responses supported my assumption that there was a sufficient lack of understanding of the phenomenon being studied to warrant an additional qualitative interpretive study that provided insight into the 'black box' of implementation of innovations in the QLs. Furthermore, the KIQNIC study was limited in that it applied a positivist approach and used only quantitative data to study the implementation phenomena. Thus, the

overarching purpose of my dissertation study is to build upon the KIQNIC study by further exploring the implementation of innovations in the QLs using qualitative data and a systems approach.

Complexity theory is the systems approach used to guide the study, which incorporates a mix of both positivist and interpretivist principles. As previously stated, this study is different from the KIQNIC study in that the goal is to observe, reflect and describe the phenomenon in a theory driven way, as opposed to trying to identify specific determinants of implementation and predict the system (Greenhalgh et al., 2010). In this way, qualitative interpretivist findings can be added to the KIQNIC findings at a later date to create a more comprehensive understanding of the implementation phenomena. Although integration of the quantitative and qualitative findings is outside of the scope of this project, next steps after this study is completed can involve integration of the data in order to move the KIQNIC project towards a mixed-methods integrative approach.

Overview of the Innovation

I chose to explore one innovation in greater depth and selected an innovation from the list of 23 innovations in the implementation section of the KIQNIC survey (appendix A). The innovation selected from the list was: "to evaluate the effectiveness of the QL." This innovation was selected as the case based on discussions with my PhD committee and the Director of Research for NAQC. We decided that this innovation would be best to explore in more depth because it was of significant interest to NAQC and also because

of its level of complexity as an innovation. By this I mean that there were many factors involved in implementing this innovation.

The selected innovation (evaluating effectiveness of the QL) was of significant interest to NAQC members and one that NAQC had been investing effort and resources into implementing. One of NAQC's main priorities was, and is still, to create and implement a standardized system for evaluating effectiveness of the QLs. This goal was considered important because it will enable the QL community to evaluate services and produce data that can be used to answer decision-makers questions to inform practice and policy decisions. Therefore by selecting this innovation to study, the dissertation project had the potential to produce results that could be used by NAQC to support their future evaluation endeavors. My understanding of the importance of this innovation to NAQC is based on my review of materials on the NAQC website and my discussions with NAQC's Research Director. Of course, one of the limitations of this approach is that my understanding is based on information from a single person. I recognize that this innovation may not be of significant interest to all decision-makers in the QLs.

A major step in progressing to this desired outcome was made in 2005, with the completion of the Minimal Data Set (MDS), which is a standardized data collection system for outcome data such as reach and quit rates. At the time of this study (2010), all of the QLs had implemented the MDS. Although the majority of the QLs collect MDS data, there has been no aggregation of the data at the QL network level. As mentioned in the CDC section above, aggregation of the MDS data for all American QLs will happen for the first time in 2011, as part of a mandate for the ARRA funding. The CDC will be collecting MDS data from these QLs and entering it into a common database that will be

made public. This process will begin in 2011 and will be the first time that American QL data of any type will be collected and entered at the aggregate level for the purpose of cross-QL comparisons. The MDS provides information at the QL level on the reach and quit rates, but it does not provide data to evaluate effectiveness or compare different service options (e.g., four counseling calls versus five counseling calls). As such, the MDS does not achieve NAQC's ultimate goal, to collect standardized data from the QLs that can be used to create an evidence-based system and inform practice and policy decisions. Another effort to achieve this goal was made in 2007 when the National Cancer Institute (NCI) partnered with researchers from UCSD (California QL) to create a 'data warehouse,' a database that housed information on the QL services. These efforts were unsuccessful in part because many of the QLs were uneasy about providing information for the database for reasons not fully known to me.

No plans were described by any of the participants to include Canadian QL data in the CDC database. Also as previously mentioned, there is one primary evaluation contractor for the Canadian QLs located at the University of Waterloo. Although the data from these QLs is in a single database, there was no evidence to suggest that any cross-QL comparisons or aggregate analysis had ever been conducted using this data.

There are several characteristics of the selected innovation that make it challenging to study, but also valuable and applicable to many of the implementation challenges in public health. Similar to many of the innovations listed on the KIQNIC survey and to innovations in public health generally, the evaluation innovation lacks a clear definition and description of its components. Although evaluating effectiveness of QLs is designated a "best practice" by the CDC, there is no specific definition provided

in the KIQNIC survey (CDC, 2004). It is difficult to assess whether or not an innovation has been implemented successfully when that innovation has not been fully defined. Furthermore, the innovation of interest (evaluating effectiveness) is a small piece embedded within a much larger effort to create a system of evidence-based cessation services. Therefore, it is necessary to study the larger picture of systems change in order to understand implementation of this innovation.

As explained in the literature review different research designs and study methods provide various advantages and disadvantages to studying implementation problems. Studies such as Greenhalgh et al. (2010) have demonstrated the value of using qualitative data to answer questions on the meanings of innovations to stakeholders and the social and technical challenges of implementation. The recent studies that explicitly recognize the complexity of health service innovations, as well as the complexity of the systems, tend to favour mixed-methods that aim to observe and reflect, as opposed to quantify and replicate (Greenhalgh et al., 2010). The following chapter provides the detailed methods of the study that I used to address my research questions.

Chapter 4. Methods

Chapter Overview

The overarching goal of my study was to explore the implementation of the evaluation innovation in the QLs using a systems approach. I used a combination of inductive and deductive techniques to explore the phenomenon (Silverman, 2000). In this case, the phenomenon of interest was the implementation of the innovation, evaluating effectiveness in the QLs. To do this, I conducted 19 semi-structured interviews with decision-makers in the QL community and analyzed the interview transcripts using a thematic analysis (Braun & Clark, 2006).

This chapter provides a detailed description of my research methods, starting with a description of my researcher perspective, location, and role. I then provide details on ethical issues including obtaining consent, confidentiality, and risk status of the participant population. In the following section, I provide details describing the participant sample and the recruitment strategies used. Next, I outline the interview process including the activities that occurred both during and after the interviews with participants. Lastly, the methods are described for the data analysis process, specifically the thematic analysis methodology (Braun & Clark, 2006). The final section of the chapter addresses the quality of the study and the analysis.

Researcher Perspective, Location, & Role

I mentioned briefly in chapter one that I started working on the KIQNIC project approximately six-months prior to starting my dissertation and that it was my interest in the implementation phenomenon that was the impetus for my research project. I had

initially been recruited to the KIQNIC project specifically to provide expertise and assist with measuring implementation. I noticed quickly that my perspective and epistemological grounding was different than some of the other researchers on the project and also different from the conceptual framing of the project. For example, one of the primary goals was to assess decision-making in the QLs as a mediator variable between network characteristics and implementation outcomes. For the numerous reasons highlighted in chapter one, I struggled with being able to identify a single mediator variable across such diverse QLs and innovations. Furthermore, the decision-making construct was being measured by the theory of planned behaviour, which is an individual behaviour change model (Ajzen, 1985) and from my perspective does not seem appropriate for assessing change in a complex system. The impetus for this project was the desire to explore the implementation phenomenon from a different research paradigm and epistemological grounding.

My epistemological grounding is a mix of approaches and methods, as the majority of my formal academic training has been in quantitative methods using a linear reductionist approach. Until my PhD program, the models of change that I used and was familiar with were primarily individual change models. However, I have also worked on several qualitative projects, starting with a nursing project during my Bachelors degree. I would not consider myself either a quantitative or qualitative researcher, but I have sufficient knowledge to work with both.

I had initially considered including both qualitative and quantitative data in my dissertation but decided to focus on just the qualitative approach for several reasons. The first is that I view this methodology as the most necessary for addressing my research

questions. The second reason is that I wanted to demonstrate the value in using a qualitative approach to study implementation phenomena. As previously mentioned, the KIQNIC study was already investing significant resources into the quantitative approach and there was no discussion of the need to add a qualitative component. I saw this as an opportunity to demonstrate the value in using a qualitative interpretivist approach to studying implementation and decided to focus entirely on this approach.

Furthermore, during my PhD program I have been studying systems thinking, particularly complexity science and have moved towards organizational and systems change models as opposed to individual change models. I also have a strong background in evaluation, with a personal bias towards conducting utilization-focused evaluation. This means that I consider the goal of evaluation to be utilization of results, which is achieved through collaborative efforts with stakeholders of the evaluation (Patton, 2008). Although this may sound obvious, the traditional approach to evaluation would suggest that the evaluator remain separate and objective from the 'evaluand' (evaluation term meaning target of the evaluation) and the stakeholders and does not emphasize utilization (Patton, 2002; Patton, 2008). I mention this because although the study is not an evaluation study per se, I am looking at an evaluation innovation and my approach to conducting evaluation will influence my interpretation of the data.

Because I had only been involved with the KIQNIC project for approximately six-months prior to starting my project, my knowledge of the QL network was limited. I was also an 'outsider' to the QL community and my position on the KIQNIC project was not sufficient for connecting with the QLs as this project was also outside the QL community. In order to get more insight into the QLs while developing my project, I

communicated extensively with the Research Director of NAQC who was also a member of the KIQNIC research team. She and I had numerous discussions about the focus of my project including what methods would be best to use and what innovation would be best to study. By 'best methods' I mean what would be acceptable to the participants and also most likely to provide useful information to NAQC. I considered these practical issues in conjunction with the potential for the study results to make a theoretical contribution to the literature. This individual was also involved in the interview guide development process and joined my PhD committee in reviewing and providing feedback on my early versions of the guide. She was particularly helpful with the language in the guide and ensuring that it was appropriate for the respondents. She was also the first person I interviewed for the study, as described in a later section of this chapter titled 'participant sample.'

The dynamics between the participants and myself was different from any other project that I had worked on. This is because unlike past participant samples, this sample consisted of all professionals with high education levels (e.g., lowest was a bachelors degree). I did not feel the same potential power issues that I have experienced on other community-based research projects. This is not to suggest that there were no power dynamics between the participants and myself. The participants in many ways felt like my peers, or in some cases fellow academics and researchers. The majority had a graduate degree and were sympathetic to and supportive of graduate research projects.

Also, because the interviews were conducted over the telephone both parties were blinded to each other's age, gender and race/ethnicity. However, in some cases age approximation could be deduced and in the majority of the cases respondents were most

likely older than me. There were two participants that were clearly younger than the majority of the participants and closer to my age. Gender was also deduced and is included in the section below on interviewee sample.

Ethics

The study was approved by the University of British Columbia's Behavioural Research Ethics Board. All participants were emailed an electronic version of the consent form during the initial email communication. They were also mailed a hardcopy of the consent form prior to the interview, along with a pre-stamped and addressed envelope to return the signed consent form back to me. The population of decision-makers was considered a "minimal risk" population as they were not vulnerable and the risk involved was part of their regular job function. It is possible that they may have felt coerced into participating or feared that refusal to participate in the study could jeopardize their job or relationship with the QL community. However, at no time did I sense the participants were participating out of coercion.

Although the participants were a minimal risk population, I was aware that some of the information that they shared was sensitive and could pose a risk if it was made public and linked to them. A few of the participants expressed some hesitation and concern regarding the information that they provided. I attempted to minimize the risk to participants in several ways. First, all participants were given the opportunity to review their interview transcript after the interview and were allowed to edit the transcript by deleting information, correcting information, and inserting additional information. Further details regarding the number of transcripts and the information that was edited

can be found in the post interview section below. A second strategy for protecting the participants was that I assigned the transcripts codes instead of putting interviewee names on the transcripts (e.g., KI-3 for the third key informant interview conducted). Lastly, all interviewees were invited to participate in a webinar session/focus group to discuss the results of the study. The details regarding the focus group process and participation are provided later in this chapter.

Despite precautions taken to protect the confidentiality and anonymity of the participants there is still a risk that participants may be identified based on their quotations. It is also possible that by describing sample characteristics, others could incorrectly or correctly deduce whom the individuals were who participated in the interviews. The reason for providing participants with an opportunity to review the results was to ensure that nothing is published that will put them at risk or that they are uncomfortable with.

Interviewee Sample

In total, nineteen interviews were completed including two with NAQC, eight with service providers, seven with funders, one with CDC, and two with third party evaluation contractors (see table 1). Descriptions of these different organizations and actors are provided in chapter three. As described in chapter three, the system can be viewed as having different levels, including organization, state/province and federal, as well as different niches including research and funding. Together this sample of interviewees represents a mix of organizations, levels and niches in the system, and together they provide a range of perspectives on the implementation phenomenon.

Table 1. Interviewee Organization Sample Summary

NAQC	Service Providers	State Funders	CDC	Evaluation contractors
2*	8	7	1	2

* One of the NAQC interviewees was also an evaluation contractor but is placed in the NAQC category because this is his/her current and dominant role

Table two provides a more detailed summary of the sample characteristics. The service provider and funder interviews together represented 24 of the 62 QLs. Two of the interviewees were from Canada and the remainder was from the U.S. The two Canadians that were interviewed were highly knowledgeable of the Canadian QLs. Interview number 17 was conducted with two participants from the same organization together. Both of these individuals had been recruited through the KIQNIC survey (described below) and emailed recruitment letters separately. They responded to my email collectively and requested to be interviewed together. They both had specific knowledge to answer different parts of the interview and together were able to address all of the topics in the interview guide (described later).

Table 2. Sample Characteristics by Interviewee

Interview	Recruitment Method	Organization	Country	Gender	Educational Level	Position
KI-1	Elite interview	NAQC	US	F	Doctoral	Director of research
KI-2	Snowball	CDC	US	F	Doctoral	Senior scientific advisor
KI-3	KIQNIC survey	Service provider (university)	US	F	Doctoral	Assistant professor
KI-4	Snowball	Service provider (university)	US	F	Masters	Project manager
KI-5	KIQNIC survey	Service provider	US	F	Masters	Director
KI-6	KIQNIC survey	Service provider (university)	US	F	Masters	Manager
KI-7	KIQNIC survey	Service provider	US	M	Doctoral	Director
KI-8	KIQNIC survey	Service provider	US	M	Bachelors	Program manager
KI-9	KIQNIC survey	Funder/health department	US	F	Masters	Cessation coordinator
KI-10	KIQNIC survey	Service provider (university)	US	M	Bachelors	Program director
KI-11	Snowball	NAQC & evaluation contractor	Canada	F	Doctoral	Evaluator & associate professor
KI-12	KIQNIC survey	Funder/health department	US	M	Bachelors	Project director
KI-13	Snowball	Evaluation contractor	US	F	Doctoral	Program evaluator
KI-14	Snowball	Service provider (university)	US	M	Masters	Director
KI-15	KIQNIC survey	Funder/health department	US	M	Masters	Tobacco cessation coordinator
KI-16	KIQNIC survey	Funder/health department	US	F	Bachelors	Tobacco treatment specialist
KI-17	KIQNIC survey	Funder/health Department	US	F F	1) Bachelors 2) Masters	1) Tobacco cessation specialist 2) Director
KI-18	Snowball	Evaluation contractor	Canada	F	Masters	Manager
KI-19	KIQNIC survey	Funder/health department	US	M	Masters	Public health specialist

Recruitment & Data Collection

The following section provides details for the data collection process including the recruitment strategies used, the interview process, and the post interview activities. Participants were recruited via two different recruitment strategies: the KIQNIC survey (purposive sampling) and snowball sampling. The total number of interviewees was determined based on feasibility, data saturation, and number of individuals available to interview. Data saturation "refers to the point at which an investigator has obtained sufficient data to feel confident that an understanding of the phenomenon has been achieved" (Corring, 2004, p. 70). Although I felt that an understanding of the phenomenon had in fact been achieved, I do not believe that data saturation had been reached. Based on the systems theory presented in this study, there are numerous diverse perspectives across the QL system. By no means can 20 decision-makers provide the necessary information to reach saturation given the number of different perspectives in the system. For this reason, it seems improbable that data saturation can ever be reached in a system study such as this.

The final recruitment numbers from all sources is listed in table 2. Recruitment emails and letters were tailored as much as possible for each interviewee. This was partly to create an environment that suggested to the interviewees that they were not just another case in a sample of many and that their perspectives were valued. The sampling process was a combination of snowball and purposive sampling to recruit expert interviewees. An expert interviewee is an individual who has expertise on the particular topic of interest (Boeije, 2010). The first interview was an elite interview conducted with the Research Director of NAQC. An elite interview is slightly different from an expert

interview as it refers to someone that is either "high-ranking or well-known" and in this case the Research Director was well known in the QLs (Boeije, 2010, p. 63). The purpose of interviewing her first was to collect contextual information to help frame the problem of interest and to learn about the history of QLs and more specifically, evaluation in the QLs. A limitation of this approach is that the problem was framed initially from this elite interview. However, obtaining her assistance with the project was both necessary and invaluable as it enabled me to make the results more relevant to the participants and also provided me with the knowledge I needed to sufficiently understand the practice and context being studied to conduct the other interviews.

KIQNIC Survey Recruitment

The purposive sampling technique was done using the KIQNIC survey. Purposive sampling is a non-probability sampling strategy where "each sample element is selected for a purpose, usually because of the unique position of the sample elements" (Schutt, 2006, p. 155). According to Macnee and McCabe (2008, p. 121) "a purposive sample consists of participants who are intentionally or purposefully selected because they have certain characteristics related to the purpose of the research." In this approach, people who are knowledgeable of the targeted issue and represent specific perspectives are selected, but they are not intended to represent the larger population. According to Schutt (2006), the goal is to get adequate representation of the sample and situation and to sample until you have achieved saturation and completeness, meaning that no new information is being collected and an overall sense of the issue has been achieved. Macnee and McCabe (2008) suggest that a key strength of this approach is that the

researcher can obtain rich data by carefully choosing individuals to interview who are knowledgeable of the topic being studied. However, these authors also suggest that a major limitation is that a researcher may prematurely focus the data collection on a specific perspective or element and miss other broader information.

For the purposive sampling, a question was added to the first wave of the KIQNIC survey conducted in July-August 2009. The principle investigator of KIQNIC, who is also one of my doctoral committee members, gave his approval for me to add the recruitment question to the survey. The question asked if the respondent would be interested in participating in a dissertation project to explore innovation implementation in the QLs. After the data from the first survey wave had been collected for the KIQNIC project all project team members were provided with an Excel file with the KIQNIC data in it. I extracted the cases that responded, "yes-interested," to the recruitment question on the KIQNIC survey and put them into a tentative sample file. In addition to the recruitment question, I also had the following variables in the file: the QL and organization they worked for, the state/province of the QL, and their responses for the stage of implementation for the evaluation innovation (Appendix B).

There were a total of 276 decision-makers recruited for the KIQNIC survey and as illustrated in the consort table (see figure 3), 192 of them completed the survey and responded to the dissertation recruitment question on the survey. Of the 192 who responded 37 agreed to be interviewed, and these 37 individuals together completed 59 KIQNIC surveys, representing 50 QLs. As described in the context chapter, a QL consists of a funder organization and a service provider organization, and in some cases, there can be one service provider for multiple QLs. For example, the American Cancer Society

(ACS) was the service provider for 10 of the QLs in the network. One of the decision-makers for the ACS completed 10 versions of the KIQNIC survey, one for each QL that they provided service to.

Figure 3. Consort Table

Initial number of decision makers recruited for KIQNIC survey	**276**			
Number that responded to the recruitment question on the KIQNIC survey	**192** Responded	**84 N/A**		
Response to recruitment survey question	**137** No	**37** Yes		
Level of implementation (N= QLs)	**1** Not aware / **1** Aware & in discussion / **2** Low Level	**1** High level	**45** Fully	
Number of potential interviewees			**1**	**32**
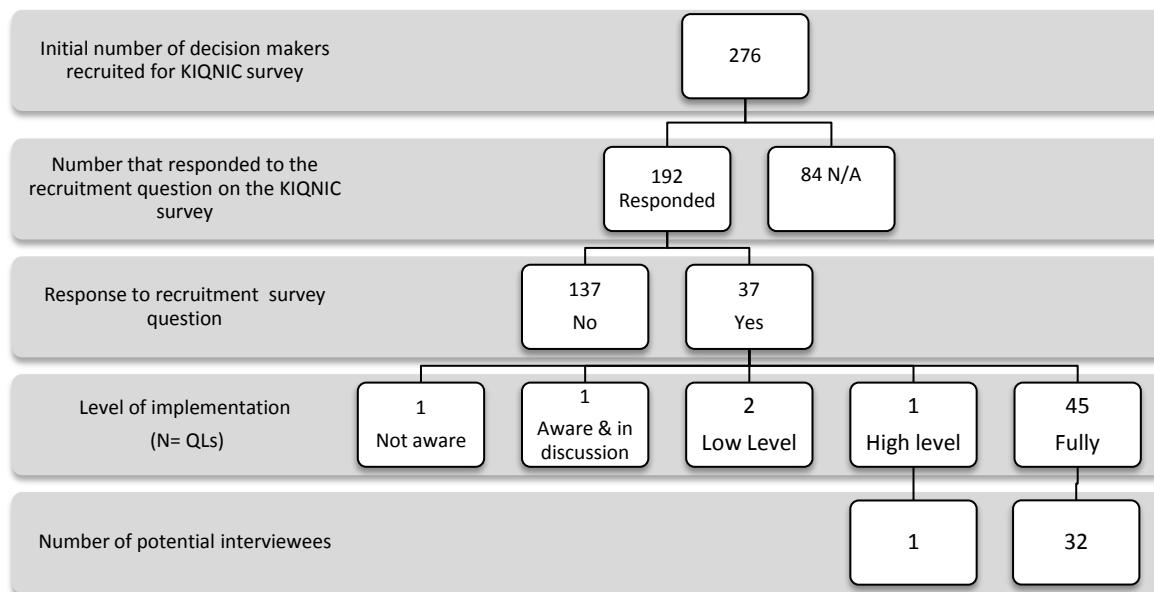
I realize that 37 out of a potential 192 respondents is a low response rate and it would have been ideal to follow-up with those who chose not to participate. However, I could not follow-up with these 155 decision-makers because I did not have their contact information since only those that responded 'yes' to the recruitment question provided their contact emails. Based on lessons learned from the KIQNIC study, I suspect that part of the reason for the low response rate is that this population already participates in a considerable amount of research and may be suffering from research fatigue. Given that they were being recruited through a research survey that was fairly lengthy, they were

probably not inclined to volunteer for yet another study. Also anecdotally, I have noticed a lack of enthusiasm for research in this population because they do not always see the benefit of it.

Of the 50 QLs represented, four were excluded due to having responded on the KIQNIC survey that their QL was at a low level of implementation for the innovation (i.e., evaluation of effectiveness). Only respondents who responded with either 'fully implemented' or 'high level of implementation' on the KIQNIC implementation question for the case innovation were included in the final sample population. This decision was made based on the advice of my PhD committee with the intention to focus the study. The rationale for this decision was that it was already necessary to have multiple interview scripts due to variation in the interviewees (e.g., funder versus service provider) and having QLs at too many levels of implementation would have made it difficult to draw comparisons across interviews and QLs. Of the remaining 46 QLs, one responded that they were at a 'high level' of implementation and the others reported to be 'fully implemented.' There were a total of 33 decision-makers who completed the KIQNIC survey for the sample of 46 QLs and these were considered the final sample of potential interviewees from the KIQNIC survey.

Nine of the 33 key informants were from service provider organizations (defined in chapter three), which together represented 21 of the QLs. One of these individuals did not complete the contact information on the KIQNIC survey question and therefore could not be reached. In addition, between the time of the KIQNIC survey and interviews for this project being conducted, two of the 33 key informants left the QLs. One of the

decision-makers worked for the American Cancer Society (ACS), which ended its service provider contracts in December 2009 and the majority of its employees were let go.

All the individuals in the initial KIQNIC recruitment sample were emailed recruitment emails (Appendix C). The invites were sent out in waves over a three-month period. The order of the invites was made based on characteristics of the individuals and information learned from previous interviews. I selected the first five individuals based on my perceived expectation that they would be able to provide information to inform the history and context of the innovation and the QLs. My objective was to get a better understanding of the innovation and context that I was exploring.

There were 19 individuals identified through the KIQNIC survey who were sent recruitment emails but were not interviewed. Of the 19, only one individual responded to confirm that he/she was not interested in participating. One individual forwarded the recruitment email to a staff member. Two emails bounced back as incorrect email addresses. The remaining 15 individuals did not respond to the recruitment emails. There was also an interview conducted with two individuals together from the same organization. Both of these individuals had been identified through the KIQNIC survey and requested to be interviewed together when they responded to my recruitment email. A second follow-up email was sent approximately one month after the first recruitment email was sent. In total, thirteen of the original 32 individuals identified through the KIQNIC survey were interviewed. And as described above, two of those thirteen were interviewed together, making a total of twelve interviews conducted through KIQNIC survey recruitment.

Snowball Sampling Recruitment

"A sampling procedure may be defined as snowball sampling when the researcher accesses informants through contact information that is provided by other informants" (Noy, 2008, p. 329). Snowball sampling is the most widely used sampling method in qualitative research and there are many advantages and also some disadvantages to this recruitment approach (Noy, 2008). An advantage suggested by Offredy and Vickers (2010, p. 139) that is relevant to this study is that "it can be an effective strategy for the identification of participants who are able to provide important insights, knowledge, understanding and information about the experience or event that is the focus of the research". However, there are also disadvantages to this approach as in my study informants often suggested other informants that they had a close working relationship and with whom they may have shared similar perspectives. Since snowball sampling "relies on and partakes in the dynamics of natural and organic social networks," it was difficult to recruit from outside the informants' social network using this approach (Noy, 2008, p. 329). It is possible that individuals in the same network may have similar experiences and values and I need to be careful not to assume that they represent all individuals in the network.

Each interviewee was asked for recommendations for individuals to interview who were knowledgeable of the topic. If they had someone to suggest, then I would send a recruitment email to the interviewee to forward on to the potential interviewee identified in the interview. The potential interviewee would then contact me via email to schedule an interview. Attempts to recruit individuals recommended for an interview by another interviewee were made for all those recommended. All individuals who responded

positively to the recruitment email were scheduled for an interview. In total, six

individuals identified through snowball sampling were successfully interviewed. Only

one of these individuals was nominated by the initial elite interview.

Another challenge to the snowball sampling approach was that I had the initial

informant send a recruitment email (provided by me) to the snowball informant and left

the onus of contacting me with the snowball informant. This approach proved to be

problematic as only four of the 16 individuals sent snowball recruitment emails

responded to the email and contacted me to participate. Also as previously mentioned, the

decision-makers are asked to participate in a lot of research and it is possible that they

opted not to participate because of research fatigue and/or being too busy with their jobs.

In the remaining two cases, I directly contacted individuals who were identified

with snowball sampling, but did not respond to a forwarded recruitment email. In both

cases, I was aware of the individual and their role in the QLs through documents

published on the NAQC website. Both of these individuals were viewed as critical to

interview because they were highly involved with evaluation activities in the QLs. I used

email addresses that were posted publicly on the Internet to send recruitment email letters

to them and both individuals responded positively and were subsequently interviewed for

the study.

Interviews

The primary source of data collection was interviews with decision-makers from

the QL network representing different parts of the system, including organizations, levels

and niches of the system. Fontana and Frey (2005, p. 697) suggest that "interviewing is

one of the most common and powerful ways in which we try to understand our fellow humans." Interviews can take multiple forms such as structured, unstructured, and semi-structured interviews. Fontana and Frey (2005) also suggest that interviewing is a subjective process rather than a neutral data-gathering tool and that the data obtained from interviews is a mutually developed story that is bounded in history, politics, and culture. Interviews for this study were semi-structured to ensure that key information was covered and to provide flexibility to allow participants to provide additional insights unknown to the interviewer and to pursue some tangential matters (Hakim, 2000).

Participants were scheduled for interviews after they had received and signed the informed consent. The interview was scheduled via email communication at a time that was convenient for the participant. The duration of the interviews ranged between 45-80 minutes in length and were conducted via Skype, an online communication tool, and recorded using 'Call Recorder for Skype.' I called the participants from my computer using Skype to their telephone so it was not possible to use two-way video conferencing. Although video would have been nice so that we could have visual contact, it would have required that all of the participants have computers with Skype software, microphone, and a video camera. Skype video conferencing is also problematic as it reduces the quality of the call and requires a strong Internet connection. For these reasons I decided to have the interviewees participate in the interview from their phone and did not use the Skype video option. With this format the quality of the calls was very good with only minimal incidents of not being able hear each other.

A limitation of the study is that I did not ask participants to complete a demographic survey prior to the interview. However, the first section of the interview

was descriptive questions (Neuman, 2006) including questions related their role and experience with the QLs. I also obtained their educational information via email signature and other documents, as well as made inferences regarding their gender.

At the time of the interview, I called the participant and began the interview, by asking if he/she had any questions about the study or the consent form. I answered any questions the participant had and informed the participant that the interview would be recorded. In hindsight, I should have asked for the participants' permission to record instead of informing them that they would be recorded, but none of the participants objected to being recorded. Next, I provided a brief overview of the study and then proceeded with the interview questions.

I had developed an interview guide for different actor groups in the system based on the implementation and systems literature, as well as the research questions (Appendix D). I shared the interview guides with the Research Director of NAQC and my committee members and revised them based on the feedback provided. The interview scripts had to be tailored for different actor groups because the same questions would not be appropriate for everyone, as they varied on several key characteristics. For example, some individuals represented service providers and others funders. Some worked with only one QL, whereas others worked with multiple QLs. An example of how I tailored the interview script is that I asked interviewees from service providers and health departments (QL organizations) about their QL funding source, but I did not ask this question of the CDC interviewee because it would not make sense to do so.

I also tailored the interview questions based on the individual interviewee's characteristics. Tailoring interview guides for specific respondents and situations is a

typical characteristic of field interviews (Neuman, 2006). For example, if the interviewee had a lot of experience and involvement in the broader QL network, then I would pursue questions about the QL network. However, if the interviewee's experiences were limited to their individual QL, then I would not ask questions about broader network issues. Although there was variation in which questions were asked and how they were asked exactly, there were consistent topic sections across all of the interviews. I had topic issues that I wanted to cover and had developed a list of questions for each issue or topic.

Below is a list of the topics in the interview script and samples of the questions under each topic issue (see table 3). I did not link exact interview questions with a specific research question because information from a single question could provide insights into both specific factors influencing implementation (question 1) and also patterns in the system (question 2) when combined with other information. For the most part, the background topic area provided background information on the topic and the other questions addressed research questions one and two. Although this distinction was sometimes blurred depending on the information provided by the interviewees. The third research question was not linked specifically to any of the interview questions and instead was addressed through the interpretation and discussion of the findings.

Table 3. Interview Questions & Topic Areas

Topic Area	Sample Interview Questions
Background	Can you describe your organization and QL? What is your role in the QL? What is your experience with evaluation?
Funding	What are the sources of funding for your QL? How stable has this funding been over the years of operation? Are there any expectations or mandates for evaluation with this funding?
Interviewees' perspective of the innovation	What does it mean to evaluate effectiveness? What is the goal of evaluating effectiveness? What does it mean to "fully implement" this innovation?
Description of evaluation in their QL	Who conducts the evaluation for your QL? How are decisions about evaluation made in your QL? What type of evaluation is conducted at your QL and organization? How are the evaluation results used? What are barriers to your QL evaluating effectiveness? Why did your QL start conducting evaluations?
QL changes over time	How has evaluation in your QL changed over time? Why did these changes occur?
QL Relationships	Does your QL share information about evaluation with other QLs? How would you describe the relationship between your organization and your partner organizations (e.g., service provider, evaluation contractor)?
Evaluation in the broader network	How has evaluation in the QL network changed over time? How can evaluation in the QL network be improved? What are some of the differences across QLs that influence evaluation?

The interview was characteristic of a field interview and not a survey interview in that I posed open-ended questions, allowed joint control over the pace and direction of the interview between myself and the respondent, and showed interest in responses and encouraged elaboration (Neuman, 2006). Similar to Chase's (2003) instructions for conducting interviews, I developed a detailed script and spent significant time revising it and working on the flow of the questions, but then with the exception of the first section of the interview guide, I did not follow the order exactly. Instead, I responded to the natural flow of the conversation and worked with the information being provided by the

interviewee, as opposed to trying to force him/her into following the order of questions on the pre-structured interview script. Respondents would often inadvertently address topics from sections later in the interview guide while answering an open-ended question near the beginning of the guide. I listened to the respondents and made notes of things to follow-up on with them later in the interview once I had an opportunity to interject without interrupting the respondent's speech or thoughts.

I allowed the respondent to jointly manage the pace and direction of the interview, although making sure that all topics in the interview guide were addressed (Neuman, 2006). There were very few times when it was necessary to bring the interviewee back to the interview topic and the majority of them were talkative and required little or no prompting to share information. There were two interviewees who answered questions with a couple of sentences and required more coaxing to elaborate on their responses. Out of respect for the participants' time I attempted to be diligent with keeping to the one-hour time limit. In a couple of cases the participants insisted on continuing past the hour mark, providing assurance that it was not problem for them. Overall the participants' response to the study was very positive. Several participants expressed a desire to contribute to understanding the problem under examination. There was also very little difficulty with the scheduling of the interviews as only one had to be rescheduled.

I also adjusted my language to be appropriate for different respondents, which is typical of a semi-structured interview process and is a way to "approach the world from the subject's perspective" (Berg, 2009, p. 107). For example, when interviewing the two respondents from evaluation contracting organizations and the two NAQC interviewees (table 1), I used specific evaluation terminology such as "process" and "outcome

evaluation," in response to their usage of this terminology in the interview. I did not use such formal evaluation terminology with all participants as many of them would not be familiar with this terminology and did not use it themselves.

There were several incidents where interviewees hesitated in stating a thought, mentioning that he/she maybe should not share this information or were concerned they were sharing too much information with an outsider. In these cases, I reassured them that they would be given an opportunity to review the transcript and to remove any statements that they were uncomfortable with after the interview. Five of interviewees made changes to their transcripts, which is described in greater detail in the following section titled "post-interview".

A limitation of the study was that the interviews were conducted over the phone and "telephone interviews lack face-to-face nonverbal cues that researchers use to pace their interviews and to determine the direction to move in" (Berg, 2009, p. 122). Although all of the participants were in their offices for the interview and there was no background noise or distractions during any of the interviews, I still felt a little bit disconnected from them. I could not see the body language or pick up on any nonverbal cues and this made it more difficult to connect with them. However, telephone interviews were necessary due to the geographical constraints on the study.

Post Interview Procedure

Interviewees were given a five-dollar gift certificate to Starbucks in return for their participation. The gift certificate was mailed directly to them immediately upon completion of the interview. The interview recordings were converted from movie files

into MP3 files and saved under a file name that corresponded with the interview number (e.g., KI-3). The MP3 files were sent to a professional transcription company where they were transcribed verbatim and saved as a Word document with the same file name as the corresponding MP3 file. The primary advantage of using a professional transcription company as opposed to transcribing them myself is that it required less of my time. The disadvantages of this approach is that it reduced my familiarity with the data, increased risks with respect to maintaining confidentiality of the data, and increased risk of errors in the transcription. I minimized the last risk by checking the transcripts for accuracy before being saved as the final transcript document. This activity also helped to familiarize myself with the data. I reviewed the transcripts within two to seven days after completing the interviews, cross-referencing potential errors with the audio recording of the interview and correcting any errors in the transcript. I minimized the risk to confidentiality by not including their name on the file, although their first name is stated in the interview.

Participants were contacted once the transcript was ready and asked if they wanted to review it. If they did, it was emailed to them in Microsoft Word. They were also asked to make any requests for changes to the transcript within two weeks and to highlight requested changes by using the track changes feature when reviewing the document. Interviewees were invited to edit the transcript by deleting information, correcting information, and inserting additional information. Twelve of the interviewees opted to review their transcript and five of those made changes to it. The majority of the changes were additions of information that the participant thought of after the interview and a few removed information they were uncomfortable sharing. All changes were tracked so that

I could see and document the changes. I considered all the deletions of information to be

benign and non-threatening to the integrity of the data in any way. For example, one

participant deleted a comment where she/he had remarked on a colleague's age, in that

they "had been around forever." Final transcripts, with participant modifications were

saved and then converted into text files. The text files were uploaded into NVIVO

version 8 (a qualitative software analysis program) for analysis.

Data Analysis & Interpretation

The data analysis and interpretation of the data was conducted through an iterative

process of coding, decision-making, and interpreting. I used the thematic analysis

approach described by Braun and Clark (2006) which "is a method for identifying,

analysis and reporting patterns (themes) within data. It minimally organizes and describes

your data set in (rich) detail" (Braun & Clark, 2006, p. 79). I used this approach because

it was flexible and it allowed me to see patterns in my data (Braun & Clark, 2006). In this

approach, analysis of free flowing text is done in large blocks or chunks of texts (Ryan &

Bernard, 2000). The unit of analysis is the chunks of text that reflect a single theme

which is a "...patterned response or meaning within the dataset" (Braun & Clark, 2006, p.

82). "Themes are abstract (and often fuzzy) constructs that investigators identify before,

during, and after data collection" (Ryan & Bernard, 2000, p. 780). Sources of themes

include the literature, the scientists' own experiences with the subject matter, and the text

itself. Braun and Clark (2006) provided six phases for thematic analysis which I generally

followed: 1) familiarize yourself with your data, 2) generate initial codes, 3) search for

themes, 4) review themes, 5) define and name themes, and 6) produce the report.

Coding

Coding refers to the analysis of the transcripts, it is a process of dissecting and differentiating the data in a meaningful way, while also combining and keeping the relations between the parts intact while reflecting on the information provided (Flick, 2006). Coding is a data reduction process (Flick, 2006) and interpretation occurs throughout the coding process (Ryan & Bernard, 2000).

The coding was done in chunks of data large enough to capture the essence of the interviewee's point. In some cases, a chunk of data may have had multiple codes which helped to identify overlap between different coding categories. The transcripts were read and coded in NVIVO. I read the first five transcripts several times, attempting to code them while keeping in mind the systems literature. I tried different approaches to coding the data that would capture the information being communicated by the participants from a systems approach.

At one point during the process of developing themes the information coming from the data reminded me of a systems change framework that I had read previously. The systems change framework developed by Foster-Fishman et al. (2007) provides four broad system element categories that they consider essential for achieving systems change: system norms, system resources, system regulations, and system operations. The framework elements are sufficiently broad that they cannot provide specific information for different systems and innovations, however the framework does provide a useful heuristic to guide systems thinking. I decided to use the framework as a heuristic and considered this a strength of the study as it allowed me to contribute to building on other

systems change work in the literature. This framework works well because it was developed based on the organizational change and systems thinking literature, which is consistent with the theoretical approach of my study. This is the deductive part of my analysis. I also reflected on a list of questions provided by Foster-Fishman et al. (2007) to help think through the relevance of the various system elements to systems change. Again, the questions are broad and not appropriate for all systems but are useful for prompting reflexive thinking during the analysis. I used both inductive and deductive approaches to identify the themes and subthemes. For example, the Foster-Fishman et al. (2007) framework presents three types of system resources - human, social, and economic. I used these resources as themes because they corresponded to what interviewees had to say and then inductively drew further subthemes from the data such as "designated staff" under the human resources theme. In this way, the analysis helped to flesh out the original framework based on the case under examination.

I developed the final codebook through an iterative process of reading, coding, revising, and refining the codes. After I had coded a group of transcripts, I would create a coding report through NVIVO which consisted of all the coded information for each theme. The coded information was copied and pasted into a word document that had all the themes and subthemes with the coded information under them. I reviewed the document multiple times and re-organized the information as patterns were identified. I adjusted and refined the codes along the way, as I read additional transcripts, and other trends and patterns started to emerge. The codebook was modified and revised in this iterative manner until I had an empirically based list of final codes that was a good fit for the data. While developing the codebook I also developed and refined definitions for the

codes. The final codebook and definitions were then used to code all 19 transcripts (see appendix E). I then drafted the results chapters and got feedback from my committee members on the findings and made revisions to the write-up based on their feedback. Four months later I analyzed the transcripts a second time using the codebook and with consideration of the revisions that had been made to the write-up of the results. I compared the results of the second analysis to the first and found that results were consistent across both analyses. There were slight differences in the tallies for the number of participants that had identified each theme. This is primarily because the transcripts had a lot of information and not all of it was captured in the coding both times. Overall, the coding results were very consistent across both of the coding analyses.

The data analysis process started after the fifth interview had been conducted and ended eight months later. As I read through the transcripts and coded, I also documented my thoughts and reflections in a Word document. I found that ideas about systems concepts came to me while I read the transcripts and I documented these analytical thoughts so that I could use them to guide the later interpretation.

Interpretation & Writing Up the Results for the Thematic Analysis

The interpreting and sense-making aspect of the thematic analysis was a long and reflective process. I spent a significant amount of time reading through the coded transcript quotations looking for meaningful links and trends in the data. I added my thoughts and interpretations as I read through the quotes, while also connecting transcript codes that were relevant. I also extracted quotations from the coded data that I thought communicated a clear sense of the theme being represented.

There were several resources guiding my interpretation of the coded data. I constantly reflected on my knowledge of systems theory and the systems literature, as well as contemplated the practical meanings and applications of what I was reading based on the interviewee's comments. This reflection guided my interpretation and narrative text. I also referred to the notes I had taken documenting the thoughts and reflections I had as I read through the transcripts.

I used the heuristic to help me with organizing the findings of the thematic analysis by arranging the results chapters by the Foster-Fishman system elements (2007). However, it is important to note that the themes and subthemes came from the data and were specific to the evaluation innovation and the QL system.

Webinar Focus Group to Discuss Study Results

Invitation emails were sent to all of the individuals who participated in interviews asking if they would be interested in participating in a webinar focus group to review and discuss the results (Appendix F). A focus group is essentially a group interview and can take numerous different forms depending on the purpose of the focus group (Fontana & Frey, 2005). Fontana and Frey (2005) propose that data from focus groups can be used in conjunction with other data or for triangulation purposes. The email stated that I would present the results of the analysis, including the major themes from the data and strategies for improving implementation in the network (e.g., leverage points in the system). In the email, I also requested that they respond within one week if they were interested in participating in the session.

The response to the focus group invitation was very positive with fourteen of nineteen interviewees responding to the email, with all but one were interested in participating. I used Doodle, an on-line scheduling program to identify the best date/time to hold the focus group and selected the time when the majority of the participants were available. Of the ten participants who planned to attend the focus group, five of them actually participated. Two of the initial participants emailed to cancel their participation the day of the session because of other work responsibilities. And three of the ten participates originally informed me that they may not be able to attend the whole session but were going to try to attend some of it, they ended up not joining the session at all.

The focus group session lasted two hours and was conducted through an on-line program called Webex. This program allows for individuals to hold conference calls and give Power Point presentations at a distance. The participants of the Webinar focus group were able to watch the Power Point presentation from their own computers and listen to what I presented and engage in the discussion with other participants via the phone. I began the focus group by having all participants briefly introduce themselves. I then thanked participants for their time and interest and described the purpose of the focus group. Next, I presented a forty-minute Power Point presentation of the study including the findings, which took approximately seventy-five minutes to present because the participants asked questions throughout the presentation, which I encouraged them to do. The remaining thirty-five minutes of the focus group were used for discussion with the participants. The presentation was meant to provide a platform for the discussion and I tried to initiate more discussion throughout the presentation as well, but it did not work very well. Similar to the interviews, one of the major limitations of the focus group

session was that it was not conducted face-to-face (Garbett & McCormack, 2001). The material that I presented and that we tried to discuss was very dense and it was difficult for me to sense the participants' reaction to it. Because I could not see the body language or pick up on any nonverbal cues, it was difficult for me to determine the reason for the lack of discussion during the presentation. When I asked them about this, they explained that the results were resonating with them but it was a lot to digest.

For the discussion part of the focus group webinar, I initially left it open to the participants' thoughts and questions. I also asked them to specifically respond to the following questions:

1) Did I interpret the information from the interviews correctly?

2) Did the themes make sense?

3) Are there any practical implications of the results for the QLs?

Overall, the general response from the participants was positive and they found the results interesting and potentially useful to the QLs. They were in agreement with how I had interpreted the interview data and did not make any suggestions for changes or corrections. The themes all made sense to them and they expressed how many of the themes resonated highly with them. For example, one of the participants particularly liked that the themes demonstrated how many of the implementation issues are outside of the control of the organization. This individual explained how the policy context often changes annually due to political issues and how this shifting policy context directly impacts implementation at the QL level. The other participants concurred with these statements and supported the notion that there are many factors that influence

implementation that are outside of the organization. The participants seemed to appreciate the results because they explained why implementation failure is not always due to organizational level factors. I sensed that the participants felt some resentment about being blamed for not implementing innovations and appreciated having the results demonstrate the inherent complexity of the issue. However, only five QL members participated in this phase, so additional research is needed to determine if these opinions are more wide spread within the system.

The participants suggested that there were certainly practical implications for the results and asked for a report of the findings to be shared with NAQC. Although they were unable to provide suggestions for specific practical applications, they did express an interest in further efforts to apply the findings in the QLs. I ended the focus group webinar by thanking everyone for his or her time and participation. I also emailed the Power Point presentation and some literature on leverage points to the participants after the session ended, per their request. Lastly, I encouraged the participants to email me if they thought of any additional questions or comments about the study.

Overall, I found the focus group session useful because it provided me with an opportunity to get feedback from at least some of the participants and too engage them more in the research process. However, in the future I would try to avoid conducting focus groups over the phone because it does not work very well. It was difficult to have an in depth conversation about such a complex topic without being able to see each other. I had opted to not send reading material to the participants prior to the focus group session because I was wary of asking for too much of their time. In hindsight, I wish that I had sent a brief report for them to read prior to the focus group session because I think

that it would have provided a basis for a more rich discussion. The participants were slightly overwhelmed with the information and the concepts and expressed how they needed some time to digest the information. Despite the challenges, the focus group session still added to the credibility of the study and improved the quality of the data analysis.

Judging the Quality of Qualitative Data

If this were a quantitative dissertation, I would need to address the rigour of the study, which would include the reliability, validity and generalizability. However, these terms are not appropriate for a qualitative study, which aims to understand the phenomenon from the social actors perspective. The quality of the results is an equally important issue in qualitative research, but it is determined with different criteria. Instead of rigour, qualitative research is assessed on its trustworthiness, which is determined through credibility, transferability and dependability of the study (Flick, 2006). Although they mean different things, credibility is the parallel term for validity, dependability is a parallel term for reliability, and transferability is the parallel term for generalizability.

Credibility refers to how well the data and process of analysis addresses the intended focus of the research (Flick, 2006). Credibility relates to making decisions about the context, participants, and approach to gathering data and analysis, including how to code and how much data to collect (Graneheim & Lundman, 2004). Two suggestions by Graneheim and Lundman (2004) for increasing credibility is to include quotations from the transcripts that represent the coding theme and second is to seek agreement from other researchers, experts, and participants as a means of confirming the results. In my

study, I conducted a focus group session with five interviewees from the study to confirm and discuss the results. In this way, I increased the credibility of the study results, even though it would have been preferable to have more participants involved in this process.

Dependability, according to Lodico, Spaulding and Voegtle (2010, p. 130) refers to "whether one can track the procedures and processes used to collect and interpret the data." These authors also suggest that: "dependability is often the difference between an experiential report that simply summarizes a researcher's conclusions and an empirical research-based qualitative study that includes a thorough explanation of methods" (p. 130). Dependability can be addressed by providing a detailed account of the methods used for data collection and analysis so that that study could be replicated by another researcher, although not necessarily expecting to achieve the same results given different study contexts (Shenton, 2004).

Dependability also focuses on the risk of inconsistency in data collection over time when judgments are made to balance maintaining consistency and evolving with the interviews and observations as needed (Flick, 2006). Changes in the data collection over time was less of an issue with my study than it would be in a larger study, because only 19 interviews were conducted over a three-month period. Although I maintained a consistent focus and interview topics to address across the interviews, there was still some evolution and change throughout the data collection process. Specifically, my understanding of the phenomenon and QL system grew so I was building on some of the analysis I had already completed and was posing questions to verify some of my findings (Neuman, 2006).

Transferability refers to the extent to which the findings can be transferred from one context to another (Lincoln & Guba, 2002). Qualitative researchers do not use the term generalizability because they are not trying to generalize from a sample to a population. For my study, I am not claiming that my results can be applicable in all situations and contexts. However, in a similar context, we could expect some of the same patterns to emerge. According to Graneheim and Lundman (2004, p. 110), "to facilitate transferability, it is valuable to give a clear and distinct description of culture and context, selection and characteristics of participants, data collection and process of analysis." I describe the context of the study in detail in the third chapter. I also provide a detailed description of my research activities in the methods chapter.

One of the key differences between a quantitative study and a qualitative study is that a quantitative researcher attempts to be objective and separate from the data and the study. In contrast, a qualitative researcher assumes that objectivity and separation between the research and the data is not possible. Instead, they embrace the research perspective and their subjective interpretation of the study, describing their role and attempting to understand the study results within their own social context.

As a qualitative researcher, I accept that my social location and my past training and experiences have influenced the analysis of the findings. I am not trying to claim that someone else reading the transcripts would have come up with the exact same interpretation. This is my point of view at one point in time given my own subjectivity, context, and location as a researcher. I recognize that someone else might review the transcripts and come up with a totally different coding scheme because they have a different background and could be using a different theoretical approach to guide them.

My perspective is guiding my analysis and interpretation, and as such, I have a subjective approach.

Of course, as I was conducting the analysis I had concerns about how I was interpreting the quotations. I wondered if I was reading in too much to the data based on my systems background and if I was capturing the participants' meaning accurately. However, I recognize that this is a normal concern for qualitative researchers. My analysis was conducted systematically and I put the necessary elements into place to ensure quality of the data analysis.

Summary

Overall the study design and methods used were appropriate for studying the targeted phenomenon and produced useful and interesting findings. The Foster-Fishman (2007) framework provided a useful heuristic for ensuring that the analysis remained anchored in a systems perspective. However, the data themselves were the primary driving force for identification of the themes found in the thematic analysis. The fact that the data produced themes that happened to fit well into the system elements proposed by Foster-Fishman provides support for the value of the framework as a guide for systems change. However, because of the breadth of the framework categories it will be necessary to still use the data to identify appropriate themes for each unique innovation and system. The following three chapters present the unique themes and subthemes identified as relevant to implementation of the evaluation innovation in the QL system. As stated above, the results chapters are organized using the Foster-Fishman system element themes, with the next chapter presenting the normative elements (chapter five), followed

by the system resources (chapter six), and lastly the system regulations and operations (chapter seven).

Chapter 5. Normative Elements

Chapter Overview

This chapter describes the findings within the normative elements theme (system norms), a concept borrowed from the Foster-Fishman et al. (2007) framework and defined as the stakeholders' worldviews including assumptions, values, and beliefs. These authors suggest that norms can either facilitate or constrain systems change, depending on if they are consistent or inconsistent with the desired change. Norms are especially powerful if shared across multiple actors, thereby creating a dominant norm. According to Foster-Fishman et al. (2007), successful systems change efforts require identifying the dominant norms and shifting them to be consistent with the desired goal. The interviewees described some of the concepts presented by Foster-Fishman (e.g., values), but also provided another concept that would fit under normative elements. They provided a great deal of information that described their 'perspectives' on the practice. Peoples' perspective is a concept mentioned in the systems literature as being important for understanding systems change (Hargreaves, 2008), but is not explicitly listed under normative elements in the Foster-Fishman framework.

In this chapter, I begin by describing the interviewees' perspective of the innovation, including how they defined the innovation and their perceived goal of evaluating effectiveness. I then present a theme that describes how the justification for funding was a major driver for conducting evaluations in the QLs. Following this, I describe from the interviewees' perspective how the QLs' focus on conducting meaningful utilization-focused evaluation varied across a continuum of high to low utilization. I then describe the variation in how interviewees viewed the relationship

between evaluation and QL operations. And lastly, I describe from the interviewees'

perspective how they valued network level evaluation and aggregation of evaluation data

across the QLs, as well as underlying fear related to cross-QL comparisons.

Definitions of the Innovation

During the interview process, I struggled with the boundaries and the elusive

nature of the innovation under investigation. It was difficult to get clear definitions from

interviewees for evaluating effectiveness when I directly asked them for their definition.

But when I asked them about the type of evaluation being conducted in their QL and why

they conducted it, I got more insights into how they defined the innovation, which is

described in the following sub-theme looking at goals of the innovation. Similarly, when

I asked them directly what it means to evaluate effectiveness, interviewees struggled to

provide a definition.

It was also difficult for many of the interviewees to separate evaluation from

research and these terms would often end-up being used interchangeably throughout an

interview. In some cases when an interviewee was repeatedly using evaluation and

research interchangeably, I would ask them to describe the difference between research

and evaluation and several distinctions were made. Below are quotations that capture

responses to the question asking how they define and/or distinguish between research and

evaluation. In all three cases this follow-up question was asked because the interviewee

had repeatedly used the terms interchangeably. One distinction was related to timing and

planning. A few of the interviewees suggested that research was something set-up in

advance, whereas evaluation was done retrospectively and based on events that have

already taken place. In conjunction with this, research was viewed by several of the interviewees as having a specific design (e.g., RCT), whereas evaluation was seen as being reactive, analyzing what had already happened, without a specific design in mind.

Quotation #1-

> Well I guess for me, they're pretty close linked. For me it's hard to tease them apart. But I guess research might be to see if a new program would work. And maybe to create some kind of artificial circumstance where you're trying out this new program. And of course there are all kinds of ethical issues around that with making sure that people are receiving service before they could do RCT and send people to one thing or another. But if we have that, and that certain things works, then you want to make sure people are getting the best service they can. But yeah, I guess to me research would be maybe asking a new question like that or hmm… Yeah, actually I really have trouble teasing them apart because I think for me also evaluation answers questions that quit lines, people who run quit lines need to use to make decisions; people who fund quit lines need to use to make decisions. So both research and evaluation could do that. (KI-18)

Quotation #2-

> For me, you know, it's more of a randomized control trial or, or yeah, something that's really set up ahead of time. I mean, and, and evaluation is more a retrospective study, based on events that have already taken place. (KI-19)

Quotation #3-

> Yeah I think that, you know, again I'm not trained in evaluation. So I'm trained in, you know, like basic epidemiology. But a lot of I guess, you know, in my mind when you're talking about, you know, randomized control trials that's more in a not, you know that's more in like an epidemiologic research realm versus, you know, I wouldn't say, I wouldn't think that. When I would see a randomized control trial, I would not necessarily think that that would be evaluation research. I think of evaluation research as more maybe at a kind of going in a like a community has implemented something or a program has implemented something and doing some, you know, study of how that implementation went and what its affects were is more what I see evaluation. And we, you know, within CDC we typically don't do, you know, randomized control trials so. (KI-2)

Goals of Evaluating Effectiveness

Another critical variation in perspectives that emerged across interviewees was the goal of evaluating effectiveness. This information was difficult to tease apart from the interviewees' definition of the innovation, as the goal and definitions of the innovation were very connected. The different subthemes that emerged relevant to the goals of

evaluating effectiveness included monitoring outcomes, quality improvement, furthering

the evidence-base, and justification for funding (see table 4). The first three subthemes

are described in greater detail below, and the latter subtheme is described in the following

section, as it was a more substantial and independent subtheme.

Table 4. Participant Identified Subthemes Relevant to Goals of the Evaluation Innovation

Subthemes Relevant to Goals	Number of Participants who Identified the Subtheme
Monitoring outcomes	19
Quality improvement	12
Furthering the evidence-base	11
Justification for funding	15

The data collected by QLs differed in relation to the different goals of the

practice. Those who viewed the goal of evaluation as monitoring outcomes only collected

basic outcome data (e.g., reach and utilization of services). Several of the interviewees,

suggested that funders are primarily interested in monitoring basic outcome data (e.g.,

quit rates) and are not interested in conducting more complex evaluations (e.g.,

comparison of services) to improve services or further the evidence-base. The quotation

below illustrates the perspective of many of the interviewees that funders are primarily

interested in basic outcome data.

> Mostly I think funders are interested in outcome, so they want to know how many people quit
> smoking. Obviously they're really interested in knowing some of the monitoring indicators like
> call volume, you know how many people are calling the quitline. Hmmm, what are some other
> things that they're interested in? I guess what kinds of services people receive… some
> characteristics of people who are calling. So demographics, they're interested in the proportion of
> the Aboriginal people that are calling the quitline. Those kinds of things. (KI-18)

One point that was made by several interviewees was that telephone counselling

had already been evaluated in RCTs and deemed to be an evidence-based practice

therefore they did not see any reason to further evaluate the QLs. This point represents

the lack of support for evaluation in the QLs and is illustrated well in the two quotations

below taken from an interview with a funder organization. The interviewee at two

different times in the quotations justified her QL's cessation of evaluation based on the

fact that QLs are an evidence-based practice.

Quotation #1-

> Okay. Well, funding has become an issue for us. We have suspended doing any evaluation calls for the QL. Now, the reason we did this is because we know it's an evidence-based service. While we're interested in quit rates, we're confident that the experience we're getting, you know, we're getting an evidence-based program, so that's, it was sort of an easy decision to make. (KI-19)

Quotation #2-

> NAQC recommends doing research very strongly but we, or doing evaluation very strongly, but we forego evaluation because the research tells us QLs have undergone numerous randomized controlled trials, so we know they're effective. So if we're gonna save money and save money for, for services, we're gonna give up the evaluation because we know it's effective. (KI-19)

In contrast, some interviewees were more supportive of evaluation and described how

they used evaluation findings for quality improvement in their QL. Half-way through the

interviews, I began to notice a strong distinction between participants that did not value

conducting evaluation and saw it as unnecessary monitoring of outcomes and those who

valued it as an important part of QL operations. During an interview with a funder, I

asked the interviewee for his thoughts on why some QLs felt that it was not necessary to

conduct evaluations. The following quotation is his response and it illustrates the

different perspectives relevant to the intent to use evaluation for quality improvement

purposes.

> Probably because they're not making any changes. One of the things that we get feedback from every single evaluation call is basic customer satisfaction; and if they could change something what would it be, and why? This feedback that we've received over the years, from our customers has been critical for process improvement. And I feel that a quit line which doesn't follow that same principle is really missing the boat and leaving themselves open for lack of innovation and lack of improvement. (KI-12)

The interview data suggested that some decision-makers in the network perceived evaluation as including comparison of services and furthering the evidence-base for QLs. This perspective was not shared by all decision-makers as several of them defined evaluation as monitoring outcomes and defined anything more complex, such as comparison of services, as research. The following quotation from a service provider illustrates the perspective that comparison of services is part of evaluating the QLs.

> Well I guess it would depend on how you define evaluation. We have our own data analyst on the staff, and he works on other things besides just the (x-state) tobacco help line, because we have other tobacco programs that we operate. But we can continually, or at any time go in and look at the data that we have to answer questions that we have about specific, just to look at the impact of specific activities that we've got going on in one county or another. Things like that. So it's not as periodic and regular as the outcomes evaluation. It's not something we report on at any set interval, but we're continually looking at the data. For example, we have a fax referral program from medical practices, and we're in the process right now of doing a pretty robust evaluation of how people who are referred that way experience the program, compared to everyone else. And so he's involved in a project that is sorting through all the data that we have since we implemented the fax referral program. But it's kind of a project. It's not an ongoing thing. (KI-8)

The data from the interviews suggests that NAQC fit into this group, as their original intention for the MDS (Minimal Data Set) was to develop a standard set of questions in order to consolidate data from different QLs and use it to answer research questions and further the evidence-base. One of the interviewees in this group explained how the call volume from a single QL was not enough to provide sufficient data on population segments, however, combining the data from different QLs would provide a tool for answering more complex evaluation questions. Thus, NAQC's goal was to collect data from all the QLs and use that data at the aggregate level to pursue more complex evaluation questions that could be used to inform policy and practice decisions in the QLs.

The challenge presented in these subthemes is that the different perspectives of the innovation were clashing, creating resistance to change towards full implementation

and institutionalization of the evaluation innovation. One group considered anything more complex than monitoring outcomes to be research. As they were not interested in conducting research, they were also not interested in collecting the necessary data to pursue more complex evaluations. In contrast, another group viewed evaluation and quality improvement as part of service provision and considered more complex evaluations as necessary to answering practice and policy questions. For those who considered the goal of evaluation to include furthering the evidence-base for the QLs, then it was necessary to collect data that could be used to compare effectiveness of different services, and answer decision-makers questions.

Justification for Funding as the Driver to Evaluate Effectiveness

Another subtheme that emerged from the transcripts is that for many of the QLs, the incentive or 'push' for evaluating effectiveness was to justify funding and a driving factor behind this was the current climate of accountability. Interviewees often mentioned how there was a push for reporting outcomes in order to justify funding (see table 4). The quotation below taken from an interview conducted with an evaluation contractor illustrates this point.

> Evaluation provides us with the ongoing monitoring and to make sure that we continue to get the quitlines funded so that you're providing those numbers that you need to provide to continue the funding. So that's really the reality of what evaluation is used for. (KI-18)

Some of the interviewees described how their states used evaluation reports for quality assurance, to check that the quit rates remained high and to make sure that the QL was reaching everybody, especially the underserved and ethnic minorities. Several interviewees also described how their QLs used evaluation results for writing

surveillance reports or fact sheets that could be sent to all interested parties involved. One funder interviewee explained how reports were tailored specifically to the people that were in positions that they needed to influence, in order to back-up funding. Evaluation was seen as a way to achieve sustainability in funding, as illustrated in the following quotation.

> Sustainability. That's all- that's what it is. It's, you know, being able to have the data you need to prove that your programs are cost effective and efficient and necessary. (KI-17)

QLs also used evaluation outcomes to justify the amount of money they spent on services they provided. For example, one QL had compared Nicotine Replacement Therapy (NRT) to a new drug called Chantix and had used the quit rate outcomes to justify the funding put into it. QLs that were grant funded needed to be able to justify the existence and continued funding for the program. Even interviewees from states that were supportive of funding tobacco prevention programs described how it was still important that the QLs provide data to justify funding. The quotation below was taken from an interview with a service provider representative who described his state as being very supportive of funding tobacco control programs.

> The state of (X) has a pretty good record for using the money from the master settlement agreement to treat and prevent chronic conditions and tobacco, as well as any other state, better than a lot of them in the country. But that doesn't mean that in challenging economic times, people aren't eyeballing that chunk of money. And so we need to be able to equip our funders and partners at the state of the Department of Human Services with the information that will allow them to continue to justify the program and all of that. (KI-8)

Based on the interview data, the pressure at the national level to justify funding in the U.S. had intensified with the recent American Reinvestment and Recovery Act (ARRA) funds ($44,500,000). According to interviewees, three major concerns of accountability were demonstrating: the number of people being served, the type of

services being provided, and that services reached all segments of the population. The

evaluation data was critical to answering policy maker's questions for accountability. As

the quotation below illustrates, evaluation data, in particular monitoring outcomes,

provided evidence that there was a demand for the QL services and that the funds were

being used appropriately.

> We're always telling policy makers we need more money so policy makers are saying, well,
> you're already spending this amount. What are you doing with that? Can you show that, you
> know, you're spending the money wisely? And that will, if we can show that then that will
> encourage them to, you know, put more money to this, to the QLs. And so that's what we're trying
> to do is advocate at a national level to get additional funding for these grossly under-funded QLs.
> And I think evaluation is critical to, you know make those connections for policy makers that there
> isn't enough money. There's more demand out there. We are using the money appropriately. We
> are serving people and providing appropriate services and we are reaching across all populations.
> (KI-2)

The interview data suggested that CDC was especially interested in accountability

and public sharing of QL data to demonstrate value for the funding. The interviewees

explained how CDC was planning to create a database with QL data (including MDS

data) that would be made public by posting it on their website. The interviewee from the

CDC explained how the CDC's goal focused on reporting for accountability reasons and

therefore the database would only include the "basic measures that people can

understand" (KI-2) such as populations being served and quit rates. CDC's goal was not

to answer any research questions or to take into consideration QL variation such as

different protocols or number of callers.

Relationship between Evaluation & QL Operations

Another key difference in perspectives on the innovation deals with assumptions

of whether or not evaluation was part of QL operations or an extra competing activity.

Several interviewees described evaluation as an extra activity that competes with the

operations of the QL. In contrast, other interviewees described the two activities as complementary and viewed evaluation as a necessary aspect of QL operations. The interviewees' perspective of the relationship between QL operations and evaluation can be placed on a continuum from completely separate to completely complementary, with interviewees dispersed across the continuum. With the complementary perspective, evaluation was viewed as practical and applicable. This distinction in perspectives was particularly strong if the definition of evaluation included anything more complex than monitoring outcomes, which was considered by many to be research. The two quotations below illustrate the perspectives at different ends of the continuum.

The first quotation is particularly interesting because the interviewee (a service provider) used the term "business product" in reference to the QL. This interviewee used research and evaluation interchangeably and clearly stated how the QL operations (i.e., business product) were separate and in competition with the research/evaluation. Although none of the other interviewees referred to QLs as a business product, many expressed the competition between research/evaluation and QL operations. The second quotation also comes from a service provider interview who viewed research/evaluation as part of QL operations. This interviewee explained that the issue was related to how stakeholders defined research and their expectations of how it should be done. He suggested that there needed to be a broader definition of research in the QLs.

Quotation #1-

> The tensions came when, when the demands of the business product were competing with the demands of whatever research we were doing. (KI-7)

Quotation #2-

> But anybody who's involved in operating anything does research on a daily basis, you know. You become aware of a problem and so you research it. You know you it's just figuring out what's going on. (KI-10)

There were also some strong opposing views on the cost and ability to conduct more complex evaluations. Some saw using an evaluation design as just not being possible or feasible, whereas others suggested that there were ways to do it at a low cost by building it into ongoing QL operations. Those who viewed it as not being feasible most often stated RCTs as the necessary evaluation design and lamented how it was not possible to randomly assign callers to intervention and control groups. Those who supported more complex evaluations countered by explaining ways to build evaluation into QL operations with little extra cost. Below is a quotation taken from a service provider interview with a respondent who felt strongly about the complementary relationship between QL operations and research/evaluation.

> And people tend to think that that's you can only do that in the context of a large controlled, you know, trial but there are lots of little ways to build in randomization. Lots of little ways that QLs can build in that kind of design to answer key questions. And so if they would do that and then disseminate their findings I think that would really help the field. (KI-10)

Conducting Meaningful Utilization-Focused Evaluation

Another important subtheme that emerged from the data is the concept of doing meaningful utilization-focused evaluation. The interviewees were all asked if and how their QL utilizes evaluation results. Similar to the previous theme, there were a range of responses across a continuum from not at all used to being a significant priority. At least six of the participants explicitly mentioned utilization of the evaluation results during the interview (table 5). Based on the interview data, there were QLs that evaluated effectiveness with the intention of utilizing the evaluation results and there were QLs that

evaluated effectiveness just for the sake of doing it because it was required. QLs that

perceived evaluation as strictly for monitoring purposes were more likely to conduct

evaluation that was not utilized. Although some of the QLs that strictly monitored

outcomes also emphasized the need to use the outcome data for reporting purposes to

policymakers, others used the evaluation results for quality improvement purposes. The

results of the analysis suggested that whether or not QLs are doing meaningful

utilization-focused evaluation is important because it directly impacted their perceived

value of evaluation and subsequently the sustainability of the innovation. The first

quotation below illustrates how there was a difference across QLs in the degree of

utilization of evaluation results. The second quotation was taken from an interview with a

respondent who explicitly expressed the need for utilization of results and even described

steps taken to make evaluation reports useful to stakeholders.

Quotation #1-

> Yeah, so I mean I think so basically, you know, the states have been collect, through their vendors, have been collecting a lot of information on callers. And some states have been, you know, utilizing that information and publishing it. But, you know, other states really have not been doing anything with their data and not really looking at it. (KI-2)

Quotation #2-

> I mean at the center there's always been a commitment to making our reports as useful as possible to our funders. So even if it was an evaluation of something that the organization funding the research or who's getting the results can actually take the report and use it…. (deleted text)…. And then there's also my own personal commitment to making evaluation results useful because that's just a really big bugaboo of mine, when evaluations just aren't tailored and aren't responsive to the needs of the people who are actually going to be using it. (KI-13)

Table 5. Participant Identified Subthemes for Utilization, Aggregation, & Fear of QL Comparisons

Themes	Number of Participants who Identified the Theme
Utilization of Evaluation	6
Aggregation of QL's Evaluation Data	12
Fear of QL Comparisons	14

There were some QLs that used their evaluation data to answer specific questions about their services, and for quality improvement. For example, one interviewee described how his/her QL compared the impact of specific activities in different counties and also compared different pharmacological treatments. In this group, one of the goals of evaluation was to answer questions and use the evaluation results for decision-making. The perspective on utilization of evaluation results in this group is illustrated well in the following quotation.

> The, it sort of waivered depending on how much staff time there was to really pay attention to the evaluation results. But a lot of the evaluation questions that we were asking were really important in terms of how big and improved the programs that we were running. And so, we deliberately designed the evaluation to help answer specific questions. Or we would see interesting patterns from one evaluation and so we would add questions or change the design slightly to be able to help us further answer those questions in the next evaluation and then change the program based on those results. So it was very much oh, an iterative process of getting evaluation findings and then changing the program accordingly to make it better, make it more effective, more cost effective. Increase caller satisfaction, those kinds of thing. (KI-1)

There were also QLs that conducted the evaluation specifically for reporting outcomes to stakeholders and they were very conscious of tailoring the reports for different stakeholders including: service providers, the public, legislators, policymakers, and governments at the funding level. One interviewee (an evaluation contractor) was explicit and emphatic about the need for evaluations to be utilized and described the QLs commitment to making their reports as useful as possible to their funders. The individual described how they put talking points into the reports and presented them as a series of

facts sheets so that a policymaker could look at it and use the information to promote

certain anti-smoking or smoking cessation policies.

> Oh well, when we were first funded it was required that we put together an annual report. And the formative annual report is something that is sort of a negotiated thing. I mean we want it to be as useful as possible for the state to be able to use in terms of providing information to the public and to the legislator. So what goes into it and how it's presented is sort of an iterative process, which really I think as an evaluator is the best. I think the integrative process can be a really good way to do it, because you want the evaluation report to be useful. I mean if we want to take a utilization focused evaluation approach on it, there's absolutely no reason to even do evaluations if you're not going to use the results. And if I just put together some, a bunch of tables that you can't read, then there's no purpose in evaluation and it's wasted money. So I think working with the stakeholders and working with our funders which is the state to provide useful evaluation results is really important. So we do that. (KI-13)

Finally, there were two interviewees who described not using the evaluation data

at all and only conducted the evaluations because they were required to. At this end of the

spectrum, the goal of the evaluation was not utilization of the results for quality

improvement or reporting purposes. In these cases, the goal was strictly monitoring

outcomes, primarily quit rates. The quotation below supports this conclusion drawn from

the data and provides some insights into the QL differences that contributed to different

rates of utilization.

> For some quitlines that don't necessarily, so it gets complicated because if you have a quitline, where you don't have a lot of staff expertise, but they've written it into their contract with the service provider that the service provider will conduct the evaluation, the evaluation might not be used for anything. But it's something that we've always done, that's part of the package that our service provider offers us, therefore we will continue to do it. (KI-1)

One interviewee described how they had "de-prioritized" the evaluation because they

were not using it. The evaluation outcomes were not being included in reports to

policymakers or used for quality improvement purposes. As illustrated in the quotation

below, the interviewee stated how they were going on the knowledge that a QL and

telephone counselling services were still a recommended best practice. She described

how they had felt solid about their service provider and that the evaluation outcomes had

been constant for 10 years. Funding had been cut and as a result they had stopped doing

the follow-up evaluation. However, she added that even if the funding were restored, the

evaluation plan may not be reinstated because it had been "de-prioritized" because they

were just not using it.

> And so, during that time we were on a very limited budget and we were not oh, conducting a follow up evaluation of callers. However, we have conducted, we had conducted 10 years worth of follow up evaluations and we, with (X-service provider) and knew that we had a very solid quit rate and we're going on the knowledge that a quitline and counseling and services provided are still a best practice and recommended and all of that. And we certainly have been justifying quite a bit why we were not conducting that follow up evaluation. Because we felt that we had the 10 years of data. Indicating that it did work and then one year that we were not evaluating we weren't expecting to see a difference. (KI-9)

The interviewee quoted above was clear about the fact that the QL had been

conducting "meaningless evaluation," collecting outcome data because they were

supposed to, not because it was of use to the QL. Interestingly, the same interviewee also

explained how with the new ARRA funding and subsequent reporting requirements, the

QL would have the extra push they needed to begin to conduct more meaningful

evaluations. The interviewee first stated that, "The only reason that we're considering or

that we are moving forward with the follow up evaluation now is because it's a

requirement under the ARRA grants"(KI-9). This statement suggested that the funding

mandates to conduct evaluation are sufficient in forcing QLs to conduct the evaluation.

The interviewee later stated that:

> The ARRA requirement to do this has been really nice for us because we've been able to tack on a couple of projects that have been building up as part of our follow up evaluation and we do have things that we'll be using the data for this time around. So the reports won't just be sitting around. (KI-9)

This statement suggests that not only would the ARRA requirement enforce

conducting evaluation, but they would also have a positive affect on utilization of results.

To clarify this assumption, I asked if the respondent was suggesting that the policy push for accountability and evaluation had been a good thing in terms of giving them more reason to do evaluation and to use the evaluation. The interviewee provided the following response, which indicates that the ARRA funding did in fact provide a push for more meaningful evaluation.

> I think more reason to do effective, more meaningful evaluation. I think that up until 2008 we were just doing evaluation under the, okay, we should be doing this. We should know what our quit rate is, but it wasn't really getting us anywhere. We had no need for it. But now that we're looking at how better to improve services and whether or not policy does need to change to support this maybe was, I'm sorry, the Medicaid contracting or with Addictions & Mental Health (AMH) requirements for their health care providers those types of things, I think now that we have some improvements we're looking to make the evaluation will be much more useful as opposed to just, oh, this is the way we've done it. This is the way we always do it. (KI-9)

Aggregation of Evaluation Data & Fear of Cross-QL Comparisons

The interviewees were specifically asked how they felt about aggregating evaluation data across QLs and creating a database with QL evaluation data. The responses indicated that in general, evaluation was viewed as important by the QLs and aggregation of evaluation data across QLs was viewed favourably. The data also suggested that interviewees perceived having evidence on the effectiveness of the QLs as valuable because it could be used to support requests for additional funding. At least twelve of the participants mentioned positive reasons and support for aggregation of the evaluation data (see table 5).

Although there appeared to be a perceived value for collecting and sharing evaluation data within the QL network, there was also an underlying feeling of fear expressed with respect to comparisons being made across QLs. At least fourteen participants mentioned fear for making comparisons across the QLs during the interview (table 5). The fear of comparison was primarily related to funding, although several

interviewees suggested that even in cases where funding was not an issue, the QLs may be fearful of looking bad relative to other QLs. The following two quotations demonstrate some of the issues related to the fear subtheme.

Quotation #1-

> Well, I think there's this push and pull among the quitlines. I mean when it comes to the minimal data set, I think we can all agree that we all benefit if quitlines are shown to be effective as a whole. We all get the opportunity from our federal dollars and we all get more buy-in from our state health departments and Health and Human Service professionals to send people our way and from the general public to call us if we collect some essential data points and prove that we work. But I think when you get to the next level, which is how are we different and who's the most effective, that's when the walls start going up. (KI-6)

Quotation #2-

> *Interviewer*: Are there any issues like that in Canada where quitlines from different provinces don't want to be compared to each other?
>
> *Interviewee*: Yes there is. There's at least one province that doesn't want to be, that doesn't want to share their data. (KI-11)
>
> *Interviewer*: And from what you can tell that's primarily because of comparison concerns?
>
> *Interviewee*: It's because of comparison not because of competing for funds but because they don't want to look bad, because, yeah, basically. (KI-11)

A key issue related to the fear of comparison, was the type of data that is collected and used to make the comparisons. The interviewees stressed the importance of collecting sufficient contextual data to describe the QL's unique differences and contexts. One seemingly root of the fear was that sufficient data on the QL context would not be collected and comparisons would be made based on quit rate outcomes, without understanding the reasons behind the outcomes. There was also general concern with the idea of using prevalence and quit rates as comparative factors, as explained by the interviewees in the following two quotations.

Quotation #1-

> Sometimes the evaluation is the problem. It's not the information that we have but it's actually what's being asked or what's being looked at in the evaluation that doesn't give the broad spectrum of the individuality. (KI-5)

Quotation #2-

> *Interviewer*: And the provinces in Canada, is there from your experience a lot of sensitivity around having their quit rates and information compared to other quit lines in Canada?

> *Interviewee*: Oh yeah I think for sure there's sensitivity around looking at data. Especially when you're looking at funding and maybe risk being compared in a negative way. I think there's even more sensitivity around being compared to some of the findings coming out of the United States just because it's such a different environment. (KI-18)

Furthermore, there was a concern that comparisons would be made that would then be tied to funding without taking into consideration the QLs' differences (e.g., populations served, budget). The fear of having comparisons tied to funding was greater for some American QLs than others based on the state context and legislative support. Commitment of the legislation for tobacco control differed greatly across states and for states that had little support or strong opposition from their legislation the fear was much greater. The following three quotations illustrate how fear of sharing evaluation data was directly related to the state context.

Quotation #1-

> I think people fear that if we start collecting information and sharing it, it could get used against them, that they'd be more than happy to participate as long as the information doesn't get shared with legislatures or decision makers that are really more interested in the funding than they are in the service being provided. And we see that in our own state. You know we have legislators in (X-state) sitting there would like to see all the funds pulled. (deleted text) It depends on which legislator but we have some legislators that don't believe that tobacco tax money should be used for anything but the general fund. You know and they're very vocal about it like right now there is a bill sitting in…it's actually being voted on and it's been approved in the house. It's sitting in the senate in (X-state) and they're looking at putting the tobacco protected funds and all protected funds back on the ballot to unprotect them so the legislature can use the money for anything they want. (KI-14)

Quotation #2-

> *Interviewer*: Why do you think it is that some of the states are worried about looking bad or being compared to other states? From what you describe it sounds like everybody has their own pot of money. Is there a competition between states or maybe you can explain it?
> *Interviewee*: I don't think it's so much a matter of competition as with any state a lot of quitline funders really have to go to battle for funding for cessation.   And there are legislators or other policy makers who are, you know, continually challenging them and saying, you know, why is this a good use of money and so anything that they feel can look, you know, reflect has the potential to look poorly on the performance of the quitline they fear will be used by these, you know, critical legislators to cut funding, you know. So for example if, I think all quitline funders fully support the idea that the quitline should be as accessible as possible and that it should be utilized by a diverse clientele. But what if the data show that, you know, even though ethnic minority smokers make up twenty percent of your state's population or your state's smokers, yet, you know, ninety percent of callers are white, well then that you know somebody could use that as ammunition to show that, you know the quitlines not serving a diverse clientele. So cut his funding. And there, I know there's been a lot of antsy-ness about reporting outcomes because, you know, a lot of people are nervous about, you know that they're overall quit rate will be lower than somebody else's and they're afraid that differences in client characteristics will not be taken into consideration. (deleted text) So, and I think that's a valid concern. (KI-10)

Quotation #3-

> Well, you know, initially I think there was, there was some concern on the part of the states as well a, who are the quitline funders, as well as the quitline providers, about just how comparison data would be used, either by the federal government in, in allocating money or that was the states concern, I think, was that if their state showed up as being less effective in the way that it, it delivered a service, perhaps they would not receive as much federal support in the future. And from the quitline providers' perspective, of course, it was a competitive business and, and there was concern that your quit rates would be used against you by one of your competitors. Now, a lot of that well, over time, that's become less of an issue to the states, because the federal government has taken an increasing role in funding quitline operations and part of the requirement of taking the federal money is transparency in terms of effectiveness and so you, you've gotta agree to share that information, if somebody wants to push it. (KI-7)

One of the interviewees suggested that part of the problem was that there was no

standardized formula for how to fund a QL. People feared that if their QL appeared to be

performing worse than other QLs than their legislators may use it as an excuse to reduce

their funding. The conclusion drawn from the data was that if there is to be cooperation

and support for network level evaluation then there needs to be boundaries defined for the

potential comparisons made. The QLs require reassurance that QL differences will be

taken into consideration and that the data collection and sharing will not be tied to

funding in anyway. They also need assurance that the information will not be shared with legislatures or decision makers who as one interviewee put it, were more interested in the funding than in the service being provided. The conclusions were supported by the following two quotations from a service provider interviewee in which she explained the necessary boundaries that must be created in order to make QLs more comfortable with sharing evaluation data.

Quotation #1-

> Well if it's out there and we have shared this information with NAQC, if it's reported in aggregate format that's okay. But if it's, if it has anything to do with funding and it's identified then probably we wouldn't like that to be exposed. But if it's not funding related then it's okay. (KI-4)

Quotation #2-

> People always compete and make comparisons and our QL is very unique…. (text deleted). So when it comes to funding and competition and comparison that could be a problem. But if it has something to do with just sharing the data and working together to achieve the goals and incorporating projects like chronic diseases, diabetes and all that, that's okay. I mean as long as we're working together it's okay. (KI-4)

Summary

The normative element theme overall provided rich and insightful information for understanding the implementation phenomenon. The findings in this chapter support the Foster-Fishman (2007) framework that suggests that normative elements are a 'deep' below the surface element that must be addressed in order to achieve systems change. For example, the fact that the participants were unable to provide a clear and standardized definition across interviews, along with the variation in the perceived goal of the innovation, provided insights into the variation in perspectives of the innovation. Such a finding has important implications for the KIQNIC study, which assumes that all survey respondents are reporting implementation on the quantitative survey with the same

innovation in mind. The differences in perspectives were important because they were related to if and how the innovation was implemented and sustained.

Of course, normative elements is also a broad category and there are many different subthemes that could be identified for different practices and systems. The normative subthemes identified for the evaluation innovation are unlikely to transfer to other innovation in the QLs. In this way, the framework provides a broad starting point that reflects that which is provided in the systems literature but it is still necessary to identify the subthemes through the data. The following chapter describes another major theme in the thematic analysis, which is the system resources.

Chapter 6. System Resources

Chapter Overview

The following chapter is the second of the three results chapters that describe the findings from the thematic analysis. This chapter presents the themes and information identified under system resources. System resources were mentioned throughout the interviews and stressed by participants as important to the implementation of the innovation. The primary question that elicited information described in this section was a question about barriers to evaluating effectiveness of their QLs. This information also came out inadvertently from open-ended questions such as how evaluation had changed over time in the QL. The interviewees mentioned financial barriers as well as lack of necessary human capacity. The interviewees also mentioned various policies and regulations that are discussed in the next results chapter. Lastly, one of the barriers mentioned by interviewees pertained to the lack of positive and collaborative relationships necessary to conduct evaluations. The idea to place relationships into the resources category came from the Foster-Fishman et al. (2007) framework. They suggest that system resources is an important element to consider because it determines what is available to enact the systems change. The authors also provide three types of system resources, which represent main subthemes uncovered in the analysis: human resources, social resources and economic resources. Allocation of relationships (under social resources) for the analysis seemed appropriate given the emphasis that interviewees placed on them as necessary for successful implementation.

In this chapter, I describe the three different types of resources in the context of implementation as explained by the participants. I will also describe the two types of

human resources that were mentioned by interviewees. In the social resources category, the interviewees presented five different types of relationships that directly and indirectly impacted implementation of the innovation in the QLs. These relationships will be described from the perspective of the interviewees, including their impact on implementation. Lastly, the two primary topics under economic resources mentioned by the interviewees will be explained: 1) cost of evaluation, and 2) funding arrangements.

Human Resources

There were significant differences in human resources across the QLs and the lack of human resources in some cases was a barrier to achieving implementation of the innovation. The interviewees identified two different issues pertaining to human resources as a barrier or facilitator to implementing the evaluation innovation (see table 6). One was having the necessary skills and knowledge to conduct evaluation. The second was having staff actually designated to doing the evaluation. The interviewees were explicit in conveying to me that just having the skills and knowledge within the QL was not sufficient to successfully implement evaluation innovations. An additional necessary factor was having staff designated to doing the evaluation activities. For this reason, these two topics were coded separately and described as such below.

Table 6. Participant Identified Human Resources

Resource	Number of Participants that Identified the Resource
Skills and Knowledge	7
Designated Staff	11

Skills & Knowledge

One of the key human resource issues related to implementation was whether or not

the QL had employees with the necessary skills and knowledge to conduct evaluation.

QLs that did not have the necessary skills and knowledge internally needed to contract

out for the evaluation services. In some cases, the funder contracted with the service

provider to conduct the evaluation in addition to the service contract. One example of

lack of skills and knowledge was with respect to developing an online electronic database

to manage QL evaluation data. In this case, I had asked the interviewee why they needed

to hire externally and if there was anybody in their organization who could develop an

electronic database. She replied emphatically:

> Absolutely not. We're talking about web-based software, absolutely not. We did not have anybody in the company with the expertise to create a database. We had to go with a company that could create and help us manage it. So when we have changes we go through them, when we want to add, when we want to delete things, when we want to tweak reports we go through them and they help us do that. (KI-4)

I was surprised by her response, not because they lacked the necessary skills and

knowledge, but because her tone suggested that it was such an impossible feat. The

difference in QLs' capacity to conduct evaluation in-house was important because it

directly impacted the cost of doing evaluation. As illustrated in the above quotation, QLs

that contracted out to external organizations for evaluation and data management were

reliant upon those external agencies. As a result, these QLs incurred higher costs and

greater challenges to conducting evaluations and revising evaluation protocols.

I was confused by this the first time this issue arose in an interview because my

assumption was that once the web-based system had been developed, the QL would be

able to make changes to it on their own without paying an external contractor. However this was not the case. As illustrated in the quotation below, the QL needed to pay the external contractor each time changes were made to the database, which can become very expensive if regular changes to evaluation questions are requested from NAQC or CDC. The interviewee in this quotation is responding to a clarification question I asked about the problems their QL encountered with implementing the MDS.

> *Interviewee*: If the development and enhancement of our database, so it requires funding every time you add a question or you want to remove a question or you want to change a question we have to pay the programmer like five thousand dollars to do that work. And sometimes we've been challenged to immediately adapt to those changes. But we're almost current we feel we try to, you know bring our self up to speed and make the changes on a timely manner. But sometimes when it comes to funding, yes, the database management has suffered there. (KI-4)
>
> *Interviewer*: So I'm a little confused. Is your database manager external to your organization?
>
> *Interviewee*: We have vended the service to…to a company that created our database because, okay, when NAQC recommended we should have the MDS and all that we were already working with the vendor to manage our data. So the same vendor has received work from us to add on to the MDS and evaluation. So yes it's costing us money. The person, who manages the software, because our database is electronical, is an outside company. Yeah they create all that language so when the counsellors and the clients or the specialists and the research evaluation person go into the database you have a log in. You have a password. It's all done electronically. For example you can access the database from home. It's password protected. But we don't do that. We do it from work. So in order for us to put our data into an electronic format we've had to hire a vendor to do that, a vendor company. (KI-4)

Designated Staff for Evaluation

A specific aspect of human resources and the internal capacity to evaluate, which was mentioned numerous times, was having designated staff for evaluation. The presence of a staff member (e.g., evaluation manager) designated to evaluation and data management was a critical human resource factor that determined a QL's ability to engage in evaluating effectiveness and other evaluation activities. Not having a designated staff person made it difficult for QLs to do major evaluations, as well as to share evaluation information with other QLs. One interviewee described how it was more difficult to formalize evaluation efforts, in the absence of staff designated to evaluation

activities. Another interviewee explained how even when data had been collected it would take a long time before it got analyzed because there was no staff designated to evaluation activities. Having a designated person on staff allowed the QL to go in and look at the data on a regular basis. And as the quotation below describes, it made it easier for QLs to conduct evaluations, as opposed to simply putting together some data to report to funders.

> The QL/state health department just hired an evaluation manager and so we have a new focus as a department and there are a lot more evaluation projects that were, are going on that we're involved in and that we're focused on as opposed to oh, well, CDC wants us to report something so we've put it together. (deleted text). Having the staff person dedicated to evaluation makes it easier to do evaluation projects and there was a cultural shift in the organization in order to have this staff person as a resource. (KI-9)

A few of the interviewees mentioned that their health department had an evaluation person who was shared across different projects including the QL. This person was usually an epidemiologist and the QL was competing for his/her time. Because interviewees' comments regarding the epidemiologist were used to illustrate the lack of staff designated to evaluation, I coded the comments under this subtheme. However, it is important to note that the skills and knowledge of an epidemiologist are not the same as an evaluator. The fact that the QL only had an epidemiologist for evaluation activities also demonstrated a lack of the necessary skills and knowledge in the QL for conducting evaluations. In the following quotation, the interviewee is responding to a follow-up question inquiring why evaluation is not conducted in her QL.

It is a topic that comes up actually a fair amount, it's funny you mention it. Because CDC when we apply for the ARRA money, one of the comments back from CDC on our plan was that we didn't seem to have a, sort of had a very vague evaluation plan for the amount of projects we had planned. And our reply to that is we don't have, we really don't have the money to fund an evaluation program. And our epidemiologist that works here, is stretched in multiple different directions, because we work on the, we also do the Youth Tobacco Survey here in the program. We help significantly with the Youth Behaviour Risk Survey. We do data collection so we, we publish an annual data book. So our epidemiologist is quite pulled in different directions. So although it certainly is recognized and highlighted that evaluation is important and necessary we really don't have a lot of the infrastructure resources, to actually conduct a full and total evaluation and what we try to do is sort of, get (their QL service provider), whatever they do in (X-state) or if they have something going on for evaluation that we asked them to do the same. Collect the same data for us. (KI-16)

Another issue pertaining to the designated staff subtheme described by several participants was the difficulty in creating evaluation positions. The task of creating a position for an evaluation person in the QL was difficult because the position required prior approval from the funders, or in the case of state health departments, approval from legislators. Interviewees such as in the quotation below described the bureaucratic 'red tape' and the time required to build a case for needing an evaluation staff member. The interviewees sounded frustrated when they described this process.

Yeah. I think, you know lack of funding. Lack of staff. They go together. Building the case very slowly was the process and the problem. You know, having limited time for both our director and myself to do some sort of post outcome evaluation of ok, we did this outreach effort. Did it work? You know, and then slowly building the case that hey, you know, if we had someone here full time to do this, then we could answer a lot bigger questions more effectively. So yeah. But it's building the case to prove that you need extra funding to have a person on task full time. That was a barrier. (KI-6)

Social Resources

The interviewees mentioned a variety of relationships that both directly and indirectly impacted implementation of the innovation. I had identified the relationship subthemes coming out of the transcripts and decided to put them under the resources theme based on the Foster-Fishman et al. (2007) framework. The authors suggest that relationships are important for reasons such as flow of information and access to resources. This was certainly the case in this study, as relationships were emphasized by

the interviewees and examples provided for how they impacted flow of information and access to resources. In this section, I will focus on describing the relationships. In the next chapter I will describe issues pertaining to flow of information and access to resources.

There were five key relationships identified through the thematic analysis of the transcripts that were relevant to systems change for evaluating effectiveness: 1) QL to QL, 2) NAQC and CDC with each other and with the QLs, 3) service provider and funder, 4) QL and the state/province, and 5) QL and third party evaluation and data management contractors (see table 7). The five key relationships and their consequences for implementation of the innovation are described in the following section.

Table 7. Participant Identified Relationship Resources

Relationship Resource	Number of Participants that Identified the Relationship Resources
QL to QL	13
NAQC & CDC	14
Service provider & funder	13
QL and the state	14
QL & third party evaluation contractors	9

Relationship 1: QL to QL

One of the interview questions posed to interviewees was whether their QL shared evaluation information with other QLs in the network. Based on the interviewees' responses, not a great deal of evaluation information was shared. This information is of interest to me because I assumed that in order for the innovation to be fully implemented across the QLs, it is necessary to have strong relationships. These relationships have begun to develop over the last ten years, in large part due to the creation and efforts of NAQC, which is described in the next section. Yet, the level of communication between

QLs seemed to vary, as not all QLs communicated regularly with other QLs. The ones that had a service provider for multiple QLs seemed to communicate with other QLs indirectly, via the common service provider.

One barrier to communication and open sharing of information mentioned by several of the interviewees was the "public-private barrier." The quotation below illustrates one interviewee's perspective of the public-private barrier to sharing information.

> This is just what I see. But you know say you go to a counsellor or you attend a webinar. There's less information about what the actual protocols are in a private quitline. And so maybe it's just my lack of exposure to a webinar where they talk about it. That could totally be it. But I don't see as much information sharing about the actual client processes, the client treatment processes in a private quitline as maybe a state quitline. (KI-6)

It was also mentioned several times that there was some concern on the part of the public QLs that the private service provider contractors might at some point bid for their contracts. The interviewees described how the private-for profit service providers had different regulations, goals, and operation structure. One issue was that the private-for profit service providers can, and have, applied for contracts outside of their home state, whereas not all of the public service providers could do that. Examples were provided of private-for profit service providers winning bids for QL contracts in other states that were previously held by a public service provider (e.g., a university). Interviewees suggested that there may be some reluctance of QLs to share information across the private-public barrier for fear that they may be in a position of competing for a contract at some point. The fact that there are private and public QLs and that there is competition for funding dramatically affected the dynamics of the network and consequently information sharing within it.

Relationship 2: NAQC, CDC & the QLs

As mentioned in the previous section, the formation of NAQC as a mediating entity was critical in building communication between the QLs. In general, NAQC's relationship with the QLs was a reoccurring subtheme, as was CDC's relationship to both NAQC and the QLs. NAQC was viewed as a positive force in developing the QL environment and the proliferation of the QLs across North America. The interviewees described how prior to NAQC and the development of the MDS, there was no standardization of evaluation and little communication across the QLs. NAQC provided a platform to facilitate communication between them and that communication helped build trust amongst the QLs. Several of the interviewees stated that prior to NAQC the lack of communication between the QLs perpetuated feelings of fear and mistrust about sharing data. As illustrated in the quotation below, NAQC was a centralized member organization that provided a forum for communication and collaboration (e.g., conferences and webinars), which has allowed people to get to know each other and build relationships, cooperation, and trust.

> I think they've provided a forum where people who didn't trust each other got to sit down and know each other and I think that's produced more cooperation and more trust. (KI-7)

The interviewees also described the significant influence that NAQC had over the QLs who were motivated to implement their recommendations because they wanted to be a part of the QL community. NAQC's power over the QLs is explained in greater detail in the next chapter, which discusses the power and authority subthemes. In this section I describe the details of the relationships between NAQC, CDC, and the QLs that created the power dynamics.

120

It became clear from the interviews that not all QLs were equally engaged with NAQC activities, nor were the relationships between NAQC and QLs equally strong across different QLs and QL organizations. One issue mentioned by several interviewees was that not all QLs could afford NAQC memberships for their stakeholders (e.g., staff, service providers, evaluator contractors). Being able to provide memberships made it easier for the QL to follow NAQC's recommendations because they could be engaged in learning forums (e.g., implementing MDS). Another issue was that some QLs and their respective decision-makers valued network engagement and following NAQC's recommendations more highly than others. There were a few interviewees who described explicit instructions in their QL to be involved in the QL community and NAQC sponsored activities. In contrast, there were other interviewees who did not convey the same level of perceived importance to be involved. The quotation below is taken from an interview with a service provider and the interviewee describes how it was easy for them to implement NAQC's evaluation recommendations in large part because of their QL's emphasis on involvement in the QL network.

> The vice president of our organization had said always be involved, be as involved as possible. And in fact (S.S.), who I referred to yesterday as being the creator, really the primary creator of the helpline concept here in (X-state), was one of the primary advocates for adopting such a standardized approach. So it wasn't anything that had to be strong armed at all. It was easy to do because it wasn't much of a transition. But it was also something that we wanted to do because of the emphasis on involvement. (KI-8)

NAQC also had a strong relationship with CDC and this relationship was described by interviewees as instrumental in achieving implementation of their recommendations. The interviewee from CDC explained that their organization had a formal relationship and contract with NAQC to foster their efforts. Based on the interview data, I identified two major benefits to the strong relationship of these two entities: 1) it facilitated a

nationally coordinated effort, and 2) it provided authority for enforcing implementation of innovations. CDC relied on NAQC's expertise and in return provided the monetary incentive for the QLs to implement their recommendations. In the absence of this strong relationship, it is possible that CDC may have requirements in their funding agreements that are not consistent with NAQC's recommendations. However, because the two entities worked closely together, there was alignment of the funding mandates with the expert recommendations. The quotation below provides a specific example of the first benefit of the relationship between these two entities for the implementation of the MDS.

> I also think that just having the NAQCS, or NAQC, and the CDC pushing for the minimum data set, I think that and just sort of the more nationally coordinated effort to bring things sort of into line, I think that helps, because that means I mean that if there's this nationally coordinated effort or nationally systematic way of doing things, then that helps. (KI-13)

As a funder for the American QLs, CDC had power to leverage change in the system, especially through mandates in funding agreements. Because CDC is funding the QLs, they were in a position to require certain innovations as a requirement for funding. For example, CDC mandated the implementation of the MDS in their collaborative agreements with QLs. This mandate was described as a tipping point in the system that quickly led to all of the QLs in the U.S. implementing the MDS. However, CDC did not have an in depth understanding of how the system operates, which made a working relationship with NAQC ideal because the expertise and financial power to leverage change is a joint endeavour. The quotation below illustrates the second value of the relationship in that CDC was able to bring the authority and financial power to enforce implementation of NAQC's recommendations in the QLs.

> NAQC has no money to provide to quitlines. NAQC has no authority over quitlines. All NAQC activities are volunteer based so, well I think it probably would have been eventually been adopted by the majority of quitlines, I think it would have happened at a much slower pace without CDC's involvement. (KI-1)

Another reason why this relationship was perceived as beneficial by the interviewees was because CDC itself has also developed a partnership with the QLs and engendered an element of trust over time, which has made the QLs more willing to share evaluation data with them. Two of the interviewees suggested that this collaborative environment was in large part due to CDC's increased funding role and support for the QLs, as well as their past history of collecting and posting data for other initiatives (e.g., the national health behaviours survey). Overall CDC was perceived by participants to have built the right type of relationship to ensure that QLs felt secure sharing evaluation data with them. The following quotation illustrates how CDC had established this partnership with the QLs over time and the subsequent atmosphere of trust that had resulted from it.

> I think that also over that period of time the, the, the CDC has taken an increasing role in funding the state QLs and I think that as you develop, you know, a partnership with people over time and you get to know them and you, you've worked with them and I, you know, I think that, that develops an atmosphere of trust. (KI-7)

Relationship 3: Service Providers & Funders

A critical relationship subtheme that emerged from the transcripts was that of the service provider and funder and the need for it to be a partnership. The relationships between these two entities varied significantly across QLs and were mentioned by at least thirteen of the participants (see table 7). For example, in one QL the funder and service provider worked closely and had even published a paper on the QL together. In a contrasting example, the service provider and funder did not collaborate at all and communicated very little. In another example, the service provider had two funders, one

was hands-off and the other was very engaged, and described by the interviewee as being more like a partnership.

Several of the interviewees expressed that Canadian funders in general tended to be less involved in the QLs, in comparison to their U.S. counterparts. This information was supported by the fact that none of the Canadian funders were recruited for interviews for this study even though they had the same opportunity as the U.S. funders. The potential reasons for this difference were not clearly stated by the interviewees. Two interviewees however, provided an example of an exception to this generalization when they described a QL in Canada that was unique in that the funder was very engaged and there was a strong partnership between the service provider and the funder. This exception suggested it was possible for Canadian funders to be more involved. The quotation below illustrates the value in having a strong relationship between the service provider and funder for this province.

> The (X) QL has got its got higher reach than any other QL in Canada. And that's a result of working very closely with the whole tobacco strategy which is funded by the government and making sure that their tobacco strategy is a package that goes out to the public. (KI-11)

The service providers who had partnerships with their funders all expressed the benefits and an appreciation for them, seeming to value the collaborative relationships. Those interviewees who described partnership type relationships, also explicitly recognized the value in addressing the different needs of both parties involved. As stated by the interviewee quoted below, in partnership situations both entities should be responsive to each other's needs.

> Well I personally believe that an evaluation should answer questions that are important to the funder as well as questions important to the operator. So if you don't have a close relationship with the funder you don't know what they're expecting from your operation. You don't know what information they're looking for so then you can't guide the evaluation. So I think everything we do is a participatory evaluation. And that you know we're trying to answer questions that the

funder has or the operator has like how do we get more people to call or, you know is this the best investment of our money from a government perspective. So where I see the QL operation having a close relationship with a government funder then I see the program is more responsive. It's more responsive to the needs of the funders there. (KI-11)

One particular interviewee from a service provider organization was very adamant that a partnership between the service provider and funder was important because it allowed the QL to be more integrated into the wider state context for tobacco control. In the quotation below, the interviewee is responding to a follow-up question at the end of the interview in which I had asked for clarification on a point mentioned earlier in the interview that states with partnerships with their QL seem to have less concern about losing their funding and more willingness to share data. I inquired because I was not sure what he meant as I had assumed that most states have partnerships with their QLs in terms of a service-funder relationship. The interviewee had specified how his QL along with another public QL had partnerships between the service provider and funder.

> They work together when they need to make changes. If the QL sees that there are things that they aren't doing effectively or could be doing more effectively they meet with them, they make some decisions about how they're gonna make the change, when the change should be implemented and they share their information very openly back and forth. To…to actually make the changes be more integrated. And I see that differently than the relationships that some of the QLs have with their funder where the funder just basically said we have this money. We need you to provide this service. Here's the money. Tell us how many people you're serving. Here's the list of criteria we need from you and this is invoices and this is how we're gonna pay you. But it's still proprietary service. It's…you're buying a package basically. (KI-14)

Other interviewees described how a lack of partnerships between the two entities presented difficulties for the funder in being able to access data from the service provider. As illustrated in the quotation below, funders who did not have a close relationship with their service provider seemed to have a more difficult time getting the data that they needed from their service provider.

> I mean, it was, it was set in the contract and, and they were very good at provide, at least telling us, yeah, if you want something we can try and get it for you, but it always was, was you know, we

never, until the year eight, we never knew, and I don't know why, but we never knew, we never had a complete data dictionary. We never, we weren't able to, to figure out exactly what it is they had. If we had some, an issue to ask for more data, they would say, yeah, we, we have a data point that sounds like that and they would provide it to us, but we never really had it in front of us to look at and, and decide what we wanted. (KI-19)

Although it came through clearly in the data that a partnership was important, what was less clear was how to create those positive partnership relations. One funder participant suggested that having multiple service providers is a barrier to forming partnerships and to using evaluation for quality improvement. He explained how having multiple service providers made quality improvement efforts almost impossible because of confidentiality issues with clients and having both service providers assigning blame of problems to the other service provider. The interviewee quoted below used the term "true partnership" and explained how this cannot be achieved if the service provider has a staunch protocol that they exchange for a set amount of money. This interviewee also provided very specific descriptions of what it means to have a "true partnership" in a QL and described it as the service provider and the funder working collaboratively to make changes if things are not working effectively. He also suggested that a "true partnership" cannot be achieved in situations where a funder is purchasing a packaged service and paying to make changes to the protocol.

Yeah. This is my definition of a true partnership is working collaboratively. Now some of the agencies would probably disagree with me, that provide quitlines services, but some of the quitline service providers have staunch protocols and they basically say if you give us this amount of money this is what we're gonna give you. And then anytime you need to make a change of course we have to renegotiate our contract, which to me is not a true partnership. It's basically you go out, you find somebody who's offering the service and you buy the service, where here, and I know (X-state) their person who oversees the Tobacco Control Program there talks about this a lot, is they see their relationship with their quitline provider as a partnership…(text deleted). And even though there may be changes down the line they would say well…this is where we work as partners. You basically say our funding is being cut by three hundred thousand dollars. Then, me, as a provider, I'm gonna say okay then I have to provide one third less of the service that you bought. I don't see as much of the collaborative effort to say, okay, if our funding is gonna be reduced by three hundred thousand what else can we do with the funding that we have. Our contract is a little more flexible and we're integrated into the whole cessation process in the state so that if, you know, like this whole thing about changing focus of communities. We're a part of that. We're included in the discussions which I think a lot of times some of the quitlines, the funders make those decisions and they just tell the quitline this is who we're focusing on so you may be getting more calls from this community. But because that quitline provider is three thousand miles away they don't really participate integrally in that community when those decisions are made. (KI-14)

In the above example, the interviewee was referring to situations where a private service provider had contracts with multiple QLs and offered the same service and evaluation protocols to all QLs. There was a split of perspectives on whether or not having a service provider with multiple QL contracts was an advantage or disadvantage for a QL. It seemed to depend on how responsive the service provider was to the requests of individual QLs. There were examples of service providers with multiple QLs that did this well and also suggestions that it was done better with service providers that only run one QL.

Relationship 4: QL with the State

The relationship between the QLs and the state was another key relationship subtheme. This relationship may be similar for the Canadian QLs with the provinces, but because not much data was collected on these QLs, this section focuses specifically on the American QL relationships with their respective states. The state in this case includes

the funder (health department) in addition to the legislation and other state tobacco prevention programs. The interviewees described varying degrees of communication between QLs, their legislators, and other tobacco control programs in the state. Based on the data, the relationship of the QL to the broader state context appeared to impact the QLs in different ways, one of which was evaluation. The relationship between the QL and the state was influenced by the level of political support, as well as the structure and communication within the state. The relationship was also important for integration of the QL into the state context.

One of the major issues presented with respect to the relationship between the QL and the state was whether or not the state (e.g., legislators) was supportive of tobacco control and prevention programs. This subtheme overlaps slightly with the fear of funding subtheme that describes how the level of fear for sharing data is related to the state context. Interviewees described vastly different political contexts with respect to levels of state support. Some states were very supportive of tobacco prevention programs, respecting that funds from tobacco Master Settlement Agreements be maintained for tobacco prevention programs. However, other state legislations were not supportive of these programs and were constantly in search of reasons to stop funding them. Additionally, QLs in less supportive states were understandably more nervous about having their evaluation data, such as quit rates or population reach numbers shared publicly. The following quotation is from an interviewee who described a non-supportive state context:

Quotation #1-

> I don't think it's so much a matter of competition as with any state a lot of QL funders really have to go to battle for funding for cessation. And there are legislators or other policy makers who are, you know, continually challenging them and saying, you know, why is this a good use of money and so anything that they feel can look, you know, reflect, has the potential to look poorly on the performance of the QL they fear will be used by these, you know, critical legislators to cut funding, you know. (KI-10)

The next quotation is intended to contrast the above quotation and was taken from an interviewee who described a supportive state context. The quotation is strikingly different from the previous quotation and provides an excellent reflection of the diverse socio-political contexts that the QLs work within.

> Cessation is a fairly popular thing. It's because it's very easy for them to understand. You know? It's one of those okay we have, they recently in 2007, they increased the state tax on tobacco products. And then it depends, that was 2007, 2008 there was a smoke free air law that took affect so they see if we do these taxes and we do these you know smoking restrictions, that you need to provide services to help people quit. And it's, again it's fairly straightforward and you can kind of see the return on investment. Oh we put this money in to provide you services, thousands of people call and you know X percent quits, so you know, yay. Whereas prevention, putting that same money to prevention is a little bit more you know, you don't see a return and you can't put it into a spreadsheet how many people you've prevented. So I think it's probably, it's a reasonably popular program among the legislators. (KI-15)

There are several reasons why the QL's relationship with the state is important that came out of the data. One reason is that communication between the QL and the state can help inform the appropriate structure of the QL (e.g., multiple QL out-of-state service provider, versus in-state single QL provider). An interviewee suggested that it is necessary to be able to look at the state context to assess how the QL fits into the state and national agenda. Results of this assessment should guide QL decisions, such as what type of service provider would be better. For example, the quotations below illustrate this point when he suggests that if the state does not support tobacco control, then it would be best to go with the easiest cheapest provider (e.g., large out of state package).

> Funders are often asking that question, what's the best fit for us? And the first question you have to ask is how does your QL fit into the overall picture of tobacco control because that will help guide

you on whether you should have somebody in your own state doing it. I mean those are some of the things I think the whole national picture would really contribute to. And I think funders would be very happy if there was just more sharing of information so that they could really assess for themselves who's the best fit for me. (deleted text) There are some states that don't have a commitment to reducing. I mean this is my own, again, personal knowing the legislatures, there are some state legislatures that are not committed. They don't care whether they provide a QL or not. They don't fund QLs in some states. Some states only funding is from the CDC. You know so when you look at that level of commitment, sure. They might just want to go with the easiest, most cost effective method with one of the large providers because they do it the cheapest. And we don't really care whether they come to meetings or not. They just need to send us a report telling us how many people they are calling, and, you know, what they're doing. (KI-14)

The last interview question asked was whether respondents had anything else to add. The same respondent as above wanted to convey the importance of having a collaborative relationship between the QL evaluation team and the state evaluation team. The following quotation captures the essence of this interviewee's response, which emphasized the need for a strong relationship between these entities.

Having an evaluation team internally is nice, but unless I can integrate with the bigger picture evaluation team I don't see us really being that contributing that much to evaluation. So for me it's real important no matter what program that they're funding and they have an evaluation team. It's that our relationship needs to be pretty tight with them about what it is that we're collecting so that what we do here with evaluation can help guide them so they have useful information from us as well for their evaluation team. (KI-14)

As this illustrates, it was important to have collaboration and integration of the tobacco control evaluation team (state) with the QL's own evaluation team, regardless of whether it is internal or an external third party. The QL evaluation needed to be integrated into the bigger evaluation picture, which required a tight relationship between the QL evaluation team and the tobacco control evaluation team. The interviewee suggested that the relationship needs to be very clear, very strong, and the outcomes for that contract need to be explicit.

Relationship 5: QL & Third Party Evaluation Contractors

The final key relationship subtheme that emerged was between third party evaluation (and data management) contractors and the QLs. This relationship differed greatly across QLs and also significantly impacted a QLs ability to engage in meaningful evaluation. It was not clear based on the interview data whether it was better overall to conduct evaluation in-house or to contract with a third party entity, as interviewees saw advantages and disadvantages to both. One clear and consistent pattern in the data was that if a third party is contracted to do the evaluation, the relationship between the contractor and the QL (i.e., funder and service provider) was of critical importance.

Individuals were interviewed from QLs who described both very positive and negative relationships with a third party evaluation contractor. In the case of the negative relationship, the third party contractor was assigned through the state health department. The relationship was considered negative based on the information provided by the service provider interviewee who described a situation of extreme frustration, an inability to meet contractual obligations, a deterioration of relationships with other entities, which almost ended in the service provider loosing the QL contract. The relationship between the evaluation contractor and the QL was described as non-existent with no face-to-face interaction, minimal discussion between the two groups, and no effort on the part of the evaluation contractor to really learn about the QL. In addition, any changes to the evaluation were expensive and the evaluation contractor made it difficult to implement NAQC's recommendations for evaluation, which resulted in a huge delay in getting the MDS implemented. Also, in this case, the relationship was described in the past tense, as the QL had in recent years been successful in changing to an in-house evaluation set-up

and no longer contracted with the problematic third party evaluation entity. The quotation

below captures a fraction of the frustration that the interviewee clearly conveyed

regarding the third party evaluation contractor situation.

> They don't know what we do. They don't know where our protocols are. And yet they're being… they're putting reports out about how effective we are. And I'm like I have a problem with this because I want to see a human being who is evaluating our program, in my program. At some point I'd like to see somebody come over here, go over the protocols, meet with the coaching staff, meet with the intake staff, meet with the management team, talk about what our priorities are here and how we do things and put an evaluation plan together. (KI-14)

Although this particular case was an extreme example of the detrimental effect a

poor relationship with an evaluation contractor can have on a QL, there were other issues

that, while less extreme, also had an impact. One of the primary challenges to having a

third party evaluator was that the QLs were often detached from the data. Many examples

were provided that described situations where the QLs were unable to obtain or use the

raw data. In most cases, the QLs were given standard reports on things such as utilization

of the QL and interviewees described challenges to getting additional information from

their evaluators in a timely manner when they needed it. The issue of access to data is

described in greater detail in the 'ownership of data' subtheme, however the next

quotation helps to illustrate the consequences of poor working relationships between

evaluation contractors and QLs. When asked how the external evaluator impacts the QLs

ability to do evaluation, the interviewee responded:

> Timing can be a problem with the 3rd party evaluator agency. The QL may need information quickly, particularly during legislative sessions, and they can't get it from the evaluator agency fast enough. (deleted text) Well it just slows it down. I mean it doesn't make it impossible, as it just takes longer to get data. The whole art of communication in and of itself is a huge challenge; for example right now we're communicating verbally; what if I might request something? You understand my request as one thing, but I really needed it in another way; and when I get my request information back from you, and it really wasn't what I wanted. So then we need to try it all over again and that happens a lot. (KI-12)

Although the above examples would suggest strong evidence for not having an external evaluation contractor and instead doing evaluation in-house, there was also an equally compelling example for having an external evaluation contractor. In contrast to the above examples, there was a QL case with a very successful third party evaluation contractor arrangement. Two individuals from this QL were interviewed, one from the funding agency and one from the evaluation contracting agency. The individual from the evaluation agency was recruited through snowball sampling at the suggestion of the funding partner. It is interesting to note that no other QL stakeholder interviewed recommended that an individual from their evaluation contracting agency be interviewed. This might have been because no other QL had as strong a relationship with their evaluation contracting agency as this particular QL had.

One of the described characteristics of the case that demonstrated the advantages of a third party evaluation contractor, over those that demonstrated the disadvantages was a collaborative relationship. Both the funder and the evaluation individual described a three-way collaborative relationship between the funder, service provider, and evaluation contractor. The collaboration was built on open and regular (i.e., weekly) communication between all three stakeholders. The following quotation is taken from the funder interview and describes his perspective of the relationship between the three entities and the process they went through when they started the QL.

> It was a very collaborative process. And you know it took us all reading studies and looking at best practices to figure out how we're going to do this, share the data, you know? E-mail it across or what was going to happen? Create agreements. It was all so yeah, it was very, very cooperative. (KI-15)

Also, the evaluator interviewed from this triad relationship described clear goals of utilization for the evaluation which included improving services and tailoring reports

for stakeholders. In stark contrast to other cases, this QL funder explained how he was "swimming" in data provided by both the evaluator agency and the service provider. The QL provided a clear example of how the funder's need for particular data had been easily addressed via communication with both the evaluation entity and the service provider. The quotation below is taken from the evaluator interview and further illustrates the strong working relationship between the three entities and also the benefit of this relationship. The interviewee was responding to a specific question asking about the nature of the relationship between the service provider, funder, and evaluation contracting agency of the QL.

> As things come up because we have a good working relationship, it's easy for us to make changes and to be flexible and responsive, which I think is really important. And I mean there is a very clear structure and system with the evaluation. It is a systematic coordinated evaluation. But the fact that we have regular contact means that it's also – it's not removed. It's not like this thing happening apart. It's objective and it's external. But it's still connected. (KI-13)

One of the primary conclusions derived from the analysis was that QL service providers with external evaluation contractors assigned through the state funder agency needed to be able to communicate with the evaluators and be involved in the evaluation. Based on the interview data, it seemed that the funders needed to take responsibility for building this relationship and facilitating communication between these entities.

Economic Resources & Opportunities

The subtheme economic resources was taken from the Foster-Fishman et al. (2007) framework, but the two subthemes within this category came from the data. The overarching economic resources subtheme was difficult to tease apart from the other themes because the economic issues were often directly tied-in to other issues such as

human resources and policies. The two subthemes that were consistently mentioned by the interviewees that were related to economic costs were the cost of evaluation and funding arrangements (see table 8).

Table 8. Participant Identified Economic Resource

Economic Resources	Number of Participants that Identified the Economic Resources
Cost of Evaluation	12
Funding Arrangements	12

Cost of Evaluation

The economic costs and opportunities for evaluation varied across QLs. For the most part, federal funding in the U.S. required that a certain percentage of a grant be allocated to evaluation. CDC for example, suggested that QLs allocate 10% of their budgets. Thus in theory, all U.S. QLs should have had a sufficient budget, although several of the interviewees did claim that a lack of financial resources was a barrier to conducting evaluation. Also, for the majority of the QLs, funding was a barrier to doing more complex evaluations. The following two quotations illustrate the points identified above with respect to financial resources and implementation of evaluation practices.

Quotation #1-

> It is a challenge and I think it has to do with a lot of economics. Every time, you know, funding is being reduced one of the first things that goes is, you know, service always takes precedent and evaluation and question asking things seems to fall, and training, all seems to fall to the side. (KI-14)

135

Quotation #2-

> *Interviewer*: Is there other evaluation that you'd want to do but can't for any particular reasons?
>
> *Interviewee*: Oh absolutely. I think funding is an issue, so typically you get your funding to run the quitline and then you get a percentage of what the overall funding is that's used for evaluation. So that somewhat limits you in what you're able to do, especially if you have ongoing monitoring. So it all costs X amount of dollars to produce the monitoring reports. Every year. I mean right now we're doing some sort of a monitoring with a recorder for them. But then they certainly have additional, more complex evaluation that'd be nice to do. (KI-18)

Overall what emerged from the interviews was that economic resources was not a simple and consistent barrier to evaluation across all of the QLs and the degree to which it was varied in relation to QL characteristics such as structure and funding arrangements. For example, QLs that contracted with external organizations for evaluation and database management services incurred higher costs to conducting evaluations and modifying evaluation protocols. In contrast, QLs that did all the evaluation activities in-house considered economic costs as less of a barrier. QLs that considered this factor to be a barrier also saw the evaluation budget as competing with the service budget. Thus, in order for a QL to allocate funding to evaluation instead of service, there needed to be a clear benefit. The same issue existed for making changes to current evaluation systems. For example, the interviewees described how a barrier to QLs adopting the MDS initially was a lack of clear benefit to changing evaluation protocols that already worked well up to that point. This is illustrated in the two quotations below, taken from two different interviews.

Quotation #1-

> It takes resources to change your data collection system, and your database and add new fields, and add new questions or change questions. And so, you know there's a resource allocation issue, do we want to spend money on, on changing our evaluation questions? Or do we want to spend money on providing services to tobacco users? And that might be very different depending on the goals and the needs of each quitline. (KI-1)

Quotation #2-

> And a lot of that will come down to funding because when programs often say we don't get any money to do evaluation. We provide service. We don't get money to do evaluation. So one of the challenges is always who's gonna pay for it, you know. (KI-14)

Funding Arrangements

A final issue related to economic costs that emerged from the data was the type of funding arrangement a QL has. The interviewees described how the funding sources and regulations that accompanied the funding varied across the different QLs. For instance, QLs had different rules for what they could spend their funding on and not all funders would pay for evaluation. The federal government in Canada for example, stopped providing evaluation funding for the QLs based on new regulations stating that major evaluations would only be done of programs that Health Canada funded and those evaluations were to be done internally by the government evaluators.

Despite CDC's increased role as a funder, there were still significant differences in funding arrangements across the QLs and these variations impacted evaluation efforts. For example, some QLs had a guaranteed set amount every year from tobacco Master Settlement Agreement funds, while others were funded through tobacco taxes. One interviewee described how his/her QL funding depended on the amount of services that they had provided the previous year. This QL was required to track and report utilization to legislation which was then used to determine QL funding amount for the following year.

Finally, some QLs used an inter-agency agreement mechanism for negotiating new contracts. This could happen with universities because it was considered a government-to-government relationship. In this mechanism, the university did not have to

bid for the contract. A past example was provided where this had a negative impact, in that the university contractor did not maintain high quality service. A current example was provided where this arrangement was extremely beneficial where despite the fact that the university was not required to compete for the funding, the service providers still made a conscious and explicit effort to maintain the quality and competitiveness of their services, as if competing for the contract.

Summary

System resources emerged as dominant themes in the data as factors impacting implementation of the evaluation innovation. Although the Foster-Fishman (2007) framework provided broad overarching categories to consider, the data provided the more specific issues that are unique to this innovation and to the system. For example, the data provided an interesting subtheme, designated staff, within the human resources subtheme.

Second, relationships are not focused on in the implementation literature, especially not positivist studies. However, it was clearly a dominant subtheme coming from the qualitative data. Relationships affected many other parts of the system and are an important leverage point for change. The Foster-Fishman framework was helpful in suggesting that the relationship themes be placed in the system resources theme, but the specific nuances of the relationship themes add to the framework when applied to the evaluation innovation.

The economic resource subtheme was the least dominant in the data and appeared to have less influence over implementation success. Anecdotally, economic issues are often cited as primary barriers to implementing innovations and particularly to

conducting evaluation. The findings of this study question this assumption and illuminate the complexity of this issue. The economic issues were most difficult to tease apart from the other themes because they were very interdependent to other system factors influencing implementation. For example, the QLs budget is directly related to the state context and support for tobacco control programs. Although economic reasons may be the most obvious and immediate barrier to implementation, change efforts must be directed towards other factors influencing the economic factors.

The next chapter is the final of the results chapters describing the findings from the thematic analysis. Two overarching themes are combined in the final results chapter, the system regulations and system operations.

Chapter 7. System Regulations & Operations

Chapter Overview

This chapter is the final of the three presenting the results of the thematic analysis. In this chapter, I describe the subthemes and information under system regulations and system operations. The first part of the chapter is dedicated to the system regulations and following that is the system operations information.

System regulations was a key feature of the Foster-Fishman et al. (2007) framework which refers to the policies, practices, procedures and routines that guide setting members to act in a unified manner. According to the authors, system change requires that the regulations align with the desired change goal. "System regulations that indicate what will be rewarded/punished are viewed as a particularly powerful regulatory mechanism and should be given careful consideration during a system change endeavour" (Foster-Fishman et al., 2007, p. 209).

The system operations concept also comes from the Foster-Fishman et al. (2007) framework. The authors suggest that there are a multitude of different operations governing system patterns, such as procedures for communication, and processes for decision-making. The power and decision-making operations are emphasized in this systems change framework. Power can take many forms and it can be both formal and informal. Power can be derived through formal policies, practices, and procedures, or less formally through an entity's reputation (Foster-Fishman et al., 2007). Power can also be derived through relationships and through an agency's ability to control information and resources. "Overall, an examination of power within a system focuses on influence-

specifically, who and what influences, how resources are distributed, how actions are carried out, and how decisions are made (Foster-Fishman et al., 2007, p. 209)."

System Regulations

Two subthemes were created based on the information from interviews that fit under system regulations: 1) funding mandate to conduct evaluation, and 2) mandate to collect MDS data. The two subthemes were both mandates in the system mentioned repeatedly by interviewees (see table 9). These mandates were described by interviewees both explicitly and implied as having an impact on the implementation of the evaluation innovation in the QLs. Similar to the above economic resources subtheme, the regulations category was difficult to tease apart from the other themes, and as a result, only two subthemes were placed in this category.

Table 9. Participant Identified Mandate Subthemes

Mandate Subtheme	Number of Participants that Identified Mandate Subtheme
Funding mandates to conduct evaluation	8
Mandate to collect MDS data	12

Funding Mandate to Conduct Evaluation

In recent years the federal government and the state funders for the American QLs have begun to mandate evaluation as part of funding. Many of the funders required that QLs allocate 10% of their budget to evaluation. In addition, QLs often needed to include evaluation results in other funding applications and were rewarded for having good outcome data (e.g., quit rates) with further grant funding. As a result, the QLs have begun to prioritize evaluation of outcomes because the funder and federal grants require it and

reward it. The following two quotations from a service provider illustrate some of the key points presented above. The quotations suggest that mandates to conduct evaluation were becoming part of federal and state funding, and that these mandates in turn were changing the climate from one that viewed evaluation as a luxury to one that considered it essential.

Quotation #1-

> Well I think it's changing with new federal grants that come out. Evaluation is a component which is built into the grant. You have to show this percentage of the grant is going to be dedicated to evaluation. How are you going to do it? But federal grants didn't always come out with that requirement. (KI-6)

Quotation #2-

> And so I think state health departments didn't always require it either. They kind of follow the federal lead. And so I think you find the climate is changing drastically and rapidly. And I think for the better, for that very reason. So we have information that we can share. But I would say that's the nature of the difference in philosophy is just that people are saying well the federal government isn't requiring it to give you money, so we're not going to require it to give you money. I think evaluation for a long time is almost seen as a luxury. (KI-6)

Mandate to Collect MDS Data

Another mandate that was perceived as a "turning point" in the system was the implementation of the MDS. The interviewees provided a picture of the important role that the development and implementation of the MDS played as it provided a common language and standardized system for evaluating outcomes. This critical mandate and its importance are illustrated in the quotation below.

> We're in a very different place now, especially with the introduction of the minimal data set. I think that was a really big, oh, turning point for the field. And for the network, just because all of a sudden with the, with the creation of a standard set of intake and follow up questions, we now had a common language to start talking about evaluation and evaluation findings, that we didn't before. And setting some standards about when the follow up survey would be done. What questions would be asked? (KI-1)

One of the relationships presented in the social resources section above comes into play with respect to this mandate subtheme. According to the interviewees, NAQC was primarily responsible for the creation of the MDS, however, because NAQC did not have the funding to facilitate or the power to enforce that QLs implement the MDS, the change would have happened more slowly had it not been for the adoption of this key regulation. The regulation, which was viewed as a tipping point in the system by the interviewee quoted below, was that CDC put the MDS into their collaborative agreements, which facilitated the institutionalization of the innovation. Once the MDS was mandated in the collaborative agreements more QLs started implementing it and the MDS quickly became a service norm, which motivated other QLs to implement it as well. The following quotation illustrates how the regulation from CDC helped shift the MDS to a service norm in the QL system.

> I think again having CDC putting into people's collaborative agreements, as the larger service providers started to adopt MDS, it started to become more of a service norm. That this is, more and more states started putting it in their contracts with their service providers. That anyone who's bidding to provider services for our state must oh, be asking all of the minimal data set questions. And so, as it became a standard in RFP's for quitlines services, it started to become much more of a norm across the quitline community. That any changes that you make as you're considering changing service providers, if you have been implementing or using the MDS, you start at that point. So that sort of became the process by which a lot of the quitlines started using it. (KI-1)

System Operations: Power & Decision-Making

The results of the thematic analysis suggest that the participants identified four entities with significant power in the QL community: NAQC, CDC/federal government, state funders, and service providers (see table 10). The following section describes the type and source of power that each of these entities has based on the information provided by the interviewees.

Table 10. Participant Identified Power Subthemes

Power Subtheme	Number of Participants that Identified Power Subtheme
NAQC	6
Conformity & Inclusion	6
CDC & Federal Government	15
State Funders	7
Service Providers	8

NAQC

NAQC had significant power and decision-making authority in the QL community, albeit primarily informal power stemming from its relationships and reputation. Interviewees described NAQC as a "centralized member organization" and viewed them as a partner and an "expert resource". As discussed in the relationships section, NAQC did not have the same type of power that CDC had as a funder, but they did have strong support from CDC. The two organizations had a strong relationship and contractual arrangements for CDC to support NAQC's efforts. It was described that CDC often defers to NAQC for many of the system decisions in recognition of NAQC's in depth understanding of the system. For example, the interviewee from CDC explained that the organization did not want to over burden QLs with data so they opted to request similar data to what NAQC recommends.

The following quotation is taken from an interview with an evaluation contractor and is in response to a follow-up question in which I asked if they considered NAQC a mediator between CDC and the QLs. This quotation illustrates how NAQC was perceived as a credible expert, as well as the power that this reputation provided.

> Not mediation but they provided they were seen as the expert, influential, organization, the industry standard you know so you...NAQC has a lot of credibility and if NAQC says this is what we think quitlines should do, most of the quitlines will do it. So that reinforced the need for standard evaluations for quitlines, for all the quitlines, but especially the quitlines that were not working with us. And they could access that standard evaluation from NAQC. They didn't have to come to us. (KI-11)

The interviewees used terminology that implied that following NAQC's recommendation was considered the path of least resistance, as well as having the greatest pay-off. For example, one explained how they were "moving with the wave", by following NAQC's recommendations. Many also suggested that following NAQC's recommendations was a protective move. Specifically, there was an incentive to follow the guidelines and what one interviewee referred to as "the gold standard of evidence-base" so that their future funding applications would have a better chance. For example, many of the interviewees explained how calculating quit rates could be controversial so QLs would choose NAQC's recommendations and benchmarks for calculating and assessing their quit rates. Also, because QLs did not want to overburden themselves with what one interview described as "too many tools," they preferred to implement NAQC's recommendations, which tended to become the standard. The following quotation illustrates several of these points.

> We wanted to have data that could be useful not only to the helpline but also to the outside community. So we wanted to make sure we were on the same page with NAQC's guidelines, that we were not left out, so as far as the North American QL Consortium recommends QL should collect data that is kind of the same across the board. That way when we're seeking funding or we're doing research or we're reporting and sharing data with the community we have equal data across the country. So we didn't want to be left out. (KI-4)

Conformity & Inclusion

A subtheme that emerged in relation to NAQC's power was that QLs followed NAQC's recommendations in part because they were afraid of being left out. QLs were

incentivized to follow NAQC's recommendations to ensure that they were included in the

QL community and conformity and inclusion were powerful motivators in the system.

When asked why her QL decided to implement the MDS when they already had their

own evaluation protocol in place, this interviewee responded with the following quotation

that exemplifies the power of conformity and inclusion in the QLs.

> … we just don't want to be left out when it comes to the QL community. We want to be on the same page with everybody else. (KI-4)

One interviewee described how his QL had changed an existing protocol to prioritize

doing seven-month, thirty-day point prevalence abstinence for a random sample of all

program participants, primarily because they wanted to conform to what other QLs were

doing. Several other interviewees explained how they also wanted to be able to

participate when it got to a point where reports were being generated and comparisons

being made, as illustrated in the quotation below.

> Well I mean the value of conformity is that, you know, when all the states provide their data then reports will be produced, you know, at some point and so we'll sort of be able to compare and so we wanted to be able to participate in that. (KI-10)

Federal Government & CDC

There were many instances in the interview data to suggest that the U.S. federal

government held significant power in the QL community, especially in recent years since

it had taken an increased funding role in the QLs. It was through their control of

resources (i.e., funding) that they had come to control decision-making and information

in the QLs. Even the decision-making at the state funder level was impacted by the

federal government's decision-making, because the state funder's requirements were

based on the federal government's requirements. In particular, CDC was viewed as an

organization with significant power to enforce regulations and many of the interviewees suggested that if there was a CDC mandate (e.g., evaluation requirements) then it would be done. This suggested that CDC had the potential to leverage change in the system as the funder. For example, the quotation below illustrates how the MDS quickly became a service norm after the CDC put it into their collaborative agreements with QLs.

> So one of the things at least in the U.S. that was very influential in most of the quitlines adopting them and using them for their data set was that the CDC put the MDS in their what do they say, call it? Collaborative agreements? So CDC is one of the oh, they provide I don't want to say a lot, but they provide funding to most U.S. quitlines. And if you want CDC funding for your quitlines you need to agree to their terms. And part of their terms I believe in 2005 or 2006, were that quitlines adopt and start using the minimal data sets. And so some quitlines took longer to do that than others but most of them, all of them at this point have implemented it. (KI-1)

More recently and as a result of the new American Recovery and Reinvestment Act (ARRA) funding, CDC was requiring U.S. QLs to start reporting MDS items and follow-up evaluation information on a quarterly basis for two years starting January 2011. The ARRA funding was clearly perceived by all interviewees as being a strong incentive for the QLs to conduct more evaluation. Even interviewees from QLs that were reticent about sharing data explained that they would conduct and report the necessary evaluation data required as part of receiving the ARRA funding. The power of the ARRA funding and subsequently the federal government was clearly illustrated in the following two quotations provided by two different interviewees. The first quotation refers to the enactment of the new evaluation data gathering and reporting mandates in the QLs. The second quotation comes from an interviewee in a QL that had stopped conducting evaluation due to budget cuts. As explained in the quotation, the QL planned to restart the evaluation because it was a requirement of the ARRA funding.

Quotation #1-

> Yeah, I mean, if, if you wanted ARRA money, you had to give up, give the CDC the data. You didn't have a choice. So if you wanted ARRA money, you had to agree to do that. So ultimately, I think that's probably the only way they could have pulled it off. (KI-19)

Quotation #2-

> The only reason that we're considering or that we are moving forward with the follow up evaluation now is because it's a requirement under the ARRA grants. (KI-9)

In addition to CDC's ability to mandate QL requirements via funding mechanisms, they also had power through relationships and trust that had been built with the QLs. At an operational level the QLs had a lot of trust in CDC, particularly with respect to data. One participant suggested that there might be greater trust that CDC would protect the QL data and "keep a tighter reign on things". This was in contrast to how people felt about some of the past efforts to aggregate evaluation data across QLs.

State Funders

The interview data suggested that the state funding agencies in the U.S. also had a great deal of power and authority over decision-making and information. Although the degree of power, as well as the level of involvement that state funders had in decision-making for evaluation, varied across QLs. One important fact that I learned from the interviews is that the state health department or other state governing body (e.g., Tobacco Evaluation Review Board) could mandate third party evaluation contractors for the QLs. In such cases, the state was selecting and hiring the third party evaluation contractor, which was one way that they had significant power over the evaluation process.

The state funders were also responsible for approving requests for evaluation staff positions, as well as final decisions on the type of evaluation that was conducted. In these ways, the state funders philosophy and perspective on evaluation was very important, as it directly influenced the QL's evaluation. The following quotation was taken from an interview with a service provider representative and the individual was responding to a follow-up interview question asking how his QL differed from other QLs in such a way that it allowed them to conduct meaningful and more complex evaluations. The interviewee responded by describing how their state funder was supportive of the QL's complex evaluation, which allowed them to do this type of evaluation. This interview, along with others, suggested that having a funder who was a champion for meaningful evaluation was of critical importance.

> And the other reason is that our state funders get it. They get the value of that. They really appreciate that, you know, we have a twin focus. One focus is on providing as good a service as we can but the other is on doing as good a research as we can so that we cannot only have some, not only so that we can have greater certitude that we're making a difference here in (X-state) but that we can influence how service is provided outside of (X-state), you know, through contributing to the scientific literature, presenting at conferences and so on. So they're, our funder, our main funder is very supportive of our research agenda and proud of it in fact. And I think that you see that also in other parts of the (X-State) Tobacco Control Program not just in the QL. (KI-10)

The funder perspective was critical because they determined funding for the QLs and allowable expenditures. Several of the interviewees suggested that funders were only interested in monitoring outcomes. However, there were some instances, as illustrated in the above quotation, where funders were also interested in understanding the context of the outcomes and conducting more complex evaluations to further the evidence-base and use it for quality improvement purposes. This interviewee also described how their funder was different because they understood the importance of working with the service provider to answer questions.

Overall, there was significant variation in the perspective of the funders described by interviewees and because funders had significant power and decision-making authority, their perspectives were vital to achieving implementation of the evaluation innovation. An example of this power was provided by one interviewee who described how a change of staff in her funder organization had allowed the service provider to move towards a more meaningful and complex evaluation. The quotation below illustrates how the positive changes in the QL evaluation had occurred as a result of changes in staff at the QL funder organization.

> I think the change has been occurring slowly for maybe the past year and a half. I would say that it is a direct result of the current staff at the Department of Health Services and their philosophy around strategic planning and program improvement. (KI-6)

There were also specific actors in the state system who had significant power and authority to make decisions that impacted the evaluation in the QLs. Some of these actors mentioned by interviewees included the governors of the state, who had the power to support or not support QL/tobacco prevention, as did legislators. At the state department level, individuals such as bureau chiefs and other higher ups could be local champions for or against the evaluation. One interviewee described how her bureau chief was very data oriented and willing to invest funds into data collection activities such as focus groups to create media campaigns. Overall, it appeared to be important to have support for evaluation from key people in the health department in positions of power and to have individuals that were prepared to fight for it.

Service Providers

The last entity identified in the interview data as having power and authority over decision-making for evaluation was service providers. The level of power they had varied across QLs, with some having significant power and decision-making authority and others having none. There were three primary means described by interviewees by which service providers garnered power in the QLs: expertise, access to resources, and ownership of data.

There were a few cases in which the individual highest up in the QL service provider's organizational chart was also an individual who had significant expertise and stature in the QLs. The position, expertise, and reputation of this individual gave power to the service provider in the QL relationship. In these cases, the service provider tended to have more decision-making capacity in comparison to service providers who were not championed by a highly reputable individual from the QL community. The power of one individual with high status in the service provider organization could push the evaluation in one direction or another. For example, if that individual identified research as a priority, then it was likely that data would be collected for research purposes.

Another way that a service provider could attain power was through ownership of the call management systems and database systems that were used to collect the evaluation data. It takes significant resources to develop the necessary infrastructure to manage telephone calls and to collect call data. Those entities who had already developed these resources were in a position of power over those who had not. There were several examples provided where one entity contracted with another entity who had a call management system or an evaluation database primarily because they did not have the

resources to develop their own. In cases where the service provider was 'for-profit', there

was an additional power issue related to how much they charged the funder for their

resources. For example, a service provider could influence the state funder regarding

what to evaluate based on the services they offered and how much they charged. This

process of negotiation between the funder and service provider is illustrated in the

quotation below which demonstrates the informal power of the service provider in the

decision-making process.

> We have, historically, that's been a separate line item in our contract with (X-service provider).
> And so they, we tell them okay we want, this is our standard evaluation. We go back and forth on
> what an average, what a realistic sample size is and they come up, they come up with what a quote
> for that would be. (KI-9)

The final means by which service providers garnered power in the QL community

and within QL relationships was through ownership of data. However, this issue was not

specific to service providers and was a strong recurring subtheme throughout the thematic

analysis. Therefore, the issue of ownership of data is discussed in an independent

subtheme below.

Information & Resources

Foster-Fishman et al. (2007) suggest that an important issue when considering

how to achieve systems change is the types of information and resources that are most

important to the system and determining who controls them. According to these authors,

power often comes from being able to control needed information and resources. Using

this understanding of power, two subthemes for this theme emerged from the analysis of

the interview data: 1) ownership of data, and 2) type of data collected (see table 11).

Table 11. Participant Identified Information & Resources Subthemes

Information and Resource Subtheme	Number of participants that Identified the Subtheme
Ownership over data	11
Type of data collected	8

Ownership of the Data

Ownership of the data collected was an important subtheme that emerged from the interviews and was not mentioned in the Foster-Fishman (2007) framework. There were many examples provided where one entity owned QL data and restricted the access of other QL partners to the data. The three entities that assumed ownership roles of data were state health departments, service providers, and third party evaluation contractors.

The majority of interviewees claimed that the states owned the QL data, but they described different regulations with respect to the service providers accessing and using the data. Some states allowed QLs open access and use while others did not. In one case, the state health department owned the QL data and did not allow the QL service provider to access, publish, or design research questions around the data. The interviewee in one case explained how his QL was not allowed to share data from evaluation results publicly, even though he described the QL as performing very well. It was a state regulation that data could not be shared, regardless of the results being positive or negative for the QL's reputation. The majority of those interviewed said their state had no such data sharing restrictions, however this regulation is hugely misaligned to the desired change as it did not allow for any sharing of evaluation data. The next quotation comes from a service provider interviewee who was responding to a question inquiring into how evaluation in the QL network could be improved. The interviewee clearly described how her QL was not allowed to share evaluation data. Earlier in the interview, she had

explained how in addition to not being able to share data, they were also not allowed to use the evaluation data or publish results based on the data. In the quotation below, it is also possible to see how there were multiple interconnected factors impeding the evaluation process including the ownership of data, a lack of formal processes for sharing data, and a lack of evaluation staff for the process.

> And maybe the problem is just that other quit lines like us you know we didn't have a problem sharing our information, but our State Health Department didn't let us for a long time. Maybe that's the hold up. Or we didn't have a problem sharing information and finally our state didn't have a problem with it, but we didn't have a formal process for reporting it yet, until we get this new quality improvement manager. (KI-6)

In cases where the ownership of the data lay in the hands of either the service provider or the third party evaluation contractor, the restricted access was achieved in a more subtle manner. The primary issue expressed by both funders and service providers with respect to the evaluation contractors' ownership of data was that it took too long for them to respond to data requests. The service provider and funder were at the mercy of the evaluation contractor in terms of analyzing data and providing them with reports because they did not have access to the databases themselves. However, this issue was not consistent across all QLs because as evidenced in the quotation below from a funder, there were some evaluation contractors and service providers who were very good about providing data to their funder partner.

> But for the most part we haven't had any problems with data gaps or requests for things that just don't exist. Though admittedly I get a lot of data from my provider, from my, both my evaluator and my service provider. So I'm swimming in data between the two of them. (KI-15)

The service providers' ownership of data posed similar and slightly different issues from the evaluation contractors. The interviewees described how many of the service providers did not share raw data with their funder and instead provided them with

summary reports. The problem with summary reports was that the funder could not manipulate the data to explore other evaluation questions. The barred access to raw data was not consistent across the QLs, as there were also many examples of situations where the service provider freely shared raw data with the funder. One service provider described how all their data were available online to allow their funders access to it anytime and to create reports as needed. In contrast, the quotation below illustrates a very different situation with accessing data where the funder conveys frustration when he responded to my questions asking how much control he had over the databases/data and making changes to it.

> None. I just request changes to data and they provide it, sometimes I get charged extra for it, sometimes I don't. (PAUSE) That's interesting, we identified a long time ago it works good for them and not necessarily well for us, as in many ways we become so disconnected from our own data. We own that data. We almost become in one way, captive to our vendors; as we pay for that data, but can't even access it. It's not a good thing. I get monthly data from the service providers and I reactively update it on a monthly basis. I have made my own database from those reports, so I'm able to work with it to a limited extent; but I can't physically manipulate their data. Typically, we don't receive data in a format that is usable for me. So I have to go to them when I need better data. I've actually requested data in several different ways this past month from both our service provider and from our evaluation provider. (KI-12)

Another issue was that the funder did not always know what data was contained in the database because they had not been given a codebook for the database or access to the database. As a result, the funder had to rely on the service provider and in some cases the evaluation contractor to suggest evaluation data to include in reports. This situation made it difficult for the funder to get the data necessary to conduct more complex evaluations. In some situations, it was also difficult for the funder to add questions for data to be included in the service providers' database.

A final issue related to power and data ownership was unique to QLs that contract with a service provider for both the service and evaluation of the service. In one interview

in particular, the funder described how they contracted with their service provider to conduct the evaluation. Although the funder would have preferred to have an external agency conduct the evaluation, they were bound by the costs that would be incurred if they attempted to switch. In this case, the service provider had a great deal of power over the funder in terms of data and resources. There was also a potential bias in having a service provider, particularly a for-profit service provider, conduct its own evaluation. In this case, the service provider owned the evaluation database and provided the funder with summary reports, but not the raw data. Overall, ownership and access to data appeared to be a significant subtheme throughout the interviews. And lack of shared ownership and access to needed data also proved a barrier to conducting meaningful evaluations.

Type of Data Collected

It is important to identify where power and authority to make decisions exists in the system because it directly influenced the type of data collected, how it was accessed, and how it was used. Furthermore, the type of data collected determined the results and subsequent potential utilization and impact. The quotation below illustrates this point as the interviewee describes how the federal government had taken over a significant funding role in the QLs and how they were interested in reporting for accountability and job creation.

> It's about funding and this is being funded by federal government stimulus funding and they are concerned about creation of jobs primarily. And so they want to measure creation of jobs and they want to be able to show that they're having an impact in the service area. So they want to know about how may quitters are you generating? And that's really all they want to know. They don't really want to know about oh, how do you improve services? They're not interested in doing experiments to find out or even natural experiments or observational experiments to be able to take into consideration the fact that there are different protocols going on, or some people get more calls than others. They just want to know how many people quit, and how many jobs are

created. And that's it. And that's what this stimulus funding is about and so, you know what gets funded, gets measured. What gets measured gets done. And so if you, this database is not being created as a research database. This database is being created as a reporting tool. And I think that's the biggest difference in my mind. (KI-1)

One of the primary problems of collecting and reporting outcome monitoring data was that reporting simple outcomes did not take the context of the QLs or their unique differences into account. Interviewees expressed the importance of considering these differences if there were to be any comparisons of results made across the QL outcomes. Examples of the types of factors that varied across QLs and should be considered for comparisons included live call rates, population demographics (e.g., Medicaid), and state context (e.g., tobacco policies). An example provided with respect to state context was that it would be unfair to expect a QL in a state with no anti-tobacco policies to have the same prevalence and quit rates as a QL in a state that has anti-tobacco policies. This general consensus for the need to consider QL context and individuality is reflected in the quotation below from a service provider.

> If an evaluation is truly going to be effective with its results it needs to take into consideration those differences within whom they're evaluating. If you answer the question what is your quit rate and you answer that collectively for all the calls you get from one state over a period of time that's one thing. If you look at what is your quit rate among lower socio economic or people with college education or whatever else then that to me would give you a lot better view of how that QL functions. So sometimes the evaluation is the problem. It's not the information that we have but it's actually what's being asked or what's being looked at in the evaluation that doesn't give the broad spectrum of the individuality. It's not how the information is collected. (KI-5)

A second issue described by interviewees was that by neglecting to collect contextual data, the data could not be used to inform quality improvement. Several suggested that a database for reporting outcomes would not be as valuable to decision-makers as a database that provided answers to questions for quality improvement. For these reasons, the data about QL services and context should be included in the database

so that it can be used to inform decisions to improve practice and policy. The quotation

below taken from a funder illustrates the perceived need for more contextualized

evaluation data.

> I, you know, without, without context, a lot of this data doesn't, doesn't mean a lot, doesn't mean what it could. You know, things that I'm interested in is, you know, who, who or what causes people to stay on QLs longer, attend more counselling sessions. You know, what kinds of protocols are creating the best outcomes. You know, I, I have no reason to doubt a five-session protocol, but is a five-session protocol needed, for example? Could it be done in two or three? And if it were two or three, would people stay on? (KI-19)

Summary

This chapter provided information on the subthemes under system regulations and

system operations. Both of the subthemes in the system regulations theme were mandates

that had facilitated the implementation of the evaluation innovation over time. Mandates

for evaluation in grant funding were creating a shift in the evaluation climate towards an

expectation of conducting evaluation. The other mandate in the system was CDC

requiring MDS as a stipulation for funding, which was perceived as a 'tipping point' for

standardized evaluation in the system. Once the MDS was mandated in the collaborative

agreements, more QLs started implementing it and the MDS became a service norm,

which motivated other QLs to implement it as well.

The second overarching theme presented in this chapter was the system

operations. The results of the thematic analysis suggested that there are four entities with

significant power in the QL community: NAQC, CDC/federal government, state funders,

and service providers. NAQC had significant power and decision-making authority in the

QL community primarily stemming from its relationships and reputation. A subtheme

that emerged in relation to NAQC's power was that QLs followed NAQC's

recommendations in part because they were afraid of being left out. The CDC/federal

government had significant power in large part because they were funders. For example, the ARRA funding had provided a strong push for evaluation. State funders also had significant power with respect to evaluation, in particular because they determined the type of evaluation that was conducted in the QLs. The last entity identified as having power and authority over decision-making for evaluation was the service providers. The level of power they had varied across QLs, with some having significant power and decision-making authority and others having almost none. There were three primary means described by interviewees by which service providers garnered power in the QLs; expertise, access to resources, and ownership of data.

A final subtheme under system operations was information and resources. Two subthemes emerged from the analysis of the interview data under this theme: 1) ownership of data, and 2) type of data collected. Ownership of data was a dominant subtheme mentioned numerous times throughout the interviews. There were three entities that assumed ownership roles of data: state health departments, service providers, and third party evaluation contractors. Ownership of data was taken by denying access to data, providing limited access to data, and by providing data in limited formats. Ownership of data was an important barrier to address in order to improve implementation of the evaluation practice.

Another subtheme under information and resources was the type of data collected. It was important to identify the power and authority in the system because it was directly tied to decision-making to guide the type of data that was collected. The type of data collected, including contextual data, had implications for making comparisons between QLs based on the evaluation data. The system operations theme in particular provided a

wealth of information and insights to improve the implementation of the evaluation practice.

In the following discussion chapter, interpretation of this information in the context of the literature review will be presented, along with strategies for improving implementation in the QL system.

Chapter 8. Discussion

Chapter Overview

In this chapter, I reflect on the study findings in relation to the literature review and the research questions. I begin by presenting an overview of how the results provide information to address the first research question, which asks about the factors influencing implementation. Next I describe how the findings provide information on the second research question, which asks about how the system structure and dynamics influence implementation of the innovation. Included in this section is an explanation of how the complex system principles discussed in the literature review facilitated a deeper understanding of the implementation phenomenon. The section following this is a summary of the findings for research questions one and two which includes a figure illustrating the results described in the previous three chapters. The final section specifically addresses the third research question by presenting various strategies including complex system management tools and leverage points. Discussion of the complexity principles is interwoven throughout the chapter. In fact, it is difficult to address each research question entirely separate from the other questions because they are inherently interconnected. As such there will be some overlap in discussing the different research questions.

The overarching goal of this study was to contribute to the quantitative KIQNIC study findings by using a qualitative systems approach to study implementation of the evaluation innovation in the QLs. The KIQNIC study applies a positivist approach to studying the phenomena through the use of quantitative survey data. The underlying assumptions of the KIQNIC project are that there are linear cause and effect relationships

between variables that represent a 'true' reality that can be identified based on mathematical logic and that these results can be predicted and reproduced in different contexts. In contrast, the interpretivist approach does not aim to identify predictable linear cause and effect relationships that represent a 'true' reality and instead views social processes as complex and grounded in unique contexts that are constantly changing. Complexity science has characteristics of both approaches. For example similar to interpretivism it does not aim for reproducibility or predictability of specific outcomes or consider it possible to identify a 'true reality'. However, similar to the positivist approach, it assumes that there are identifiable causal relationships that create patterns in the system, but the difference is that these relationships are nonlinear in complex systems.

This qualitative study if combined with the KIQNIC study at a later date would reflect a mixed-methods integrative approach to studying the implementation of the evaluation innovation (Van de Ven & Poole, 2005). However, the integration of the quantitative and qualitative results has not been done yet as it is viewed as beyond the scope of this study and proposed as next steps in the 'Recommendations for Research' section. This study aims to observe, reflect, and describe the phenomenon in a theory driven way as opposed to attempting to identify specific determinants and predicting outcomes (Greenhalgh, 2010). In doing so, the findings of this study provide valuable insights into the implementation of the evaluation innovation to address the three research questions.

Research Question 1

What are the factors influencing implementation of the innovation?

The findings of the thematic analysis suggest that there are numerous factors influencing implementation of the evaluation innovation in the QL system. These factors come from different levels and niches of the QL system as well as the innovation itself. For example, there are factors related to individual actors and organizations, as well as the social, cultural, and political context. An example of an individual level factor that influenced implementation of the evaluation innovation was norms about evaluation. This finding is consistent with other literature, particularly with Foster-Fishman's (2007) suggestion that there are deep structures in the system, which are critical for systems change. Perspectives of the innovation (i.e., goals and definitions) in particular were an important normative element, which is consistent with literature on other systems change efforts (Hargreave, 2008). Moving to the organizational level, my findings align with previous work, suggesting that additional factors influencing implementation involve having staff designated to evaluation and availability of economic resources for conducting evaluation (Gibbs, Napp, Jolly, Westover & Uhl, 2002). Finally, my findings suggest that there are many different contextual factors such as social, cultural, and political contexts that influence implementation of the innovation. Similar to previous research (Kreger, Brindis, Manuel, & Sassoubre, 2007; Suppovitz & Snyder, 2005; Netting, O'Connor & Fauri, 2007), I found that the funding culture was part of the cultural context that impacted implementation and funding was viewed as a driving force in conducting evaluation. Thus, it is important to consider these contexts and how differences between QLs and their contexts influence implementation, as these differences were a key factor influencing implementation. Importantly, the individual and

organizational factors are interconnected with the broader contextual factors as well as the underlying power and authority in the system. The interconnectivity of these factors is a complex system principle discussed in the next section under research question two which looks at the structure and dynamics of the system.

In addition to system characteristics discussed above, the innovation itself can also influence successful implementation (Greenhalgh et al., 2005). In my study, the characteristics of the innovation were important factors influencing implementation outcomes. Specifically, implementation of the innovation is more complex and multidirectional as a result of the innovation being high in complexity (Greenhalgh et al., 2005). Evaluation innovations inherently have a high degree of dynamic complexity in part because it is necessary to involve members in the evaluation process and to include a process of reflection and adaptation to ensure that the evaluation is relevant and useful (Patton, 2002; Patton, 2008). The evaluation innovation was also deeply embedded within a larger systems context and the perspectives on the innovation varied significantly across QL stakeholders. Systemic change is needed to implement the evaluation innovation because it is "embedded within dominant system norms, resources, regulations and power operations and their interdependencies" (Foster-Fishman, 2007, p. 213). Another factor that increased the complexity of this innovation was that there were two levels of evaluation: the individual QL level and the QL network level. These levels were interconnected because individual evaluation is necessary for the network evaluation and network level evaluation reinforces the individual QL evaluation.

The characteristics of the innovation that influenced implementation are not unique to this study and in fact overlap with two key issues in the evaluation literature, one of

which is the long-standing debate on the differences between research and evaluation (Mathison, 2008). The contribution of this study with respect to this topic is primarily practical, as it illuminates how decision-makers' definitions of evaluation can be barriers to implementing evaluation innovations. Although there is no universally agreed upon distinction between research and evaluation, there is a key characteristic of evaluation that is generally accepted in the field, which is not reflected in all of the decision-makers' definitions of evaluation. The key characteristic has to do with the 'intent', which should be to inform decision-making about the worth of a specific program, policy or other evaluand (Mathison, 2008). The absence of this intent for evaluation was a barrier to successful implementation of the innovation because they did not see the value in conducting evaluation given that it was not being used to inform practice and policy. There are certain types of research (e.g., applied) that have this characteristic, but also types that are intended to create and generalize knowledge about how the world works (Mathison, 2008). The fact that many of the decision-makers considered anything more complex than monitoring outcomes to be research highlighted a specific perspective of evaluation that was not consistent with institutionalizing the evaluation innovation. Similarly, the fact that many of the decision-makers perceived evaluation as being something that was done after data had been collected and without a pre-planned design illustrates a limited perspective of evaluation, which was a barrier to implementation.

Research Question 2

How do system structure & dynamics impact implementation of the innovation?

Identification of various factors influencing implementation does not provide a full picture of the phenomenon. These factors alone are not sufficient and it is necessary to look at the interconnections between different system parts. As described in the literature review, the reductionist approach aims to act on specific actors and parts of the system, whereas the systems approach acts on the relationships and patterns in the system (Olson et al., 2001; Zimmerman, 2001). A systems approach would look at how other system parts are interacting with the norms to create patterns that sustain norms that are incongruent with the systems change goal (Foster-Fishman, 2007).

In this study, the factors influencing successful implementation were highly contextual and interacted in complex ways (Plsek, 2003), which is why it is not possible to generalize the results of this study to other innovations and settings. However patterns in the data may be similar if the contexts of the other settings under examination are also similar (Lincoln & Guba, 2002). The different parts are interconnected and relationships at the micro level produce patterns of complexity at the macro level. The goal when using the Foster-Fishman (2007) framework is to look at these system parts across levels, niches, organizations, and actors to determine differences between system parts or interactions that create patterns in the system. The patterns can be used to identify leverage points that can shift the system towards the desired state of change.

Similar to other studies, this study demonstrated that there was a dynamic local context and wider policy environment that was influencing the systems change process (Greenhalgh et al., 2009). The findings of this study contribute to other studies that

provide evidence for the innovation-context interaction and the inherent complexity of the implementation process (Greenhalgh et al., 2010; Greenhalgh et al., 2009).

Dynamic Complexity

The use of complexity science was a useful guide that provided insights into both theoretical and practical implications for implementation. The study results demonstrate the dynamic complexity of the system, the innovation, and the implementation process. As mentioned in the literature review, many implementation studies use a positivist approach, which assumes the determinants of implementation and the environment are static (Bucknall, 2007). This study demonstrates however, that the system is highly dynamic and determinants are continuously changing. For example, funding and evaluation policies are constantly changing and influencing implementation of the innovation. The results of the study also suggest that similar to other studies (Bucknall, 2007), the behaviours of decision-makers and other actors in the system vary, depending on the characteristics of the context at a particular point in time.

The results also demonstrate the concept of interconnectedness, a key complex system principle discussed in chapter two. Interconnectedness is a structural characteristic that greatly impacts implementation efforts in the QLs (The National Academies Keck Futures Initiatives, 2009). There are a multitude of interconnections across the QL system and because of the dynamic quality of the system, different system parts are constantly interacting and creating patterns of change. A specific example was provided by two of the interviewees in which change towards the implementation goal occurred as a result of the interaction of different system parts. In this example, the QL

service provider was a university with a third party evaluation contractor hired by the state health department. Initially, the service provider did not have access to the evaluation data or authority over evaluation decision-making. The funder perspective was focused on monitoring outcomes and there was virtually no communication between the evaluation contractor and the service provider. However, changes to two system parts exerted influence on other parts, resulting in the QL making a dramatic shift towards the implementation goal. The first change occurred in the research niche while the second change occurred in the funder perspective. The service provider was at a university and consequently was experiencing mounting pressure from the institution to conduct evaluations and publish results. The pressure from the university provided justification for the service provider to request changes from the state funder to the evaluation structure and regulations. In conjunction to this, the state health department experienced a change in staff and these individuals had a different perspective of evaluation in that they valued conducting more complex evaluations and sharing the results to further the evidence-base for QLs. The changes in these different system parts interacted to create a dramatic shift towards the implementation goal and systems change.

The complexity theory, together with the qualitative data and interpretive approach, provided a unique understanding of the complexity of the implementation phenomenon, as well as the underlying patterns in the system. The complexity theory facilitated a deeper understanding of the phenomenon by illustrating how different factors in the system interact to create patterns at the macro level, although these patterns differed depending on the context.

Structural Differences Across QLs

Another structural issue of the QL system that influenced implementation of the innovation was the differences across the QLs. Differences between system parts are important for understanding systems change because the differences create patterns in the system (Hargreave, 2008; Olson 2001; Foster-Fishman, 2007). For example, the fact that there are private and public QLs and that there is competition for funding, dramatically impacts the dynamics of the network and consequently, information sharing within the network. Another critical difference between QLs is their funding, including both the source and amount of funding received. There is significant variation in the degree of stability of the funding across QLs, which is partly attributed to the source of the funding and partly to the political context of the state. The different funding sources across states also come with different reporting requirements, as well as different regulations about what the funds can be used for. To my knowledge, these funding differences have not been identified or discussed in the literature. However, this finding is not surprising given anecdotal information that within the QL network disparities in funding sources and amounts are well known. Nevertheless, the huge variations in funding (e.g., source, amount, stability) result in very different QL contexts that inevitably impact the implementation of evaluation innovations in the QLs and should be considered when developing implementation interventions. Lastly, a critical difference between QLs is whether or not their evaluation is conducted in-house or by a third party evaluation contractor. This difference is important as it directly influences the QLs capacity for and economic costs of evaluation. Whether or not a QL conducts its own evaluation was determined by two factors: 1) evaluation capacity within the QL and 2) presence or

absence of a mandate from the state funding agency. In many cases, the QL was assigned a third party evaluation contractor by the state. The QLs that are not assigned an evaluation contractor by the state may or may not conduct their own evaluations. Some of the QLs choose to hire external entities to conduct their evaluations or to do the database management because they do not have the internal capacity to conduct it themselves. In some cases, a QL funder will contract with the service provider to do both the service and the evaluation of the QL. There were however, examples of QLs that conduct all aspects of the evaluation in-house. The important point here is that there is significant variation in the capacity to conduct evaluations and evaluation related tasks (e.g., database management) across QLs. The variation in QL capacity in turn impacts a QL's economic costs related to evaluation, as well as their control over the evaluation, which in turn impacts the implementation of the evaluation innovation.

The information provided in this section provides valuable theoretical and practical insights into the heterogeneity of the QLs and the impact of QL differences on implementation outcomes. This section also illustrates how the implementation process is highly complex and influenced by context interaction (Cheng & Van de Ven, 1996; Greenhalgh et at, 2010; Plsek, 2003). These QL differences are important to consider for implementation of all new innovations in the network.

Summary of the Findings

The following section is intended to provide a summary of the findings described in the three results chapters and the above sections of this discussion chapter. As

illustrated in figure 4, the various factors (e.g., resources, regulations) identified in the

results chapters interact in a dynamic way with the innovation characteristics across the

different levels of the system (e.g., organizational, QL network). The evaluation

innovation is embedded within the different system levels, as well as the factors,

including the system norms, regulations, resources, and operations. The innovation

characteristics are interacting with the other system parts and levels. For example, the

norms interact with evaluation characteristics and the norms are different across different

system levels and parts. Specifically, there were decision-makers that perceived the goal

of evaluation to be strictly monitoring outcomes and therefore did not value conducting

evaluation. In contrast, those decision-makers that considered the goal of evaluation to be

quality improvement for the QL placed a higher value on evaluation which improved the

sustainability of evaluation in their QLs. This is an example of how the norms in the

system interact with innovation characteristics to influence implementation of the

practice.

The interdependencies in the system depicted in figure 4 can be viewed as a

network with a multitude of interconnections of such magnitude that it is impossible to

describe every connection in the system. Furthermore, as the system changes over time

the connections within the system change as well. It is this dynamic complexity that

makes it impossible to predict or replicate outcomes in a complex system. Another

example of the interconnections in the system that also illustrates the dynamics over time

is the relationships between resources, regulations, and implementation of the innovation.

The participants described how the policy context was rapidly changing and how this

context established priorities for the innovations that the QLs should implement. They

also described how there was currently a strong political climate for accountability which was resulting in greater emphasis being placed on evaluation and consequently more funding being made available to conduct evaluations. Prior to the shift in the political climate towards one of accountability, evaluation was viewed more as a luxury and was conducted less.

One last example of the interdependencies in the system and how they influenced implementation of the innovation draws from the system operations theme, which includes the power and authority for decision-making in the system. Based on the findings of this study, both NAQC and the CDC had significant power in the system. The power from these two entities interacted with each other as well as with both the resources theme and the regulations theme. NAQC gained much of its power through its relationship with the CDC, and the CDC primarily gained its power in the system through its control of funding. One way in which the CDC exerted its power in the system was through the regulations, for example placing mandates in their funding contracts for the QLs to implement certain practices. This example demonstrates the complexity of the implementation process that results from the multitude of interdependencies in the system.

Also depicted in figure 4 is the non-linear nature of the relationships in the system. Specifically, there is a circular loop around the evaluation innovation, and this non-linearity adds to the dynamic complexity of the system. Complex systems are comprised of negative and positive feedback loops and as such there are too many in the system to describe. One example of a feedback loop identified in the QL system is the utilization of evaluation results. There were examples of QLs that had positive

(reinforcing) feedback loops for evaluation through utilization of evaluation in that using the evaluation data resulted in the QLs placing greater value on conducting evaluation. In contrast, QLs that did not use the evaluation data viewed the innovation more negatively because it depleted QL resources that could have been used for services and this created a negative (dampening) feedback loop.

In the following chapter, next steps for application of the study findings are discussed and include using a systems approach to create intervention strategies. Part of these next steps may include identifying key feedback loops relevant to the specific intervention strategies. For example, a system dynamics model could be developed to map and identify the different feedback loops for each intervention strategy.

Figure 4. Summary of the Findings
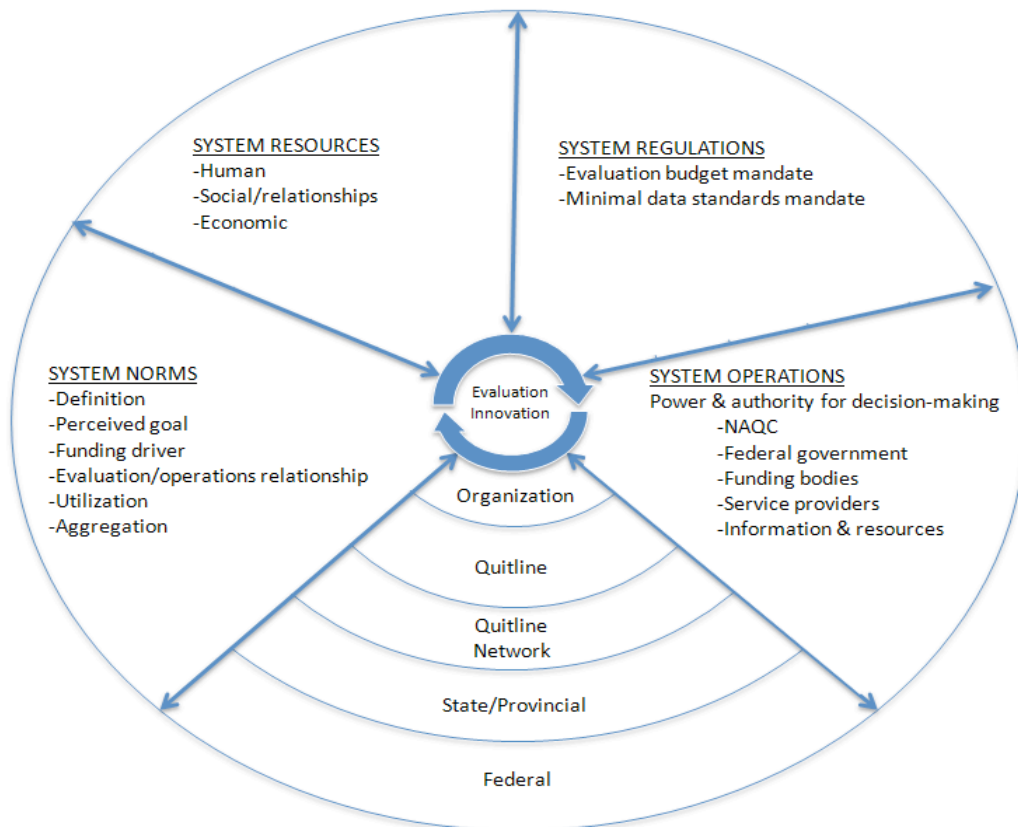
Research Question 3

What strategies can be used to achieve successful implementation of the evaluation innovation?

The following section describes some strategies for managing implementation of the evaluation innovation in the QLs developed based on the results of the thematic analysis and review of the literature. This section builds on the results presented for research questions one and two and considers the complex systems principles presented in the literature review. In this way, the following section builds on and contributes to current thinking about strategies for implementation.

In consideration of the interconnectedness of the system, one important implication related to implementation interventions is that efforts should not always be focused on the most immediate or proximal barriers. Instead, interventions could target the more distal factors that are interconnected and impacting the factors more proximal to the targeted barrier. For example, the findings of this study suggest that a factor influencing implementation outcomes for the evaluation innovation is whether or not there is staff designated to doing evaluation. However, in order to change this issue it is necessary to first change the norms related to evaluation at the funder level, since this entity is responsible for determining funding for evaluation. Also regarding intervention efforts, it is essential that there is cohesion in change efforts across all system parts (Suppovitz & Snyder, 2005). For example, goals, rules, and norms of the system must all cohere to the same desired outcome or change. Lack of coherence in efforts across system parts will constrain systems change. Note that coherence does not imply rigidity or homogeneity of implementation efforts across different system parts, but rather coherence in the goal of the change efforts.

The final implication is related to the dynamic quality of the QL system, which makes implementing innovations and developing interventions for implementation very challenging. One example provided by a focus group member was that it was difficult for his/her QL to implement new policies because those policy mandates could change annually as a result of the highly unstable political climate. This is not an unusual challenge encountered when working in complex systems. It is in large part because of the rapidly shifting context that a strategic/Newtonian type management approach does not work well in complex systems (Olson et al., 2001). According to Olson et al. (2001), Newtonian management methods work when: 1) systems are closed, 2) change is slow; 3) interdependencies are low; 4) certainty is high; and 5) variability is low. An alternative recently emerging management approach is Organizational Development, which places a unique emphasis on participative change action and complexity science (Olson et al., 2001). Organizational Development recognizes that in a fast-paced interconnected society, management strategies need to be equally quick, reflexive, and adaptive to the changing context. One such strategy employed in Organizational Development is the utilization of 'simple rules'.

Simple Rules as Strategies for Achieving Implementation

Simple rules are "the minimum set of guidelines or norms that circumscribe behaviour in a system. If all of the agents in a system follow the same simple rules, then each one adapts to his or her immediate and local circumstances effectively, while remaining a part of the larger system" (Olson et al., 2001, p.106). Simple rules are

important because if change focuses on a specific part, it could change quickly and become obsolete.

Based on the findings of this study, examples of two simple rules that could be applied to the QL system to improve implementation of the evaluation innovation are: 1) building strong partnerships and communication between actors and organizations, and 2) creating environments that facilitate learning. The first simple rule is primarily drawn from the many examples provided by interviewees of how poor partnerships and lack of communication was a barrier to conducting meaningful evaluation and the contrasting examples of how strong collaborative relationships facilitated meaningful evaluation and sustainability of the innovation. This simple rule would be appropriate across the diverse QL contexts and across a changing political climate. The second simple rule stems from both the findings and the systems literature. The interviewees described a system that is dynamic and constantly changing, as well as a lack in positive (reinforcing) feedback loops for evaluation. An increase in organizational and network learning would allow the QLs to adapt faster to the changing contexts (e.g., political, social) making it easier for them to implement and sustain the evaluation innovation.

Both of these rules will facilitate implementation of the innovation regardless of the context, such as policies, funding, or human resources. An underlying component in both of these rules is information flow and particularly feedback into the system. As stated previously, learning requires feedback and without sufficient learning in the system the QLs will not be able to produce, share, implement, and improve upon new practices and policies. Therefore, the two simple rules to improving the implementation of

innovation in the QL system are to increase learning, both organizational and network, and second is to build partnerships and increase information flow.

Simple Rule 1: Promote Organizational & Network Learning

Complex systems such as the QL network are often slow at learning, which creates policy resistance (Sterman, 2006). An increase in learning capabilities of the organizations and the QL network will facilitate the creation of practice-based evidence and also the implementation of new evidence (innovations). Therefore, one of the recommended strategies to improve implementation in the QLs is to increase feedback in the system to promote greater organizational and network learning.

The concept of organizational learning is one based on a holistic view of organizational functioning and change and as such aligns well with complexity science. The learning organization has become an important concept in the organizational literature because of its role as a mediating variable between system and lower level factors (Schechter, 2008) in increasing uptake of innovations and improving organizational performance. The findings of this study demonstrate how factors at multiple levels of the system are impacting implementation, thus organizational learning is a strategy for mediating between these levels. The assumption in the literature is that a higher degree of learning organization characteristics will predict a higher level of adoption and innovation (e.g., climate of innovation) (Jiménez-Jimenez, Valle & Hernandez-Espallardo, 2008). Therefore a strategy for improving innovation implementation in the QLs that is consistent with complexity science could be to increase learning.

According to researchers in this field, learning can occur at the individual, group, organizational, and system level; the latter being the interorganizational network (Sterman, 2006). Learning, both at the organizational and network level are viewed as a social process and a nonlinear transformation that follows the guiding principle of self-organization which along with nonlinearity, are complex system concepts described in chapter two. According to Rycroft and Kash (2004, p. 1), self-organizing is the result of an internal logic of the network which allows a network to learn and be innovative and "innovation networks are organized around constant learning". The authors also suggest that self-organization most often occurs in inter-organizational collaborative activities, not single organizations and it allows for the combining of research and practice to develop innovations (Rycroft & Kash, 2004).

Much of the network learning and innovative network literature looks at competitive markets or for-profit markets. But similar principles can be applied to public health inter-organizational networks such as the QLs. Currently the QL system is self-organizing around the goal of evidence-based practice (EBP). The QL network needs to instead self-organize around the goal of learning and innovation. The QL is considered a 'strategic network'- an organized group engaging in collective action – and should therefore exhibit processes (such as learning) that resemble processes in complex organizations (Rycroft & Kash, 2004) and should be striving for network learning. Organizational and network learning is of particular importance in the QL system because of the lack of clearly defined best practices and innovations. The strength of the QL organizations and the QL consortium will depend on the capacity to learn and adapt quickly to new information and external influences. Because of the constantly changing

environment, it is important to build the capacity of QLs to problem solve and be creative.

Simple Rule 2: Build Partnerships & Increase Information Flow

Another simple rule based on my findings is related to the relationships within a QL and the need to build partnerships and increase information flow between the different entities including: the service provider, the funder, and the third party evaluation contractor. The results of the thematic analysis suggest that this is very important because it was apparent that the QLs that had the best information flow between these three entities were most able to conduct meaningful evaluations. There was significant variation across QLs in terms of the relationships between these three entities. There were examples of QLs that had a very estranged relationship with their third party evaluation contractors, which resulted in significant challenges to the QL being able to conduct meaningful evaluations. In contrast, there were also examples of QLs with close partnership-like relationships between the QL and the third party evaluation contractors. Unfortunately, I was not able to interview a service provider, funder, and third party evaluation contractor for any one QL and I would recommend that this be done in future research. However, my interviews with participants attesting to having close partnership-like relationships with the other organizations in their QL suggest that the different entities were able to collaborate and create meaningful evaluations that answered questions that were useful to all three entities.

One of the major issues to address with this strategy is data ownership. This subtheme from the analysis had interdependencies with many other system parts. There

were several examples provided by interviewees of situations where one entity could not access necessary data that was owned by another entity in the QL. In some cases, it was the third party evaluation contractor who was not providing access to evaluation data that the funder needed. There were also examples of service organizations not providing the necessary data to the funder. In many of the cases, the data would be provided in the form of a report, but not in the raw data form. Although a report is useful to some extent, it does not allow the other entity to work with the data and pursue additional evaluation questions. One of the challenges to requesting data was a lack of knowledge regarding what data was being collected and the format of the data. For this reason, it is important that the different QL entities be provided with a comprehensive codebook of the database.

Overall, it was apparent from the analysis that the levels of partnerships within QLs, varied significantly across QLs. It was also apparent that the relationships and the flow of information between the different entities were critical in terms of evaluating effectiveness, as well as for the overall success of the QL. The results of the KIQNIC social network analysis can provide useful quantitative figures that measure the connections between these different entities. But again, the quantitative data is limited in being able to explain 'why' the relationships exists as they do, the consequences of the relationship ties, and a deeper level of understanding of the relationship ties. In this way, the qualitative data provides valuable information and additional insights that can be used to complement the quantitative survey data.

Leverage Points as Strategies for Achieving Implementation

Leverage points are discussed in the literature review as a method for shifting complex systems towards a desired outcome. Foster-Fishman et al. (2007) in their framework propose identifying leverage points as the final step in achieving systems change. However, they provide little information on how to identify leverage points in the system. Systems change requires coherent effort at multiple levels (Suppovitz & Snyder, 2005). Malhi et al.'s (2009) work that takes Meadows (1999) twelve leverage point levels and collapses them into five system intervention levels can provide a framework to guide this coherent change and build on the Foster-Fishman framework.

The qualitative interpretive approach used in this study has provided insights into the interdependencies in the system, which can then be used to identify different leverage/intervention points in the system. A key strategy for improving implementation of the evaluation innovation is to identify leverage/intervention points in the system and develop coherent strategies across these different points (Suppovitz & Snyder, 2005; Malhi, 2009). Below are examples of two potential leverage/intervention points in the system that were identified based on the results of the thematic analysis.

Leverage Point 1
Paradigm Shift from an Evidence-Based Practice to Practice-Based Evidence

The most effective leverage/intervention level proposed by Malhi et al. (2009) is a paradigm shift. The paradigm of the system and the mindset of the system are similar to the norms theme (deepest held beliefs) from the Foster-Fishman et al. (2007) systems change framework. A critical finding from my first research question was the perspective

from stakeholders that telephone counselling was an evidence-based practice and as such they saw no need to evaluate the QLs. This finding is important because it suggests deeply held beliefs of evaluation and evidence-based practices that are misaligned to the implementation of the evaluation innovation. This perspective assumes that evidence-based practice is one directional and that evidence only comes from research which is a traditional, linear approach to KT (Best et al., 2008). The problem with this perspective in the context of this study is that evaluation is not valued from this perspective. This key finding connects with an argument posed in the literature by Green and Glasgow, (2006, p. 126) that if "we want more evidence-based practice, we need more practice-based evidence."

One of the main reasons why the creation of practice-based evidence is so important is because the QL contexts vary significantly. The initial findings to support the effectiveness of telephone counselling for tobacco cessation was produced at the University of California, San Diego (UCSD) in a randomized clinical trial (Zhu et al., 2002). This context is very different from other QL contexts and it would be incorrect to assume the evidence is equally effective and appropriate for all contexts. For example, the QL in Alberta, Canada uses a call centre as the service provider. In contrast, the California QL service provider is UCSD, which is managed by PhD level researchers and uses only Masters level counsellors. To assume that the findings from the clinical trial setting can be applied equally to all QLs across North America neglects the context-intervention interaction (Plsek, 2003).

The findings from this study support the notion that evidence production cannot be a one directional process where knowledge is transferred from one place to another

(e.g., from research) (Ramstad, 2008; Best et al., 2008). Evidence about health systems should instead be viewed as a dynamic concept that is created through a constant action and reflection process connecting research and practice settings. In order for QLs to be able to adapt research to their local context and be innovative, they need to employ a "participation and discursive strategy, where innovation activities are understood as a joint and open learning process" (Ramstad, 2008, p. 2) involving all actors involved in the network. In this way, the practice setting can both create and adapt evidence to be appropriate for the unique and dynamic context of each QL and also produce practice-based evidence that can be used to further the evidence-base and inform policy and practice decisions throughout the QL network. Given these findings, a key leverage/intervention point in the system is the overarching paradigm shift from a belief in evidence-based practice to practice-based evidence and innovation.

Leverage Point 2
Create Feedback Loops through Utilization of Evaluation Results

Another outcome of the study that overlaps with a key topic in the evaluation literature is the need for utilization of evaluation results. The evaluation field has evolved over the last half century to have an increased emphasis on utilization of evaluation results (Patton, 2002; Patton, 2008). Utilization-focused evaluation is a specific approach to research that involves stakeholder engagement with a goal to use the evaluation results (Patton, 2008). Part of the assumption of this approach is that evaluation that is not used is meaningless and is not a sustainable practice. The findings of this study support the notion of utilization-focused evaluation and contribute to it by connecting it to systems theory and particularly feedback loops.

Sustainability of evaluation innovations requires utilization of the evaluation results. Utilization of evaluation results is a potential feedback mechanism in the system and a fundamental premise in systems thinking is that all learning depends on feedback (Sterman, 2000; Senge, 1990; Meadows, 1999). Sterman (2000) suggests that missing feedback is a common cause of system malfunction. There is currently not enough feedback into the system with respect to evaluation results and "…information feeding back into the system is vital for the system to self-regulate or improve" (Foster-Fishman et al., 2007, p. 211). The evaluation and utilization feedback loops already exist in some but not all of the QLs and the system as a whole lacks the necessary utilization and feedback of evaluation results. For example, the MDS has been implemented by the majority of the QLs but the results of this standardized evaluation are not being used. Two of the interviewees had described how evaluation had gone stale because the results never changed. As a result, the assumption was that there was no point in continuing to conduct follow-up evaluation because they were not getting any new or different results. A stronger feedback loop for evaluation results at all levels of the QL community is necessary because it allows for self-regulation of the system.

Summary

Overall, the qualitative approach provided additional insights into the various factors influencing implementation of the evaluation innovation. The findings demonstrate how characteristics of both the innovation and the system influence implementation outcomes. For example, factors in the system that influenced implementation outcomes included the norms, resources, regulations, and operations.

Factors related to the innovation included its dynamic complexity as well as its embeddedness in the system. The various system parts and factors influencing implementation are interdependent and dynamic which creates patterns in the system. The findings also demonstrate how the structure and dynamics of the system provide insights to explain implementation outcomes, as well as provide information to inform intervention strategies for improving implementation outcomes. For example, differences in QL characteristics and structure influence implementation outcomes and should be taken into consideration when assessing and intervening on the system.

Complexity science was a valuable theoretical approach for framing the data collection and analysis because it facilitated identification of the system patterns and dynamics. Lastly, my findings provide strategies for better managing implementation of the evaluation innovation in this complex system, including the use of simple rules, learning, and leverage points.

The next and final chapter will present lessons learned, contributions, strengths and limitations, and recommendations for future practice and research.

Chapter 9. Conclusions & Recommendations

Chapter Overview

In this final chapter, I begin by presenting the lessons learned related to using the Foster-Fishman framework and systems thinking. I then provide a discussion of the methodological contributions with respect to how this study has complemented the KIQNIC study. Next I describe how this study has contributed theoretically to understanding the connection between implementation of the evaluation innovation and systems change. Finally, I reflect on the study's strengths and weaknesses and conclude by presenting recommendations to practitioners followed by recommendations for research.

Lessons Learned

There are three main lessons that I have learned from conducting this study. Each of these lessons has broadened my understanding and appreciation of the complexities of conducting research using a qualitative interpretive systems approach. Ultimately, these lessons will shape the way I approach future research.

Lesson 1: Application of a Systems Approach

From a theoretical perspective, the systems approach and specifically complexity theory, was appropriate for studying this phenomenon although I found it difficult to apply at times during the study. There is a paucity of literature available to guide this approach and much of the systems literature is theoretical in nature with few practical examples to draw from. Furthermore, the complexity principles such as self-organizing

and feedback loops are relatively abstract concepts and it was difficult to maintain this theoretical approach in the data analysis and interpretation. Although I had not intended to use the Foster-Fishman et al. (2007) systems change framework at the start of the study, I found it a helpful heuristic to assist with maintaining a systems approach while conducting the thematic analysis. As previously mentioned, the framework was developed based on the organizational change and systems thinking literature.

Although the framework was useful it also presented several challenges. The primary challenge was that some aspects of the framework were not sufficiently described. I found that the system element categories provided by the framework (e.g., normative elements) worked well, especially as overarching codes to guide the thematic analysis. However, I also noticed that there was some overlap in the information for different categories. For example, the mandate to allocate 10% of a QL's budget overlaps with both the economic resources and the regulations theme. In general, the four system elements covered a lot of information that was sometimes difficult to parse. My study findings helped to demonstrate the nuances that need to be taken into account under each category and this is where one of the theoretical contributions of the study lies, as will be discussed later in this chapter. Although I used the framework as a heuristic, the subthemes were still very much driven by the data. In this way the findings add to the framework by providing specific information relevant to the evaluation innovation.

Lesson 2: Engaging Study Participants

Although my research was not participatory, I did attempt to engage the participants in the research process and in doing this I learned several lessons relevant to

this research approach. From a practical perspective, I found that engaging participants in the research process with geographical barriers such as in this study was difficult to do. For example, I attempted to engage participants by holding a focus group over the Internet to discuss the study results. It was difficult to have an active discussion over a conference call in part because it is not possible to see non-verbal cues (Berg, 2009). As a result, the focus group was not as engaging as I had hoped. Despite the challenges however, I learned a lot about the value of engaging participants in the research process and using qualitative methods. The participants of this study seemed genuinely appreciative of the opportunity to participate in the research process and to be given the opportunity to review their transcripts and provide feedback on results. I also learned that one of the values of this approach is that participants seem to enjoy having an opportunity to be heard and to share their thoughts on a problem relevant to their daily work. As mentioned previously, the decision-makers are asked to participate in a lot of research, primarily in the format of quantitative surveys. The interview process was more interactive and provided them an opportunity to share their thoughts and be heard.

Lesson 3: Systems Thinking

The phrase, 'the whole is greater than the sum of the parts' is often used in the systems literature (NCI, 2007). This phrase reflects an underlying belief in the systems paradigm that properties emerge at the macro level that cannot be seen at the micro level. The assumption of emergent properties means that parts of a system must be studied in consideration of their interconnectedness with the rest of the system. This is in contrast to the positivist reductionist approach that assumes a mechanistic perspective, dissecting

individual parts to study independently from the rest of the system. Although it is not possible to study every part of a system, the belief that the parts function in interdependent ways that create emergent phenomena at the macro level underpins the systems approach.

I recognize that my findings are based on only 19 perspectives in the system and that 'holism' in the sense of data collection was not the goal. I also recognize that the knowledge produced by this study is always partial because it is socially constructed, complex, and constantly shifting (Denzin & Lincoln, 2005). However, my study is guided by the underlying belief in emergent properties at the macro level and the need to consider parts as interconnected with the broader system. The findings of this study provided me with a deeper understanding of the phrase, 'the whole is greater than the sum of the parts,' in the context of implementation. Based on my findings, I could see how interconnected the system was and how changes in one part of the system could create changes in multiple parts of the system. For example, the relationship between the funder norms, the state context, and the individual QL's human capacity to conduct evaluation was all interconnected. In order to understand change in any of these system parts it is necessary to also consider the dynamic relationships between them.

Researchers in the implementation literature have invested much time into identifying the various parts of the system in an attempt to put them together to create a universal formula for implementation. For example, one study may identify that adopter skills are a key determinant of successful implementation and another may identify that a receptive context is a key determinant. The findings of this study demonstrate the limitations of this approach and the need to recognize how these different parts are

interconnected and the emergent properties that develop as a result of the interconnectivity. Determinants of implementation cannot be extracted and studied as individual parts because what matters are the interactions between them (Greenhalgh, 2005).

In this way, the study also contributes evidence to support the perspective in the literature that it is not possible to develop a universal formula for successful implementation, in large part because of context interaction (Plsek, 2003). In this study factors were highly contextual and interacted in complex ways (Plsek's, 2003). For example, human resources was a factor that impacted implementation outcomes and it was directly influenced by other factors such as funder perspective. Systems change requires consideration of the contextual factors as an active ingredient (Suppovitz & Snyder, 2005; Netting, O'Connor & Fauri, 2007). The qualitative interpretive approach can consider contextual factors as an active ingredient as it assumes everything is context specific and the meaning varies across contexts.

The findings of this study also provided me with a better understanding of the value of using case studies to explore implementation interventions. Based on the study findings, I have a deeper understanding of why KT interventions that focus primarily on changing attitudes and beliefs of adopters generally are not successful and why a systems approach is needed. A research paradigm based on a systems approach recognizes that the KT intervention interacts with the context and therefore cannot be simply extracted and generalized to other settings. It also recognizes that the attitudes and beliefs of adopters are interconnected to broader organizational and socio-political factors so changing those factors at the micro level does little without consideration of the larger

context. For me, these points provide a strong argument to try using case studies and qualitative data in future studies of KT and implementation. This approach allows us to describe the context and the system in detail so that findings can be transferred to similar contexts (Graneheim & Lundman, 2004).

Contributions of the Study

There are several theoretical and methodological contributions of my study. Each of these contributions will be discussed in the following sections.

Contribution 1: Example of Systems Thinking Application

Despite Foster-Fishman's framework for assessing and creating systems change, achieving it remains difficult. One of the main reasons is a paucity of literature to guide such efforts. As mentioned in the literature review, Greenhalgh et al. (2005) identified only one large-scale program (Riley, Taylor, & Elliott, 2001) that was designed around a 'whole systems' approach in their comprehensive review of the literature. Thus, it is not surprising that scientists trained in the linear reductionist approach have difficulty moving to a 'whole systems' approach to implementation because it requires a different type of thinking. Furthermore, the majority of the systems literature is at a theoretical level and it is often difficult to translate the theoretical concepts into methods and practice. One of the primary practical contributions of this study is that it provided a detailed and comprehensive example of how to apply systems thinking, and specifically complexity science, to studying implementation of evaluation innovations. For example, this study provided tangible examples of concepts such as interconnectedness and

leverage points. By doing this, it provides a platform for discussion at a less abstract level.

Contribution 2: Providing a Qualitative Approach to the KIQNIC Study

Although integration of the quantitative and qualitative data is outside of the scope of this study, I will reflect on some aspects of the KIQNIC study to illustrate the contributions of the qualitative interpretivist approach to studying this phenomenon.

The qualitative data provided information and perspectives that both paralleled and added to that which was provided by the KIQNIC's quantitative survey. For instance, the findings of this study paralleled KIQNIC's findings related to the role of NAQC as a mediator organization (central hub) in the network (Moor et al., 2010). These parallels are demonstrated through the results of the social network analysis (SNA) developed using the quantitative data.

One of the differences however, is that the qualitative data also provided insights into the reasons for these results as well as a deeper understanding of the phenomenon. With the open-ended interview questions I was able to probe further and have respondents elaborate on responses, which provided different information to complement the quantitative data. For example, the qualitative data provided information on the connection between NAQC's role and the relationship with CDC, as well as information on how the decision-makers felt about NAQC's role. This additional information is needed for a more comprehensive understanding of the phenomenon and for developing successful implementation interventions.

Another example of the contribution of the qualitative data to the KIQNIC study is related to the implementation question (Appendix B). As mentioned in the context

section, one of the challenges on the KIQNIC study is that there was significant inconsistency in responses to the implementation section of the survey for respondents from the same organization. For example, four different respondents from the same organization selected four different options for implementation level of the evaluation innovation (e.g., aware, decided not to implement; aware, in discussion; and fully implemented). There was also no information provided by the KIQNIC survey to explain why there was so much inconsistency in the implementation responses. The results of this qualitative study however, provided significant insights into the reasons for the inconsistent responses. The qualitative findings show how the decision-makers have very different definitions of the innovation and are responding to the quantitative question from these different perspectives.

Another contribution of the qualitative approach is that it allowed the participants to provide information on the factors influencing implementation that the researchers may not already be aware of and therefore did not include in the quantitative survey. For example, ownership and access to data was an important factor influencing implementation that came out of the thematic analysis. A related factor is the type of data that is collected, as this is directly connected to the fear social actors had for sharing evaluation results across the QLs. The quantitative survey for decision-making, which is intended to assess the different factors influencing decision-making for adopting and implementing innovations, did not provide any information with respect to power and authority issues, which were uncovered in the qualitative analysis.

Another benefit of this methodological approach is that it allowed me to uncover and explore differences in the system, such as across QLs. The QLs have significant

variation in structure and characteristics that should not be ignored. The differences in QLs is a key practical implication that provides evidence for why it is not possible to develop a single standardized intervention for implementation that will be equally appropriate for all QLs (Plsek's, 2003). The qualitative data were able to provide insights into the inherent complexity and context-interaction of the implementation process that was not identified through the KIQNIC quantitative survey.

In addition to providing information to guide future efforts to improve evaluation in the QL network, this project has also produced findings to help guide efforts to make the QL system more innovative and evidence-based. Because evaluation is key to feedback and learning in the system, improving evaluation is directly connected to increasing overall innovation in the system. Therefore by improving the implementation of the evaluation innovation, it is also possible to improve the implementation of all new practices and policies in the QLs.

Contribution 3: Humanizing the Data

As mentioned in the literature review, implementation is often not a linear rational process. Instead, it results from an exchange of ideas, interactions, and mutual sense-making (Greenhalgh et al., 2005). The qualitative approach provided insights into the social process and human factors that influence implementation outcomes. For example, the normative elements, including the perspectives, values, beliefs and assumptions of the stakeholders have strong implications for the implementation of the evaluation innovation. In this way, one of the major contributions of the qualitative interpretive approach is that it 'humanized' the data in a way that the quantitative survey cannot and

provided insights into the important social process and human factors involved in implementation.

This approach also has the potential to make the study findings more relevant to the participants. For example, many of the interviewees stated at the end of the interview that they enjoyed the interview, or that they had learned something from the process and from being able to express their thoughts. Ensuring that the study is relevant to them is particularly important in order to continue to build a partnership between research and practice. It is also important from an ethical perspective to ensure that the research benefits the participants. One of the challenges on the KIQNIC project is being able to make the study results of relevance to the decision-makers (participants) and being able to ensure continued participation in the three-year annual survey.

The value of adding the qualitative approach to the KIQNIC survey is that it has the potential to make the study relevant and beneficial to the participants. The qualitative approach provided an opportunity to build relationships with the participants and have them more engaged in the research process. These advantages are critical in terms of later integration and application of the study findings.

Contribution 4: Insight into Systems Change & Implementation

This study also provides insights into the connection between systems change and implementation. Specifically, this study illustrates how the characteristics of the innovation and the system together can create a situation where systems change is necessary in order to achieve implementation of an innovation. Greenhalgh et al. (2005) suggests that the more complex the innovation, the more complex and iterative the

implementation process will be. The evaluation innovation explored in this study was high in dynamic complexity and the findings demonstrated how this increased the complexity of the implementation process. Furthermore, the QLs vary significantly and as a result it is necessary for the innovation to be adapted to different settings. As a result of the dynamic complexity of both the system and the innovation, systems change is needed to achieve implementation of the evaluation innovation.

Not all innovations require systemic change for implementation and can instead be addressed through incremental change and/or change at the organizational level. Such innovations are unlikely to have a multitude of interconnections with other parts of the system outside of the organization or QL. Implementation interventions and models, such as Graham et al. (2006)'s model, focus on identifying and developing interventions for specific barriers to implementation. This type of change is incremental and works well in certain situations; specifically, systems that are closed and lower in dynamic complexity as well as when there is a single or standardized context for implementation. In a highly dynamic open system incremental change does not work well, especially when the implementation contexts vary significantly as they do in the case of the QLs.

Stacey (1996) provides a matrix for aiding decision-makers in the management field in deciding whether management decisions are complex or not. Complexity is determined based on the degree of certainty in the outcome and level of agreement in the approach. Zimmerman et al. (2001) build on this concept in their book intended to aid health care leaders in managing change in a complex systems. A similar approach might be useful for determining the level of complexity for innovations on the KIQNIC list and determining whether or not a complex systems approach is needed for implementation

over an incremental focused approach. As demonstrated by the findings of this study, the complexity of the innovation and the system are key to understanding the challenges of implementation.

Contribution 5: Addition of Subthemes to the Systems Change Framework

The final theoretical contribution of my study is that it helped to demonstrate the nuances that need to be taken into account under each element of the Foster-Fishman (2007) systems change framework. Although I used the framework as a heuristic, the subthemes were still very much driven by the data. In this way the findings add to the framework by providing specific information relevant to the evaluation innovation. The various subthemes that emerged (e.g., designated staff, ownership of evaluation data, and QL to QL relationship) help to flesh out the framework and can serve as a comparison point in future evaluation and other innovation case studies. For example, future studies can explore whether or not these same subthemes arise even with different stakeholders that have different perspectives. If new sub-themes do emerge than a relevant follow-up issue would be the reasons for the different subthemes. The fleshing out of these subthemes given the specific evaluation innovation is an important theoretical contribution of my study.

Study Strengths

One of the main strengths of the study is that it provided new information and insights to a problem currently being studied. Specifically, the study demonstrates how a qualitative interpretive approach provides valuable information for understanding implementation problems related to the evaluation innovation. Another strength of the

study is that it was theoretically grounded and rooted in the systems literature. As stated in the above section, there is a paucity of literature available to guide practical application of systems concepts to change efforts. As such, it is important to build on and further current available literature in this area. This study has contributed significantly to the literature by providing a detailed example of the application of systems thinking to an implementation problem. Another strength of this study is the potential for the results to be used by the QL community. The results of this study will be useful to the QL network to directly inform intervention approaches to improve implementation of the evaluation innovation, as well as other innovations in the QL network. Overall, this study successfully contributed to implementation sciences at a methodological, theoretical, and practical level.

Study Limitations

There were also limitations to this study that should be noted. The primary limitations of the study were that: 1) only one data source was used, 2) there was limited participant participation, 3) there was a potential for sample bias, and 4) only one innovation was studied.

As this was a dissertation, I was limited in how many sources of data could be incorporated and the number of interviews that could be conducted. Ideally, data would have been collected and integrated from multiple sources. Also, it would have strengthened the results if all organizations in the QL network and their stakeholders could have been included in the interviews. For example, the current interview sample did not include anyone from Free and Clear, which is the largest American QL service

provider in the network. This highlights a potential bias in my recruitment in that not all QLs are equally represented in the sample. I also did not recruit very many Canadian interviewees and as a result, the focus of the study is primarily on the American QLs. It would have strengthened the study to have more information to represent the Canadian QLs.

Another potential limitation of the study was that I did not have as much participant involvement as I would have liked throughout the research process. It would have been difficult to involve the participants more due to time constraints on the dissertation process and also constraints on their time to participate in the project. I addressed the above limitations to some extent by providing the participants with an opportunity to participate in a focus group to discuss the results of the study. Although only five of the participants were able to participate in the focus group their feedback was positive and provided greater credibility to the study results (Flick, 2006; Graneheim & Lundman, 2004).

The final limitation of the project is that I only studied one innovation from the list of 23 innovations being used by the KIQNIC study. It is important to acknowledge that the findings from this current study are specific to the evaluation innovation. As previously mentioned, the evaluation innovation is highly complex and may not be similar to other innovations on the list. As such, the findings of this study may be unique to the evaluation innovation and may not be applicable to the other 22 innovations.

Recommendations for Practitioners

The findings of this study provide several recommendations for practitioners including next steps to take in order to successfully institutionalize the evaluation innovation in the QL system.

Recommendation 1: Application of a Systems Approach to Implementation

A systems or whole systems approach should be taken to improve implementation and institutionalization of the evaluation innovation (Kitson, 2009). There are several reasons why a systems change approach is needed, primarily related to the dynamic complexity of both the innovation and the system. A systems approach should include identifying leverage points in the system and creating interventions at the different leverage points (Meadows, 1999; Malhi et al, 2009). The leverage points should target different levels and parts of the system and should be coherent (Suppovitz & Snyder, 2005). As described in the literature review, targeting change in one component of the system is unlikely to achieve success. Two leverage points were provided in the previous chapter (i.e., paradigm shift and positive feedback loop) that can be used as examples to facilitate discussion on leverage points and developing coherent intervention strategies in the system.

As application of the findings using a systems approach is difficult, the following information has been provided as potential next steps to develop intervention strategies based on the findings from this research. This example pertains to next steps for developing interventions to address the second proposed leverage point, which is to

create and strengthen a positive feedback loop for evaluation via utilization of results. It should be noted that the interventions should be developed through participatory methods in collaboration with the decision-makers in the QL system.

In keeping with a systems approach the interventions should target different parts and levels of the system, as well as consider the factors influencing implementation that were identified in the results chapter (e.g., human resources, norms, regulations). One of the first steps would be to identify which QLs are and are not utilizing evaluation results, as well as to what degree and in which way results are being used. This is important because as demonstrated in the findings, there are significant differences across QLs and these differences must be taken into consideration. To do this, each QL must be considered as a unique entity and strategies must be adapted to each unique QL context.

The intervention should target the norms at both the organizational (i.e., service provider) and state (i.e., funder) level. Such an approach could involve educating the decision-makers from both the service providers and the funders regarding what they should expect from their evaluation contractors and the characteristics of a collaborative relationship. For example, they should expect their evaluation contractors to communicate with them, share evaluation data, and work with them to develop evaluation plans that produce data that can be utilized and answer questions to inform practice and policy. It would also be necessary to educate some of the decision-makers on how evaluation data can be used to improve their QL services (i.e., quality improvement). This intervention is specific to those decision-makers who consider the goal of evaluation to be strictly monitoring outcomes. For those that are only monitoring outcomes, it would

be necessary for them to conduct more complex evaluation in order to have result that they can be used.

I would suggest developing case studies of QLs that are successfully using the results and in the case studies describe how evaluation data is used while also highlighting the benefits of utilization. The purpose here is to share information across QLs that can help change the norms regarding the perspective of the innovation, as well as the perceived goal and value put on evaluation. It would also be necessary to identify the decision-makers that do not see the need for evaluation because they consider QLs already an evidence-based practice. This norm needs to be addressed as well through education strategies in order to have evaluation results utilized. All of these points will help develop the positive feedback loop for evaluation utilization by changing the norms at different levels of the system.

One specific strategy I would use would be to employ a team of evaluation consultants to guide the process and provide assistance to the QL decision-makers. Specifically, the evaluation consultants should interview the decision-makers in the QLs that are not utilizing results and attempt to identify the barriers to implementation so that plans can be tailored to individual QLs. They could use the findings from this study as a framework for identifying barriers and facilitators to evaluation implementation. The evaluation consultants would also be available to assist the QLs in applying and tailoring existing strategies being used by QLs that already successfully use evaluation results.

The strategies must also consider the regulations, resources, and operations in the system, as described in the results section. For example, it will be necessary to identify QLs that have regulations (e.g., policies) that are barriers to evaluation utilization. QLs

that have state policies that prohibit service providers from accessing and utilizing the evaluation data must be identified and this policy barrier addressed. This may require ongoing communication between the service provider, funder and NAQC to convey the importance of being able to access the data to the state. It may also be necessary for CDC to apply a mandate that evaluation data be made available to service providers as a stipulation to the state receiving QL funding. CDC could also create incentives in the system by rewarding QLs that demonstrate utilization of evaluation results through quality improvement or furthering the evidence base with funding.

Lastly, the resources of the system must also be addressed as part of the intervention strategy to act on this leverage point. QLs that do not have the human capacity to utilize evaluation results must be identified and assisted to overcome this barrier. To do this, I would follow the plan described in the following section to reduce fragmentation of evaluation resources in the QL system and have CDC develop and provide a central evaluation system that can be used by QLs that do not have sufficient resources themselves to successfully conduct evaluation. It is critical that this barrier be addressed in order to create and strengthen the positive feedback loop for evaluation.

The proposed plan described above is only a snapshot of what could be done to begin to address the feedback loop leverage point. The plan illustrates how the intervention strategies must act on different levels and parts of the system. It also integrates the factors influencing implementation, and the system structure and dynamics presented in the results. As previously stated, a full plan to intervene at the leverage points should be developed in collaboration with the decision-makers in the QL system.

However, this provides an example of the next steps that could be taken to begin to develop a systems intervention to improve implementation of the evaluation innovation.

Recommendation 2: Reduce Evaluation & Resource Fragmentation in the System

Currently, evaluation efforts and resources are fragmented across the system. Based on the interview data, the lack of internal evaluation capacity is a barrier to conducting more complex evaluations for many of the QLs. Lack of consistent evaluation capacity across the system poses several barriers to systems change. For instance, one barrier relates to the disparity in evaluation capacity across different QLs. Although some of the QLs have the internal capacity to conduct more complex evaluations, many of them do not. Another barrier is related to differences in economic costs that result from the disparity in evaluation capacity. QLs that lack the internal capacity for evaluation have to contract with external agencies. QLs that depend on external database management companies and external evaluation contractors incur higher costs to conducting evaluation and modifying evaluation protocols. A systemic barrier is the overall inefficiency in the use of resources due to redundancies in the system for conducting evaluation. The QLs are currently working in isolation from each other and directing resources into independent efforts such as creating database systems and QL software. The QLs would likely benefit from a more cohesive and coordinated effort to build evaluation capacity.

For example, CDC could potentially provide a centralized online database and data management system that QLs could use. Such a resource would be of greater value to the QLs that have smaller budgets and less internal capacity to evaluate. For QLs with

larger budgets, and an existing evaluation structure in place, the CDC efforts may not provide an initial benefit. However, access to this centralized system would help level the playing ground between QLs that have varying degrees of funding and capacity for evaluation. It would also reduce costs to the QL network as a whole by eliminating redundancies in the system and providing an opportunity for direct data collection and aggregation. It would be important that such a system allow for intra-QL variation, while also providing some standardization across QLs. For example, there may be a core data collection framework (e.g., MDS) that all QLs collect data for while also allowing QLs to tailor the system for their unique QL. This system would have several benefits that would facilitate the implementation of the practice. The system would allow QLs to look at the aggregated data, and benchmark or compare their results against the overall QL community results, without identifying individual QLs. The QLs would also be able to see what type of data is being collected by other QLs, which would serve as a knowledge sharing opportunity.

Recommendation 3: Strive for a Paradigm Shift Toward Innovation

There are several concepts that are relevant and intertwined with achieving a paradigm shift toward innovation in the system. The need to self-organize around innovation and focus on organizational and network learning were already discussed in the previous chapter. Also, the need to focus on practice-based evidence and view evidence as a dynamic concept that emerges from research and practice is critical for achieving successful implementation of the evaluation innovation. Many of the participants stated that QLs are an evidence-based practice and therefore they did not see

the need to evaluate them. This is an interesting finding that illustrates a linear one directional view of evidence. I would suggest creating an initiative in the QLs to change the view of evidence to be more of a dynamic concept that includes practice-based evidence. There has been significant attention put on the need for evidence-based practice and this message has been received by decision-makers. A similar message should be communicated for the importance of practice-based evidence and the need to use evaluation to create this evidence and to be innovative.

Recommendations for Future Research

The findings of this study provide the basis for several recommendations related to future research. The most important recommendation is for implementation researchers to be explicit about the paradigm they work from and the underlying assumptions of the research. It is important to recognize that different epistemological groundings will provide different findings. As mentioned in the literature review, systems thinking is fundamentally different from reductionist science as it emphasizes the need for consideration of emergence and the interactions between individual parts of a system.

Another recommendation for future research is to conduct more studies using complexity theory. This approach is useful for guiding case studies of implementation and systems change (Anderson, 2005; Murphy-Smith, 2004). Complexity science is also useful because it incorporates elements of both positivism and interpretivism (Vogel, 2009). As previously mentioned, an integrative approach has been shown to have an advantage over a strictly interpretivist or positivist approach in implementation and organizational change studies (Saberwhal & Robey, 1995; Poole & Van de Ven, 1989).

Future research on implementation should also include mixed-methods case studies, as this is a valuable approach for studying implementation and addressing the context interaction issue. Case studies should include detailed descriptions of the innovation, the context and the implementation process. In this way, the findings and lessons learned from a case study can be transferred to other similar settings (Graneheim & Lundman, 2004). Although it is not possible to generalize results from one implementation study to another, patterns in the data may be similar if the contexts under examination are also similar. Case studies allow for the detailed description that is necessary for transferring results across settings. Case studies should capture the differences in implementation settings such as the differences in QL characteristics and not assume homogeneity of the QLs and the actors in the system. For example, there is significant variation across QLs and organizations, as well as across the actors. Future research should explore these unique differences and the subtle nuances in the system, as well as how they impact implementation outcomes. Future research should also study factors at different levels of the system including the organizational and socio-political context and the interdependencies between these levels. It is important to consider the factors outside of the organization that influence implementation.

Although these findings have pertained specifically to the evaluation innovation in the QLs there are certainly lessons to be learned that can be applied to other innovations in the QLs, as well as implementation of innovations in other systems. In addition to applying this approach to other innovations, future research projects should also apply similar methods to the implementation of innovations in other systems, such as health authorities. In this way, the literature in the systems change and implementation

fields can be expanded upon providing greater detail on methods and theories to help guide future efforts to achieve implementation of innovations through systems change.

There are also several next steps for my own research. One next step is to begin to integrate the qualitative findings from this study with the quantitative findings from the KIQNIC study. The qualitative findings can be used to help interpret the results from the KIQNIC survey and to guide analysis of the social network data. Together the quantitative and qualitative data can provide a more comprehensive picture of the implementation phenomenon.

The integration of the qualitative and quantitative data will be a challenging next step requiring substantial review of the literature and planning to guide this process. However, a few details for how I intend to begin to go about the integration process are described below. Logistically, I intend to work closely with the PhD student on the KIQNIC project that is primarily responsible for the social network analysis. Based on previous conversations, we have already begun to identify ways in which the two types of data complement and inform each other.

Social network analysis is the primary source of quantitative data and it describes patterns and structure of relationships between actors and organizations and simplifies the relationships into numerical data. The SNA data does not provide contextual information to aid in accurate interpretation of the ties identified nor does it explain why those ties exist or not exist. A main role of the qualitative data will be to provide that contextual information to improve accuracy of the interpretation (Conti & Doreian, 2009). Chiu and West (2007) suggest that the SNA provides 'outsider's' view including the structure of the network, but the qualitative data derived from sources such as semi-structure

interviews, can provide an 'insider's' view including information such as the processes of the network.

The process will be iterative looking at how the qualitative informs the quantitative and how the quantitative informs the qualitative. The first step will be for me to become familiar with the results of the SNA. Examples of the findings of interest to me will include, what ties exist, how ties differ in strength, and what type of clustering affect is seen in the network. I will also want to know the results of the implementation and decision-making sections of the survey. For example, which decision-makers consider cost to be a significant factor in deciding to implement the innovation? I will then pull findings from the qualitative data to help interpret these findings and relate the quantitative findings to the qualitative themes.

I will use the qualitative data to provide information on how ties or lack of ties may be influencing implementation. For example, one issue I would explore is the role of these relationships in control of information and resources. This is one of the subthemes from the qualitative analysis that is of relevance to the quantitative findings. The SNA will provide information on the number of ties, but the qualitative data will be used to inform questions about the quality of the ties and how it effects access to resources (Bidart & Lavenu, 2005). The qualitative data will also be used to identify the cultural contexts, and other contexts that the ties are embedded in and to better understand the nature and meaning of the networks (Edwards, 2010). One step towards achieving this is to produce network maps illustrating different network characteristics and structure, and use the qualitative data to write an interpretation paragraph under each one.

I will also look at the results of the decision-making survey and use the qualitative

data to examine the complexity of some of the results. For example, for the respondents who reported that cost was not a significant factor in decision-making to implement a practice, what were some of the other characteristics of the QL drawing from the qualitative findings? Does the QL have a supportive state context? Does it have strong ties to other QLs and to NAQC? In this way, the qualitative data provides questions to guide interpretation and a deeper analysis of the quantitative data.

The next steps provided are realistic and doable without additional data collection. However, if further data collection were an option, then it would be of interest to use the SNA and other quantitative results to identify and select more decision-makers to interview who could add different perspectives to the current qualitative study. For example I would identify organizations in the network that are in different positions in the network (e.g., central, removed). I would also interview respondents with a variety of scores on the implementation section of the survey (e.g., high, medium, and low). Lastly for the decision-making survey it would be of interest to breakdown the responses at the individual level and interview the actors to learn more specifics about how certain factors did or did not influence decision-making to implement specific innovations.

Also, based on the results of the focus group session, there is interest in exploring ways to use these results and consider the implications for other innovations in the network. I would like to work on translating the findings of this study into practice by conducting participatory research with decision-makers in the QL to further develop and refine the leverage points and interventions to improve implementation. However, one of the lessons learned from the focus group session is that systems concepts tend to be very foreign and field specific (e.g., leverage points) and a barrier to engaging stakeholders in

meaningful discussions about systems change. In order to engage stakeholders in a discussion on system change and move towards developing intervention strategies it might be helpful to first provide them with a training in systems thinking and particularly the terminology used in this field. Therefore another next step is to work on developing a systems change workshop for decision-makers in the QL network.

References

Agency for Healthcare Research and Quality. (2001). *Translating research into practice (TRIP)-II* (AHRQ pub No. 01-PO17). Fact Sheet (pp. 1-6). Rockville, MD: Agency for Health Research and Quality.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In *J. Kuhl & J. Beckman (Eds.), Action-control: From cognition to behavior* (pp. 11-39). Heidelberg: Springer.

Anderson, R., Crabtree, B., Steele, D., & McDaniel, R. (2005). Case study research: the view from complexity science. *Qualitative Health Research*, *15*(5), 669-685.

Behrens, T., & Foster-Fishman, P. G. (2007). Developing operating principles for systems change. *American Journal of Community Psychology*, *39*, 411-414.

Benowitz, N. (2009). Pharmacology of nicotine: addiction, smoking-induced disease, and therapeutics. *Pharmacology Toxicology*, *49*, 57-71.

Berg, B. (2009). *Qualitative research methods for the social sciences* (7th ed.). Boston: Allyn & Bacon.

Best, A., Hiatt, R., & Norma, C. (2008). Knowledge integration: Conceptualizing communications in cancer control systems. *Patient Education and Counseling*, *71*, 319-327.

Bidart, C. and Lavenu, D. (2005). 'Evolutions of personal networks and life events'. *Social Networks*, 27:359-376.

Boeije, H. (2010). *Analysis in Qualitative Research.* Thousand Oaks, CA: Sage.

Braun, V., & Clark, V. (2006). Using thematic analysis in psychology. *Qualitative Research Psychology*, *3*, 77-101.

Bucknall, T. (2007). A gaze through the lens of decision theory toward knowledge
    translation science. *Nursing Research*, *56*(4S), S60-S66.

Carlson, J., & Doyle, J. (2002). Complexity and robustness. *Proceedings of the National
    Academy of Sciences of the United States of America*, *99*(S1), 2538-2545.

Center for Disease Control. (2005). Annual smoking-attributable mortality, years of
    potential life lost, and productivity losses--united states, 1997-2001. *Morbidity
    and Mortality Weekly Report, 54*(25), 625-628.

Center for Disease Control. (2009). Cigarette Smoking Among Adults and Trends in
    Smoking Cessation - United States, 2008. *Morbidity and Mortality Weekly
    Report, 58*, 1227-1232.

Center for Disease Control. (2004). *Telephone Quitlines: A Resource for Development,
    Implementation, and Evaluation*. Atlanta, GA: U.S. Department of Health and
    Human Services, Centers for Disease Control and Prevention, National Center for
    Chronic Disease Prevention and Health Promotion, Office on Smoking and
    Health.

Chase, S. (2003). Learning to listen: Narrative principles in a qualitative research
    methods course. In *Up close and personal: The teaching and learning of
    narrative research* (pp. 79-99). Washington, D.C.: American Psychological
    Association.

Cheng, Y., & Van de Ven, A. (1996). Learning the innovation journey: Order out of
    Chaos? *Organization Science*, *7*(6), 593-614.

Chiu, L.F. and West, R.W. (2007). 'Health Intervention in social context: Understanding
    social networks and neighbourhood'. *Social Science and Medicine*, 65,1915-1927.

Canadian Institutes for Health Research. (2005). *About knowledge translation.* Retrieved

from http://www.cihr-irsc.gc.ca/e/29418.html

Cilliers, P. (1998). *Complexity and postmodernism: Understanding complex systems*.

London: Routledge.

Conti, N. Doreian, P. (2009). Social network engineering and race in a police academy: A

longitudinal analysis. *Social Networks*, 32(1), 30-43.

Corring, D. (2004). Ensuring a client perspective in evidence-based rehabilitation

research. In Hammell, K.W. & Carpenter. *Qualitative Research in Evidence-

Based Rehabilitation*. Philadelphia: Elsevier.

Cummings, G., Estabrooks, C., Midodzi, W., Wallin, L., & Hayduk, L. (2007). Influence

of organizational characteristics and context on research utilization. *Nursing

Research*, *56*(4S), S24-S39.

Davies, H., Nutley, S., & Walter, I. (2005). *Assessing the impact of social science

research: Conceptual, methodological and practical issues: A background

discussion paper.* UK: Research Unit for Research Utilization, School of

Mangement University of St Andrews.

Denzin, N., & Lincoln, Y. (2005). Introduction: The Discipline and Practice of

Qualitative Research. In *Handbook of Qualitative Research* (3rd ed., pp. 1-28).

Thousand Oaks, CA: Sage Publications.

Dopson, S. & Fitzgerald, L. (Eds): *Knowledge to Action? Evidence-based health care in

context*. Oxford: Oxford University Press; 2005.

Easterby-Smith, M., Lyles, M., & Tsang, E. (2008). Inter-organizational knowledge

transfer: Current themes and future prospects. *Journal of Management Studies*,

45(4), 677- 690.

Eccles, M., & Mittman, B. (2006). Welcome to implementation science. *Implementation Science*, *1*(1), 1-3.

Edwards, G. (2010). ESRC National Centre for Research Methods Review paper: Mixed-method approaches to social network analysis. Accessed online, April 6, 2011: http://eprints.ncrm.ac.uk/842/1/Social_Network_analysis_Edwards.pdf

Egede, L. E., & Zheng, D. (2002). Modifiable cardiovascular risk factors in adults with diabetes: Prevalence and missed opportunities for physician counselling. *Archives of Internal Medicine*, *162*, 427-433.

Eoyang, G. H. (1998). *Coping with Chaos: seven simple Tools*. Circle Pines, MN: Lagumo.

Flick, U. (2006). *An introduction to qualitative research.* (3rd ed.). Thousand Oaks, CA: Sage publications.

Flood, R. L. (2010). The relationship of 'Systems Thinking' to action research. *Systemic Practice and Action Research*, 23, 269-284.

Fonesca, J. (2001). *Complexity and Innovation in Organisations*. London: Routledge.

Fontana, A. & Frey, J. H. (2005). The Interview: From Neutral Stance to Political Involvement. In *Handbook of Qualitative Research* (3rd ed., pp. 695-728). Thousand Oaks, CA: Sage Publications.

Foster-Fishman, P. G., Nowell, B., & Yang, H. (2007). Putting the system back into systems change: a framework for understanding and changing organizational and community systems. *American Journal of Community Psychology*, *39*, 197-215.

Frenken, K. (2000). A complexity approach to innovation networks. The case of the

aircraft industry 1909-1997. *Research Policy*, *29*, 257-272.

Gay, B., & Dousset, B. (2005). Innovation and network structural dynamics: Study of alliance network of a major sector of the biotechnology industry. *Research Policy*, *34*, 1457-1475.

Garbett, R., & McCormack, B. (2001). The experience of practice development: An exploratory telephone interview study. *Journal of Clinical Nursing*, 10, 94-102.

Gibbs, D., Napp, D., Jolly, D., Westover, B., & Uhl, G. (2002). Increasing evaluation capacity within community-based HIV prevention programs. *Evaluation and Program Planning*, 25(3), 261-269.

Graham, I., Logan, J., Harrison, M., Straus, S., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: Time for a map? *Journal of Continuing Education in the Health Professions*, *26*, 13-24.

Graneheim, U., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*, 105-112.

Green, L., & Glasgow, R. (2006). Evaluating the relevance, generalization, and applicability of research. *Evaluation and the Health Professions*, *29*(1), 126-153.

Greenhalgh, T. Glenn, R., Bate, P., Macfarlane, F., & Kyriakidou, O. (2005). *Diffusion of Innovations in Health Service Organisations: A Systematic Literature Review*. Australia: Blackwell Publishing Ltd.

Greenhalgh, T., Potts, H.W.W., Wong, G., Bark, P., & Swinglehurst, D. (2009). Tensions and paradoxes in electronic patient record research: A systematic literature review using the meta-narrative method. *The Milbank Quarterly*, *87*(4), 729-788.

Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Mohammad, Y., & Russell, J. (2008). Introduction of shared electronic records: multi-site case study using diffusion of innovation theory. *British Medical Journal*, *337*(a1786), 1-10.

Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Russel, J., & Potts, H. (2010). Adoption and non-adoption of a shared electronic summary record in England: a mixed-method case study. *BMJ*, *340*(c3111), 1-11. doi:10.1136/bmj.c3111

Greenhalgh, T., Charlotte, H., Hughes, J., Macfarlane, F., Butler, C., & Pawson, R. (2009). How do you modernize a health service? A realist evaluation of whole-scale transformation in London. *The Milbank Quarterly*, *87*(2), 391-416.

Grimshaw, J., Thomas, R., MacLennan, G., Fraser, C., Ramsay, C., & Vale, L. E. A. (2004). Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technology Assessment*, *8*(6), 1-72.

Hakim, C. (2000). Qualitative research. In *Research design. Successful designs for social and economic research* (pp. 34-45). London, UK: Routledge.

Hargreaves, M. B. (2010). Evaluating Systems Change: A Planning Guide. Accessed December 2010, www.mathematica-mpr.com.

Health Canada. (2008). *Smoking cessation in the workplace*. Retrieved from http://www.hc-sc.gc.ca/hc-ps/pubs/tobac-tabac/cessation-renoncement/index-eng.php

Health Canada. (2010a). *Canadian Tobacco Use Monitoring Survey*. Retrieved from http://www.hc-sc.gc.ca/hc-ps/tobac-tabac/research-recherche/stat/_ctums-esutc_2009/w-p-1_sum-som-eng.php

Health Canada. (2010b). *Role of governments*. Retrieved from http://www.hc-sc.gc.ca/hc-

ps/tobac-tabac/about-apropos/role/index-eng.php

Hirsch, G. B., Levine, R., & Miller, R. L. (2007). Using system dynamics modeling to understand the impact of social change initiatives. *American Journal of Community Psychology*, *39*, 239-253.

Holland, J. (1992). Complex Adaptive Systems. *Daedalus*, *121*(1), 17-30.

Hughes, J.R. (1995). Combining behavioral therapy and pharmacotherapy for smoking cessation: an update. *NIDA Research Monogram*, *150*, 92-109.

Jallinoja, P., Absetz, P., Kuronen, R., Nissinen, A., Talja, M., UUtela, A., & Patja, K. (2007). The dilemma of patient responsibility for lifestyle change: Perceptions among primary care physicians and nurses. *Scandinavian Journal of Primary Health Care*, *25*, 244-249.

Janzen, R., Nelson, G., Hausfather, N., & Ochocka, J. (2007). Capturing systems level activities an impacts of mental health consumer-run organizations. *American Journal of Community Psychology*, *39*, 287-299.

Jiménez, D., Valle, R.S., Hernandez-Espallardo, M. (2008). Fostering innovation: The role of market orientation and organizational learning. *European Journal of Innovation Management, 11*(3), 389 - 412.

Kash, D., & Rycoft, R. (2000). Patterns of innovating complex technologies: a framework for adaptive network strategies. *Research Policy*, *29*, 819-831.

Kitson., A.L. (2009). The need for system change: reflections on knowledge translation and organizational change. *Journal of Advanced Nursing*, *65*(1), 217-228.

Kitson, A.L., Harvey, G., & McCormack, B. (1998). Enabling the implementation of evidence based practice: A conceptual framework. *Quality & Safety in Health*

Care, *7*, 149-158.

Kitson, A.L., Rycroft-Malone, J., Harvey, G., McCormack, B., Seers, K., & Titchen, A. (2008). Evaluating the successful implementation of evidence into practice using the PARiSH framework: theoretical and practical challenges. *Implementation Science*, *3*, 1.

Knight, L. (2002). Network learning: Exploring learning by interorganizational networks. *Human Relations*, *55*, 427-454.

Kreger, M., Brindis, C., Manuel, D., & Sassoubre, L. (2007). Lessons learned in systems change initiatives: benchmarks and indicators. *American Journal of Community Psychology*, *39*, 301-320.

Lincoln, Y. S. & Guba, E. G. (2002) Judging the Quality of Case Study Reports. In *The Qualitative Researcher's Companion* (pp. 205-217). Thousand Oaks, CA: Sage Publications.

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in Educational Research: From Theory to Practice*. San Francisco, CA: Jossey-Bass.

Macnee, C.L. & McCabe, S. (2008). Understanding nursing research: using research in evidence-based practice (2nd ed.). United States of America: Lippincott Williams & Wilkins.

Makomaski-Illing, E., & Kaiserman, M. (2004). Mortality attributable to tobacco use in Canada and its regions, 1998. *Canadian Journal of Public Health*, *95*(1), 38-44.

Malhi, L., Karanfil, Ö., Merth, T., Acheson, M., Palmer, A., & Finegood, D. (2009). Places to intervene to make complex food systems more healthy, green, fair and affordable. *Journal of Hunger and Environmental Nutrition*, *4*, 466-476.

Marble, R.P. (2000). Operationalising the implementation puzzle: an argument for eclecticism in research in practice. European Journal of Information Systems, 9, 132-147.

Mathison, S. (2008). What is the difference between evaluation and research and why do we care. In *Fundamental Issues in Evaluation.* New York: The Gilford Press.

Meadows, D. (1999). *Leverage Points: Places to Intervene in a System.* The Sustainability Institute.

Meyers, P., Sivakumar, K., & Nakata, C. (1999). Implementation of industrial process innovations: factors, effects, and marketing implications. *Journal of Product Innovation Management*, *16*, 295-311.

Moor, G., Leischow, S., Provan, K., Bonito, J., Saul, J., Beagles, J., Terpstra, J. L., Best, A., Ruppel, E., Walsh, M. 2010. KIQNIC: Knowledge Integration in Quitlines - Networks that Improve Cessation. 3rd Annual NIH Conference on the Science of Dissemination & Implementation, March 16, 2010, Bethesda, MD.

Moss, N. (1983). An organization-environment framework for assessing program implementation. *Evaluation and Program Planning*, *6*, 153-164.

Murphy-Smith. (2004). *Journal of Public Health Management Practice*, *10*(2), 109-115.

National Cancer Institute. (2007). *Greater Than the Sum: Systems Thinking in Tobacco Control*. NCI Tobacco Control Monograph Series (Vol. 6). Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

Netting, F., O'Connor, M., & Fauri, D. (2007). Planning transformative programs: Challenges for advocates in translating change processes into effectiveness

measures. *Administration in Social Work, 31*(4).

Neuman, W. (2006). *Social research methods. Qualitative and quantitative approaches* (6th ed.). Needham Heights, MA: Allyn & Bacon.

North American Quitline Consortium. (2009). *All Quitline Facts: An Overview of NAQC Annual Survey of Quitlines* (pp. 1-6). North American Quitline Consortium. Retrieved from http://www.naquitline.org/resource/resmgr/ql_about_facts/2008_quitline_facts_qa _final.pdf

Noy, C. (2008). 'Sampling knowledge: The hermeneutics of snowball sampling in qualitative research'. *International Journal of Social Research Methodology*, *11*(4), 327-344.

Offredy, M. & Vickers, P. (2010) *Developing a Healthcare Research Proposal: an Interactive Guide.* United Kingdom: Wiley- Blackwell.

Olson, Eoyang, Beckhard, R., & Vaill, P. (2001). *Facilitating Organization Change: Lessons from Complexity Science*. San Francisco: Pfeiffer.

Patton, M. (2002). Utilization-focused evaluation. *Evaluation in Education and Human Services*, *49*, 425-438.

Patton, M. (2008). *Utilization-Focused Evaluation* (4th ed.). Thousand Oaks, CA: Sage Publications.

Plsek, P., & Greenhalgh. (2001). Complexity: The challenge of complexity in health care. *BMJ*, *232*, 625-628.

Plsek, P. (2003). *Complexity and the adoption of innovation in health care. Paper presented at Accelerating Quality Improvement in Health Care: strategies to*

speed the diffusion of evidence-based innovations. Washington, DC: National Institute for Health Care Management Foundation and National Committee for Quality Health Care.

Poole, M., & Van de Ven, A. (1989). Using paradox to build management and organization theories. *Academy of Management Review*, *15*(3), 562-578.

Ramstad, E. (2008). Developmental evaluation framework for innovation for innovation and learning networks: Integration of the structure, process and outcomes. *Journal of Workplace Learning*, *21*(3), 181-197.

Repenning, N. P. (2002). A simulation-Based Approach to Understanding the Dynamics of Innovation Implementation. *Organization Science*, *13*(2), 109-127.

Riley, B., Taylor, S., & Elliott, S. (2001). *Health Education Research*, *16*(4), 425-441.

Rogers, E. (1995). *Diffusion of Innovations*. New York: Free Press.

Ryan, G., & Bernard, H. (2000). Data Management and Analysis Methods. In *Handbook of Qualitative Research* (2nd ed., pp. 769-802). Thousand Oaks, CA: Sage Publications.

Rycroft, R., & Kash, D. (2004). Self-organizing innovation networks: implications for globalization. *Technovation*, *24*, 187-197.

Saberwhal, R., & Robey, D. (1995). Reconciling variance and process strategies for studying information system development'. *Information Systems Research*, *6*, 303-327.

Schechter, C. (2008). Organizational learning mechanisms: The meaning, measure, and implications for school improvement. *Educational Administration Quarterly*, *44*(2), 155-186.

Schutt, R.K. (2006). *Investigating the social world: the process and practice of research.* Thousand Oaks, CA: Pine Forge Press.

Senge, P.M. (1990). *The fifth discipline: the art and practice of the learning organization*. New York, New York: Bantam Doubleday Dell Publishing Group, Inc.

Sesselberg, T. S., Klein, J. D., O'Connor, K. G., & Johnson, M. (2010). Screening and counseling for childhood obesity: Results from a national survey. *Journal of the American Board of Family Medicine*, *23*(3), 334-342.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative reearch projects. *Education for Information*, *22*, 63-75.

Shortell, S. M. (1981). Suggestions for improving the study of program implementation. Paper presented at the 109th Annual Meeting of the American Public Health Association, Los Angeles, November 1981.

Silverman, D. (2000). Analyzing Talk and Text. In *Handbook of Qualitative Research* (2nd ed., pp. 821-834). Thousand Oaks, CA: Sage Publications.

Stacey, R. (1996). *Strategic Management and Organizational Dynamics.* London: Pitman Publishing.

Sterman, J.D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. United States: McGraw-Hill/Irwin.

Sterman, J.D. (2006). Learning from evidence in a complex world. *Public Health Matters*, *96*(3), 505-514.

Stetler, C.B., Ritchie, J.A., Rycroft-Malone, J., Schultz, A.A. & Charns, M.P. (2009). Institutionalizing evidence-based practice: an organizational case study using a

model of strategic change. *Implementation Science, 4*(78).

Suppovitz, J., & Snyder, T. (2005). Systemic education evaluation: evaluating the impact

of systemwide reform in education. *American Journal of Evaluation*, *26*, 204-230.

The National Academies Keck Futures Initiatives. (2009). Task Group Summary 3: How

can we enhance the robustness via interconnectivity? In *Complex Systems: Task

group summaries*. Washington, D.C.: The National Academies Press. Retrieved

from http://books.nap.edu/openbook.php?record_id=12622&page=R1

Timmreck, T. C. (1997). Health services cyclopedic dictionary (3rd. ed.). Boston, MA:

Jones and Bartlett Publishers.

Titler, M., Everett, L., & Adams, S. (2007). Implications for implementation science.

Nursing Research, *56*(4S), S53-S59.

Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings.

American Journal of Community Psychology, *39,* 217-228.

United States Department of Health and Human Services. (2004). *The Health

Consequences of Smoking*. Atlanta, GA: Centers for Disease Control and

Prevention, National Center for Chronic Disease Prevention and Health

Promotion, Office on Smoking and Health.

Van de Ven, A., & Poole, M. (2005). Alternative approaches for studying organizational

change. *Organization Studies*, *26*, 1377- 1404.

Vancelik, S., Beyhun, N., Acemoglu, H., & Calikoglu, O. (2007). Impact of

pharmaceutical promotion on prescribing decisions of general practitioners in

Eastern Turkey. *BMC Public Health*, *7.*

Vogel, M.P. (2009). Understanding emergent social phenomena comparatively: The need

for computational simulation. *European Journal of Social Sciences, 7*(4), 84-92.

W.K. Kellogg Foundation. (2007). *Designing Initiative Evaluation: A Systems-oriented Framework for Evaluating Social Change Efforts*.

World Health Organization. (2009). *Cancer Key Facts* (Fact sheet No. 297). Fact sheet.

Willy, C., Neugebauer, E. A. M., & Gerngroß, H. (2003). The concept of nonlinearity in complex systems: An additional approach to understand the pathophysiology of sever trauma and shock. *European Journal of Trauma*, *29*, 11-22.

Wingo, P., Ries, L., Giovino, G., Miller, D., Rosenberg, H., & Shopland, D. (1999). Annual report to the nation on the status of cancer, 1973-1996, with a special section on lung cancer and tobacco smoking. *Journal of the National Cancer Institute*, *91*, 675-690.

Zhu, S., Anderson, C., Tedeschi, G., Rosbrook, B., Johnson, C., Byrd, M., & Gutiérrez-Terrell, E. (2002). Evidence of real-world effectiveness of a telephone quitline for smokers. *The New England Journal of Medicine*, *347*(14), 1087-1093.

Zimmerman, B. (2001). *Edgeware: insights from complexity science for health care leaders*. USA: CreateSpace.

Appendices

Appendix A. KIQNIC Survey List of Innovations

1. Provide proactive (outbound) telephone counseling
2. Provide reactive (inbound) counseling
3. Provide a multiple call protocol (two or more calls for the same quit attempt)
4. Provide free (or discounted) NRT to callers without requiring counseling
5. Provide NRT to callers but require that they register for counseling
6. Conduct mass media promotions for the mainstream population
7. Conduct mass media promotions for targeted populations
8. Provide self-help materials to proxy callers
9. Provide self-help materials to callers calling for self
10. Provide self-help materials to all callers who register for services
11. Provide counseling immediately to all callers who request it (either through real-time staff capacity, or on-call staff capacity)
12. Staff the quitline with counselors who meet or exceed Masters-level training
13. Conduct an evaluation of the effectiveness of the quitline
14. Serve callers without insurance coverage
15. Obtain Medicaid or other insurance reimbursement for counseling provided to callers
16. Refer callers with insurance to health plans that provide telephone counseling
17. Use text messaging to provide tailored support in conjunction with or instead of telephone counseling
18. Combine or link telephone counseling to face-to-face counseling programs
19. Combine or link telephone counseling to internet-based or eHealth programs (or components of the quitline program)
20. Fax-to-quit or fax-referral program
21. Re-contact relapsed smokers for re-enrollment in quitline services
22. Supplement quitline services with Interactive Voice Response (IVR) services (e.g., automated check-in IVR calls for relapse prevention)
23. Train provider groups on 2A's or 3A's and refer (with or without a fax referral program)

Appendix B. KIQNIC Survey Implementation Stage Question

Innovation: Conduct an evaluation of the effectiveness of the quitline

1. Are you aware of this practice? ○ Yes   ○ No

2. Where are you in the decision-making process?
(If your quitline has initiated any actions towards putting a practice into action, it is assumed that a decision to implement the practice has been made)

○ Have not yet discussed   ○ In discussion   ○ Decided not to implement   ○ Decided to implement   ○ Not sure   ○ N/A

3. Where are you in the implementation process? Please indicate your response using the following five-point scale, where:

 1 = No progress has been made yet

 2 = A low level of implementation has been reached (e.g. some discussion, staff informed, someone assigned to lead the process, etc.)

 3 = A medium level of implementation has been reached (e.g. formal plan for implementation, resources committed, training begun, etc.)

 4 = A high level of implementation has been reached (e.g. pilot project has been implemented, or other testing has begun)

 5 = Fully implemented (the practice has become part of the quitline's policy or standard operating procedures for all eligible callers).

 No progress ⇒⇒⇒⇒ ⇒ Fully implemented
<practice> ○1 ○2 ○3 ○4 ○5 ○ Not sure   ○ N/A

Appendix C. Example of Recruitment Email

Dear <u>Participant</u>,

I hope you are doing well.

I'm sending you this email because you expressed an interest in participating in an interview for a dissertation research project, on the recent KIQNIC survey you completed. The study I'm conducting will look at how innovations (i.e. new practices/policies) are disseminated and implemented in the North American quitlines. This project is unique in that it will collect qualitative data and will apply a systems thinking lens to data collection and analysis.

Because innovation is a broad topic, I have chosen to explore in more depth, the implementation of an evaluation practice, specifically. Evaluation was chosen as a focus after much consideration and feedback from NAQC members. I anticipate that this project will provide information that can be used to improve innovation implementation generally, as well as specifically to improving evaluation in the quitline network by collecting data on issues relevant to implementation, such as: knowledge transfer, barriers and facilitators to implementing and maintaining evaluation and, system level patterns and characteristics.

For your participation in this study, you will be asked to give an interview that will take approximately 45-60 minutes. The interview will be scheduled for a time that is convenient for you, and would be conducted over the phone. In return for your assistance, you will be provided with a $5 gift card for Starbucks. The gift card will be mailed to you upon completion of the interview.

If you are still interested in participating in the interview, or have any questions, then please email or call me at the number or email address listed below.

Thank-you in advance for your time and support. I look forward to hearing from you!

Sincerely,
Jennifer

Appendix D. Interview Script

I. **Introduction**

My name is Jennifer Terpstra and I'm a Doctoral Student at the University of British Columbia in Vancouver, BC, Canada. I'm conducting this study for my PhD dissertation, which will look at how new and innovative practices are implemented in the North American quitline network. Because this is a fairly broad topic, I've decided to focus my study by looking specifically at the implementation and institutionalization of evaluation, which is what I'll be asking you about today. The findings from this study will be provided to NAQC and hopefully used to help improve networks capacity for evaluation.

The interview will take approximately 45-60 minutes depending on the length of your answers. If at anytime during the interview you feel uncomfortable and wish to not answer a question or stop the interview, you are free to do so, and that will be fine. In return for your assistance and any inconvenience the interview process may cause, you will be provided with a $5 gift card for Starbucks. The gift card will be mailed to you upon completion of the interview, at a mailing address that you can provide at the end of the interview.

II. Consent Form

Before I start the interview I am going to review a few details. The interview will be tape recorded and transcribed. I will not put your name on the transcript and I will provide you with an opportunity to review the transcript prior to presenting any of the results.

Did you receive and have an opportunity to review the consent form that I mailed to you? (if no, then will review with the participant). *Do you have any questions or concerns?*

III. Interview questions

1. Can you tell me about your quitline?
Probes:
- Number of people, history, date of startup, funding source(s)
- One quitline, or multiple
- Funder, service provider
- Stable/consistent employees

2. Describe your role in the organization. And the quitline community?
Probes:
- Responsibilities
- Length of time in organization and role
- Role with respect to the quitline
- Role in evaluation

3. Describe the QLs funding
Probes:
- Expectations from legislators
- Stability of the funds
- Ongoing commitment from state to fund the quitline

- Competition for funds, risk
- Any CDC funds?

4. Relationship with __(partner organizations)_____
Probes:
   - Stability
   - Involvement with decision-making

Evaluation Overview
5. What does it mean to "evaluate effectiveness" of your Quitline?

6. On the KIQNIC survey you said that you had "fully implemented" the practice of evaluating effectiveness. What does that mean?

7. What type of evaluation data do you collect?
Probes:
   - Operational data
   - Process
   - Outcomes

8. Who does the evaluation for your quitline? Has it always been done this way?
Probes:
   - Collects the data?
   - Analyzes it?
   - Uses the information once it is collected?

9. When and why did you begin evaluating your quitline?

10. What is the goal of the evaluation?
Probes:
   - Mandate
   - Incentives
   - Funder

11. How are decisions about evaluation made?
Probes:
- Factors considered?
- Who decides, who's invovled
- Who's needs need to be met?

12. What type of evaluation is expected from your quitline, if any?

13. Are there any challenges or barriers to doing evaluation in your quitline?
Probes:
- Skills and knowledge
- Capacity

- Financial

14. Are evaluation results used, and if yes, how?
- The state reports, legislators
- Funding applications
- Publications

15. How has evaluation in your quitline and organization changed over time? Why have these changes occurred?
Probes:
- Key events
- Actors involved
- Perspectives, system,
- Funding
- Type of data collected and used

16. Is there other evaluation that you would like to do but can't?

17. How has funding impacted evaluation in your organization?

18. Do you share information about evaluation issues with other quitlines and quitline organizations? Why or why not?

19. How have you seen evaluation change in the broader quitline community over the years? Why?

20. Do you think evaluation and data collection is being used to its maximum potential in the quitline community?

21. Will the CDC stimulus funding reporting requirements affect evaluation in your quitline?

22. How do you feel about CDC's plans to collect QL evaluation data and put into a common database for sharing and comparisons?

That was the last question. Is there anything else that you would like to add?

Snowball Sampling Questions
1. During the interview you mentioned __(name of person)__. Would you recommend that I interview this person as well? Is there anyone else that you would recommend I interview?
2. Would you be willing to forward the study information email to them for me?

IV. Concluding Statement
Thank you for your time, the information you've shared with me today is very valuable for my dissertation. As I mentioned before the interview, I will provide you with an

opportunity to review the interview transcripts once it's been transcribed. At that time, if there is anything that you want removed from the transcript, I can accommodate that request. I anticipate that the transcription will be complete within the next 3 weeks.

If you have any questions in the meantime, or think of additional information that you'd like to give me, please feel free to contact me either via email or at the phone number listed on the consent form.

Appendix E. Final Codebook

Overarching Themes	Subthemes	Sub-subthemes
Descriptive Information (Refers to information describing the interviewees, QLs, and the system)	**Interviewee characteristics**	Position Involvement in network Length of employment Evaluation experience
	QL characteristics	Date of operation History Partner organizations Funding
	System description	Funding Actors System levels Niches
System norms (Information describing interviewee's worldviews including perspectives, assumptions, values, and beliefs)	**Evaluation Definition** (Information describing perceived definition of the practice)	
	Goal (Information describing the perceived goal of the practice)	Monitor outcomes Quality Improvement Furthering Evidence-base
	Justification (Information pertaining to justification of funding as a driver for evaluation)	
	Relationship (Information describing the perceived relationship between evaluation and QL operations)	
	Utilization (Information pertaining to utilization of evaluation results)	
	Aggregation (Information that describes thoughts on data aggregation and cross-QL comparisons)	Fear

Overarching Themes	Subthemes	Sub-Subthemes
System Resources (Information describing the resources available to implement the practice)	**Human resources**	Skills and knowledge Designated staff
	Social resources (Relationships that influence implementation)	QL-QL NAQC & CDC Service provider & funder QL to state QL – evaluation contractor
	Economic resources	Cost of evaluation Funding arrangements
System regulations (Information describing the regulations in the system that influence implementation)	**Mandate: Eval budget** (Information describing mandates to allocate budget for evaluation)	
	Mandate: MDS data (Information describing mandate to collect MDS data)	
System Operations: Power & decision-making (Information describing power and decision-making authority in the system)	**NAQC**	Conformity & inclusion
	Federal gov't & CDC	
	State funders	
	Service providers	
	Information & resources (Information describing power issues related to information and resources)	Ownership of data Type of data collected

Appendix F. Focus Group Invitation

 Dear <u>Participant</u>,

Hope you're doing well. I'm emailing you to invite you to participate in a focus group/Webinair session to discuss and validate the results of my dissertation. I recently completed the analysis of the 19 interviews I conducted with stakeholders in the quitlines and would like an opportunity to present the results and get feedback from some of my participants. I'm looking for about 5-8 individuals to participate in a two hour focus group/Webinar session, during which time, I will present the results of my study (approx 30min).

The focus of my dissertation was the implementation of new practices in the quitlines, and I specifically looked at the implementation of the practice: to evaluate effectiveness of quitlines. The results of my study suggest that both the practice and the system are high in dynamic complexity, and as a result, successful implementation requires systems change. In my presentation, I will discuss the qualitative themes that emerged from the analysis in relation to systems change for implementing this evaluation practice. I will also discuss the potential "leverage points" in the system that could be used to achieve systems change for implementation of this practice.

I'd like to hold the focus group in mid-June, date yet to be determined. If you are interested in participating in the focus group/Webinair session, please let me know via email within the next week.

Thank-you in advance.

Sincerely,
Jennifer