A THEORY OF PROGRAM EVALUATION PRACTICES

IN DISABILITY MANAGEMENT

by

Patricia Louise Swenson

B.A., Simon Fraser University, 1980
M.A., Simon Fraser University, 1986
M.A., The University of British Columbia, 2004

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

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Measurement, Evaluation & Research Methodology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2014

Abstract

This grounded theory study developed a theory of evaluation in disability management programs. Disability management involves managing the interactions between health condition impairments and their environments to overcome functional barriers. A sample of four sites was selected each site representing a different paradigm of disability management practices: biomedical, labour, biopsychosocial or insurance. Data collection included semi-structured interviews with 9 participants, including an administrator and practitioner from each site, the Readiness for Organizational Learning and Evaluation Instrument, and documents from each site were analyzed. There were five major findings of the study. 1) Meaningful disability management program evaluation requires insight into how impairment environment interactions are being managed by the program. 2) The presence or absence of collaboration among stakeholders contributes significantly to the variability in disability management and disability management evaluation. 3) Understanding how disability management programs are adapting to contextual influences contributes significantly to an explanation of variability in disability management and disability management evaluation. 4) There are five primary disability management evaluation criteria: return to work, cost savings, timeliness of services, client satisfaction, and client functioning. 5) Disability management evaluation followed a consumer working logic approach, and was predominantly concerned with usefulness of services, and secondarily framed from perspectives of multiple stakeholders. Additionally, disability management programs and their funding organizations are increasingly using technology to develop new data management systems for future use in evaluation.

Preface

This dissertation is original, unpublished, independent research by the author, Patricia Louise Swenson.

The research study reported in Chapters 1, 3, 4 and 5 were covered by:

1. University of British Columbia Behavioral Research Ethics Board Certificate of Approval # H09-02993
2. Vancouver Coastal Health Authority Clinical Trials Administration Office Approval Research Study # V10-0051

Table of Contents

List of Tables

List of Figures

Acknowledgements

I express my sincere gratitude to all the faculty members who have guided and supported me throughout my studies at the University of British Columbia, including during this research study my committee members Dr. Sandra Mathison, Dr. Izabela Schultz and Dr. Nand Kishor. I am grateful to Rebecca Trainor of the Faculty of Graduate and Postdoctoral Studies for her encouragement during the dissertation process. I am very appreciative of the camaraderie shared with my fellow MERM students during our time together at UBC, and thank my employer WorkSafeBC and Dr. Robert McKenzie for their support. Last but not least I thank my daughter Missy Stone for her faith and Lionel Webb for supporting me in so many ways.

Dedication

I dedicate this research study to Paramahansa Yogananda who inspires appreciation for

lifelong learning.

CHAPTER 1: INTRODUCTION

1.1 Background

A recent analysis of the state of evaluation in Canada concluded there is too little research on evaluation and the authors called for studies to evaluate the relevance, performance, outputs and contributions of any evaluation (Gauthier et al., 2009). Research on evaluation practice can "keep current problems in evaluation in better historical perspective, provoke thoughtful consideration of present options, and enable us to create more effective alternatives for the future" (Smith and Brandon, 2008, p. viii). A growing evidence base would contribute answers to questions such as, "Which approaches to evaluation, implemented how and under what conditions, actually lead to what sort of improvements" (Mark, 2008, p. 115).

In response to this call for more research on evaluation the current study explored evaluation practice in the field of disability management. Over the past two decades little has been documented about disability management evaluation, however, during this same period the disability management field has expanded into a multi-billion dollar industry worldwide. By examining evaluation practices in disability management programs, this study will to some extent fill the void in research on evaluation by developing a theory regarding disability management evaluation.

Originally disability management was based on a biomedical perspective, where medical professionals oversaw treatment of impairments, and employers were considered responsible to provide jobs that were suitable to accommodate disabilities. As different stakeholders' interests became more prominent in the field (such as workers, employers, insurance funders, multi-disciplinary treatment teams, unions) disability management

expanded and evolved to include those various different perspectives. Multiple paradigms of practice emerged in the field representing that diversity of stakeholder perspectives.

With the expansion of paradigms it was recognized that research (and presumably program evaluation) on disability management needed to examine levels of complexity that had developed in relation to multiple stakeholder priorities (Pransky, Gatchel, Linton and Loisel, 2005), "An ideal model of [return to work] should make sense from multiple stakeholder perspectives, and incorporate a range of their priorities – especially sustained employment, worker productivity, and costs, as well as key features of the [return to work] process. It should serve to bridge the gulf between traditional biomedically-driven practices, and empirically supported biopsychosocial approaches that are more acceptable, and perhaps more effective in driving meaningful change in [return to work]-related practices" (p.456).

Given that evaluation and disability management programs often operate within complex and dynamic organizational systems, related factors identified in the literature on evaluation and disability management can be identified. This study of disability management evaluation examined the roles of diversity, cultural competence and organizational learning within evaluation, and focused on four disability management paradigms: biomedical, labour, biopsychosocial and insurance.

1.2 Conceptual Underpinnings

A social constructivist epistemology underlies this research. Constructivism holds there is not one objective external truth that exists and is waiting to be discovered, but that meaning is made through our conscious engagement with our world and we do not discover knowledge so much as construct it (Crotty, 2003). A constructivist epistemology

does not strive to build a theory of one single truth, but assumes that knowledge that is real is based on multiple perspectives (Nagy Hesse-Biber, 2007). "People do not invent the world anew each day. Rather, they draw upon what they know to try to understand what they do not know" (Corbin and Strauss, 2008, p. 75). As individuals are exposed to new information they continually develop understandings, and throughout this process construct new meanings and new values.

Social constructivism refers to a sociocultural and historical dimension of this construction (Schwandt, 2007) where knowledge is shared among people. Social constructivism holds that meaning is made through our conscious engagement with the world (Crotty, 2003) and historical and sociocultural shared understandings of ideas or facts (Schwandt, 2007). To understand how people are forming interpretations the enquirer must enter into their situation to see it from their perspective, consider what they take into account, and how they interpret information encountered.

Theoretically this study assumed a subjectivist approach to valuing. The sample of programs and participants were selected to maximize perspectives as diverse as possible. This included sampling four disability management sites that each represented a different paradigm of disability management, and selecting participants from each site that included a practitioner and an administrator. The study focused on the unique combination of influences faced at each disability management site, including exploration of the site's context, diversity, organizational learning and cultural influences.

This study used a grounded theory research methodology. Grounded theory involves data collection and simultaneous analyses employing techniques of induction, deduction and verification (Schwandt, 2007). Grounded theory methods included two

levels of data coding. Open coding involves the researcher naming events and actions in the data, constantly comparing them to one another to analyze how they relate (Harry, Sturges and Klinger, 2005). Conceptual coding reflects grouping of open codes with similar properties, and the researcher identifying meaningful themes from analyses of the conceptual findings. Themes are tested and interrelated as an explanation emerges of how the substantive model operates, culminating in development of a theory.

1.3 Purpose of this Research

Little research has been done on evaluation of disability management, and much of the research on disability management evaluation focuses on outcome studies that measure economic and social cost savings, and return to work. Prior research has adopted a narrow definition of disability management and a narrow perspective of evaluation.

Seeking a single narrow perspective of a phenomenon rather than exploring diverse perspectives is a potential loss of important information (Guba and Lincoln, 1989). Consideration of context and diversity within disability management evaluation offers the potential for new insights.

Disability is a constructed variable relative to its context (Smart, 2001). For example, various disabilities can be seen as best managed by medical experts, or disabilities can be seen as the responsibility of those with disabilities. And, how the individual manages their disability varies a great deal based on both personal and contextual factors. Recent research gives central consideration to contextualization of the individual, and how the personal system of the individual (physical, cognitive, affective and social factors) interacts with health, workplace, and compensation systems (Loisel et al., 2001).

Characteristics of organizations in which disability management programs exist, including how learning is perceived within the organization, also contribute to understanding disability management programs and their evaluation. Organizations that encourage learning offer an opportunity for change and renewal (Kaufman and Senge, 1993) and foster cultures likely to build evaluation capacity (Taut, 2007).

Learning and evaluation in organizations have a synergistic relationship. Evaluation can be a mechanism to build learning within organizations: "For organizational learning to occur, it is critical that an environment for learning be established and maintained. This involves creating processes that support employees' efforts to reflect on their experiences, discussing and analyzing how their efforts contribute to the organization's strategic plan, and assessing current work systems to determine their effectiveness in meeting customer needs and expectations" (Preskill, 1994, p. 292).

In an area where little research has been conducted, as is the case with disability management evaluation, research on current practices fills a void in understanding evaluation within a particular domain.

Four research questions guided this study:

#1. What is the extent and nature of evaluation practice within the disability management program?

#1a. How does disability management evaluation practice vary depending on whether the organization is a learning organization?

#1b. How does disability management evaluation reflect diversity and cultural constructions?

#1c. What evidence is there that disability management evaluation is grounded in a particular paradigm of disability and return to work?

1.4 Assumptions and Limitations of the Study

This research examined evaluation at four sites each representing a different paradigm of disability management: biomedical, labour, biopsychosocial and insurance. "The boundaries among the models are somewhat arbitrary as they share many common themes and factors" (Schultz, Stowell, Feuerstein and Gatchel, 2007, p. 313) but the distinctions are robust enough to categorize disability management programs and thus to examine potential differences in approaches to evaluation. Evidence of multiple paradigms may be present within any one site to greater or lesser degrees; however, each site was selected with the understanding that services at the site were predominantly representative of one paradigm.

Each site was selected as representative of one particular paradigm based on my familiarity with each program, reviews of documentation provided by each site, initial conversations with participants, reviews of the literature on disability management paradigms, and my analysis of this information based on my knowledge gained from over 20 years experience working with disability management programs. Although practices at each site primarily reflect one paradigm, I recognized that individuals from the same site participating in this study may have aligned their beliefs with different paradigms, as each of them holds unique understandings about what is valued and how it should be evaluated. Each individual's understandings may also reflect characteristics of multiple paradigmatic orientations, and their views could change over time.

During the study I made a conscious effort to recognize the influence of my own understandings. Corbin and Strauss (2008) argued, "more than one story can be derived from data" (p.50). Different analysts vary on what they focus on, interpretations they make, meanings they relate to and conclusions they draw, "furthermore the same analyst might look at the same data differently at different times" (p.50). I have attempted to maintain awareness of my perspectives and clarify potential influences on interpretations.

CHAPTER 2: LITERATURE REVIEW

2.1 Program Evaluation and Planning

Strategic program planning starts with identification of needs and the development of program objectives that link the program to serving those needs (Taylor-Powell, 2006). Ideally, acceptable standards of services are explicitly stated. Programs conceptualized at senior levels of an organization or where funding is being decided, rely on the expertise of experienced administrators to lay program foundations, and input from all levels for support throughout program implementation (Curtis and Scott, 2004). Stakeholder buy-in is a process that requires building consensus without exception on an ongoing basis (Rankin, 2001). Integrated, pluralistic approaches to leadership ensure that efforts throughout the organization, point in one direction toward a shared vision (Preskill and Torres, 1999).

It is generally accepted that evaluation should be conceptualized during program planning, including formative evaluation to create opportunities for program feedback intended to support the process of improvement, and summative evaluation to assess the degree to which program outcomes meet targets (Wholey, 1996). During the planning phase of a program, consideration should be given to establishing regular program evaluations. This includes conducting needs assessments, establishing criteria and standards expected to meet needs, developing methods to evaluate whether program objectives are being met, and providing up to date information to understand program successes and changes, "Ongoing evaluation is really the only objective way of knowing what aspects of your program are working and what aspects need reviewing" (Rankin, 2001, p. 129).

Those closely involved with day to day programmatic operations can be sources

of knowledge regarding program functioning that may contribute to evaluative insights

for redirection or ways to mine previously undiscovered opportunities (Mayne, Divorski

and Lemaire, 1999; Sonnichesen, 1999). Whether evaluation is conducted by non experts,

internal evaluators or external evaluators, learning should be promoted through

evaluation practices, as well as fostering further evaluation development and utilization

of results (Lemaire and Boyle, 1999). To achieve quality standards, evaluation findings

should be communicated to all levels of the organization for use to address deficits and

make improvements (Harder and Scott, 2005; Nickerson, 2000; Strasser, 2004).

2.2 Foundational Issues in Research on Evaluation

There are a number of recurring foundational issues identified in the research on

evaluation literature: the role of the evaluator, stakeholder participation, establishing

meaningful criteria and standards to evaluate, and exploring ways to ensure findings are

useful (Smith and Brandon, 2008). Additionally, other emerging issues in the evaluation

literature are: the importance and role of cultural competence, promotion of a

transformative theme which strives to address power discrepancies, identification of

interacting variables that can place some individuals at a disadvantage (Mertens, 2008),

and, how organizational openness to learning, change and evaluation capacity building

can contribute to sustained and dynamic evaluation benefits (Preskill and Torres, 1999;

Senge, 2006). Consideration of context in which evaluation occurs can contribute

relevance, rigor, and improved opportunities for evaluation findings to be generalized

(Rog, 2012). Several of these issues seem particularly salient to disability management

program evaluation: cultural competence, diversity, organizational learning and context

analysis. These issues have the potential to contribute multi-dimensional perspectives of issues that impact disability management and its evaluation.

2.2.1 Cultural competence. Sensitivity to cultural diversity in program development, implementation and evaluation involves the understanding and valuing of multiple dimensions, perspectives and world views of diverse stakeholders. Cultural competence in program evaluation "… rests on active awareness, understanding, and appreciation for the context at hand, and it uses responsive and inclusive means to conduct evaluation" (SenGupta, Hopson and Thompson-Robinson, 2004, p. 12). Cultural competence in evaluation has been defined as, "…systematic, responsive inquiry that is actively cognizant, understanding, and appreciative of the cultural context in which the evaluation takes place; that frames and articulates the epistemology of the evaluative endeavor; that employs culturally and contextually appropriate methodology; and that uses stakeholder-generated, interpretative means to arrive at the results and further use of the findings" (SenGupta, et al., 2004, p.13).

Madison's (2007) review of publications focusing on cultural competence in evaluation over the prior twenty years reported that evaluators found cultural responsiveness and cultural competence make a positive difference in evaluation outcomes and utilization, and are worth the increased investment of time. Cultural competence involves evaluators seeking awareness of their own culturally-based assumptions, understanding worldviews of culturally-different participants, and using appropriate evaluation strategies and skills in working with culturally different groups. Culturally significant factors include,

The shared experiences of people, including their languages, values, customs, beliefs, and mores. It also includes worldviews, ways of knowing, and ways of communication. Culturally significant factors encompass, but are not limited to race/ethnicity, religion, social class, language, disability, sexual orientation, age and gender. Contextual dimensions such as geographic region and socioeconomic circumstances are also essential to shaping culture….Cultural groupings can refer to...organizational culture, gay culture, or disability community culture. Culture also refers to the institutions (such as government, education, family, and religion) and economic systems that shape and preserve shared patterns of thought, behavior, and beliefs. (American Evaluation Association, 2012, p.3).

Culturally competent evaluators are encouraged by professional standards, such as those offered by the American Evaluation Association, to see cultural categories as fluid, and to avoid reinforcing cultural stereotypes and prejudice (for example when working with data organized by cultural categories).

Cultural competence in health care, and by extension disability management, involves sensitivity to and understanding of individuals' beliefs and values in relation to their heritage. Cultural competence for health care providers has been defined as: awareness of one's self without having undue influence on those from other backgrounds; demonstrating knowledge and understanding of clients' culture, health needs and views of health and illness; accepting and respecting cultural differences; not assuming clients' and providers' beliefs and values of health care are the same; resisting judgment; being open to cultural encounters; and consciously adapting to be congruent with the client's culture (Purnell and Paulanka, 2003).

2.2.2 Diversity. Disability management contexts include the simultaneous interests of multiple stakeholder groups (Franche, Baril, Shaw, Nicholas and Loisel, 2005; Young, Wasiak, et al., 2005). Overlooking the potential impact of multiple variables, risks making incorrect assumptions, and missing important information regarding complex program contexts and program participant motivations (Loisel et al., 2005). Research on evaluation of disability management has not explored potentially valuable information regarding multiple perspectives and, "…their social origins in workplace interaction or in particular institutional policies and administrative structures" (Eakin, Clarke and MacEachen, 2002, p.8).

Viewed from a systems theory perspective, disability management involves understanding the diversity of multiple stakeholders, including their motivations, interests and concerns. A systems theory perspective, "…maintains that people with disabilities and their life outcomes are influenced by the family, school, peer, independent living, employment, health and rehabilitation service, and social – political- economic environments" (Young, Wasiak, et al., 2005, p.544). Conceptualization of disability has been changing to a paradigm where it is thought to involve an interaction among the individual, the disability and the environment (Smart and Smart, 2006). As disability management considers the interactions between impairments and their environments, the disability management process requires understanding diverse psycho-social factors within the environment, and interactive processes between diverse stakeholders: employee, employer, insurer, health care provider (Franche and Krause, 2005).

2.2.3 Organizational learning and evaluation. Within the current era where knowledge is a valuable commodity, collaboration and learning are keys to organizational

sustainability. In learning organizations "people are always enquiring into the systematic consequences of their behavior rather than just focusing on local consequences" (Kafman & Senge, 1993, p. 16) and what they can learn is more important than what they already know. However, people can be resistant to change (Beer, 2009) and a great effort must be made for organizations to achieve a cultural shift to become learning organizations. Constructivist learning theory considers learners as active not passive where behavior is mediated by the social environment (Preskill and Torres, 1999). The process of evaluative inquiry, when grounded in a constructivist theory of learning, builds organizational learning through the following steps: "…(a) the collective creation of meaning, (b) action, (c) the development of new knowledge, (d) an improvement in systemic processes, and (e) the overcoming of tacit assumptions" (Preskill and Torres, 1999, p. 49).

Constructivist learning theory holds that learning is about making meaning, and is built upon the belief that all knowledge is based on experience and that meanings are arrived at by continually seeking order in these experiences. Rather than just reacting to whatever they encounter in the world, people are purposive and confront issues so that they can make meaning of one another's actions (Schwandt, 2007). Constructivist learning theory is useful for understanding learning in organizational environments, "Adding the sociocultural variable to learning, social constructivism theory views learners as active agents in the construction of outcomes and stresses that the social setting itself is an evolving construction. When members of a social setting (e.g. an organization) share their social constructions, the cycle of learning is renewed" (Preskill and Torres, 1999, p. 20). Preskill and Torres argue that a learning culture grounded in

social constructivist theory assumes collective creation of meaning and development of new knowledge, overcoming of assumptions, and dissemination of knowledge throughout the organization.

Organizational learning has been enthusiastically embraced as a means of enhancing capacity for change and renewal. In learning organizations people are encouraged to be open, to let go of assumptions and certainties, and risk learning about complex issues. This takes vision and courage to look past the usual stability and examine possible systemic consequences of one's actions (Kofman and Senge, 1993; Senge, 2006). Organizations that encourage evaluation, and that foster cultures that appreciate learning from evaluation, are likely to build evaluation capacity, and evaluation is likely to have impact within those organizations (Taut, 2007). Success in becoming a learning organization involves individuals making changes by creating, acquiring and transferring knowledge (Owen, 2005). The process is dynamic, where individuals from all levels of the organization take responsibilities for the creation and transfer of learning, as they are considered the experts of their own learning needs. This requires basic shifts in how we think and interact beyond the individual within the corporation, penetrating our assumptions and habits (Kofman and Senge, 1993).

2.2.4 Context. The environment or setting in which a program functions is what evaluators most commonly view as context (Rog, 2012), and understanding the social realities within the context contributes to a more valid interpretation of program evaluation findings. Conner, Fitzpatrick and Rog (2012) recommend "placing context among the primary considerations that are involved in the evaluation process" (p.89)

including doing context analysis during evaluation planning, during implementation and during utilization of findings.

Rog (2012) proposed five areas of context analysis to consider: the phenomena and the problem; the nature of the intervention; the broader environment/setting; the evaluation context; and the decision making context. Within each area of context analysis Rog proposed considering physical, organizational, social, cultural, tradition, historical and political dimensions. As circumstances are dynamic, any of these five areas may be relevant at given times, and other significant influences may also be discovered.

Context is a complex phenomenon, where most contexts have multiple layers and multiple dimensions that can be interacting in important ways (Greene, 2005). In evaluation, information can lose meaning if decontextualized, and "good evaluation is responsive to, respectful of, and tailored to its contexts in important ways" (p. 84).

2.3 Evaluation Theory

There is no single theory of evaluation. Rather, evaluation is comprised of many different models or approaches that explain activities and processes of evaluation applied in relation to specified goals and depending on particular sets of circumstances and assumptions. Evaluation theory is "that aspect reflecting our thinking about how and why we engage in evaluation; whether evaluation is done for purposes of validation, accountability, monitoring, or improvement and development; whether evaluation is a form of knowledge production, client service, social reform, or political control" (Smith and Brandon, 2008). Multiple classifications have contributed to understanding the formalization of evaluation.

One of the most notable classifications of evaluation practice was the seminal work of Shadish, Cook and Leviton (1991) that described stages of evaluation and significant developments within the field. Evaluation had primarily borrowed methods from other social sciences until the 1960s when growth of social programs led to substantial focus on their evaluations. Theories at this stage emphasized scientific rigor in solving social problems, and focused on cause-effect relationships. Evaluation theorist Scriven developed a four step logic of evaluation to generate value statements about any entity: select criteria of merit; set standards of performance; measure performance; and synthesize results in to a value statement. This approach was directed to consumers, and remains a main logic within evaluation. Campbell clarified traditional scientific experimental methods versus quasi experimentation, and internal and external validity.

During the 1970s stage two theories criticized the scientific approach for having had an inadequate focus on evaluation use to improve social programs. Theoretical focus shifted to ways of increasing use, involving for example identification of intended users of evaluation findings, determining information needs of decision makers, and providing information on why programs had failed and how they could succeed. Theorists including Weiss, Wholey and Stake focused more on enlightenment evaluation (for long term policy changes) rather than instrumental evaluation (for incremental improvements). These theories emphasized pluralistic approaches and multiple methods compared to stage one theories, and considered questions about program description, explanation, generalization and discovery, rather than just questions about causation.

Stage three theories synthesized work from the preceding stages, including use to improve social programs and systematic methods to obtain valid knowledge. The focus of

these theories was how evaluation could influence policy, primarily concerned with enlightenment rather than formative or summative program evaluation. These evaluation theories continued the focus on using evaluation to justify and improve future programs, with consideration of the conditions or contexts under which programs occurred and potential evaluation impacts.

In an effort to overcome weaknesses of earlier generations of evaluation theory that had been based on objectives, description and judgment, Guba and Lincoln (1989) introduced fourth-generation evaluation that focused on intensive stakeholder participation. This approach assumed a constructivist epistemology with an objective of uncovering multiple values and used mixed methodologies, fitting the method to the question (Lincoln, 2005). A range of other participatory and collaborative approaches have developed, including: practical participatory, transformative participatory, democratic, developmental, and empowerment evaluation (Cousins and Whitmore, 2007).

A seminal classification of evaluation theory developed by Christie and Alkin (2013) used a tree metaphor to explain the roots and branches of evaluation theory. The evaluation theory tree has three roots: social inquiry, epistemology and social accountability, each contributing to development of the field in different ways. The social inquiry root contributes systematic, methodical and justifiable evaluation procedures for being accountable. The epistemology root enables arguments on the nature of knowledge. The social accountability root has been an important motivation for evaluation to improve programs and society. There are three branches on the evaluation tree: methods, valuing and use. The methods branch of evaluation grows predominantly from the social inquiry

root, and is primarily guided by research methodology. The valuing branch grows predominantly from the epistemology root, and is either objectivist (evaluator driven) or subjectivist (pluralistic attending to multiple stakeholders' values). The use branch grows predominantly from the social accountability root and focused on decision making. The tree metaphor is three dimensional, situating theories and theorists on the tree in relation to all roots and branches.

On a conceptual level, critical features contribute to evaluation theory throughout its ongoing development. These features include theories of knowledge construction, valuing, purposes, practices, and use, which along with variables such as types of data or evaluation logic, guide evaluation practices and comprise evaluation theories.

Epistemology, the theory of knowledge, is embedded in theoretical perspectives that guide social science, including: objectivism (that assumes meaning and reality exist apart from any consciousness, and an object exists whether or not anyone is aware of it); or, constructivism (that assumes there is no objective truth waiting to be discovered, truth comes into existence through our engagement with the world as we construct meanings about things we encounter) (Crotty, 2003). Like social sciences, epistemological perspectives underlie the perspectives of evaluators and evaluation.

It is valuing that distinguishes evaluation from general social science research. Evaluation theory addresses questions about valuing, such as: whether evaluation should compare programs to each other or compare programs to established standards, or whose criteria and standards should be considered for judging programs. Valuing theory also specifies the nature of metaevaluation, including justification, validation and verification (Mathison, 2005).

Valuing within evaluation theory has generally been classified as either descriptive or prescriptive. Descriptive refers to "a set of statements and generalizations that describes, predicts, or explains evaluation activities-such a model is designed to offer an empirical theory". Prescriptive refers to "a set of rules, prescriptions, prohibitions, and guiding frameworks that specify what a good or proper evaluation is and how evaluation should be done" exemplars generated by knowledgeable members of the evaluation field (Alkin, 2013, p.4). Prescriptive program evaluation "consists of an explicit theory or model of how the program causes the intended or observed outcomes and an evaluation that is at least partly guided by this model" (Rogers, Petrosino, Huebner, and Hacsi, 2000, p.5).

Differentiation among the various purposes for conducting evaluations is also instrumental to evaluation theory. Scriven (1991) identified two main evaluation purposes: s*ummative,* to judge whether a program has met its objectives, and therefore to verify its merit or worth; and, *formative,* to inform program adjustments toward improving program implementation and outcomes. Patton (2008) summarized four other main evaluation purposes: *monitoring*, to contribute internally to routine program management; *accountability*, for external decision making and resource management; *developmental*, to contribute information for making strategic systems changes within dynamic environments; and, *knowledge*, to generate information toward overall incremental accumulation of information for design, planning, theorizing, research and policy making. Chelimsky (1985) suggested there are three main purposes of program evaluation: *policy formulation, for* development of new programs; *policy execution, for* assessing existing programs; and, *accountability, for* determining program effectiveness.

Evaluation theory is also characterized by evaluation practices. Smith and

Brandon (2008) explained evaluation practice refers to "the immediate world of politics,

clients, resources, role ambiguity, and changing field conditions; the practical concerns of

getting the work done well and of making a difference (p. ix). Mathison (2005) explained

a theory of evaluation practices includes the evaluator's role, the nature of the evaluand

and the program, the nature of evidence, identification of stakeholders, how stakeholders

will participate in the evaluation including conceptualization of power, the nature of

normative discourse (cause and effect), and ways of synthesizing. "There are ways of

doing things that are a part of evaluation, and although methods for evaluation are drawn

from the social sciences, there are questions that must be addressed in relation to the use

of those methods by evaluators for the purposes of assigning value. What is evidence, and

how do we make sense of it? What is the relationship between generalizations and

evaluation? How do we conceive of evaluands? What are the interpersonal, political, and

social components of evaluation" (p. 142).

Another significant feature of evaluation theory is identification of the ways

evaluation can be used. The theory of evaluation use is one of the most researched areas

of evaluation, and has been defined as "the effect the evaluation has on the evaluand and

those connected to the evaluand" (Christie, 2007, p.8). Use of evaluation findings has

traditionally been classified into three categories: instrumental use, conceptual use and

symbolic use (Johnson et. al. 2009; Patton, 2008). *Instrumental use* refers to when

evaluation knowledge is directly used to inform a decision or contribute to problem

solving. *Conceptual use* occurs when no direct action is taken, but an evaluation

influences people's understanding. *Symbolic use* is when token or rhetorical support is

given for an evaluation to maintain appearances, with no intent to take either the process

or findings seriously.

Patton (2008) provided a nuanced explanation and categorization of different

types of use including: direct intended use, longer term more incremental influences,

primarily political uses, misuses, non-uses and unintended outcomes. *Instrumental use*,

which is direct and intended, can be differentiated into conceptual use or process use.

Conceptual use influences how key people think about a program or policy, but no action

flows from the findings. *Process use* refers to when changes result from engagement in

the evaluation process. Longer term, more incremental influences of evaluation can be

intended or unintended, and can flow from the evaluation process or results.

Enlightenment refers to influences that can occur when new ideas from evaluation

contribute to new understandings and in the long term to policy making.

Uses that are primarily political are referred to as *symbolic.* Symbolic use is

defined broadly as "the use of evaluation to maintain appearances, to fulfill a

requirement, to show that a programme or organization is trustworthy because it values

accountability, or to legitimate a decision that has already been made" (McNulty, 2012,

p. 496). Multiple symbolic uses have been differentiated: legitimate use, persuasive use,

imposed use, or mechanical use (Patton, 2008). *Legitimate use* refers to using an

evaluation to support a decision that was made prior to the evaluation. *Persuasive use*

refers to using evaluation findings, often selectively, to support one's position in funding

decisions or political debates. *Imposed use* occurs when those at a higher level of power

mandate a particular form of evaluation use by those at a lower level, for example a

governmental condition for funding a program. *Mechanical use* refers to going through

the motions to meet an evaluation requirement, where the motivation is compliance and implementation is mechanical.

Misuses of evaluation refer to "calculated and intentional suppression, misrepresentation, or unbalanced use of evaluation findings to influence opinions and decisions" (Patton, 2008, p.113). *Inadvertent misuse,* also called mistaken misuse, occurs when those using findings lack the competence or spend too little time to understand findings, or are swayed by the evaluator's status, expertise or personality rather than the findings. *Overuse* occurs when too much emphasis is placed on weak evaluation results, or there is a lack of attention to local conditions such as when supposed best practices are universally mandated.

Nonuses have been differentiated as: due to misevaluation, political nonuse or aggressive nonuse (Patton, 2008). *Nonuse due to misevaluation* can be justified when evaluation results in weak evidence, a late report, poor evaluator performance or other failures such as not adhering to professional standards. *Political nonuse* occurs when findings are ignored because they conflict with a potential user's values, prejudices or preferences. *Aggressive nonuse* is calculated and refers to situations where use is undermined because results conflict with or raise questions about a preferred position. *Unintended effects* of evaluation are any use of findings or evaluation processes that were not planned, predictable or were unforeseen.

Another area of significance to evaluation theory involves differentiating the purposes and processes of evaluation, from the purposes and processes of monitoring, accountability, performance management and auditing. Monitoring systems are intended to assist internal managers with information on where management strategies would be

beneficial. Accountability involves holding someone accountable to someone else to justify or explain what has been done, and traditionally attends to external stakeholders that a program is responsible to, or to funders (Patton, 2008). It has been argued that accountability is primarily political, and does not provide sufficient information for decision making. Performance management is the "production of information about an organization's actual outputs and results (outcomes) as measured against its mission, goals, objectives, and targets" (Julnes, 2013, p. 82). Auditing compares the degree of correspondence between what a program reports and what is considered proper (Chelimsky, 1985).

Performance management and evaluation have been considered complementary tools to measure and manage performance (Lahey and Nielsen, 2013). "Evaluation is necessary to validate performance-monitoring data and, of course, to assess impact. But it can be costly and time consuming, and often the results may not be available in time to inform the next stage of portfolio development or other decision making. Performance measurement, in contrast can provide real-time data useful in day-to-day decision making" (Boris and Winkler, 2013, p.76). Where programs lack the capacity to undertake formal evaluations, they can collect information through internal performance monitoring, and as long as the program appears to be achieving reasonably positive results, performance management can suffice, at least for the short term (Boris and Winkler, 2013).

Performance management has been described as "the set of self-correcting processes grounded in real-time data measuring, monitoring, and analysis, that an organization uses to learn from its work and to make tactical (front line, quotidian) and

strategic adjustments to achieve its goal and objectives" (Hunter and Nielsen, 2013, p. 10). However, it has been argued that performance indicators are useless for problem solving, decision making or resource allocation because they do not explain why results are as they are (Patton, 2008). The principle shortcoming of performance management is that the validity of these data can be questioned, and they do not demonstrate, in the way evaluations do, that changes observed were caused by the program or intervention" (Boris and Winkler, 2013, 76). Performance measurement emphasizes storing aggregated information, while program evaluations disaggregate information explaining on a more detailed level why performance was high or low, contributing to an understanding of how to make improvements (Hatry, 2013).

Chelimsky (1985) argued that both auditing and program evaluation are useful for program formation and accountability, and are complementary. Program evaluations have long borrowed accounting methods for cost-benefit or cost-effectiveness analyses. While auditing asks normative questions, comparing what a program reports to established criteria, program evaluation does ask normative questions, but more frequently asks descriptive questions, and involves systematic research of a program's design, implementation and effectiveness. Auditing supports deductive reasoning, while program evaluation reasons deductively, but also reasons inductively (probabilistically). Both approaches are retrospective, systematic, focus on relevance to users, and are concerned with objectivity. While auditing tends to "record and store data in a linear, chronological way", program evaluation "tends to group data with an eye toward the ensuing analysis and the demonstration of patterns and relationships in the data" (p. 497). Auditors are

independent from clients but evaluators tend to work closely with clients collaboratively (Wiser, 1996).

Evaluation theory involves a body of principles that explain and provide direction to the practices of evaluation (Mathison, 2005). Many areas of disagreement existed throughout the field's history (Shadish, Cook and Leviton, 1991) and still exist today (Smith and Brandon, 2008). Questions include: why should evaluation be done; is the purpose of evaluation validation, accountability, monitoring, improvement or development; is beneficial social change best accomplished by changing present programs or creating ideas for future programs; what is the role of the evaluator; how are stakeholders best involved; should evaluation be managed primarily by evaluators or in collaboration with stakeholders; should evaluators should focus on users, and if so which ones; whose criteria of merit should be considered for judging programs; should programs be evaluated compared to each other or to absolute standards; whose values should be represented in evaluation; what questions should be asked; what is acceptable evidence for making evaluative decisions; which methods would best be used to answer which questions; what can the evaluator do to facilitate use; and what are possible risks of oversimplifying social knowledge. These questions remain unresolved today and are addressed by the body of principles that comprise evaluation theory.

2.4 Defining Disability Management

Having a clear definition of a domain facilitates understanding evaluation practice within that domain. Disability management involves multi-disciplinary health, safety and return to work processes, which are proactively applied within organizations to minimize the economic and social costs resulting from time off work due to illness or injury.

During the 1980s and early 1990s costs of disability in the workplace reached crisis

levels (Dunn, 2001). Early disability management programs, "emerged in the 1980s as a

response of self-insured employers in the USA to rising costs of disability and injury. The

management of employees with disabilities is now an issue in countries around the

world" (Westmorland and Buys, 2002, p. 746). Since the mid 1990s disability

management, "has come into its own as a profession and as a viable workplace strategy to

reduce the human and economic cost of disability" (Galvin, King, Knuelle and Rushby,

2005, p. 1).

Disability management has been defined as combining, "the clinical and case

management practices of vocational rehabilitation counseling, the multi-disciplinary team

approach of rehabilitation, and principles of organization development and program

administration into a comprehensive framework that is managed and coordinated within

the firm" (Tate, Habeck and Galvin, 1986, p.5). Critical to disability management

programs are return to work processes, which when successfully applied, include:

teamwork, management support, written policies and procedures, education and

communication, and comprehensive job evaluations (Strasser, 2004). Facilitation of early

and suitable transitional and long term employment is achieved through the work of

multi-disciplinary teams wholly committed to the same goal. Team participants typically

include management, workers, unions, health and safety personnel, occupational

therapists and employee health nurses.

Case management has increasingly been identified as a core function of disability

management (Rosenthal, Hursh, Lui, Zimmermann and Pruett, 2005). During the early

years of disability management case management was grounded in a biomedical focus,

assisting workers to return to work after medical treatment in order to minimize

(financial) costs associated with extended time off. More recently disability management

has increasingly embraced using multidisciplinary data to assist case managers who are

responsible to make decisions (Rosenthal, et al., 2005).

In addition to managing disabilities, disability management has evolved to include

health promotion and prevention of illness and injury. Job task analyses, ergonomics,

health incentives, and employee assistance programs have emerged within disability

management as means to improve overall well being (Bruyere and Shrey, 1991; Dyck,

2002). As these other disability management practices have emerged, programs no longer

limit their focus primarily to medical treatment, return to work and economic cost benefit

assessments (Young, Roessler, et al., 2005) and now take into consideration social,

psychological, motivational and educational orientations, systems, and organizational and

management structures, including preventative interventions.

Disability management programs have evolved over the past two decades into

rich team based and educational processes, which to succeed depend upon effective

communication systems and the participation of multiple stakeholder groups (Currier,

Chan, Berven, Habeck and Taylor, 2001). Programs are no longer preoccupied with

medical based teams and systems that facilitate timely return to work after disabilities

have occurred, and rely more on organizational awareness and commitment to contextual

integration. Typical stakeholder motivations include: (a) worker (health, financial

stability, happiness); (b) employer (financial viability, productivity, safety/security); (c)

health care providers (financial viability, client health); (d) payer (financial viability,

profitability, public image); and labour (rights of workers, job accommodations) (Franche, Baril, et al., 2005; Young, Roessler, et al., 2005; Young, Wasiak, et al., 2005).

Benefits from employment beyond financial remuneration include work environment bonding, through which the formation of an individual's self-concept is often deeply rooted in one's occupational identity (Shrey, 1991). Work activity regulates life activities (Galvin, et al., 2005). Unemployed individuals stand to lose their social network, self worth and positive identity, which are often tied to their ability to function as valued participants within the labour market, "Work provides more than a task, it provides meaning" (Curtis and Scott, 2004, p. 298). The unemployed worker's existence has become provisional and in a certain sense he cannot live for the future or aim at a goal (Frankl, 1963). Unemployed persons can come to feel unemployable.

Human rights legislation in Canada requires employers to accommodate persons with disabilities in the workplace, provided that in doing so employers do not sustain undue hardship. Undue hardship refers to either financial costs that would make an organization insolvent, or, outcomes that could lead to health or safety risks (Eakin, et al., 2002).

The World Health Organization (WHO) has developed an International Classification of Functioning, Disability and Health to define and measure disability (World Health Organization, 2001), the overall aim of which is, "…to provide a unified and standard language and framework for the description of health and health-related states" (p. 3). In developing the WHO Disability Assessment Schedule, "A series of systematic field studies was used to determine the schedule's cross-cultural applicability, reliability and validity, as well as its utility in health services research" (Ustun,

Kostanjsek, Chatterji and Rehm, 2010, p. v). The classification, "…has moved away from being a 'consequence of disease' classification (1980 version) to become a 'components of health' classification" (World Health Organization, 2001). Smart (2005) pointed out that this new perspective is an individual-driven rather than diagnosis-driven system of interdisciplinary collaboration, where physicians will no longer be the sole authority on disability, and where accommodation and rehabilitation are advocated.

In summary, return to work outcomes have the potential to impact many different individuals and organizations: workers, employers, payees, health care providers and society, and factors related to those individuals and organizations complicate the disability management process (Franche, Baril, et al., 2005; Loisel et al., 2005; Young, Wasiak, et al., 2005). Distinct models have been conceptualized in the field of disability management reflecting the various stakeholder group perspectives. The following section describes disability management paradigms that have emerged.

2.5 Models of Disability and Return to Work

Based on a systematic analysis of the theoretical and empirical literature on disability related to musculoskeletal pain, Schultz, Crook, Fraser and Joy (2000) identified five main conceptual models of diagnosis and rehabilitation in occupational disability. These were a biomedical model, psychiatric model, insurance model, labour relations model, and biopsychosocial model.

The biomedical model has been and continues to be the predominant framework for many health care professionals. In this model impairment is related to anatomical tissue damage. The mind and body are separate entities, and psychological, social and behavioral dimensions are relatively unimportant, and often are identified as functional

overlay. The physician is considered responsible for control and relief of the problem, and the one upon whom the patient can rely. This model offers a scientific approach valuable in cases of ruling out serious medical conditions, but can be restrictive in diagnosing pain when there is limited evidence of patho-anatomical defects.

The psychiatric model holds three fundamental beliefs: pain is either organic or psychological in origin; pain that cannot be attributed to physical causes must be psychological; and persons with undiagnosed intractable pain are a psychologically homogenous group. This model supports that persons either respond normally or abnormally to pain, where abnormal responses grossly out of proportion to the organic pathology can be evidence for diagnosing a mental pain disorder. Diagnosis of psychological pain disorder can become a chronic problem. This model is valuable for persons diagnosed with psychiatric disorders.

The insurance model is also referred to as a forensic or compensation model. The major tenet of this model is that persons who are claiming financial benefits through compensation or litigation may be dishonest about their symptoms, for purposes of financial gain or to be relieved of their workload. This model shares with the biophysical model the need for objective evidence of biopathology.

The labour relations model is a systems based model where work injury is primarily understood and managed within the sociopolitical context of the work place, rather than in terms of medical management. The premise is that employment security is critical for workers with disabilities, and the employer is responsible to provide work place accommodations, and physical, psychological and social preventative education

programs. This model requires supportive policies and procedures and effective communication system wide.

The biopsychosocial model views disability as integrated and multifaceted, "The model recognizes that the relationship between pain, physical and psychological impairment, functional and social disability is far from simple: pain and response to injury are complex and interactive phenomena" (Schultz, et al., 2000, p.281). One tenet of this model involves a conceptual distinction between impairment and disability. Impairment is loss of function, and disability occurs when contextualization of that impairment results in a decreased capacity to meet related demands or to perform intended functions. Another tenet is that organic pathology alone does not predict impairment or disability, and psychological and social cultural factors play major roles in responding to disability.

The biopsychosocial approach, "… has been modified in many different forms and is generally the most commonly considered and consensual framework for understanding the multidimensional aspects of many health problems" (Schultz, et al., 2007, p. 329). The biopsychosocial approach is best classified as including both a systems and an individual focus (Schultz, et al., 2007). More than other models (except perhaps the ecological/case management model) the biopsychosocial model takes a broader psychosocial perspective and best explains the disability continuum. The disability continuum involves the individual (physical and psychological impairment related to structure and function) and the contextual system related factors (such as workplace and treatment programs). Basic tenets of the model also include underlying

values, cognitions, stage of readiness for return to work, self-efficacy, targeting of psychosocial factors in treatment, and interdisciplinary psychosocial prevention factors.

The labour relations and biopsychosocial models have the capacity for enhanced interdisciplinary and functionally oriented assessments and workplace multi-specialty interventions that have proven to reduce the risk of chronic pain (Schultz, et al., 2000).

Loisel and Durand's (2002) conceptual model of disability management, the Sherbrooke model, is a most comprehensive model where actions and attitudes of key stakeholders, and health care and compensation systems are critical (Schultz, et al. 2007). Central to the Sherbrooke model is the importance of situating work rehabilitation in the workplace (Loisel and Durand, 2002). The strategy includes an early work site based rehabilitation process graded to match improvements in the worker's capabilities, with progressive augmentation of work demands, and simultaneous ergonomic intervention to permanently reduce excessive work demands. The goal is to return workers to regular work, rather than striving to cure a disease. The usual medical and worker participants at a clinical treatment site are replaced in the Sherbrooke model with groups of participants (worker, rehabilitation multidisciplinary teams, employer, attending physician, union) at the actual work site but with reduced duties.

In summary, disability management was originally medically focused and aligned with labour relations, and emphasized the employer's responsibility to provide work accommodation. Due to multiple stakeholder motivations, different paradigms of disability and return to work emerged and have continued to evolve. The predominant medical focus has declined, and more disability management has incorporated aspects of biopsychosocial approaches, which consider multiple factors.

2.6 Evaluation of Disability Management

  2.6.1 Literature reviewed. This study is an exploration of evaluation in the field

of disability management to develop an explanation of the nature of evaluation as it is

currently practiced. Over the past two decades little has been published regarding

evaluation of disability management, while much has been published about the expansion

of this industry worldwide during the same period. I based my initial conclusion that little

had been published about disability management evaluation on my literature review that

included sources listed in Table 1.

Table 1   Sources of Literature Reviewed at the Onset of this Study

Database	Pubmed (also known historically as Medline and Index Medicus) – produced by the US National Library of Medicine (covers all aspects of medicine, including disability evaluation; approximately 40% foreign coverage outside North America)
Database	PsychINFO – produced by the American Psychological Association (psychological aspects of disability)
Database	NIOSHTIC – produced by the US National Institute for Occupational Safety and Health (disability from an OSH perspective; primarily US coverage, but some international)
Database	HSELINE – produced by the UK Health and Safety Executive (disability from an OSH perspective; includes European literature)
Database	CISDOC – produced by the International Occupational Safety and Health Information Centre (CIS), International Labour Organization (disability from an OSH perspective, world-wide coverage)
Database	ERIC – the Education Resources Information Center, an online library of education research and information sponsored by the Institute of Education Sciences (IES) of the US Department of Education.
Evaluation Journals	I reviewed three evaluation journals for up to 18 years prior
Books	Edited and authored books on disability management
Conferences	Disability management conference manuals & compendiums
WorkSafeBC Library	Internal reports and non indexed holdings
Note: I have subsequently reviewed 6 other evaluation journals over 10 years and there were no articles specific to evaluation of disability management related to employment.	

I was assisted to access some sources by the Head Librarian at the WorkSafeBC Library. Keywords that were used related to disability management and evaluation. There was no time limit placed. The language chosen was English. Review of these sources turned up publications on disability management dated between 1986 and 2007, from Canada, US and outside North America. A total of 35 publications were obtained. I reviewed all publications in depth. A total of 19 included views on or recommendations for research or evaluation of disability management programs.

In a subsequent review of a 2005 volume of the Journal of Occupational Rehabilitation dedicated to disability management, and the edited Handbook of Complex Occupational Disability Claims (Schultz and Gatchel, 2005), a further 25 publications related to disability management were identified, with 6 of those referring to a need for program evaluation or research studies on disability management programs.

2.6.2 Disability management evaluation. While disability management programs have evolved and the industry has expanded over the past two decades, little has been published regarding evaluation of disability management. Evaluation methods reported within the disability management literature have commonly focused on summative measures which relate to economic outcomes, such as return to work rates, incidence and duration of absence, lost productivity, and benefit cost reduction (Akabas, Gates and Galvin, 1992; Currier, et al., 2001; Dunn, 2001; McMahon, et al., 2000; Pransky, Shaw, Franche and Clarke, 2004). The complex issues associated with work related disabilities have been analyzed in somewhat of a simple manner (Shrey and Olsheski, 1992).

Throughout the 1990s there was expanding recognition of issues related to employment for persons with disabilities, in part due to the American Disabilities Act being signed into law during 1990 (Akabas, et al., 1992). Disability management focus shifted from return to work outcomes to incorporating services such as transitional work, ergonomics, worker perspectives, health and wellness promotion, and building bridges between employees, employers and the community stakeholders.

Additionally, the early 1990s medical model of disability and return to work was giving way to contextualization and consideration of psychosocial issues. With this redirection there was a greater need for rehabilitation professionals (other than medical doctors) to work in disability management, to contribute their expertise (Gottlieb, Vandergoot and Lutsky, 1991). Rehabilitation professionals were increasingly expected to contribute to policy and program decisions, and to monitor effectiveness of program activities toward making improvements in prevention and treatment.

The focus of disability management, which had typically only monitored workers until they returned to their original work (Roessler, Schriner and Fletcher, 1991) shifted toward collaborative partnerships among different stakeholder groups (employers, workers and rehabilitation professionals), who were working toward providing services that not only restored work abilities, but also led to durable return to work and job satisfaction. Examination of disability management from the perspective of human services was recommended (Tate, 1987, p.65). There was a growing awareness of outcomes related to disability that had seldom been addressed. These included non-monetary costs to employers and direct costs to workers such as: individual loss of self-esteem derived from work and interaction with one's peers, emotional and psychological

distress affecting the disabled worker's family and friends, and a sense of worthlessness and of being deprived of one's occupational role in society (Tate, et al., 1986). However, evaluations of disability management programs did not explore individual experiences regarding return to work, or personal perceptions of workplace issues (Eakin, et al., 2002).

In the 1990s, evaluation of disability management was on occasion considered important, but few actual program evaluation studies were reported, a trend that continued into the 2000s (Currier, et al., 2001; Dunn, 2001; Pransky, et al., 2004). Disability management evaluation has mainly been conducted by economists, and evaluations of private non public sector organizations were practically nonexistent (Dunn, 2001). Program evaluation methods for disability management were poorly understood (Dyck, 2002; Dyck, 2009). Dyck (2009) suggested that evaluation should identify gaps between the current state of a program and the desired outcomes such as achieving the goals of cost effectiveness and return on investment, and focus on development of program improvement strategies relating to service quality and delivery.

In summary, over the past two decades little has been reported on the nature and extent of evaluation in disability management. However, the literature illustrates a sense of importance of evaluation. Evaluations of disability management programs were needed, ones that "…involve all relevant stakeholders, consider legal, professional, administrative and cultural environments and aim at developing new global return to work strategies that are effective, efficient and have potential for successful implementation" (Loesel, et al., 2005, p. 518). Quality disability management was seen as

relying on evaluation that links interventions to outcomes, that examines efficiency and accountability, and that emphasizes improvement (Nickerson, 2000).

2.7 Research and Evaluation

While both evaluation and research have consistently been conceptualized as important to disability management, the literature has not always clearly differentiated between the two. Publications have recognized either or both as needed for their potential to provide the opportunity to monitor, understand and improve disability management. There has been a resultant tendency in the literature for the purposes of doing evaluation versus doing research to become blurred in their uses to monitor disability management.

Evaluation's universal focus on stakeholder perspectives is an essential element in any evaluation, a feature not shared by social science research (Mathison, 2008). Evaluation's focus on stakeholder perspectives is particularly relevant for the study of disability management.

Recognizing that stakeholders have competing goals and varying definitions of disability and what constitutes return to work, researchers need to consider what matters to stakeholders (Young, Wasiak, et al., 2005). Young, Roessler, et al. (2005) recommended improvements to return to work research that includes: (a) development of a set of consistent conditions of key terms applicable across all contexts and stakeholders; (b) improved understanding of outcomes that matter to stakeholders; (c) addressing complexity of return to work outcomes by exploring the relevant variables related to various disability management paradigms; and (d) understanding return to work as a dynamic process.

A limitation of prior return to work research had been due in part to a narrow focus on outcomes (time off work or performance deficits), rather than focusing on issues that could predict successful return to work. Recognizing how prior studies on return to work had suffered conceptual and methodological limitations, expansion from biomedical to broader biopsychosocial understandings had been recommended (Pransky, et al., 2005). The paradigm shift from biomedical to biopsychosocial perspectives transfers, "…responsibility for outcomes from the health care provider – patient perspective, to a multi player decision making system influenced by complex professional, legal, administrative and cultural (societal) interactions" (Loisel, et al., 2005, p. 511). Understanding is required of the perspectives of multiple decision makers, the patient, physician, employer, occupational health staff and third party payers, each with their own values, objectives, interests and training.

Noting that communication between healthcare providers and workplaces rarely occurred and may be difficult to practice, Loisel, et al. (2005) suggested, "Future studies should involve all relevant stakeholders, consider legal, professional, administrative and cultural environments and aim at developing new global [return to work] RTW strategies that are effective, efficient and have a potential for successful implementation" (p. 518).

Conceptualization of research that could address preventative developments in disability management would require special skills to face difficulties working in, "…complex fields with many stakeholders with various interests, important intervention costs, ethical issues and system variations" (Loisel et al., 2005, p. 518). Given that evaluation methods are consistently grounded in consideration of stakeholder

motivations, it may be inferred the role of program evaluation is more ideally suited to achieve this type of investigation, than the role of research.

Research studies have recommended increased evaluation and accountability. In a survey of 1500 disability management specialists, three knowledge domains were recommended: (1) program development, management and evaluation; (2) disability case management; and, (3) disability prevention and workplace intervention. Each of these areas was thought to have the potential for harm if practiced incompetently. To accomplish this disability management managers must, "…truly understand research and encourage their employers, unions and other interested stakeholders to become involved in collecting data, applying research standards, and using data via superior data management strategies to make valid decisions that positively affect productivity and lead to more efficacious cost-containment approaches." (Rosenthal, Hursh, Lui, Ison and Sasson, 2007, p. 83).

2.8 Disability Management Evaluation Contextualized within the Evaluation Field

Reported evaluation of disability management to date has mainly focused on summative measures related to economic outcomes such as return to work rates, incidence and duration of absence, lost productivity and benefit cost reduction (Akabus, Gates and Galvin, 1992; Currier et. al., 2001). These evaluation criteria reflect values of the employers who are the funders of labour (Westmorland and Buys, 2002). That approach may have made sense at the time the field of disability management was starting to develop over two decades ago, when the primary stakeholders were the worker and the employer. At that time disability management followed a medical paradigm, where management of a worker's impairment was largely under the control of the

physician, including deciding when and under what conditions a worker could return to work, and the employer was responsible for accommodations with suitable job tasks. Workers had little control over how they managed their impairments or their presence at the work place (Smart, 2001).

During this same period through the late 1980s and 1990s that disability management was starting to develop from being based primarily on a medical paradigm with evaluation conducted primarily from the perspective of employers or funders (Westmorland and Buys, 2002), evaluation theory had progressed to include pluralistic approaches. Evaluation had advanced away from being based primarily on the 1960s scientific approach measuring cause and effect, and evaluation that was prescribed predominantly from the perspective of evaluator values. During the 1970s and 1980s evaluation theory had evolved to include pluralistic approaches that incorporated the values of multiple stakeholders, to make evaluation more relevant to its primary stakeholders and with a goal of increasing evaluation use (Shadish, Cook and Leviton, 1991).

During the 1990s disability management started to shift away from a predominantly medical paradigm to incorporate the interests of multiple stakeholders, (i.e. workers, unions, safety personnel, medical practitioners other than physicians etc.) (Galvin, et. al., 2005). Typical stakeholder values expanded to include: workers' health, financial stability, rights at the workplace, job satisfaction, organizational learning about safety, prevention of injuries and illnesses, financial viability, ergonomics, etc (Dyck, 2002). Reported evaluation of disability management did not advance to include multiple stakeholders' values during that same period of disability management expansion.

It has been most recently, during the 2000s, that increased consideration has been given to shifting disability management away from a medical paradigm where physicians managed impairments, to biopsychosocial approaches, where workers manage their impairments at the workplace ideally (Loisel, et. al., 2005). Advances in this direction have been supported in the field, including by the World Health Organizations which in 2001 implemented a new International Classification of Functioning, Disability and Health focusing on the person and placing emphasis on impairment interactions with their environment.

Ideally, disability management evaluation will progress to meet the expansion of stakeholders' interests in the field, and to meet advances in the evaluation field in general. No doubt disability management evaluation will face the same ongoing struggles as evaluation has regarding: questions about how to best include pluralistic stakeholder values in evaluation, what purposes of evaluation to prioritize, which stakeholder values to prioritize, what evaluation criteria to use, funding concerns, and how to maximize use of findings.

2.9 Summary

Many different operational definitions and paradigms have been reported in the body of literature on disability management as it has emerged and continually developed over the past two decades. Reporting on evaluation of disability management programs has been scarce. The present study will explore disability management evaluation within a sample of four sites, each representing a different paradigm: biomedical, labour, biopsychosocial and insurance.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Obtaining and Maintaining Ethical Consent for this Study

The procedures and timeline that I followed to obtain and maintain ethical consent are outlined in Table 2.

3.2 Research Questions

#1. What is the extent and nature of evaluation practice within the disability management programs?

#1a. How does disability management evaluation practice vary depending on whether the organization is a learning organization?

#1b. How does disability management evaluation reflect diversity?

#1c. Is disability management evaluation grounded in a particular paradigm of disability and return to work?

3.3 Selecting Which Paradigms to Include in this Study

This study focuses on evaluation at four sites each representing a different dominant disability management paradigm: biomedical, labour, biopsychosocial and insurance. In part the study explores the extent to which evaluation practices at the sites may reflect the interests associated with different paradigms. Much has been published on the history of disability management, including how originally the field was based on a biomedical perspective. As other stakeholders' interests became more prominent in the field (workers, employers, insurance funders, multi-disciplinary treatment teams, unions) disability management expanded and evolved to include those multiple perspectives. Paradigms have emerged in the field representing the diversity of stakeholders, and may reflect different approaches to, or emphases in, evaluation practices.

Table 2   Timeline Followed to Obtain and Maintain Ethical Consent

Date	Procedures for Ethical Consent
November 2, 2009	Obtained Certificate of Completion for having completed the: *Interagency Advisory Panel on Research Ethics' Introductory Tutorial for the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS)*
November 24, 2009	Submitted ethics application to the University of British Columbia Office of Research Services Behavioral Research Ethics Board for authorization to conduct research involving human participants.
November 30, 2009	Received provisos from UBC Behavioral Research Ethics Board required for modification of the ethics application.
December 6, 2009	Resubmitted ethics application to UBC Behavioral Research Ethics Board with required provisos incorporated.
December 21, 2009	Completed final revision of 4 page informed consent form to be provided to potential participants of study (see Appendix A).
December 23, 2009	Received University of British Columbia Behavioral research Ethics Board Certificate of approval # H09-02993 to conduct research for one year to November 29, 2010.
January 2010	Met at three sites of proposed research organization and obtained three signatures of approval required before submitting application to Vancouver Coastal Health Research Center Authority Clinical Trials Administration Office requesting approval to conduct low risk research within that organization.
February 10, 2010	Submitted application to Vancouver Coastal Health Authority Clinical Trials Administration Office for approval to conduct low risk research.
March 15, 2010	Received Vancouver Coastal Health Authority Clinical Trials Administrative approval to conduct research Study # V10-0051.
November 2010	Submitted application to UBC Behavioral Research Ethics Board for a one year extension of the ethics approval November 2010 to November 2011.
December 2010	One year ethics extension approved by UBC Behavioral Research Ethics Board to continue study November 30, 2010 to November 29, 2011.
December 2011	Submitted application to UBC Behavioral Research Ethics Board for a one year extension of the ethics approval December 2011 to November 2012.
December 2011	One year ethics extension approved by UBC Behavioral Research Ethics Board to continue study December 8, 2011-December 7, 2012.

Four sites were selected for this study to represent the disability management paradigms: biomedical, labour, biopsychosocial and insurance. These paradigms are representative of the differing approaches common to disability and return to work. While these models share common characteristics, they are unique in important ways, (Currier et al., 2001; Franche, Baril, et al., 2005; Young, Roessler et al., 2005; and Young, Wasiak et al., 2005).

In selecting the sample of paradigms to include in this exploratory study, a main objective was to include paradigms that had the potential to contribute perspectives of disability management and it's evaluation that were as diverse as possible. Four of the five paradigms identified by Schultz et. al (2000), biomedical, insurance, labour and biopsychosocial, are each identifiable as a main paradigm within many disability management programs in operation today, and were included in this study. The fifth paradigm defined by Schultz et. al. (2000), psychiatric paradigm, was not included in this study. Based on my twenty plus years experience practicing in the disability management field I am not familiar with any disability management programs that are based primarily on a psychiatric paradigm. I will give examples of two types of disability management programs (neither of which would be primarily oriented to a psychiatric paradigm) that would serve clients that would fit the definition for psychiatric pain disorder as described by Schultz, et. al (2000) for the psychiatric paradigm.

For clarification, the biomedical paradigm assumes a medically diagnosed organic pathology with symptoms proportional to the pathology, and that are objectively identifiable. The psychiatric paradigm makes the same assumptions, but the psychiatric paradigm also assumes that people with symptoms that are not in keeping with the

organic pathology are a homogenous group with a diagnosable mental disorder (understood to be a pain disorder due to exhibiting symptoms that exceed what would be expected for that pathology).

There are at least two types of disability management programs (that I am aware of) that serve clients with diagnosed psychiatric pain disorders as defined by the psychiatric paradigm (Schultz, et. al., 2000), but neither of these types of programs serve only these clients, or even primarily these clients.

The first type is usually called a *Pain Program*, and is offered by various providers in the community. These programs involve multi-disciplinary clinical services with a high ratio of clinicians (from areas that include psychology, physiotherapy, occupational therapy, vocational rehabilitation, physiatrist, medicine, etc.) providing education and treatment to assist clients cope with excessive pain and maximize their functioning. These programs serve both: (1) clients whose pain symptoms are commensurate with their pathology and would be categorized under the biomedical paradigm; and, (2) clients with symptoms that exceed what is expected for their pathology and would be categorized under the psychiatric paradigm; as defined by Schultz, et. al. (2000).

The second type of program that serves clients whose pain symptoms exceed what is expected for their pathology, and would be categorized under the psychiatric paradigm, is typical of the Labour site program included in this study. These programs are often offered by employers for all of their staff who have sustained illnesses or injuries, including physical or mental, resulting in disability. These programs serve clients with physical disabilities, mental illnesses and who would be diagnosed as having a pain

disorder according to the definition by Schultz, et. al, (2000). As described in this study, clients with mental disorders, including illnesses or pain disorders, have often faced stigmas due to their disability not being visible to others within their environment especially at the work place, and efforts are usually made to integrate them into the workforce without differentiating them from clients whose disabilities are physical and visible to minimize these stigmas.

3.4 Evidence of Different Paradigms within the Sample of Sites

To identify sites that were representative of the four paradigms included in this study, the key elements of each were reviewed. Table 3 summarizes the key elements of the four paradigms as described in the literature.

Table 3   Key Elements of Paradigms of Disability and Return to Work

Biomedical	- Illness is due to a physical pathology - Symptoms are directly proportionate to physical pathology - Physician is responsible for diagnosis and treatment
Labour	- Work injury is managed best within workplace context - Needs of workers and employers can be complementary - Employer is responsible to accommodate return to work - Medical diagnosis is secondary to matching job demands to functional capacities
Biopsychosocial	- Interdisciplinary whole person approach - Conceptual distinction between impairment and disability - Organic impairment does not reliably predict disability
Insurance	- People who anticipate secondary gain are likely to magnify disability - Objective medical proof of impairment and disability must be proven
Based on: Schultz, I.Z., Joy, P.W., Crook, J., & Fraser, K. (2001). Models of diagnosis and rehabilitation in musculoskeletal pain-related occupational disability. *Journal of Occupational Rehabilitation, 10*, 271-293.	

I drew upon my own knowledge regarding sites from the past 20 years in my professional interactions with disability management programs, to identify sites that had characteristics mainly representative of one of the four paradigms. The particular programs selected for possible inclusion in this study were initially identified based on my understanding that the program services matched characteristics for one paradigm as documented in the literature. The four paradigms do overlap to some extent, and each program may have some characteristics of other paradigms, but less so than the predominant paradigm that it was selected to represent in this study.

I was familiar with one person at each of the four sites selected for this study prior to initially contacting them. Confirmation that the sites were predominantly representative of the paradigms was obtained during the interviews conducted with participants from each site.

The purpose of selecting sites from different paradigms was not to analyze the degree to which fidelity of paradigms was present within the sites, but was to maximize variation and the opportunity to access as diverse a sample of participant perspectives as possible, representing different priorities and values, with the potential to contribute variation and depth of understanding regarding disability management evaluation.

3.4.1 Biomedical paradigm site. This site is an independent for profit organization contracting to provide assessments and return to work (or return to school) services after a disabling injury or illness. Services are provided to external clients from any referral source. Clients may be of any age, but most are adults and formerly employed.

All services are provided to individual clients, there are no group programs. Services include: functional capacity evaluations, job demands analyses, ergonomic job site modifications, and gradual return to work plans. The agency contracts to work for external referring funders. They first establish a client's functional capacities, and critical job demands of their regular work, and second develop and monitor a return to work plan that integrates ergonomic adjustments. The main goal is to provide verifiable information on functional capacities, based on standardized measurement tools. One of the participants interviewed at this site referred to their services as "part of what they call health care services".

This site illustrates characteristics of the biomedical paradigm of disability management where impairment is predominantly related to anatomical tissue damage, and where other dimensions (psychological, social, behavioural) are relatively unimportant (Schultz et al., 2000). Demonstration of maximum effort to objectively verify functional capacities is measured with a combination of data sources including body mechanics, muscle wasting and comparisons of measurements such as bilateral grip strengths etc. Conceptualization of disability according to this paradigm, "…conforms to a positivist approach in which the disease is an obvious and observable biological problem that needs correcting with the scientifically proven techniques of biomedicine" (Schultz et al., 2007, p 332).

3.4.2 Labour paradigm site. This site is an in-house program available to 2-3,000 employees of a large organization providing services to stay at work, return to work, or adjust to permanent disabilities after they have sustained injury or illness.

This program was developed during the late 1990s when employers in British Columbia were increasingly required to comply with duty to accommodate legislation, which was one of the underpinnings of the current program. At the time, many employees in this organization were off work and in receipt of long term disability benefits through an outside carrier with no way to return to work. This program was originally developed to assist them to transition back. Currently, participation in this program is voluntary, and the only requirement is that the employee has a medical condition (physical or cognitive) as the program is not intended for resolution of labour relations issues.

The labour site is guided by a joint labour-management steering committee for all policies and procedures, and operated by four coordinators. Program coordinators are responsible for two main services: liaison between all other parties, most notably employees and their managers while employees are off work being treated for medical conditions; and assisting employees to either stay at work, return to work in suitable job duties, or adapt to withdrawal from the workforce due to being functionally unemployable.

None of the program's four coordinators or the employees' managers is advised of the employees' medical diagnoses, only functional limitations. Nor do they decide when a client is able to return to work. During medical treatment program coordinators provide a written summary of the critical job demands of the employee's regular job to the treating physician (e.g. multi-tasking, concentrating, working independently, managing a large volume of work, working with a complicated new computer software system). The treating physician then completes a *Functional Abilities Evaluation Checklist* outlining functional limitations and timeframes for returning to work. Program

coordinators pass this information on to the employee's manager, who is then responsible to identify suitable job duties. Program coordinators work collaboratively at the job site overseeing the employee's participation in accommodated duties, assessing and making adjustments as required.

This site is characteristic of a labour paradigm as described by Schultz et al. (2000) where disability management is primarily understood and managed within the sociopolitical context of the work place, with no program involvement in management of medical treatment. Schultz et al. (2007) explained a labour relations paradigm focuses on, "…workplace characteristics such as climate, culture, organization, job demand and accommodation, policies, procedures, and practices" and "effective disability/case management" (p.335).

3.4.3 Biopsychosocial paradigm site. This site is situated within a large government funded multidisciplinary health service organization that provides inpatient, outpatient, outreach and clinical support services. The site operates four unique programs: Acquired Brain Injury, Spinal Cord Injury, Arthritis, and Neuromusculoskeletal, and has a program for Adolescents and Young Adults. A broad range of in-house clinical services are offered. This organization serves, (1) inpatient clients that have medical referrals, and (2) outpatient clients referred by any source including allied health professionals, other organizations, or self referred. Clients are adults or youth, but not children or adults over approximately age 60, as both of those groups are served by other organizations.

A biopsychosocial paradigm program recognizes injuries are complex and interactive including physical, psychological and social aspects of disability (Schultz et

al., 2007). A feature of this site is that it is multidisciplinary (problem solving draws from multiple disciplines) and interdisciplinary (disciplines work together), where clients are almost always imbedded in multiple treatment programs. At the same time that clients are accessing the biopsychosocial site for vocational rehabilitation, they are also accessing other services available to them within this organization, including: psychological counselling, social work, sexual health, drug and alcohol counselling, pastoral services, occupational therapy, physiotherapy, nursing care, specialized surgical support, speech language pathology, dietary counselling, assistive technology, orthotics, physiatrist, recreation therapy, peer mentoring, music therapy, art therapy, pet therapy, adolescent young adult program services, and spinal cord or brain injury education.

The services of this disability management program are related to employment. Some clients have never previously worked and want to, these are often adolescents and young adults. Some are trying to maintain a job, and due to effects of an injury or illness may need work site accommodations. Some are returning to work after an injury or illness and require assistance communicating with the employer regarding required job accommodations or alternate positions within the same company. Some have to change jobs completely, and may be looking into new occupational directions that require training. Some want to switch to self employment. Some clients cannot return to work due to disability or illness, and this program helps them adjust to the realization they will need to access disability benefits rather than earning employment income, and will have to explore new ways to find quality of life.

Many clients of this program are in receipt of benefits such as long term disability, and require assistance from this program to liaise with external insurance carriers, or to understand how to access whatever assistance they may be eligible for.

3.4.4 Insurance paradigm site. This site is the investigation unit of a large public organization that handles injury claims within one of their divisions. The insurance site conducts investigations into disability claims, employers, health care service providers, vendors the organization deals with, and employees of the organization within which the insurance paradigm site exists.

The insurance site investigators provide information to the division of the organization that manages injury claims. Injury claims are adjudicated by case managers who obtain the information required to make decisions and adjudicate claims from multi-disciplinary internal and external medical professionals. Case managers seldom attend meetings with people outside their offices, rely mainly on telephone contact, and can lack effective interviewing skills. When information that case managers have accessed is confusing, conflicting or incomplete, in order to minimize the risk of making inaccurate assumptions they can request assistance from the insurance site investigators who are skilled interviewers. The insurance site investigators are often called field investigators due to the time they spend obtaining objective information in the field. All internal staff participating in these case management processes are bound by provincial government ethical practices for working with the public.

The insurance site receives internal referrals from the disability claims division of the organization to investigate claimants. External referrals can come from any source, such as employers, neighbours of claimants, or through an anonymous tip line. Referrals

regarding investigation of employers, service providers or vendors similarly can come internally or externally. Referrals to investigate internal staff of the organization come from upper management.

The goal of the program is to obtain accurate information to alleviate misunderstandings so that fair adjudication of entitlement to benefits can be conducted, and to protect the integrity of the accident fund. Some services such as surveillance or investigative accountants are contracted out. Each case referred to the program is overseen by one of the 18-25 investigators providing services for the insurance site program.

The insurance model of disability management shares with the biomedical model an emphasis on verifiable medical evidence of impairment. From an insurance paradigm perspective disability symptoms, "…may be a manifestation of attempts to receive compensation for disability" (Stowell and McGeary, 2005, p. 122). This model assumes some, not all, individuals are faking disability, and has a strong moralistic element where it is necessary to clearly differentiate between honest and dishonest clients, to differentiate those faking disability for benefits, from those actually disabled and entitled to benefits (Schultz et al., 2000).

3.5 Other Characteristics of the Disability Management Programs

Disability management programs differ in a number of respects other than the paradigms described above. Understanding nuances of each program is necessary for exploration of evaluation practice.

Table 4 summarizes estimated numbers of incoming referrals and workloads at each site of this study.

Table 4   Annual Incoming Referrals and Caseload Sizes

Site	Number of Caseloads at Site	Annual referrals to program	Approximate Caseload Sizes
Biomedical	2	BM1 50-75 BM2 100-125	BM1 20 BM2 3-4
Labour	3	181	60
Biopsychosocial	2-3	395	50-60
Insurance	18-25	1,600-1,700	20
Note: All figures are estimates that can vary year to year. BM1 managed long term cases lasting up to 6 months, BM2 managed short term cases lasting several hours to several days.			

Other program characteristics at the sites of this study include: types of services provided, definitions of disabilities, funding sources, referral sources, in house or external clients, other programs that may coexist within the organization, and client demographics, such as age. Table 5 summarizes characteristics of the four sites in this study.

Disability management occurs at different points along a continuum of treatment and recovery after onset of disability has been acquired: (1) stay at work assistance is

Table 5   Characteristics of the Sites Participating in this Study

	BM	Labour	BPsy	Insurance
Program Objectives				
Assessments	x			
Stay at work	x	x	x	
Return to work	x	x	x	
Adjust to being unemployable	x	x	x	
Investigations for information				x
Referral sources				
Client self-referral	x	x		
Funding agencies	x		x	x
Health providers	x		x	
Employers	x			x
Anonymous				x
Clients				
In house employees		x		x
In house clients			x	x
External funded clients	x		x	
External individual clients	x		x	
Types of Disabilities				
Physical illness or injury	x	x	x	x
Mental illness		x		

provided preventatively before a worker or student has to go off work or leave school to avoid his/her having to go off, through interventions such as modification of activities to meet the new physical or cognitive limitations associated with the disability; (2) to assist a worker or student's timely return to meaningful and suitable work or school after his/her condition has reached workable but not full recovery; and (3) to assist a person to adapt to meaningful and satisfying life activities where consequences of the disability will preclude a return to his/her usual work or school activities.

Programs may serve in-house employees/clients, for example, employees of an organization that has an internal staff disability management program, or clients of a treatment agency that includes disability management among its services. Alternatively, programs may contract out services to external clients for a profit, for example, clients who are sponsored by funding agencies such as insurance companies and referred for specific treatment or assessment to these community programs, clients who are referred by representatives such as lawyers or the courts, or self referred clients.

3.6 Participants at the Study Sites

3.6.1 Soliciting participation in the study. In selecting multiple sites for inclusion in a study, within case issues need to be examined, "Within-case sampling is almost always *nested*…" (Miles and Huberman, 1994, p. 29). Within case sampling is iterative, where "We observe, talk to people, and pick up artifacts and documents….At each step along the evidence trail, we are making sampling decisions to clarify the main patterns, see contrasts, identify exceptions or discrepant instances, and uncover negative instances-where the pattern does not hold. Our analytic conclusions depend deeply on the within-case sampling choices we made" (p.29).

For this study, in order to maximize variation of insights, participants were selected that would be most likely to contribute diverse perspectives. Key personnel were first identified within the sites, including: (1) those whose jobs included activities that may in some way involve program evaluation, and (2) those who were program administrators, and (3) those who were practitioners providing direct client disability management services. Potential participants at each site were initially contacted by email or in person, with informed consent forms describing the study provided (see Appendix A). Those originally contacted by email were subsequently contacted in a follow up telephone call.

The first three contacts were with individuals at sites representative of biomedical, labour and biopsychosocial paradigms of disability management. Each person contacted agreed to participate in this research study, and expressed an interest in contributing to an understanding of evaluation within disability management. Five different insurance paradigm programs were contacted before one agreed to participate. I was familiar with the fifth site I contacted and the response from the person initially contacted was positive.

Through a snowball sampling procedure initial contacts at the sites recommended other individuals as potential participants in the study. The initial person contacted at the biomedical site, BM1 of this study, recommended inclusion of a colleague at the biomedical site, who became BM2. The initial person contacted at the labour site, L3 of this study, recommended inclusion of two colleagues from the labour site, who became L4 and L5. The initial person contacted at the biopsychosocial site, BPsy6 of this study, would not recommend other current employees of the site, so I then contacted a colleague who had been a former employee at the biopsychosocial site who agreed to participate,

and became BPsy7. At the insurance site the initial participant I8 recommended inclusion of his colleague, I9. I had professional relationships with one person from each of the sites prior to this study: BM1, L3, BPsy7 and I8.

All participants except BPsy7 participated in this study during hours of their paid employment with their employer's permission. BPsy7 volunteered his personal time, and after completion of one long interview and a second shorter interview with him, without his prior knowledge I provided him a gratuity of $100.00.

3.6.2 Participants at the biomedical site. The two participants from this site are both administrators/business owners as well as practitioners providing client services at the program. Both are trained Occupational Therapists. These were the first and second interviews of the study, referred to as BM1 and BM2.

3.6.3 Participants at the labour site. The three participants from the labour site are program coordinators and one is also the program manager. They were all formerly employed in other roles within the same organization as a Vocational Rehabilitation Consultant, Office Staff/Union Activist and Human Resources Manager (now manager of the labour site program). These were the third, fourth and fifth interviews of this study, referred to as L3, L4 and L5.

3.6.4 Participants at the biopsychosocial site. The two participants from this site included the current senior counsellor/team leader/administrator of the program and a counsellor who formerly worked for the program but was laid off several years ago due to loss of funding. Both are Vocational Rehabilitation Consultants by profession. These were the sixth and seventh interviews of the study, referred to as BPsy6 and BPsy7.

3.6.5 Participants at the insurance site. The first participant from the insurance paradigm site is one of eighteen investigators of this program who conduct external investigations related to disability claims. The second participant from this site is the program manager who conducts internal investigations of staff of the organization. Both hold a Certified Fraud Examiner designation. These were interviews eight and nine, referred to as I8 and I9.

3.7 Methodology

Research is rooted within paradigms, epistemological and theoretical, that reflect the basic belief systems that guide researchers, including their choice of methodology. The researcher's understanding of epistemology informs the theoretical perspective taken, and the methodology chosen to gain knowledge (Guba & Lincoln, 1994). My own social constructivist epistemology informs the methodological decisions for this research study.

Social constructivism claims there is no one objective external truth that exists to be discovered, and that meaning is made through our conscious engagement with our world (Crotty, 2003). Social constructivism "means that human beings do not find or discover knowledge so much as construct or make it. We invent concepts, models, and schemes to make sense of experience, and we continually test and modify these constructions in the light of new experience. Furthermore, there is an inevitable historical and sociocultural dimension to this construction" (Schwandt, 2007, p. 38). 'Knowledge' is the result of how the 'knower' constructs reality from his or her experiences and perceptions (Hanley-Maxwell, Al Hano and Skivington, 2007).

This research is grounded in an interpretivist theoretical perspective, one that assumes that "reality is socially constructed, filled with multiple meanings and interpretations, and that emotions are involved. As a result, interpretivists see the goal of theorizing as providing an understanding of direct lived experience instead of abstract generalizations" (Hurworth, 2005, p. 209). In the interpretive human sciences, the meaning of social action is inherent in that action, and the task of the researcher is to interpret, unearth, and understand that meaning (Schwandt, 2007).

This research study adopts a social constructionist epistemology, an interpretivist theoretical perspective, and uses grounded theory methodology to analyze data. This study assumes there are multiple different constructed realities of what is valued within the field of disability and return to work, and the potential for various approaches to evaluation. As with all grounded theory, the intent of this study is to develop a theory of evaluation practice in disability management.

3.7.1 Grounded theory methodology. The purpose of this research is to understand the nature and extent of evaluation that is occurring in disability management, about which little has been previously published. Grounded theory allows the experiences and knowledge of information to be gained from participants assuming them to have expertise, providing an empirical basis for developing theoretical understandings of a phenomenon. This research was initially informed by an etic perspective based on research literature, for example what was recommended to be evaluated within disability management programs. As participants presented their emic first hand perspectives, insights are gained regarding the ways participants experience social phenomenon, the

meanings they put on them, and how they interpret what they experience (Richards and Morse, 2007).

Programs exist within complex contexts, and research methodology should capture as much of that complexity as possible (Corbin and Strauss, 2008). This means obtaining multiple perspectives, and recognizing, "…that experience must be located within and can't be divorced from the larger events in a social, political, cultural, racial, gender-related, informational, and technological framework and therefore these are essential aspects" (p.8) of successful exploratory research.

Grounded theory methodology is based on emerging analyses throughout data collection, and culminates in substantive theory regarding social phenomena (Bryant and Charmaz, 2007). The procedures of data collection and analysis employ techniques of induction, deduction, and verification (Schwandt, 2007). Analysis of data generates insights, and hypotheses are developed and tested through further data collection and analysis. Grounded theory refers to both a method of inquiry and a product, enabling researchers to focus their data collection through successive levels of data analysis and conceptual development, toward production of theory (Charmaz, 2005).

There are several interpretations of what grounded theory methodology is. Glaser and Strauss developed grounded theory during the 1960s as a systematic methodology of scientific inquiry, originally assuming positivistic and objectivist characteristics, but the methodology developed in different ways when the founders went their separate ways in the late 1980s (Bryant and Charmaz, 2007). Glaser's position remained positivistic (Charmaz, 2000) emphasizing the use of constant comparative methods to generate concepts, not to make descriptive generalizations (Glazer, 2007). Strauss, joining with

Corbin, maintained an assumption of an objective external reality, but moved grounded theory into post-positivist directions (Charmaz, 2000), emphasizing complexities of the world in data collection, and maintaining multiple perspectives on events to build variation into the grounded theory analytic scheme (Corbin and Strauss, 2008; Bryant and Charmaz, 2007). Subsequent to Strauss's death in 1996, Corbin's perspective was influenced not only by the methodological contributions of Strauss, but also contemporary feminist, constructionist and postmodern perspectives. She denounced a post-positivist underpinning, emphasizing that individuals give meaning to events in light of their own experiences (Corbin and Strauss, 2008). At the same time Charmaz (2000) was also developing a constructivist informed grounded theory that stressed analytic strategies rather than data collection methods, where multiple sources of data are selected based on conceptual sampling to fill gaps in the literature.

   3.7.2 Grounded theory methodology for this study. Grounded theory methodology includes the following: (a) simultaneous collection and analysis of data, (b) a two-step data coding process, (c) comparative methods, (d) memo writing aimed at the construction of conceptual analyses, (e) sampling to refine the researcher's emerging theoretical ideas, and (f) integration of data into a theoretical framework" (Charmaz, 2000, p. 511). The four sites were selected simultaneously to each conceptually represent a different paradigm of disability management: biomedical, labour, biopsychosocial and insurance. Participants at each site were consecutively interviewed starting at the biomedical site, then at the labour site, followed by the biopsychosocial site, and finally at the insurance site.

3.7.2.1 Simultaneous collection and analysis of data. From the start of data collection the researcher analyzes what things mean, noting "…regularities, patterns, explanations, possible configurations, causal flows and propositions" (Miles & Huberman, 1994, p.11). Meanings that emerge from the data are tested for validity, by comparing and contrasting findings among the cases, and with the literature. As meanings become increasingly explicit and grounded, they may become robust components of the theory.

For this study, data analyses were simultaneous with data collection, including during document reviews, interviewing, discussions of the ROLE with participants, during transcription that occurred as soon as possible after each interview, during open and conceptual coding of transcripts, creation of charts, figures and tables, reviews of figures by participants, and throughout comparisons of data from all sources and comparison of data to the literature.

The data collection and analysis began with the biomedical site and iteratively added the other disability management paradigm sites. Collection and analyses of multiple sources of data from the biomedical site were first conducted. Collection and analyses of multiple sources of data from the labour site were then conducted, and data from the labour and biomedical sites were analyzed together. Collection and analyses of multiple sources of data from the biopsychosocial site were conducted, and data from the biopsychosocial, labour and biomedical sites were analyzed together. Finally collection and analyses of multiple sources of data from the insurance site were conducted, and data from all four sites were analyzed together.

3.7.2.2 Two-step data coding. Grounded theory data analysis involves a two step coding processes. Open coding involves line by line coding of data, in the case of this study transcribed interviews, staying as open as possible to capture the multiple perspectives that may be assumed. This level of coding is close to and descriptive of the data. The second step is conceptual coding, which involves collapsing open codes into a smaller number of higher level codes, based on conceptual understanding gained through constant comparisons between different sources of data. The coding process is analytic and facilitates transcending the detail and striving for higher levels of abstraction in coding.

Grounded theory is not about accurate description, but is an analytic approach to develop conceptual abstraction (Holton, 2007). It must theoretically explain not merely describe what is happening in a social setting. Coding is therefore not a distinct stage as it is in some research methodologies, but a continuous, iterative process of collecting and analyzing data. "Substantive coding is the process of conceptualizing the empirical substance of the area under study: the data in which the theory is grounded….The process proceeds from the initial open coding of data to the emergence of a core category…." (Holton, 2007, p. 275).

For this study the process of coding started with open coding of each interview transcript facilitated by the use of NVivo software. This involved reading each transcript line by line, identifying descriptive codes, and labelling those sections with code names. Up to 80 open codes were initially identified in all. Some open codes overlapped and several sections of the transcripts fit into more than one open code. A chart was

developed to facilitate grouping open codes into themes to develop an initial descriptive display of results.

Interview transcripts and documents were then re-coded to identify conceptual themes that emerged from the initial open coding of the data. This included re-coding of data from each site and across all sites in an iterative manner, including identification of which cases illustrated which themes.

Conceptual analyses involved returning to data from sites previously coded to analyze concepts among sites in an iterative manner. Data were collected first at the biomedical site, then the labour site, followed by the biopsychosocial and finally at the insurance site, however, analyses involved a process of returning to the data from all sites to understand concepts. When new conceptual codes were identified, analysis included returning to previously coded data to investigate possible evidence that may have been overlooked or possible coding that may need to be collapsed in ways relating to new meanings that were developing.

3.7.2.3 Constant comparative methods. The constant comparative method is, "…a method of analysis that generates successively more abstract concepts and theories through inductive processes" (Bryant and Charmaz, 2007, p. 607). Open codes are compared to open codes, conceptual codes are compared to conceptual codes, open codes are compared to conceptual codes, themes are developed and are compared to other themes. These comparisons constitute stages of analysis through which conceptual understandings and theory develop. As the researcher codes all sources of data a core category begins to emerge. "This core variable can be any kind of theoretical code: a

process, a typology, a continuum, a range, dimension, conditions, consequences, and so forth" (p.279).

For this study interview transcripts and documents were coded conceptually, the literature was consulted for theoretical evidence to integrate with development of the conceptual findings. Outlines were developed describing conceptual findings within cases and across cases. Through analyses of raw data, open codes, conceptual codes, and returning to the literature, themes were developed and tested, and core categories emerged that analyzed together explained the processes of disability management evaluation practices in context.

3.7.2.4 Memo writing and visual displays. Memo writing is a strategy to focus the researcher's thinking throughout the data analysis: "It is not the form of memos that is important, but the actual doing of them"(Corbin and Strauss, 2008, p.118). Writing memos begins with the first analysis and continues throughout, beginning, "as rather rudimentary representations of thought and growing in complexity, density, clarity, and accuracy as the research progresses" (p.118). An example of two memos is displayed in Figure 1.

Analyses involve complex cumulative thinking and memos are used to keep track of this thinking. Memos are conceptual not just descriptive and "provide a storehouse of analytic ideas that can be sorted, ordered and reordered, and retrieved according to the evolving analytic scheme" (p.120). Memos facilitate judgments about when a category is saturated, and properties and dimensions are well developed.

Figure 1 Examples of Memos

_____

Memo 1

Degrees of learning related to levels of abstraction and sharing information in
evaluation. Synthesis of issues into categories of evaluation served as a conceptual
organizer, but other categories also would fit. Constructivist learning developed at
higher levels of abstraction and information sharing, where there were
opportunities to access diverse perspectives, multiple sources of information,
responsiveness, accountability and democracy. An implication is that to
accomplish disability management successfully as it occurs within complex
contexts, with power discrepancies and lack of awareness of critical issues,
developmental and formative evaluation that attends to diversity, responsiveness,
democracy, and cultural competence, can facilitate the necessary social learning,
awareness and change.

Memo 2

Biomedical – organizational learning at program level, politics negative impact on
evaluation.
Labour – evaluative inquiry, organizational learning at program level, politics
balanced for program evaluation, motivation, team dynamics, organizational
processes, collaboration and communication.
Biopsychosocial – organizational learning, evaluative inquiry, politics impact
funding, motivation, team dynamics, organizational processes, collaboration and
communication.
Insurance – political influences vary, intelligence versus experience discrepancies.

_____

Visual displays of data were also developed during analyses. I began with a large

wall mounted chart on which open codes were grouped into themed categories to

organize and display descriptive codes. Four individual figures illustrating

communications, information flow and evaluation at each of the four sites were

developed and refined. These figures were presented to one participant from each of the

respective sites to obtain feedback, which was used to modify the figures thus increasing

their accuracy and validity as descriptions for each site. As categories of conceptual

codes were synthesized into themes, one conceptual framework was visually illustrated

showing how all the core concepts, categories and themes interrelated. This visual display

was analyzed and modified enabling hypotheses to be tested during formulation of the

best explanation of disability management evaluation.

3.7.2.5 Theoretical sampling. Theoretical sampling refers to gathering data

based on concepts, "Rather than being used to verify or test hypotheses about concepts,

theoretical sampling is about discovering relevant concepts and their properties and

dimensions" (Corbin and Strauss, 2008, p.144). Theoretical sampling helps to: define the

properties of categories and the contexts in which they are relevant; to specify the

conditions under which they arise, are maintained and vary; and discover their

consequences (Charmaz, 2000).

Theoretical sampling begins with concepts from the research literature, and then

is responsive to the data as it is collected. The researcher is guided by an initial

understanding of the phenomenon in selecting the first sample, which is based on

understanding from reviews of the literature. Data are collected and analyzed from the

first sample, "…concepts are derived from data during analysis, and questions about

those concepts drive the next round of data collection" (Corbin and Strauss, 2008, p.

144). The researcher's increasing understanding of categories within the data and of

developing theory with each sampling, direct subsequent sampling (Bryant and Charmaz,

2007). Information from each successive sample is compared to information from the

previous samples and to information from the literature. Data saturation occurs when

meaningful themes emerge, about which the researcher has been able to develop a depth of conceptual understanding, or when substantive theory has been formed.

For this study, theoretical sampling involved selection of a sample of four disability management program sites, each representative of a different disability management paradigm: biomedical, labour, biopsychosocial or insurance. Selecting sites based solely on a literature review of the disability management literature is one approach described by Charmaz (2000). This approach differs from the grounded theory method attributed to Glazer, who selected sites in response to the data analysis toward saturation of concepts that are being developed.

3.7.2.6 Theoretical integration. A theoretical explanation can result in confirmation of information in the literature; identification of new information or meanings not in the literature; or, identification of information from the literature that is not found in the data collection (Miles and Huberman, 1994). This can include re-synthesizing old information in new ways.

A definition of a theory according to the Canadian Oxford Dictionary theory is, "a supposition or system of ideas explaining something, esp. one based on general principles independent of the particular things to be explained." (Barber, 1998, p.1504). "A more formal understanding common in the natural and social sciences is that theory is a unified, systematic causal *explanation* of a diverse range of social phenomena. Theory of this kind is evaluated in terms of the familiar criteria of parsimony, completeness, predictive power, and scope" (Schwandt, 2007, p.292).

Theory understood this way is the proper goal of social sciences (Schwandt, 2007). Empirical and theoretical structures build upon one another, so current studies

extend earlier work, referred to as theoretical cumulativeness (Kline, 2009). In many cases theories are neither corroborated nor refuted, and no cumulative knowledge is built, often in soft areas where empirical research is difficult, or on hot topics where interest just fades away. For topics where there has been little or no prior research documentation in the literature, no scientific breakthroughs and no theories to build upon, exploratory research methods can be used to form initial conceptual understandings and from that to build theory.

While theories give oversight of a topic, they differ from hypotheses. Hypotheses may be deduced from a theory, and focus on smaller aspects of the topic that are amenable to empirical investigation (Meltzoff, 1998). For example, hypotheses that are testable and deduced from a theory may explore possible rival explanations, explore negative cases, or attempt to replicate a finding.

Schwandt (2007) suggested there are different levels of theoretical sophistication, organization and comprehensiveness. At the simplest level are theoretical ideas, concepts that function as analytic tools, pointing the inquirer in a general direction without specifying what is expected to be discovered. At a level up are theoretical orientations or perspectives, social theories that serve as approaches to explain social reality, for example, "functionalism, symbolic interactionism, behaviourism, phenomenology, hermeneutics, feminism, social constructionism, and poststructuralism)" (p.292). Crotty (1998) explained that in clarifying his/her theoretical orientation, a researcher elaborates on the assumptions being made, explaining the context for the process and grounding its logic. Schwandt (2007) suggested at a still higher level are substantive theories that differ from these theoretical frameworks because they are about a specific or

behavioural phenomenon (e.g., a social constructionist theory of a particular cancer). Substantive theories can develop into formal theories that are generalized or extended to cases other than the one studied.

Exploratory research methods applied to build theory in areas where no former theory has been developed can start with building a conceptual framework, "explaining, either graphically or in a narrative form, the main things to be studied – the key factors, constructs or variables – and the presumed relationships among them" (Miles and Huberman, 1994, p. 18). Grounded theory involves inductive processes of analysis, from data to concepts to explanation, where abstract concepts are developed and the relationships between them are specified (Bryant and Charmaz, 2007). Theoretical concepts "result from iterative processes of going back and forth between progressively more focused data and successively more abstract categorizations of them" (p. 25). The most significant categories are increasingly analyzed and raised to concepts in the emerging theory. Theoretical saturation in grounded theory is "the point in the analysis when all categories are well developed in terms of properties, dimensions, and variations. Further data gathering and analysis add little new to the conceptualization, though variations can always be discovered" (Corbin and Strauss, 2008, p. 263).

Corbin and Strauss (1990) summarized canons and procedures for grounded theory. Data collection and analyses are interrelated processes starting with the first bit of data collection. Analysis addresses all seemingly important issues that are then incorporated into the next collection of data and observations. Concepts are the basic units of analyses and "each concept earns its way into the theory by *repeatedly* being present in interviews, documents, and observations in one form or another – or by being

significantly absent" (p. 7). Consistency is achieved by seeking indicators of concepts in all subsequent data collection and analyses. Concepts are then categorized, where categories are at a higher level in the explanation of the phenomenon studied. Grounded theory uses constant comparisons among data, concepts and categories during analyses. Patterns and variations are identified, and processes are specified, such as "breaking a phenomenon down into stages, phases, or steps" (p. 10). Writing memos is integral to developing a grounded theory to elaborate ideas during analyses and integrate details. Hypotheses about relationships among categories are developed and tested, to revise and build the explanation of the phenomenon. Conditions that seem to immediately impact the phenomenon are analyzed, and broader contextual conditions such as "economic conditions, cultural values, political trends, social movements" (p. 11) are also brought into the analysis and integrated into the theory where relevant by showing specific linkages.

Corbin and Strauss (1990) suggested that four areas should be addressed in judging a grounded theory (and cautioned against positivistic connotations): 1. validity, reliability and credibility; 2. plausibility and value; 3. adequacy of the research process; and, 4. empirical grounding of the research findings. Grounded theory is designed "to develop a well integrated set of concepts that provide a thorough theoretical explanation of social phenomena under study. A grounded theory should explain as well as describe. It may also implicitly give some degree of predictability, but only with regard to specific conditions" (p.5).

3.8 Data Collection Methods

Data collection methods at each site included semi-structured interviews, review

of documents, and completion the Readiness for Organizational Learning from Evaluation Inventory (ROLE), which assesses perceptions of learning within an organization.

3.8.1 Scheduling interviews and completion of the ROLE. Following ethical guidelines, potential participants were initially contacted in person or by email (not by telephone). Initial contact included providing each potential participant with copies of (1) the informed consent form, and (2) the ROLE inventory, for their consideration.

Participants who agreed to participate were scheduled to participate in an interview with me. Interviews were scheduled to take place within the week following my initial contact with them, and their agreement to participate. Dates of interviews are listed in Table 6. Participants all chose to participate in their interviews with me at their worksites, where I met with them.

Participants were asked to complete the ROLE inventory prior to their interview, and to submit their completed ROLE to me at the beginning of their interview, which each person did. I obtained signed consent forms and completed ROLE inventories from each participant at the interview.

Table 6   Dates of Participants' Interviews

Participant	Date of Interview
BM1	March 31, 2010
BM2	March 31, 2010
L3	June 16, 2010
L4	July 14, 2010
L5	July 20, 2010
BPsy6	July 27, 2010
BPsy7	December 2010
I8	May 5, 2011
I9	May 6, 2011

3.8.2 Semi-structured interviews. Semi-structured interviews were guided by several pre-planned questions, which are included in Table 7. The interviews started with clarification of the interviewee's past evaluation experience and the role of the participant within the disability management program he/she was attached to. Open ended questions probed the interviewee's understanding of disability management evaluation practices and organizational learning. Participants were asked to describe their experiences guided by the interviewer to maintain focus on evaluation. Each participant was interviewed for 60-90 minutes.

This interview protocol was pilot tested in June 2009 to: (a) practice interviewing that successfully maintained a focus on disability management evaluation, (b) practice interviewing that led to substantial information reporting on disability management evaluation practices, (c) test pre-written questions to determine how useful they were to meet the objectives of the interview, and so that new questions could be composed that would be considered more likely to have the potential for improved data collection, (d) practice recording an interview to ensure the equipment worked properly, and (e) investigate what types of unexpected information regarding disability management evaluation may be forthcoming during interviews.

Based on the pilot interview the questions were revised to be more focused on the research questions and to increase the likelihood of eliciting information from the personal experiences of the participants. The microcassette recorder malfunctioned during the pilot interview, and parts of the interview were muted, therefore new equipment was obtained before the actual study, and tested to ensure the equipment was in sound working condition.

Table 7   Grounded Theory Semi-Structured Interview Questions

What is your background or experience with program evaluation in general?
What has been your experience with the disability management program we will discuss?
How do you know that your program is doing what it is intended to do?
How is your program evaluated?
Who is responsible for evaluation of the program?
What do you think is most important to be evaluated in disability management programs in general?
Based on responses to the ROLE, how do you think the characteristics of your organization strengthen receptivity to learning?
How do you think characteristics of your organization need to change to be receptive to learning?
How would learning relate to the disability management program?

In the pilot interview the interviewee provided significant insights into psychosocial interpretations of disability management outcomes. For example, the interviewee provided a graph he had composed and described how the graph depicted longitudinal psychological adaptation to disability. Also, the interviewee looked beyond simple return to work outcome statistics in his interactions with two different employer sites that provided similar services in the same industry, and explored underlying psychosocial issues that he concluded had resulted in return to work outcomes that were uncharacteristic of each of those two employers. These examples suggested this data collection strategy had the potential to lead to rich information describing interacting

variables that exist within complex organizational contexts, taking into consideration multiple stakeholder perspectives.

      3.8.3 Documents analyzed. Table 8 illustrates who at each site provided documents, or whether I as researcher obtained them, and which documents were acquired. All documents were reviewed in depth, with a primary objective of analyzing any information related to evaluation.

      Document review and analysis initially involved identifying which documents included information related to evaluation and which did not. Documents that focused on evaluation were: *Client Satisfaction Surveys (*from the BM, Labour and Insurance sites), *Key Performance Indicator* (from the BM site), the *Manager Joint Return to Work Program Survey* (from the Labour site), and the handwritten summary of BM2's criteria for success. Other documents did not reveal any information on evaluation at the sites. Participants were asked for any documents related to evaluation of their program, and none other than those noted were provided.

      Documents pertaining to the programs were open and conceptually coded (at the time that interview transcripts were coded so that lists of codes included codes from documents and interviews). Analyses of documents included triangulation of documents to documents from the same site, documents to interviews from the same site, and documents in relation to all sources of data together.

      3.8.4 The ROLE Inventory. The Readiness for Organizational Learning and Evaluation Instrument (ROLE) assesses the perceptions of personnel about their work environment in relation to learning from evaluation, leadership, structures,

Table 8　　　　Documents Obtained on the Programs

Site	Participant provided or Researcher accessed	Documents
BM	BM2	- Handwritten summary of BM2's criteria for success *Client Satisfaction Survey* *Key Performance Indicator*
	Researcher	Organization's web page
Labour	L3	- Client *Satisfaction Survey* - Manager *Joint Return to Work Program Survey* - Physician *Functional Abilities Evaluation*
	L5	Return to Work Program Manual
	Researcher	- Hard copy of a power point presentation describing the program - Brochure describing the program - Program Coordinator job description - Organization's annual reports for 3 prior years - Organization's mission statement
BPsy	BPsy6	Email outlining the data base categories BPsy6 was establishing
	Researcher	- Program Counsellor job description - Organization's web page - Brochure describing the program
Insurance	Researcher	*Client Satisfaction survey* - Program Investigator job description - Magazine published by the organization - Organization's annual reports for 3 prior years - Organization's mission statement
Note: I asked each participant for any documents related to evaluation, and formal evaluation reports were not available from any of the sites.		

communication and culture (Preskill and Torres, 2000). (See Appendix 2 for the ROLE

Inventory.) The ROLE was developed because "an organization must have certain

elements of its infrastructure in place if it is to truly support and encourage organizational

learning" (Preskill and Torres, 2000, p.429). The ROLE items reflect those organizational elements shown to significantly influence the extent to which evaluation supports learning and decision making.

The ROLE is comprised of 78 items, grouped into six major dimensions: culture, leadership, systems and structures, communication, teams, and evaluation. There are not right or wrong answers, and responses are on Likert Scales from 1 (strongly disagree) to 5 (strongly agree).

The ROLE is scored by calculating the mean of responses for each of the dimensions of the survey. The authors recommended when the ROLE is administered within one site, organization or department, that scores be aggregated and reported in summary form, however in this study individual participant scores were analyzed separately. Results can be used for discussion or feedback with an organization.

An example of interpretation of a ROLE score might be, "If a department or organization were to score low in one or more of the dimensions, this would indicate that learning from evaluation might not be supported or allowed to succeed. Likewise, it would indicate that the department or organization isn't prepared to engage in other kinds of organizational learning practices. These kinds of results can help the organization determine where to focus its improvement efforts if its goal is to become a learning organization." (Preskill & Torres, 2000, p.9).

   3.8.4.1 Uses of the ROLE in this study. In this study the ROLE was not used in the typical way intended by the authors. The ROLE was sent to all nine study participants prior to their semi-structured interviews and returned completed during each interview with signed consent forms.

The ROLE was used initially to stimulate discussion on organizational learning and evaluation during the interviews. Content from the ROLE was discussed during each interview to stimulate participants' thinking regarding organizational learning and evaluation, and to elicit participants' perceptions regarding strengths and weaknesses within their organizations' or programs' infrastructures, in relation to learning and evaluation.

The ROLE was selected for use in this study as the preferred method to gather information on participants' perceptions of learning and evaluation within their organizations. Administration of the ROLE guaranteed each participant would have an opportunity to contribute rigorous data on the subject by having the time to focus on each ROLE question and to complete the entire inventory. The ROLE was preferable to questioning participants on these issues during the semi-structured interviews, as during interviews there was a possibility of inconsistent or limited focus on the topics that were covered in the ROLE. This was especially true because the interviews were semi-structured with open ended questions as part of the grounded theory method, and there was no guarantee topics covered in the ROLE would be discussed in the interviews.

A third benefit of using the ROLE rather than asking questions during the semi-structured interviews was that development of the ROLE had been based on research into organizational learning and evaluation processes and practices. The ROLE items were developed based on the understanding that an organization must have certain elements of its infrastructure in place if it is to truly support and encourage organizational learning. The 75 Likert Scale ROLE items that were created, according to statistical confirmation of high coefficient alphas, represented those elements considered necessary.

Additionally, using the ROLE enabled a third data source, and triangulation among three different data sources. After completed ROLE inventories were scored, scores among participants from each site were compared and contrasted; scores among all nine participants were analyzed together; and, ROLE scores were triangulated with other data sources including the interviews and documents. ROLE results were referred to during documentation on their own merit, and in support of explanations and hypotheses being made based on other data sources.

3.8.5 Triangulation of multiple data sources. Triangulation across different data sources or multiple researchers is perceived as a means for validation of evidence, error reduction, and dismissal of rival explanations (Mathison, 2005). Triangulation among multiple sources of evidence should address consistencies, inconsistencies and contradictions, contributing opportunities for the researcher to construct plausible explanations, rather than using triangulation as it sometimes is for validation alone (Mathison, 1988).

Administration of the ROLE was not part of the grounded theory data collection, however as noted by Bringer, Johnston and Brackenridge (2010), Glazer and Strauss had pointed out in 1967, "Grounded theory does not, however, preclude the use of quantitative data such as survey data that can be used at the later stages of a project to support or further explore the initial analyses." (p.248).

The ROLE was an ideal source of data to be integrated with the other sources: interviews and document analyses, to answering the research questions of this study. In particular, the ROLE contributed to exploration of evaluation practices and learning at the sites, and to analysis of how evaluation varied depending upon whether the

organization was a learning organization. The ROLE provided detailed information on the participants' perspectives of the degree to which their organizations or programs were receptive to and appreciative of learning and learning from evaluation.

The three sources of data triangulated together contributed to identification of patterns, political perceptions, social interactions, power discrepancies, stakeholder values, evaluation purposes and procedures, and overall to conceptual and contextualized understandings.

3.9 Issues of Validity and Reliability

Methods for enhancing validity and reliability in exploratory research discussed by Appleton (1995) have been followed in the collection and analysis of data in this study. Purposive sampling was used. A pilot interview was conducted to develop research interview skills and test the interview questions. Audiotaping and transcribing interviews verbatim fostered the accuracy of data analysis. My views from prior professional involvement in disability management were explored and scrutinized, to understand and minimize any potential for biases during data collection and analysis. During data analysis I referred back to interview, document and ROLE data while developing themes, to confirm accuracy, challenge assumptions, and consider alternate explanations. Models illustrating disability management evaluation and communications were developed and reviewed by participants to verify accuracy.

3.10 Reflexivity

Researchers do not divorce their research from who they are and therefore, need to be self-reflective (Corbin and Strauss, 2008). Researcher journals are one strategy for documenting reflections, which then can be used as one source of data in the analysis.

Reflective writing refers to, "critical self-reflection on one's biases, theoretical predispositions, preferences, and so forth" (Schwandt, 2007, p. 260), and is important for establishing validity.

My analysis during this study has been influenced in a number of ways by my experience. First, I have practiced vocational rehabilitation for twenty five years, which has involved interacting with a multitude of disability management programs and different stakeholder groups (workers, employers, unions, treatment and assessment facilities, funding agencies, medical practitioners, schools, etc.). Also I have experienced in my work how some program functions and stakeholder behaviors represent different paradigms of disability management described in the literature (biomedical, labour, biopsychosocial, insurance, etc.). Third, I have experienced how stakeholders who are willing to communicate and be receptive to each other's perspectives can catalyze learning and bring about growth. I have also observed how, in contrast, those who steadfastly hang on to their perspective or withdraw into cynicism are less likely to contribute in ways that facilitate learning, can contribute to stagnation of team work, and can increase the possibility of reaching invalid conclusions with short and long term consequences. I have seen that depending on the composition of case management team participants, potential results for the same case can be entirely different. And finally, over the past several years I have worked within a group of vocational rehabilitation professionals several of whom have been accessing a disability management program like the labour paradigm site in this present study. Prior to this study I had never discussed this disability management program with anyone and had very little familiarity with its operations.

3.11 Summary

This grounded theory study involved theoretical sampling of sites representative of four paradigms of disability management: biomedical, labour, biopsychosocial and insurance. This research considered how evaluation practices at the sites may reflect what is important to the respective disability management paradigms. A total of 9 participants, including administrators and practitioners, were involved, 2 or 3 from each site. Data collection methods included: semi-structured interviews, review of documents and administration of the ROLE instrument.

CHAPTER 4: RESULTS

In this chapter the evaluation of disability management programs at four sites representing primary paradigms in the field are described. Data analyzed are interviews with nine participants, documents, and the results from the Readiness for Organizational Learning from Evaluation Inventory (ROLE) from the four sites. The contextual framework for each program is described and evaluation practices at the sites are discussed by 1) values and goals of evaluation, 2) stakeholder involvement, 3) evaluation data collected, and 4) the use and reporting of evaluation data. These features of evaluation are summarized in a diagram and discussed in relation to the study's four research questions. The ROLE results are analyzed and triangulated with interview data to characterize each organization's readiness to learn.

A cross case analysis is presented in Chapter 5, as the basis for the grounded theory of evaluation of disability management programs.

Disability Management Program Evaluation by Sites

4.1 Biomedical Paradigm Site

4.1.1 Context and framework of this program. The biomedical site was a small business owned and operated by two occupational therapists (referred to as BM1 and BM2 in this study). Services included functional capacity evaluations, job demands analyses, and development and monitoring of return to work plans.

This site primarily received referrals from two large public nonprofit disability insurance companies in British Columbia that oversaw compensation for injury or illness sustained at work or in motor vehicle accidents. BM1 received approximately the same number of referrals from the two, and BM2 received virtually all referrals from one of the

organizations, making this organization the main referral source of the biomedical site. Other referrals came from different insurance organizations, employers, unions, lawyers and client self referral.

The referring agencies administered treatment plans based on medical and functional capacity evidence. Assessment information provided by the biomedical site was one source of that evidence. BM2 clarified the services they provided,

I evaluate, I recommend solutions, I may actually implement a solution and assess its effectiveness and hopefully dissolve the situation in terms of returning the person successfully to a job. But primarily I am a supplier of information.

Toward this end, the program staff highly valued accurate information and useful reports. Figure 1 illustrates evaluation at the BM site as described to me by the BM participants of this study. In developing this diagram I forwarded a first draft to BM2 and requested feedback, and recommended modifications were incorporated.

The main referring agency providing the BM site with 75% of its referrals was a large organization that managed disability claims related to injuries or illnesses. This agency established a network of providers that it contracted with for limited periods of up to several years to provide the various biomedical services they require (functional assessments, job demands analyses, graduated return to work plans, etc.). The BM site had been one of the providers in that network since its inception over fifteen years prior. Contract lengths were uncertain lasting from year to year for each individual service. The bidding process could open with little warning, and an updated network of providers would be selected from among the applicants. There were no guarantees of contract

renewals, leaving providers such as the BM site in somewhat of a precarious situation

given the proportion of their incoming referrals that came from that one source.

Figure 2 Communication Among Program Stakeholders as Described by the Biomedical

Participants

Clients' Physicians

Clients

Clients' Employers

Self Referred Clients

Funding Referral Agencies

DM Program

S
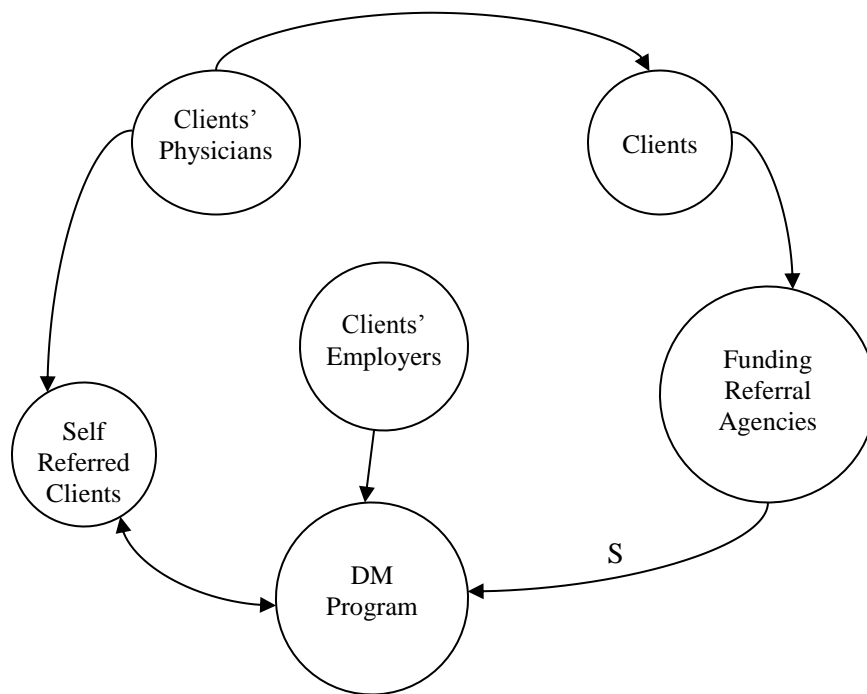
Note: S = Summative evaluation

Communications with large referring agencies were unidirectional, with little

opportunity for the BM site to decide what information was shared. There were no

iterative processes of communication between the BM site and the referring agencies, and

no opportunities for dialogue or for the BM site to solicit feedback from individuals at the

referring agencies. There were no collaborative communications on individual cases among the BM site, the referral agency and the client. BM2 explained,

> There is a difficulty in contacting a referral source directly. Most of the time they do not like speaking directly with the provider. There are actually no communications going on any more. It's all through fax and reports. Rarely will I get an actual call saying you did a great job or I have questions. Nothing. It has become a very distant relationship...So in that case I do not feel comfortable the way it is right now in actually picking up the phone, calling the referral source and saying, 'Hey how did you like my service?' Because [they are] already supposed to measure that. I'm not supposed to go and ask for that information.

There were also no communications with the main referring agency regarding the outcomes of services provided,

> We pretty much never see them again. Or it's rare that we see them again down the road. You know, how impactful was the evaluation, or how the evaluation results effected what direction the case manager or the voc rehab went in, we don't usually find out that information (BM1).

4.1.2 Evaluation practices within the biomedical site.

4.1.2.1 Value of evaluation. Given most of their referrals came from one referring agency, the biomedical site staff were acutely aware of how important evaluative feedback from this organization was, to understand whether their services were valued and whether continued referrals would be likely. The staff would have liked descriptive feedback on the quality of their reports and on the value of their assessments and job site interventions. However, the staff reported this was not available, so they

made inferences based on the only information available: that repeat referrals from the same referring agency indicated their services were valued.

The staff would have preferred an open dialogue and iterative processes with all parties on each individual case, with opportunities for everyone to share information and learn, but they rarely had an opportunity to discuss cases with the referral agency prior to, during or after completion of their services. BM1 characterized communications between the parties as so unreliable that they often received medical background documentation, which was intended to prepare them for a given referral, after they had completed their services and submitted their final report.

Additionally, due to many short term services they provided, the BM site staff was often working with new clients and their employers over brief periods of only several hours to a couple days, leaving little time to develop meaningful dialogue. The exception to this was when they would oversee return to work plans that continued several weeks to months, during which time they were able to monitor events and could recommend and incorporate changes to their interventions based on new information acquired.

Instead of relying on evaluative feedback from referring agencies, program staff supported each other, collaborating on cases and sharing information where relevant to support each other's success. Both valued feedback and ongoing learning, as BM2 explained regarding the importance of evaluation,

> You should have ways to know that what you do is effective and also to improve
> your ability where you are learning or your effectiveness on an ongoing basis….
> or even questioning myself as to what would I judge to be effectiveness.

4.1.2.2 Goals of evaluation. Services provided by the BM site mainly involved

data on how each client's physical impairment had impacted functioning, and ergonomic

return to work assistance. BM1 reported that most services involved a physical

assessment, which was based on a functional capacity assessment at the site's clinic or at

a job site.

The BM site received summative evaluative feedback from the main referring

agency that emphasized timeliness in their provision of services. The BM site attempted

to minimize the time taken to provide services, but did not believe timeliness was the

most important factor to evaluate, "… quality is first, timeliness is second. But you can

have a very good report and if it's late it will be useless" (BM2).

The BM site staff would have preferred evaluative feedback on a case by case

basis, working collaboratively with others, so they could learn how their services and

reports were impacting the clients or others (referring agency, employer, union, health

service providers) and to incorporate improvements in handling similar cases.

4.1.2.3 Stakeholder involvement. Stakeholders of the BM site included clients,

referring agencies, employers, and clients' health professionals, lawyers, union reps, and

co-workers.

The BM site was the only site of the four participating in this study that was

primarily externally evaluated. Summative external evaluation of BM site services was

conducted by the main referral agency, which specified that the BM site use two

strategies for evaluation: a *Key Performance Indicator* that generated four statistical

measures as feedback on services, and a *Client Satisfaction Survey* that gathered clients'

opinions of the program.

The referring agency summarized detailed information from the *Client Satisfaction Survey* into a single numerical rating. An average of the single rating across all client surveys was provided as feedback to the BM site, and the BM site staff did not have access to the disaggregated data from the *Client Satisfaction Surveys*. BM2 found the single numerical feedback superficial and potentially less meaningful than descriptive feedback that could have been communicated. How the evaluation was conducted was one source of dissonance between the program and the referring agency.

The staff sought formative evaluative feedback regarding their services from other sources. They had employers and workers review job demands analysis reports to ensure accuracy (BM2 advised that employers and workers often had "different views about what the job requires." so feedback from both sources was sought). They had clients and their employers, physicians, and referring agencies review graduated return to work plans to confirm feasibility. Both BM1 and BM2 conducted ongoing work site visits when overseeing return to work plans to elicit feedback from the perspectives of multiple stakeholders at the job site (employer, union, workers) and then incorporated new information and modified plans where needed.

BM1 and BM2 held debriefing sessions with clients at the end of their services, so they could make improvements where appropriate. All services at the BM site included at least one final report, and some also included interim reports. BM1 and BM2 compared their reports to those of other providers when available to incorporate better reporting strategies.

4.1.2.4 Evaluation data collected. On the *Key Performance Indicator*, four criteria were rated on a scale of 1 to 10: (1) time from referral to first client contact,

(2) report turnaround time, (3) client satisfaction, and (4) percentage of client surveys completed and returned. The main referring agency compiled scores on the *Key Performance Indicator* criteria for all agencies in their network (including the BM site) so that scores of all the provider agencies could be compared to each other. (At the time of this study there were fifteen provider agencies from across British Columbia listed on the *Key Performance Indicator.*) The referring agency provided all provider agencies the target scores (standards) to aim for on each criterion of the *Key Performance Indicator.* A summary of comparisons of scores for all the provider agencies was distributed quarterly as feedback to all agencies in the network of providers.

Completion of the *Client Satisfaction Survey* was required by the main referring agency, having replaced the BM site's own client satisfaction survey, and was administered on behalf of the referring agency by the BM site staff at the completion of services. The completed satisfaction surveys were submitted directly to the referring agency, where results were summarized. The survey included: four statements with Likert scales from 1 to 10 pertaining to satisfaction with the service, one yes/no question, and two open ended questions on strengths/weaknesses of the service. Results from the *Client Satisfaction Survey* were summarized into a single rating from 1-10. This numerical rating was then averaged across all clients for the BM site, and this mean score was the indicator of client satisfaction, one of the four criteria on the *Key Performance Indicator*.

BM2 advised "report turnaround time" was important to the main referring agency. The referral agency provided no feedback on the quality of work being done

(services or reports) or on outcomes (such as return to work) which would have been the more significant criteria for evaluation in the opinions of BM1 and BM2.

> …report turnaround time is a huge thing for [the referring agency]. When in the end it doesn't really reflect on the quality of what you think or the effectiveness of your service. But that's to tell you how [they] see you as a supplier of information and they want that information now. (BM2)

Accuracy was important to the BM site. They had employers and workers read job demands analyses reports to ensure accuracy. They had workers, employers, clients' physicians and referring agencies review return to work plans to judge feasibility and sought ongoing formative evaluations from workers, employers and unions at the job sites to make improvements while monitoring return to work plans.

4.1.2.5 Use and reporting of evaluation data. Two of four criteria on the *Key Performance Indicator* focused on timeliness: time from referral to first client contact and report turnaround time. Because the referring agency valued timeliness, the biomedical site attempted to complete these services in the minimum time necessary.

The BM site administered the required *Client Satisfaction Survey* and completed surveys were submitted directly back to the referring agency for analysis. The other criteria on the *Key Performance Indicator* were client satisfaction (as measured by the result of the *Client Satisfaction Survey)* and the percentage of *Client Satisfaction Surveys* returned. As the BM site did not find this information very useful, BM2 conducted exit interviews with clients. She described her "criteria for success" as "Debrief a client at the end of a service, to know how they felt about the whole situation, explaining what now, what I will do, and to give me some information on whether I have done a good job or

not." Feedback from these exit interviews provided information more in line with the BM site's former client satisfaction survey that had elicited clients' views.

> It was for the worker to assess our service. In terms of whether they had been treated in a respectful manner, whether they had learned something during the evaluation that I provided, or through the service that I provided.

Regardless of the attention given to satisfaction of the client, BM2 considered satisfaction of the referral agency to be a more significant indicator in the evaluation of services.

> It's not the satisfaction of the injured worker. Let's put it that way. It's important all right, but it's not the most important….Because my service was requested by an entity … that's the entity that should judge the effectiveness of my service.

BM2 said their site had tried unsuccessfully to obtain feedback from the specific person within the referral agency and was told "I'm not supposed to go and ask for that information….we had also started sending a satisfaction survey to the referral source, and we were told to stop that."

Based on the *Key Performance Indicators*, the BM site had not initially met the required target on the criterion: client response rate. This led to clarification that the rating was based on the percentage of client questionnaires returned. Subsequently the BM site staff ensured every client was administered a satisfaction questionnaire, and the following quarter their score on this criterion of the *Key Performance Indicator* rose from 71% to 100%.

The BM site minimized any negative impact on *Client Satisfaction Survey* scores that could be caused by misunderstandings the clients may have had during their

involvement in services. BM2 explained how this had been achieved by screening
completed surveys prior to submitting them,

> Sometimes I will actually ask the client not to seal the envelope. I want to read it.
> So once in a while I will pull one out and I am not supposed to do that. I am
> supposed to send them, and I am not. I just want to read what they said, because
> there is a comment section that some people were quite extensive on [saying for
> example] I don't like the fact that this is happening.

BM2 then decided negative comments could be minimized if she actually completed the
survey with the client.

> Well then I started doing the survey with the client by actually asking them the
> questions first to know if there were areas where they were not reading me very
> well on. In that case I felt that being there directly and listening to their response
> may not be the most objective measure.

> The staff felt they had little feedback on the usefulness of their reports, as
referring individuals were nearly impossible to reach. On occasion BM1 and BM2 had
access to the reports of other provider agencies when those reports had been disclosed as
part of the background information that was forwarded to the BM site with each referral.
Whenever possible they compared the formats of their own reports to the reports of other
agencies, and made improvements on their own where appropriate.

> BM2 clarified that she was interested in the formats of others' reports, but did not
want to be influenced by their opinions or findings, so she postponed reading the prior
reports until after completing her own assessment. On occasion a client would complain
to her about a prior assessment they had undergone, and in those cases BM2 would read

that prior report. Access to other providers' reports allowed BM2 to critique methods others had used, and on occasion enabled discovery of new procedures that the BM site would research for potential incorporation into their own services.

4.1.2.6 Summary of evaluation practices. Table 9 summarizes evaluation practices at the BM site in terms of three components, (1) evaluation criteria, (2) data sources, and (3) use of findings. These components serve as a guide to understand what is to be evaluated and how, including, "…evaluation issues, questions, indicators of success (qualitative and quantitative), appropriate data sources and methods to be used to collect data" (Cummings and Paulmer, 2010, p. 5). Data were not consistently available on indicators of success or standards, but this information is included when available.

The biomedical site staff perceived weaknesses in evaluation in the absence of communications and dialogue with the referring agencies. These included errors and omissions, lack of meaningful information, misunderstanding of findings, and disregard for accuracy, all of which had the potential to influence decisions for individuals or organizations. The biomedical site staff wanted to collaborate with referring agencies to identify relevant criteria and accurate data collection strategies.

4.1.3 Evaluation practices and organizational learning. BM1 and BM2 emphasized that, "We do a lot of learning." This includes taking formal training courses, attending conferences, collaborating among professionals, ongoing research (primarily using web-based sources) to maintain up to date knowledge in their field.

Asked during our interview whether there was anything within their organization that could be done so that they could benefit more from evaluation or that could enhance learning, BM2 advised,

Table 9   Evaluation Activities at the Biomedical Site

Criteria	Data Sources	Use of Findings
*Time from referral to first client contact	Program staff reported dates to referral organization *Key Performance Indicator*	Expedited client contacts to meet standards
*Report turnaround time	Referral organization data base *Key Performance Indicator*	Expedited submitting of reports to meet standards
*Client Satisfaction	Referral agency *Client Satisfaction Survey*	Conducted exit interviews with clients to gain more detailed information
*Number of *Client Satisfaction Surveys* completed and returned	Referral organization data base *Key Performance Indicator*	Administered *Client Satisfaction Surveys* to all clients and returned surveys
Client assessment of service	Exit interviews with clients	Modified interventions with clients
Quality of reports	Program staff analyses comparing their reports to the reports of other agencies	Modified subsequent reports
Accuracy of job demands analyses	Employers' and workers' reviews of reports	Corrected reports
Quality of measuring equipment	Participant research of current publications	Upgraded equipment
Note: * = Denotes the four categories listed on the *Key Performance Indicator.*		

It is always good to do this sort of review. I am an ongoing learner….I am a

seeker. I review literature. I Google my own field. I check for new tools. I check

for how things are done. I check on studies on the validity and reliability of the

tools that I use. I take certification whenever I can. I try to keep up with that….I

am constantly changing. I am not doing the same things that I was doing last year.

Throughout their work staff collaborated as much as possible. BM1 explained,

We are always chatting about how this can be done differently, or I have this

problem has it happened to you before? How did you deal with it? What is going

to work for that situation? That happens all the time….Any type of problem or

question I'll ask and we will come up with how we did this before….She'll say oh

yea you know I worked with this person and we made this type of splint, or we

got this type of keyboard or mouse and that made a difference.

Referring to collaboration BM2 explained,

It is important…because often we provide the same service, that we are seen as

consistent. We work under the same roof. [We have] to produce something that

looks very similar….So often we will talk about how we would see a certain

situation. What we would do. We are always right next to one another. So of

course there is an ongoing consultation between the two of us. We always share.

BM1 and BM2 adapted to the main referring agency's evaluation design that

limited them to the *Key Performance Indicators*, but neither found it included what they

most valued. The perceived weaknesses of evaluative judgments from the main referring

agency and the consequent difficulty in trying to ensure services were improved to meet

the needs of the referring agency resulted in an unexpected and sudden loss of a contract.

The main referring agency opened competitions to update their network of providers and the BM site applied to renew their contracts on all services, most of which they had provided since the onset of the network in 1995. The BM site, and in particular BM2 who received nearly all referrals from the main referring agency, lost the contract to provide one of their main services, overseeing return to work programs. The program was advised that insufficient detail of what their service would entail was provided in the application. This decision was made without dialogue between the referring organization and the BM site staff, in spite of the fact the site had been providing this same service to this referring agency for nearly two decades without complaint.

The referring agency's publication comparing scores of all providers in the network on the *Key Performance Indicators* compelled the staff to modify their strategies for reporting timeliness to increase the percentage of *Client Satisfaction* Surveys returned. These changes were made to improve their ranking in future comparisons with other providers' scores.

4.1.3.1 Summary of ROLE scores at the BM site. Overall ROLE response patterns for the two BM participants indicate they valued feedback and learning from evaluation within the BM program. Their goal was to use evaluative feedback to make improvements to their services and reports that would maximize their value to stakeholders (referring individuals and clients). They believed the evaluation conducted by the main referring agency lacked significance, but consciously adapted to it to minimize the risk of losing referrals. At the same time they were motivated to build evaluation that was more meaningful.

4.1.4 Organizational cultural context. Consideration of cultural issues in evaluation practice is not limited to language, ethnicity or nationality, but also extends to stakeholders' diverse or unique frames of reference. These may be perspectives shared by some stakeholders that are important for making programs more effective or fair.

4.1.4.1 Insurance system culture. The biomedical site staff described an insurance system culture. BM1 explained how this culture was manifest.

> I mean there's always the tangible, clients sometimes see us as [worker's compensation] so some people feel adversarial a little bit. They see that [worker's compensation] is forcing them to do something that they may not necessarily want to do. Some people are very angry. They think you know their employer is at fault that they're hurt. And sometimes that gets redirected to the person that they're working with….[or]…a lot of the clients we see with [government auto insurance] they were not at fault. They were driving doing everything right and someone hit them. So they feel that it is not their fault that they're in this situation and they feel that the therapist is a representative of how [government auto insurance] is trying to push them to go back to work. Sometimes before they feel that they are ready to go back to work….. I think employers too have a bit of a fear sometimes, that well this person has been hurt in a car accident and what happens if they get hurt at the work place. Is that going to affect the [worker's compensation] injury and does that then affect my [worker's compensation] premiums?

BM2 elaborated how clients may express hostility, transferring their anger from the referring organization to the BM site,

I may think oh that that person already was biased when they walked in the door.

It wouldn't matter if I had given 150 percent they will always find wrong with

anything…. The claims management system. The way that their claim has been

handled to that point. They view me as pretty much, even though we are separate

entities from the [referring organization] and are contracted out, the client will

always ask you if you work for [them]. They might even perceive that you still

work for them because you are on contract. So right away they have a bias when

they walk in the door.

To deal with these client perceptions in some cases BM2 did not submit the client's

satisfaction survey to the referring agency, predicting it could work against the

biomedical site if this misdirected anger resulted in a negative client satisfaction score.

BM2 withheld the data, even though she knew this would lower the key performance

indicator score for percentage of client questionnaires returned.

[The return of questionnaires should] be ideally 100 percent….But

sometimes…the client is so upset the whole time you think the only thing they

will do is destroy you, even though you did your best. You might think okay I'll

let that one go because that one will not affect the quality of my service.

The BM site staff had adapted to insurance culture anger by manipulating post service

client surveys, and by developing strategies of positive communications with each client

from the onset to completion of services. From the moment a client walked in the door

certain socialization tactics were used. BM2 explained her rationale,

What could I do to change things? So should I talk less? Should I observe more?

Should I make the person more at ease? Should I offer them a coffee in the

morning right at the start to kind of diffuse any kind of situation that is from their

past experiences that they may want to apply to this? Just make them more

comfortable, feel respected, someone cares, but also at the same time you don't

want to be overly caring because that's not my role. I am there to provide

objective data on a person's ability to function physically and in the workplace.

So I have to still maintain a very professional attitude. I am not going to be taking

the role of a friend. I am still very much an occupational therapist, a professional

that needs to gather information and who expects the client to do their very best

during the whole time.

The biomedical site responded to the potential impact they perceived the insurance

culture would have by selective data reporting to protect themselves with the expectation

this would affect the likelihood of having their contract renewed with the main referring

agency.

     4.1.5 Evidence of evaluation grounded in a biomedical paradigm. The

biomedical paradigm was the first conceptual model applied in disability management, at

a time when disability was primarily understood to involve physical events (Franche,

Frank and Krause, 2005). The biomedical perspective assumes the body and mind are

separate entities and the focus is on a physical condition and its treatment. A physician is

responsible for control and treatment and clients are dependent recipients, a model

predominant among health care professionals (Schultz, et al., 2000). The core of the

biomedical model is, "careful observation, the systematic collection of information, and

objectivity" (p.272), which create evidence based practices influencing treatment

decisions. "The espoused core value of the biomedical model is scientific truth, based on

scientific evidence" (p.272), suggesting that accuracy of evaluation findings about services may be most important to the biomedical paradigm.

The BM site provided four main services that could be evaluated: functional evaluations of clients' physical capacities, job demands analyses, development of graduated return to work plans, and overseeing return to work plans. The BM site's evaluation by the main referring agency did not address accuracy of these services, nor did it address usefulness of the BM site's reports on these services.

Not only did the main referring agency not address accuracy, but the BM site reported that the referring agency passively condoned inaccurate information documented in other providers' reports. BM2 reported that multiple reports of other providers forwarded to her by the main referring agency as background information on clients, included conclusions that deviated from evidence, for example misleading conclusions based on the details of client functioning provided in the reports. BM2 inferred that the other providers may have done this to meet the main referring agency's expectation that programs achieve an 85% success rate on finding workers able to return to their jobs after treatment, in order to continue to be eligible to receive ongoing referrals. BM2 explained inaccuracies in those reports,

> I found that a lot of clients are discharged as fit to return to work without [limitations] or fit to return to work with limitations when they are still not, because that is an important for them to keep their contract. To get more referrals, or to get viewed by the [referring agency] as effective, they need to have a success rate of returning people to work of I think it is above 85%.

BM2 confirmed the pressure was always there to meet the referring agency standards, and to discharge clients as fit to return to their regular jobs, "Yes it is always the other way around. The picture is always rosier than really what it is." BM2 suggested interests of the referring agency, rather than the interests of the clients or employers, were being served by evaluations that overlooked these discrepancies.

> [Reports] are serving the referral source, not the client directly. Of course they are providing the best service they can to the client, but their measure of success does not come from the client himself saying to them I feel really much better from your intervention. It's from the [referring agency] saying hey you guys are doing great. At least 85% of the clients you discharged are fit to return to work.

If these reports were as described by BM2, in passing them on to other providers in their network as valuable, the main referring agency risked vicarious learning on the part of network providers that could result in increased errors in reporting.

External evaluation of the BM site conducted by its primary referring agency did not reflect the focus one would expect within a biomedical paradigm since the criteria of timeliness and client satisfaction could be relevant across all disability management paradigms. However, the referring agency's evaluation used standard indicators and measurements, which is a characteristic of the biomedical paradigm.

The BM site's internal review of the quality of their reports, accuracy of measuring equipment, and employer and worker feedback on the accuracy of job demands analyses, did reflect grounding in the biomedical paradigm that valued objective evidence.

4.1.6. Discussion. Formal evaluation at the BM site was conducted externally by the referring agency giving the BM site no input on criteria, standards or indicators. Four program outcomes were evaluated: timeliness of first client contact, client satisfaction, number of *Client Satisfaction Surveys* submitted, and timeliness submitting the final report. Feedback was provided to the BM site every three months on the *Key Performance Indicator.*

Learning from evaluation at the BM site was pragmatic and directly a consequence of evaluations provided by the external main referring agency on the four service outcomes. Evaluation feedback was a driving force for the BM site, resulting in changes to their practices so they met the agency's standards. The BM site learned to meet the timeliness standards for both client contact and submitting reports, ensured all clients were administered the *Client Satisfaction Survey*, and submitted most of the completed surveys.

The BM site also learned to manipulate administration of the *Client Satisfaction Survey,* and on occasion did not submit responses that reflected poorly on their program, where BM2 believed negative feedback had been biased against them unfairly due to insurance claim biases. This adaptation did not necessarily generate improvements in the services provided, but was a response to the referral agency's organizational culture (Patton, 1996). Adaptation by the BM site was a conceptual theme that emerged throughout discussions with participants about evaluation.

The BM site participants did not believe the evaluation criteria were the most important criteria, but they had no opportunity to provide this feedback to the referring agency. Whereas external evaluation could normally be characterized as maintaining

objectivity, transparency, and perspective of the evaluand (Barrington, 2005), the

referring agency's evaluation was alienating given the absence of dialogue between the

organizations.

The referring agency's evaluation was organized and efficient, provided clear

indications of the criteria and standards, and gave regular feedback, leaving the BM site

with the perception that if they were meeting standards contracts would be secure. For the

most part this was the case however not always, as an unpredictable and sudden loss of a

contract did occur. However, lack of communication between the referring agency and

the program staff resulted in incomplete information, misunderstandings, invalid

outcomes and lack of data beneficial to the BM program and their clients.

4.2 Labour Paradigm Site

4.2.1 Context and framework of this program. The labour paradigm site was an

in-house disability management program available to 2-3000 employees of a large

unionized multi-site organization. Most employees of the organization were office

workers and a small number worked in trades to oversee the physical operations of the

facilities throughout the province. The goal of the program was to assist employees to

stay at work, return to work, or adjust to permanent disabilities and not being able to

work after they had sustained injuries or illnesses. The one condition for voluntary

participation in this program was having had a disabling injury or illness, physical or

mental. Clients may have had insurance claims, such as long term disability insurance

claims, or be off work with paid sick days from the employer. The program did not assist

with labour relations issues.

A joint steering committee comprised of equal numbers of senior union and management representatives oversaw the program, including determining policy and guiding procedures. The joint committee selected program coordinators from among union or management applicants that applied for the jobs posted within the organization. The program employed four full time coordinators. Participants 3 and 4 (L3 and L4) were union members working as program coordinators, and served clients who were union employees. Participant 5 (L5) was the labour site program manager and also a program coordinator who served clients who were management employees.

The main responsibilities of the program coordinators were development and implementation of return to work plans for clients, and acting as liaison among all other parties. Their primary communications were with the client (i.e. union or management employee accessing the program) and the manager of the client's department. The program coordinator, client and manager worked toward successful reintegration of the client back into the workplace. Throughout this process there were open and iterative communications among stakeholders and ongoing feedback that contributed to modifications of the return to work plan as needed.

L5, the labour site program manager was a consultant to the other program coordinators in addition to being the program coordinator of return to work services for management clients.

> I am responsible for the day to day guidance [and] management of the coordinators and if there were any issues regarding individual files that were complex or perplexing, that [they] needed to chat about, maybe the direction that a particular file should go, then they would come to me for that. (L5)

Throughout development and implementation of return to work plans the program coordinator met regularly with clients to discuss the plan, as well as with the client's manager to negotiate appropriate adjustments to the plan if any physical, psychological or social obstacles to return to work were identified. The team consultation was to ensure dialogue, collaborative participation and support, working together to identify and manage any unforeseen issues toward a successful and durable return to work.

Figure 3 illustrates communication among program stakeholders at the Labour site as described to me by the Labour participants of this study. In developing this diagram I forwarded a first draft to L3 and requested feedback, and his recommended modifications were incorporated.

4.2.2 Evaluation practices within the labour site.

4.2.2.1 Value of evaluation. Program coordinators were aware that satisfaction of both the client and the manager of the client's department were necessary to successfully integrate the client back into the workplace. Throughout provision of services program coordinators ensured there was open dialogue among the three stakeholders to make any necessary modifications to the return to work plan based on new information. To formalize evaluation and make long term improvements to the program, the steering committee administered a *Client Satisfaction Survey* and a *Manager Joint Return to Work Program Survey* to every client and manager at the end of services.

Figure 3 Communication Among Program Stakeholders as Described by the Labour

Participants
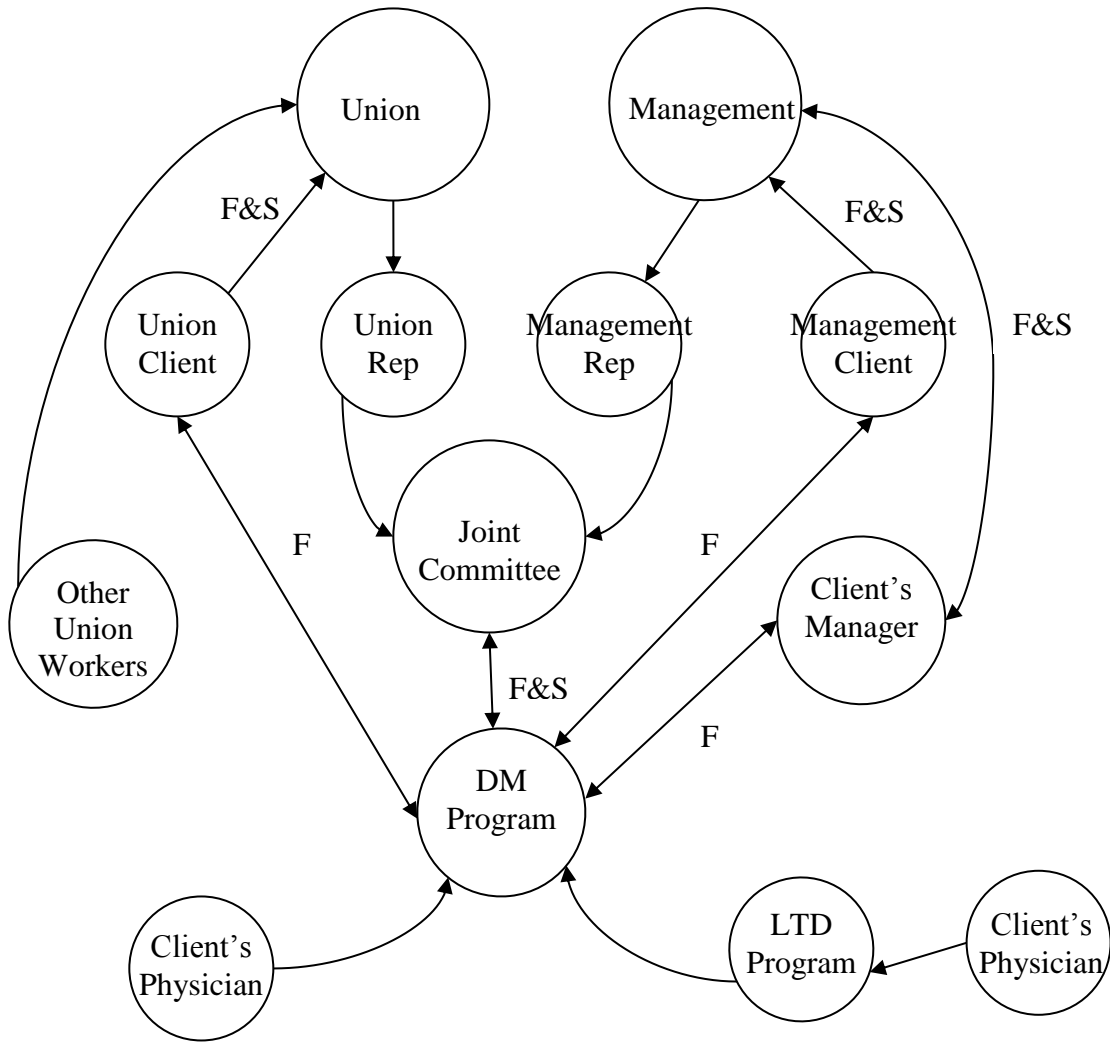
F=Formative ongoing evaluation; S=Summative evaluation feedback checklist;

4.2.2.2 Goals of evaluation. Services at this site were to accommodate each employee who was able to return to work after injury or illness, with suitable and meaningful employment, and doing so with management and union collaboration. The program relied on the program coordinators' informal reports on the outcomes of individual cases, and on completion of *Client Satisfaction Surveys* and *Manager Joint Return to Work Program Surveys* to assess the success of the program.

> By way of formal measure we don't have anything in place. But certainly the successes that each of the coordinators experience when working with the individuals and getting them back in the workplace. (L5)

L5 suggested meaningful criteria for evaluation should include return to work success rates and cost savings (long term disability or sick day costs) that resulted from interventions of the program. The organization was developing a data management system to measure program outcomes such as early initial contact by the program versus late contact and return to work outcomes for those who had accessed the program versus those who had not. L3 believed the most important evaluation criteria were whether the program was getting people back to work and the organization's responsiveness to employees' needs, especially flexibility in work accommodations for employees.

4.2.2.3 Stakeholder involvement. Stakeholders at the labour site included the joint committee, program coordinators, clients, clients' managers, union and management employees, and other professionals treating clients.

During initial provision of services the labour program coordinator acted as liaison among all other participants. The program coordinator, the client and the client's manager collaborated on development of a suitable return to work plan, and over time

dialogue continued among this team for feedback and to make modifications to improve the plan as necessary.

After completion of services the program coordinators administered satisfaction surveys to each client and manager. Clients submitted their completed surveys to designated union administrators and management submitted their completed surveys to designated managers of the organization. Completed surveys were then forwarded to the program steering committee. According to L3 the joint steering committee gave program coordinators feedback based on a review of the survey results "…if anything sort of jumps out…any trends."

After consulting with the joint committee to ensure that both union and management views on the most useful data were considered, a computerized data base was being customized by the human resources department, to enable future data analysis for evaluation of the program.

4.2.2.4 Evaluation data collected. Two years prior to this study, steering committee members created the *Client Satisfaction Survey* and *Manager Joint Return to Work Program Survey,* which are the primary formal evaluative data collected. The program coordinators participated in creation of the surveys.

Once client services were complete each client was administered a two page summative *Client Satisfaction Survey* to determine the client's level of satisfaction with services. The steering committee was the primary audience for the survey results, which were used for continuous improvement of the program. The survey included 12 items: four items addressed the referral process, four addressed return to work planning and implementation, and four addressed the client's overall satisfaction with the program.

Items' responses included: yes/no, choosing among listed alternatives, Likert scales from 1 to 4 (not acceptable, needs improvement, good or excellent), and narrative comments. Each client was also offered an opportunity to discuss concerns they may have had regarding the program with a member of the joint return to work steering committee.

At the end of services the manager of each client was administered a four page *Joint Return to Work Program Survey*, comprised of 26 items: eight items addressing the referral process, six addressing return to work planning and implementation, six addressing the manager's role and responsibilities in the return to work process, and six addressing the manager's overall satisfaction with the program. Items responses included: yes/no, choosing among listed alternatives, Likert scales from 1 to 4 (not acceptable, needs improvement, good or excellent), and narrative comments. Both the *Client Satisfaction Survey* and the *Joint Return to Work Survey* once completed went to the steering committee for analysis of results, although the identities of survey respondents were not disclosed.

L5 described data in the new data management system that would be used to evaluate the labour program's outcomes and milestones. These included: administrative data (listing of all referrals to the labour site program and assignments to coordinators), referrals (information on individual cases and assignments to coordinators), return to work files (historical and demographic information on clients, ongoing memos on the current case, and case closure information), and records of accommodations made during the labour site service (case outcome information, program milestones, permanent and temporary accommodations). "We are now using the new data base to manage the statistics and trying to look at the first day that [clients] were expected to return, and did

we actually meet that timeline" (L5). The data system would also be used to analyze cost savings through reduction in disability benefits and sick days, as well as return to work statistics of program clients compared to staff who did not access the program.

4.2.2.5 Use and reporting of evaluation data. The steering committee analyzed the client and manager surveys and used the findings to direct and improve the program and provide feedback to the coordinators. Program coordinators regularly attended steering committee meetings to ensure ongoing communications between those who operated the program and those who oversaw it.

The program coordinators relied on this evaluative feedback to understand what was most important to clients and managers. L3 and L5 both learned how important it was that managers be "kept in the loop" and informed about their workers who were not working in order to facilitate staffing of their departments. L3 explained when clients are returning to work, "maintaining communications is hugely important for the managers so that everyone knows exactly what the plan is and where things are progressing".

L4 also found feedback from the survey results to be frustrating. She felt complaints were sometimes misdirected toward the program when perceived problems did not originate within the boundaries of the program. For example, on occasion managers wanted to have more control and complained that a program coordinator should have contacted a client's physician because the manager did not agree with the functional limitations the physician provided. However, program coordinators saw this as a medical decision and not disputable by them.

L3 described how a most useful source of evaluative information for the labour site program was the ongoing formative feedback that drove decision making and

modifications to the plan as needed, throughout implementation of each individual return to work plan. L3 advised this was achieved by labour site coordinators acting as a liaisons among all participants whenever needed, and commitment by the program coordinator, client and client's manager to collaborate.

The labour site was the only site of the four participating in this study that conducted summative evaluation on the criterion return to work. This site also used the client and manager satisfaction surveys and ongoing dialogue among the three stakeholders, program coordinator, client and client's manager, for formative evaluation.

4.2.2.6 Summary of evaluation practices. This program emphasized the balance between union and management perspectives manifest in the joint steering committee. All program staff deferred to the joint committee for decision making, and all evaluation was sanctioned by them. Table 10 summarizes evaluation practices at the labour site in terms of three essential components, (1) evaluation criteria, (2) data sources, and (3) use of findings.

A data maintenance system created by the organization's human resource development department in consultation with the joint committee was being tailored for summative evaluation of how well the labour site was meeting goals. In the future, the expectation was that systematically collected data would be extracted to evaluate program outcomes and impacts.

4.2.3 Evaluation practices and organizational learning. L4 confirmed how learning was valued at the labour site, "Learning is very promoted, and the opportunity to learn…is promoted" and that the labour program had grown due to the organization's commitment to learning,

Table 10   Evaluation Activities at the Labour Site

Criteria	Data Sources	Program's Use of Findings
Return to work plan success	Iterative communications among program coordinator, client and client's manager	Modified return to work plan
*Client Satisfaction	*Client Satisfaction Survey*	Modified services
Client Satisfaction	Joint Committee exit interviews with clients	Modified services
*Manager Satisfaction	*Manager Summary Questionnaire*	Modified services
Program Functioning	Joint Committee feedback to program	Modified services
Program Outcomes	PeopleSoft data base	Joint Committee and organization analyzed findings

Note: *= Satisfaction surveys administered to all clients and client's managers

> I believe we are the organization that others will come to learn from. I think we
> have only become that because we have learned from others. You know we are
> taking the good bits and creating something really great. (L5)

"Feedback prompts some process changes and adjustments in the way we do things.
Through learning we think, okay, we need to make a change in the program. So that is
how we use learning" (L5). All three participants from this site valued feedback from the

steering committee and the client and manager surveys so they could collaboratively plan and implement processes that met stakeholders' needs.

L5 advised the organization encouraged ongoing education and learning for all staff. L4 and L5 had been sponsored by the organization to complete a three year part-time on-line disability management program offered through the labour based organization NIDMAR that familiarized them with the basics of disability management.

4.2.3.1 Summary of ROLE scores at the Labour Site. Overall, ROLE responses of the three labour site participants confirmed information provided in their interviews. All three described how the labour program achieved balanced collaboration between labour and management, open communications, and had integrated processes to learn from evaluative feedback. ROLE scores of L4, who referred to the organization, were consistently lower than scores of both L3 and L5, who referred to the labour site program. ROLE scores patterns of all three participants taken together suggest the labour site program was more attuned to readiness for learning and evaluation practices that support learning, than was the organization as a whole. This is a reasonable assumption given the efforts the labour site program had made to develop and maintain evaluation that was inclusive and used for learning. While the organization as a whole was in the process of developing data management systems to eventually be used in evaluation (including program outcomes), the labour site program had already integrated evaluation into the program and used findings to make service improvements.

4.2.4 Organizational cultural context. The labour site coordinators described adapting to both union-management and mental illness cultural perspectives, and that each perspective had the potential to impact services to clients.

4.2.4.1 Union management culture. L3 advised, "labour relations speak to the program and ways in which coordinators have learned from experience to act as conduits".

> We are a unionized environment and so if somebody is off sick, particularly if it is a mental health sort of thing, if a manager phones the client just to see how they are doing and how things are going, there is some concern. Sometimes the employee can be very sensitive about that and can take it the wrong way. They can look at it and think my manager is harassing me and doesn't understand I am not ready to come back, even when the manager's contact is well intended.

To minimize this possibility the labour site program coordinators maintained regular contact with clients when they are off work, and proactively informed managers about clients' progress. Sensitivity to this union-management culture was one of the things L3 was guided by.

The union environment culture of the program was apparent when clients argued for their rights within the collective agreement, which in turn conflicted with progress being planned for a graduated return to work schedule. L4 described how a union management relationship could create divisiveness, and found it "unfortunate" that management employees worked only with the manager program coordinator and union employees worked with union program coordinators, "I think it is [unfortunate] because it perpetuates the me and them. And I am supposed to be everybody's return to work coordinator, right?"

There was no negotiation of the union management structure and evaluation was imbedded in the dual perspectives. The satisfaction surveys were an example of assessing

the program from these two different perspectives and the joint committee took on the task of analyzing results from a third perspective, the whole program.

4.2.4.2 Stigma of mental illness. Mental illness was an impairment within disability management that program coordinators recognized as sensitive and stigmatizing. L5 explained that compared to physical illnesses there were taboos in the workplace surrounding talking about mental illnesses.

> Mental health related issues are I think hush hush. I think we have made some ground generally…as a society to be more accepting, and it is okay, it is just another disability. But I think we have quite a far way to go.

L5 explained how working with clients with mental illness was further complicated when co-workers were ignorant of the functional limitations mental illnesses could create for a worker being reintegrated into the workplace.

> I just think over time that existing staff become a little frustrated especially if it is maybe a mental health related condition. Where the person looks fine but only has to do half the work that [co-workers] have to do. And it gets back to the whole mental health and the stigma and just how accepting or not accepting organizations are.

Establishing a return to work date was more ambiguous for mental illness than physical illness, making it "hard from an operational perspective." (L5) "The whole purpose of duty to accommodate is so people cannot be discriminated against, because of their disability, religion or anything else." (L3).

4.2.5 Evidence evaluation was grounded in a labour paradigm. At the labour site disability was not treated as a medical disease. Program coordinators were provided

with the clients' functional limitations and residual abilities, and then worked to identify employment that matched these. The program focused on integrating workers within the work setting so that impairments were not necessarily barriers to work and the evaluation criteria matched that of the labour paradigm for return to work.

The labour paradigm in general includes disability management that values a team approach, involving collaboration among the worker, employer, health care team, worker's treating physician and union (Loisel and Durand, 2002). Evaluation at the labour site of this study was grounded in this paradigm as it addressed evaluation criteria identified by Loisel and Durand (2002) (return to work, time off work, financial costs, quality of life, learning in the workplace, ergonomics, and multidisciplinary at-work interventions), and did so collaboratively with all parties, including equal representation of union and management. L3, L4 and L5 valued returning clients to suitable and durable employment and recognized the importance of evaluation to address how successful they were at achieving this goal. Return to work was the main objective of this program and a new data management system was being customized by the organization to evaluate the program's return to work outcomes and milestones.

4.2.6 Discussion. The joint steering committee represented equally by union and management oversaw all program policy and procedures, including development of evaluation. There were two main forms of evaluation at the labour site. During development and implementation of return to work plans program coordinators collaborated regularly with individual clients and their managers for analysis of the unique combinations of client and contextual factors at play, and for formative evaluative feedback to guide the process. After completion of services, program coordinators

administered *Client Satisfaction Surveys* to every client, and *Manager Joint Return to Work Program Surveys* to every client's manager. Clients were also offered an opportunity to meet with a member of the steering committee to voice any concerns. Completed surveys were forwarded to the steering committee for analysis. If anything "jumped out" (L3) from the survey results, the steering committee provided feedback to the program coordinators who relied on feedback from the steering committee to understand what was most important to stakeholders and to guide them in meeting program goals. Coordinators agreed the organization valued learning and was developing new sources of data management for evaluation.

A main conceptual theme at the labour site was the value placed on open communications and collaboration among participants involved in each case. Feedback was an opportunity to learn about individual cases and program procedures in general. Coordinators acted as liaisons among all stakeholders, ensuring input from multiple perspectives, including the steering committee, clients, clients' managers and treating professionals. Equal representation of union and management interests created a system that operated with respect, where the steering committee was available to mentor during resolution of conflicts, freeing stakeholders to present their own perspectives without fear of opposition.

The labour program was in the workplace, which offers ideal opportunities for multidisciplinary interventions to resolve multi-factorial problems that can arise relating to the individual (medical, psychological, affective and social) and the individual's interaction with contextual factors (Loisel et al., 2001). Evaluation within a system at the

workplace had the potential to proactively address specific individual and contextual situations that arose, contributing constructive feedback attending to that diversity.

Barriers for the program included rigidity from strict collective agreement job descriptions that did not allow for modifications of duties to meet clients' needs. Misunderstandings arose among clients' coworkers when clients were perceived to be getting special treatment with modified job duties, but due to confidentiality rules, clarification was not possible and clients' social success sometimes became threatened.

4.3 Biopsychosocial Paradigm Site

4.3.1 Context and framework of this program. The BPsy site provided vocational rehabilitation services within a large government funded multidisciplinary health services organization that offered inpatient, outpatient, and clinical support services. The organization operated four programs for clients: brain injuries, spinal cord injuries, arthritis and/or neuromusculoskeletal conditions, and a division for young adults. Any allied health professional could refer clients, or they may be self-referred. Clients could be any age except children or adults over 60.

One of the unique characteristics of this organization was its multidisciplinary approach to service provision. The multidisciplinary services drew from psychology, social work, pastoral services, occupational therapy, physiotherapy, nursing, speech language pathology, physiatry, orthotics, dietetics, music therapy and art therapy. Services were also included in the areas of sexual health, drug and alcohol treatment, specialized surgical teams, assistive technology, recreational activities (from playing cards to sky diving), peer mentoring and spinal cord and brain injury education.

As well as being multidisciplinary, this organization was interdisciplinary. At the same time that clients received vocational services of the BPsy program, they also may have been receiving services from any number of the other departments within the organization. Practitioners from the various programs were expected to engage in dynamic communications about services provided to individual clients. This communications among departments was evidence of a learning culture, described by Preskill (1994) as, "a culture that encourages employees to engage in reflection and dialogue believing that individual learning leads to organizational learning" (p. 296). The BPsy site employed two full time and one part time counsellor to provide vocational rehabilitation services. BPsy6 was one of the counsellors and the team leader. BPsy7 had formerly been a counsellor at this site, having worked there for a couple years several years previously. The BPsy model highlighted substantial communications among program service providers described by BPsy7 as a process of cross-pollination. Figure 3 illustrates communications among the interdisciplinary programs available to clients at the BPsy site as described to me by the BPsy participants of this study. In developing this diagram I forwarded a first draft to BPsy7 and requested feedback, and his recommended modifications were incorporated.

4.3.2 Evaluation practices within the biopsychosocial site.

4.3.2.1 Value of evaluation. The organization on the whole lacked formal outcome measures.

We don't have satisfaction surveys, we don't have follow up, as to outcome. If we have made a specific recommendation we don't necessarily know if that took place. It is an area that we are lacking. We make follow up phone calls [to find

out whether] the person did get connected to where we thought they were going to be…we do have sometimes a review where we will phone somebody back…a couple of months later but that's not done too consistently….or the person will phone us back….We just sort of try working with the client up to a certain point and when they have found work or connected to another agency we close the file. But we don't have a long term understanding of what happened. (BPsy 6)

Evaluation at the BPsy program was predominantly informal and formative. BPsy6 had been trying to develop a system to maintain thorough records of cases the program handled and program outcomes. This effort was initiated in part due to her perception that potential funding cuts in the organization might immediately impact the vocational rehabilitation program. (This concern seemed warranted since BPsy7 reported there had been 6 counsellors working in this program in 2004 and BPsy6 reported the number was down to 2.7 full time equivalency positions in 2010.)

The organization had a quality control committee, but according to BPsy6 "it has never had any impact or done anything noticeable." Recently, however, a patient services manager, also a member of the quality control committee, requested information on the BPsy data maintenance system, so BPsy6 had been developing a data record system for her department.

There is no organization wide system for collecting information about what we do, which the whole organization is aware is very poor. So they are actually sort of struggling to find some way to do that now. And there is a committee that just asked me what data we collect about patients, because they are trying to create some more general system. They are now going to all these groups like me who

Figure 4 Communication among Interdisciplinary Programs Available to Clients as

Described by the Biopsychosocial Participants
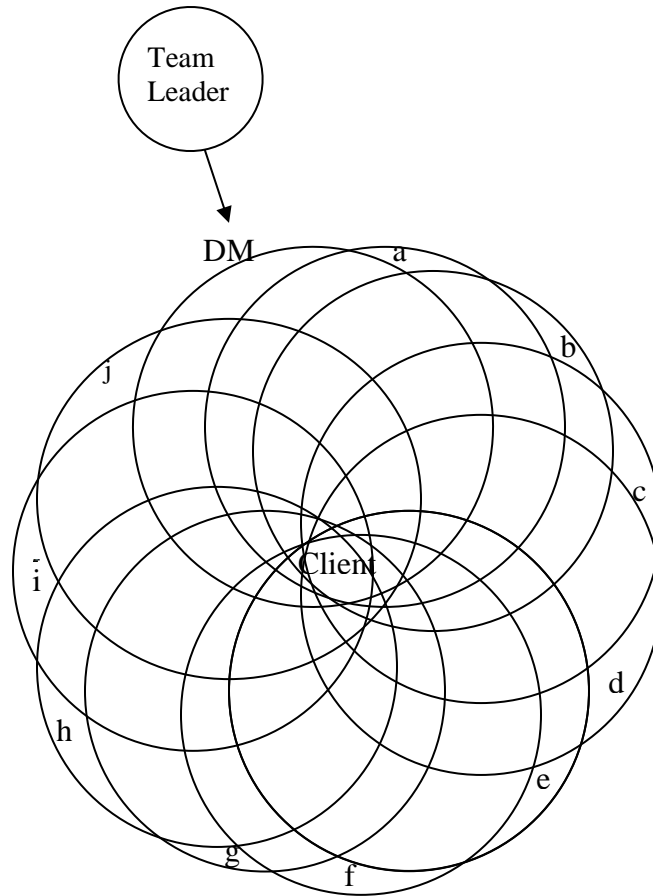
A sample of the services available to clients include: a=Psychological Counselling;
b=Social Work; c=Sexual Health; d=Drug & Alcohol Counselling; e=Pastoral
Services; f=Spinal Cord or Brain Injury Education; g=Occupational Therapy;
h=Physiotherapy; i=Nursing Care; j=Specialized Surgical Support.
(Others services include: Speech Language Pathology; Dietary Counselling;
Assistive Technology; Orthotics; Physiatrist; Recreation Therapy; Peer Mentoring;
Music Therapy; Art Therapy; Pet Therapy; and Adolescent Young Adult Program
Services.)

made up our own methods… to track what we are doing….So we have all these little idiosyncratic systems within different departments.

4.3.2.2 Goals of evaluation. The mandate of this program was to provide support, counselling, referrals, guidance, and career exploration during the earliest stages of acute conditions. BPsy7 felt that ideally the goals of evaluation should focus on early interventions and formative evaluation. BPsy7 suggested that timeliness of services was not as important a criterion as was continual availability of services. Inpatient clients at the beginning stages of medical treatment could be referred to the BPsy program for provision of support long before actual return to work was being considered. BPsy7 provided an example.

> We might go in three weeks after injury when they are still in their traction bed and dealing with all kinds of life issues and way before activities of daily living….when they are still wearing a metal halo and trying to figure out which part is paralyzed. We might go in early on to let them know there are services available, you are not alone, there is vocational assistance that will happen, here are some of the occupations that people do who have your [type of] injuries…it was supportive counselling. And it was often times provided as [the client] requested it.

In contrast, BPsy6 felt that given the limited time the program had to provide services, instead of early interventions the focus should have been on developing return to work plans and summative evaluation of whether those goals were met. Very early vocational counselling involvement may be supportive overall, but it was more important to wait until clients were ready to develop a realistic return to work plan.

…a huge amount of time [would be] spent talking to people…about something

that is so far down the road, [when] we find that many people's return to work

abilities or goals shift from the acute time to the time they [are ready for return to

work assistance].

BPsy6 explained how the organization established goals for each of the clients

and evaluated progress in meeting those goals.

…For each patient the chart has a section ….smart goals…written in language

that is very specific. Will walk fifteen feet, will know about benefits for people

with disabilities and be able to apply, or will be able to eat independently

….Whatever these goals are, is all listed in the [patient's] chart. And then they are

ticked off as achieved or not. So I would say that is evaluation….That is probably

what I would be going for, trying to get our goals a little more succinct at the

beginning of working with a client and then evaluating if the goals were achieved.

The program used a system called SMART goals that frame goals positively, "in terms of

something a person can learn to perform well" (Latham, 2009, p. 171). The acronym

SMART refers to goals that are *s*pecific, *m*easureable, *a*ttainable, *r*elevant, and have a

*t*imeframe.

4.3.2.3 Stakeholder involvement. Stakeholders at the BPsy site included staff

and volunteers from all of the multidisciplinary departments within the organization,

clients, and client's external treatment professionals, families and supporting friends.

Referrals to the BPsy program were triaged by the team leader, who selected

some clients to accept for services and referred other clients directly to other services.

Only some referrals were accepted due to high volumes that exceeded the program's

capacity to provide services. During the time client services were provided the team leader (BPsy6) recorded each client's involvement in the BPsy program into the data base she was creating. This data base was intended to eventually be used for evaluation of program functioning, although that was not yet being done. BPsy7 advised there had been no formal evaluation procedures in place regarding the BPsy program including no client satisfaction surveys.

BPsy7 reported counsellors in the program received performance feedback from the team leader and an organizational supervisor. The team leader was responsible for overseeing the work of the program counsellors and provided evaluative feedback related to the vocational rehabilitation profession on specific cases during mentoring sessions as needed. Each counsellor in the BPsy program had also been evaluated by an organizational supervisor every 6-12 months, on topics not specific to vocational rehabilitation but to more general skills related to the organization such as record keeping.

The assistant to the director of the organization advised that administration within the organization frequently changed, and most recently there had been only two levels of management, an operations director overseeing the entire organization and managers of the individual programs (including BPsy6). The assistant said that if performance reviews were still conducted they would now be administered by program managers rather than organizational supervisors.

4.3.2.4 Evaluation data collected. Although there was no formal program evaluation at the BPsy site, there were several data systems in place that could eventually be accessed as sources of data for evaluation.

The first was the team leader's data base that categorized the following information on the BPsy site's clients: file status, urgent or regular, name, program, VRC, phone number, referral date, date file opened, last chart update, date file closed, comments/outcome, address, referral source facility, referral source clinician, diagnosis, birth date, referral month, year, fiscal year, and wait time in days. BPsy6 noted this information was not coded, and therefore not used for evaluation.

The BPsy team leader also maintained records of the needs assessments conducted on all incoming client referrals, information she used to triage whether the referrals were accepted into the program, or were referred directly to other organizations that offered more appropriate vocational services.

BPsy advised a third data management system within the organization may eventually be accessed to evaluate the BPsy and other programs. The quality committee had recently collected data management systems that were in place within all the programs of the organization, and in time intended to integrate all the various systems into one standardized organization wide data maintenance system.

4.3.2.5 Use and reporting of evaluation data. Feedback provided to the BPsy program counsellors by the team leader was informal and focused on mentoring the counselors in their work; no records were maintained. BPsy7 described how mentoring was provided in relation to vocational rehabilitation,

> "…in the same way any boss would monitor an employee in terms of feedback about specific areas of practice, providing information, or …encouragement, or whatever you needed around how you were delivering services…."

BPsy7 reported feedback that had been provided by the organizational supervisor on performance involved administration of a generic checklist to all professionals in the organization, in a one on one interview format every 6-12 months, focusing on general organizational standards including ethical practices and documentation.

> I would also report to a supervisor in the nursing department who was a hospital supervisor, hospital administrator. Her training was nursing and she wouldn't gainsay me on vocational things in terms of the vocational services I was providing because that wasn't her expertise. But she would provide mentoring and sponsorship or supervision basic work practices. So things like entering logs, you know keeping my records up to date, you know general work performance evaluation. (BPsy7)

According to BPsy7 professionals from all departments within the organization maintained regular dialogue with each other and informally shared feedback about the provision of services. Feedback from clients, however, was anecdotal.

4.3.2.6 Summary of evaluation practices. Table 11 summarizes evaluation practices at the BPsy site in terms of three essential components (1) evaluation criteria, (2) data sources, and (3) use of findings.

BPsy7 reported that several years prior, the program had forfeited provincial government funding equivalent to the salary of one full time counsellor, because funding was contingent upon introduction of outcomes based evaluation, an approach the program disagreed with. BPsy7 indicated that program counselors did not believe evaluation of the program should be limited to "statistical outcome calculations." They believed focusing

Table 11   Evaluation Activities at the Biopsychosocial Site

Criteria	Data Sources	Use of Findings
Formative feedback during provision of services	Informal communications among BPsy counsellors and with those operating other services within the organization	Modified services
Quality of vocational rehabilitation counselling services	Meeting with team leader	Modified services
Quality of general services to organizational standards	Hospital supervisor administered questionnaire	Modified services
Client progress on smart goals	Smart goals documented in client chart	Planned client interventions
Client satisfaction	Communications with client	Modified services

on outcomes would cause the program to diverge from its intention to be continuously

available for vocational consultation to clients at any stage of recovery, from early on at

the intensive care stage to later stages when clients were ready to establish return to work

goals.

> The government…wanted to change to a performance based model and …have an
>
> outcome based performance evaluation... [with funding based on] outcomes,
>
> putting people through programs, and being paid a certain money for planning, a
>
> certain amount once they are in job training, paid once they are in job search, and
>
> paid when they actually find a job…. In good conscience the program …couldn't

ask a voc rehab to suddenly go start doing job development and place people. They were convinced that it would be such a different role for a voc rehab…. they actually agreed to lose a position worth of funding …because they weren't ready to tool up and add a new role.

4.3.3 Evaluation practice and organizational learning. The foundation of the BPsy site and the programs within this organization was the value placed on inter-disciplinary communications, continuous overlap of services, and collaboration on every individual client case. This approach was described as "holistic" by BPsy7. Each program offered a professional specialty that on its own would not have accomplished the goals of the organization, making it imperative that the disciplines worked collaboratively. Evaluation of this inter-program collaboration was informal, continuous, and formative, achieved through discussion among the multi-disciplinary professionals and with clients.

The BPsy program counsellors were supportive of each other, holding weekly team meetings to discuss cases and share information. Individual counsellors discussed their needs and were directed to resources, such as recommendations of books to read. According to BPsy7, counsellors from the BPsy site were expected to attend vocational rehabilitation association meetings and to maintain their professional designation. They were expected by the organization to conduct ethics presentations, and to contribute to the organization by hosting educational sessions and providing information and feedback to other professionals.

The BPsy program counsellors also contributed to clients' learning, "For example we had a rehab rap night where all of the spinal cord folks would invite the vocational rehabilitation counsellors to talk to the whole group, as opposed to talking to individuals"

(BPsy7). Counsellors in the BPsy program were expected to contribute in groups and "to provide collateral support."

> And there was a lot of learning and a lot of challenge because many times you would think oh I have an idea about this disability, and then you would be confronted about the reality of it. Be confronted with reality of somebody with completely different mentality and lived experience of whatever they were dealing with. And then you would have to try to integrate that into your understanding. (BPsy7)

BPsy7 described the organization's peer mentor program:

> A peer mentor, a spinal cord injured fellow who had the least education in the room, sometimes less education than the client, but he had the most to teach in terms of life experience with a spinal cord injury. He was an amazing fountain of knowledge and wisdom in terms of how to manage spinal cord injuries and how to live with the consequences of the changes. Because the consequences of his own life were so profound, and he had moved through them so well.

> Cross training between groups within the organization was common. BPsy7 described the communications among all of the disciplines that occurred as they worked together on individual cases as part of evaluation. He used the term "cross pollination" to describe this interdisciplinary nature of the organization. BPsy7 reported learning a great deal from these interactions and feedback he received, working together with other professional and peer workers. He described being immersed in a system where over time he learned to view disability through the eyes of peer mentors and lived experiences of disabilities.

One of the largest professional groups was the recreation department, where programmatic activities ranged from playing cards to sky diving. Activities such as skydiving originally amazed many professionals from departments other than the recreation department, but they quickly learned from positive firsthand accounts of the clients. BPsy7 described a client who had gone skydiving after just stabilizing from a catastrophic disability that resulted in his having to use a wheel chair, and sustained leg fractures while skydiving.

> He was actually ecstatic because he had had such a high and felt like he was alive again. On all these levels it had been a wonderful experience for him, and the fact he had broken his legs was an inconvenience he really didn't care much about.

It was this program that many clients seemed to value most of all for its potential to return them to the high risk behaviors that they had loved prior to their injuries.

Discharge meetings were another opportunity for professionals from the various disciplines (doctor, nurse, physio, peer mentor) to share their perspectives on the program.

    4.3.3.1 Summary of ROLE scores at the BPsy Site. Overall the ROLE scores for both BPsy participants confirmed information they provided during interviews that the organization lacked formal evaluation procedures, but the organization was starting to build evaluation capacity. Ratings on the ROLE category Communication of Information were low for both participants as there was currently no organizational data management system in place. However, ratings were high on the category Evaluation, as these participants were aware of the potential contributions evaluation could make to learning and to improving program services and outcomes. The organization was in the process of

developing a standardized data management system organization wide that could eventually be accessed for evaluation. The high ROLE scores for both participants on categories Culture, Leadership, Systems and Structures and Teams reflected the value placed on organization's multi-disciplinary inter-disciplinary systems.

4.3.4 Organizational cultural context. The BPsy site could be characterized by its "peer mentoring and lived experiences" culture. There were no stigmas related to disability within the organization, but there was recognition that outside of the organization their clients would likely experience stigmas, including "stigmas of invisible disabilities."

4.3.4.1 Peer mentoring and lived experiences. A perceived strength of the organization was that it employed many counsellors with physical impairments who brought to the services a genuineness that spoke to the clients more than professional knowledge could. Peer mentoring was perceived to be the most effective strategy within the organization.

> So here was a counsellor who had double Masters in psychology and education, but he was also in a wheel chair and he would not put his Masters degrees on the wall. And I said to him at one point, you really got to get your education up on the wall, you know I mean you worked so damn hard for it and he said, no that taught me the tools to do the job, but I think in my current situation it would alienate the people I am trying to work with. That they see me as the job guy who is also in a wheel chair, and the credibility I have comes from the wheel chair, not the Masters degrees. And so I took his lead and I took my degrees down, because I appreciated the fact that a 17 year old kid would feel safe hanging out with these

two guys that knew about jobs, one in a wheel chair one not, and neither of us

looked like big official guys because we didn't have our degrees on the wall.

(BPsy 7)

BPsy7 gave several examples of the extraordinary abilities of counsellors who had

themselves experienced catastrophic spinal cord injuries, in working with clients who had

recently become similarly injured. BPsy7 stressed that rapport with clients and among

multi disciplinary coworkers had the greatest worth within the system and was what he

believed to be the most critical criterion to be evaluated.

4.3.5 Evidence evaluation was grounded in a biopsychosocial paradigm. The

BPsy site used a multidisciplinary approach, where services were integrated as clients

accessed different programs at the same time. The organization intervened early on after

catastrophic injuries while clients were adjusting to serious life changes in relation to

suddenly acquired disabilities. Within that process the BPsy site program provided one-

on-one counselling to clients throughout their cognitive, physical and social adjustment,

and toward their eventual readiness to return to work if possible. BPsy7 believed the

highest worth of their program and other programs within the organization was the open

communications and rapport offered to clients.

This site reflected a biopsychosocial paradigm where services take an

interdisciplinary approach, and where impairment and disability are differentiated, so that

impairment alone is not a predictor of disability, as context makes a deciding

contribution. The BPsy paradigm in general conceptualizes disability as, "…an

interaction among biological, physical, behavioral/psychological factors, and social

phenomena (Schultz, et al., 2007, p. 339).

Evaluation at the BPsy program was not formally conducted. There were no client satisfaction surveys, no outcome based assessments, no data management systems in place that could be accessed for evaluation. However, there was a growing awareness of the potential value of developing data management systems that could be accessed for evaluation. BPsy6 had been trying to develop her own data management system for the program and the organization was progressing in developing a system of data management that would be consistent across all the programs of the organization.

BPsy program counsellors received evaluative feedback about their performance from two sources: feedback related to the vocational rehabilitation profession from the team leader and a performance review regarding work habits in general. There was no record keeping of the team leader's evaluation, and no mention of the performance review instrument being used beyond the feedback sessions when it was administered.

The BPsy program of this study served clients early on after catastrophic injuries. However, the BPsy paradigm in general can be applied throughout services from the onset of disability, beyond the early adjustment period, right through to the client's eventual return to work and resumption of life activities outside the hospital. BPsy6 favored focusing the program's services on preparing clients for employment nearer to the time they would return to work rather than prioritizing psychological adjustment, whereas BPsy7 favored services throughout psychological adjustment.

4.3.6 Discussion. There was no formal evaluation in place within the BPsy program or at the organization level. However, communication among stakeholders within the organization was highly developed, with short and long term foresight that benefitted clients as they adjusted to their disabilities.

Although the organization lacked formal evaluation methods, informal ongoing evaluative feedback occurred among stakeholders as they contributed to services. The BPsy program team leader provided professional feedback through mentoring sessions with counsellors on individual cases relating to vocational rehabilitation practices. As the organization was becoming increasingly aware of the potential benefits of evaluation, data management systems were being developed that could eventually be accessed for formal evaluation. Development of data management systems is consistent with one of the first steps of building evaluation capacity, where baseline analyses of processes and needs are undertaken (Taylor-Powell, 2008).

4.4 Insurance Paradigm Site

4.4.1 Context and framework of this program. The insurance paradigm site was an in-house investigation unit of a large public organization that managed disability claims within one of its divisions. The program conducted investigations into possible fraud or misrepresentation by clients who had sustained injuries or illnesses at work. The program also conducted investigations into possible fraud by external stakeholders including health care providers, vendors or employers, and internal investigations into possible fraud by employees of the organization. Participant 8 of this study (I8) conducted external investigations for the insurance site program. Participant 9 of this study (I9) was the manager of the insurance site program and conducted internal investigations. Sources of referrals originated internally from staff of the organization, or externally from any source including an anonymous tip line. The program had 18 investigators located throughout the province, with oversight by a union member supervisor and a nonunion manager (I9) located in the organization's main office.

Half of the investigators were long term employees of the program each with backgrounds that included extensive police work. I8 was among that group. He had worked for the program for 25 years, and was hired at a time when he and all other investigators in this program had prior employment experience of approximately 20 years working as municipal and RCMP police officers in Canada. I8 believed that former police experience was the most important qualification for an investigator.

The other half of the investigators were relatively new employees, with different backgrounds that involved conducting investigations (Canadian border services agency, private insurance companies, financial insurance, provincial regulatory bodies, and policing within other countries). One of the newest recruits was hired specifically to conduct internet investigations and data mining, including identification of clients through *Facebook* . I9, the program manager of the insurance site, was among this second group. His background was typical of the newer recruits, and included a BA in Criminology, a couple years experience as an auxiliary officer with the RCMP, and investigative experience with the provincial government's criminal injuries services.

As the program administrator I9 hired new recruits whose backgrounds mainly included university degrees. Proven ability to succeed at university was what he considered a main predictor of success on the job, rather than the former policing experience that had been the background of the more long term investigators. The perspective held by I9 regarding long term success on the job has been supported by research (Schmidt, 2009).

Figure 5 illustrates communication among stakeholders at the Insurance site as described to me by the Insurance participants of this study. In developing this diagram I

forwarded a draft to I8, requested feedback, and his recommended modifications were incorporated. Communication of information regarding disability management clients flowed both ways among parties except out to external referral sources, as due to freedom of information and privacy laws no information could be released to the public.

The investigators used multiple techniques including interviewing, reviewing documents, internet data mining, investigation of documentation authenticity, and contracted out for services including surveillance, videotaping and specialized investigations such as auditing by accountants. Cases resulting in prosecutions and court cases involved Crown Counsel.

4.4.2 Evaluation practices within the insurance site.

4.4.2.1 Value of evaluation. When asked the main service of the insurance site program that should be evaluated, I9 advised,

> We protect the integrity of the accident fund and just make sure that those people who should be having access to the funds have access to them, and those that are legitimately owed anything don't get anything beyond what they are entitled to. So if they are entitled to a particular benefit then great. If they are not or they are somehow attempting to fraud the [agency] in some way that is our job to protect the fund.

I8 explained that most of their work led to clarification of miscommunication and validation that the person being investigated had been truthful.

> The vast majority of material we receive to investigate is such that we do not end up affecting the outcome of that particular file to any great degree. And we establish that the person really is being truthful in which case there is no problem.

Figure 5 Communication among Program Stakeholders as Described by the

Insurance Participants
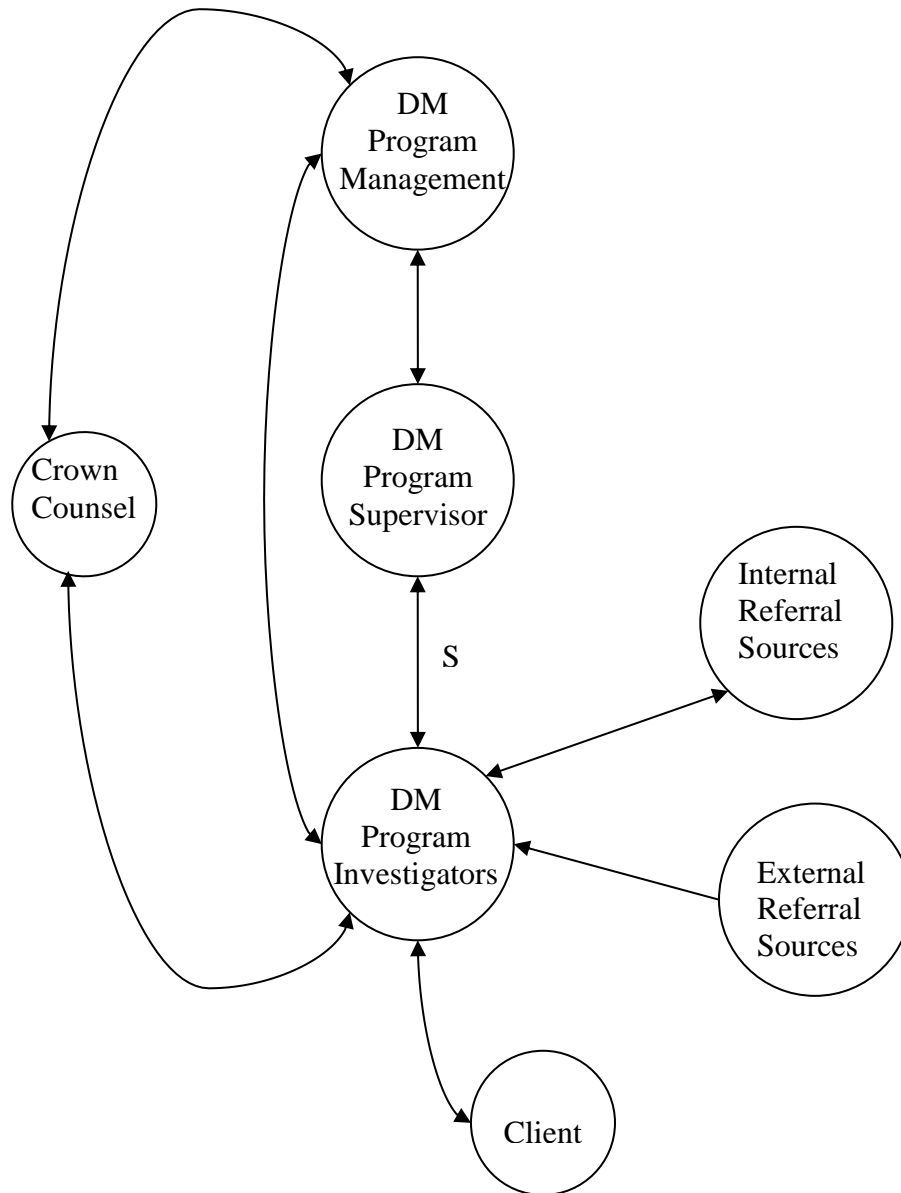
S= Summative evaluation; Internal referral sources included disability claim
managers or vocational rehabilitation consultants; External referral sources
could include anyone for example a client's employer or neighbour; Crown
Counsel was involved only on occasion when required to prosecute for
criminal purposes.

And if there is a question of misunderstanding we will clarify that…. If there are

cases of fraudulent activity then again I will deal with that and depending upon

the size of the fraud that can go anywhere from a recovery of an overpayment,

termination of benefits, to an actual criminal charge…. We tend to find that a lot

of this is really a very small problem which is primarily one of misunderstanding.

The goal of this program was to approach each investigation in a balanced manner,

oftentimes providing information that clarified misunderstandings and bias on the parts of

internal or external referral sources. I8 explained regarding investigations,

A lot of times when dealing with someone you create a bias….And so when we

get stumped you have to kind of back off and take a look at it through a different

set of eyes. And determine whether or not the facts you have are correct or

whether or not they are slanted as a result of a bias by the person who is

submitting the referral.

Their investigations could result in settling disputes and overcoming biases, and these

outcomes were highly valued by the insurance program.

4.4.2.2 Goals of evaluation. A main goal of the insurance site program was

protection of the accident fund, so evaluation of the program emphasized the extent to

which this occurred. The program maintained records of referrals and outcomes of cases

to estimate cost savings that resulted from their services, based on a formula created by

the organization's accounting department.

To evaluate quality of services I9 advised two criteria were examined: time limits

for completing each type of service, and confirmation that services have been completed

to an acceptable standard based on the professional opinion of the program supervisor. In

the opinion of I8 the most important issue for evaluation was the extent to which services

delivered by the insurance program had been sufficient to make an impact that

contributed to having support be allocated appropriately, "are the people [insurance

clients] getting the support they need, be it financial, be it psychological, or physical".

      4.4.2.3 Stakeholder involvement. Stakeholders at the insurance site included

program investigators, supervisor and manager, clients, internal and external referral

sources, other departments of the organization, external service providers, and crown

counsel.

      Incoming referrals were assigned to an investigator who worked independently

and within established timelines for the particular service. Upon completion of services

investigators were not permitted to close their own cases. When services were completed

investigators emailed the insurance site supervisor who reviewed each file for quality

control. Once the supervisor confirmed that requested services had been completed to

quality standards, the case was closed, and, if the referring source was internal,

investigative information was forwarded to the referring source.

      The supervisor maintained a data base on all cases handled by the program. Data

included identifying information on each case handled by the program, the number of

files handled, projected amount of money saved, and details on criminal prosecutions

(based on the actual numbers that were successful in laying charges and where the person

was found guilty). Organizational accountants accessed data on cases handled by the

program to calculate estimated cost savings.

      It was the understanding of I8 that the Director and Vice-President of the division

overseeing the insurance site should ultimately be responsible for evaluation of the

program, however, these higher levels of management had not been involved in any evaluation.

   4.4.2.4 Evaluation data collected. Two criteria used in evaluation of services were: timeframes for completion of services and quality of completed reports acceptable to the program supervisor. All of the referrals from internal sources had due dates (standards) attached to them. For example, a request to conduct an interview to obtain or clarify information was to be completed and the report submitted within seven days. A request to conduct surveillance had an initial due date of 30 days, which could be extended to 60 or 90 days.

   Cost savings were estimated by the accounting department of the organization, using a formula they created to compare actual costs on client claims where the Insurance site services had been accessed, to estimate costs had these services not been accessed. Actual costs were based information from a data base the program maintained on their referrals and outcomes. Estimated costs without services were based on costs for similar cases not accessing the insurance site services.

   Cases of fraud could result in recovery of benefits and/or termination of further benefits, and could lead to cost savings. I8 explained how cost savings were evaluated.

   We have statistics showing the number of criminal charges. We show the

   projected rate of savings…based upon a formula [the] audit section has created

   whereby if a file is altered or terminated using this particular computation they

   will come up with a projected dollar savings. So in the course of the year this unit

   might save 5 or 6 million dollars in projected savings. That is not necessarily

   dollars brought into the kitty so to speak, but these are just projected based upon

the fact that something occurred which caused that file to be either terminated or altered and the amount of money expended on it reduced accordingly.

The insurance site program administered a *Client Satisfaction Survey* to all staff within the organization's claim division (past and potential internal referral sources). This survey was not administered routinely, only once every several years. The purpose of the survey was to improve customer service and the variety of services provided by the program. The survey included eleven items: two on the respondent's demographics, eight Likert scale questions (knowledge of the insurance site program, number of past referrals, accessible and helpfulness of the investigator, timeliness of services, information provided, respect and courtesy of the investigator, client satisfaction, and overall rating of the service provided by the program), and one open-ended question on what the program could do to improve the services it delivered.

4.4.2.5 Use and reporting of evaluation data. I8 felt strongly that the organization was overlooking its fiduciary responsibility to publicize accomplishments of the insurance site and related cost savings that had resulted from their work, as an achievement in successful management of the accident fund.

Quite honestly I believe the [organization] has a veneer but underneath that veneer there is little in the way of substance….The organization will profess they have a fraud strategy. But they seldom will, upon the successful conclusion of the prosecution, ever present that as public knowledge, or even knowledge to the employees….The [organization] has a magazine that they publish. In that magazine there will be this little insert which says that fraud is everyone's business, which is real nice. And at the back they will show the penalties levied

against employers [for unsafe work practices]. But when we [the insurance site program] prosecute someone and get a conviction you will never see that mentioned.

When asked what positive impact he believed publishing results could have, I8 explained,

I suppose that the [organization] has a fiduciary responsibility to the employers and to the employees to make certain that any criminal act against the fund is actually pursued. Well, we do but they don't show it. So I think deterrence is a big factor….Deterrence has a value because a number of the employees know that if they send concerns forward to be investigated they in fact are. And that things are taken seriously and we will pursue them. Right now they don't know that. They see nothing in the way of information coming from head office that head office takes fraud seriously….The employers see nothing in the way of results….They are told yes we do this and all the rest of it, but they never see anything in the way of concrete results coming forth and in the form of prosecutions and stuff like that. The general public same thing. They seem to think that you know the [organization] is a gravy train and no one is ever concerned about fraud….But the thing is we have to be judged that the organization approaches these investigations fairly. And by showing that the courts have acted upon these and then determined that the information provided by the [organization] and of course its investigation was accurate and fair and that they found that yes in fact the [organization] is being correct in pursuing these prosecutions because there are people out there that are taking monies from the fund.

I8 contended that the organization's magazine did not recognize publicly the "investigative work of the insurance site program, and cost savings that result and serve as deterrence to committing fraud," and indeed a review of the September-October 2011 edition of the magazine confirms this. The content of the magazine explained the value the organization places on publication of penalties that result from investigations they conducted of employers with unsafe work practices. In the latter case, penalties were presumed to contribute positively to the organization's safety objectives as employers could learn vicariously and be motivated to comply with safety regulations when they observed other employers being penalized for not doing so.

I8 believed publication of outcomes regarding the insurance site services was the most critical evaluation criterion of the program's worth, and the goal of evaluation should be to share information to further learning among stakeholders.

4.4.2.6 Summary of evaluation practices. Table 12 summarizes evaluation practiced at the insurance site.

The supervisor was available to the investigators for consultation and feedback, and at the end of services the supervisor determined whether time lines have been met, and whether reports had been completed to an acceptable standard. The supervisor maintained records of all referrals and case outcomes, and the accounting department had developed a formula to estimate cost savings resulting from successful investigative work. The insurance site also administered a satisfaction survey to all possible internal referral sources every several years.

Table 12   Evaluation Activities at the Insurance Site

Criteria	Data Sources	Use of Findings
Report turnaround time	Program data base	Expedited submitting of reports to meet standards
Service completion	Supervisor review of report	Modified services
Quality of final report	Supervisor review of report	Modified report
Cost savings	Program data base	Justified organizational funding of program
Referral sources satisfaction	Prospective internal referral sources *Satisfaction Survey*	Modified services

4.4.3 Evaluation practice and organizational learning. According to I8 and I9

there was minimal collaboration among insurance site investigators. Most collaborative

learning was between individual investigators and the program supervisor. I8 advised that

previously, when all investigators had come from Canadian policing backgrounds there

had been ongoing collaboration and consultation among them. In his opinion the newer

recruits did not have the required skill sets or understanding to conduct investigations,

whereas investigators with policing backgrounds did.

> The majority of the senior investigators are quite able to pick up the phone and
>
> contact people. What you have here however within this group is a percentage of
>
> senior level investigators who know what is going on, how to do things…and
>
> which individuals to contact that have the information, the intelligence, or the

knowledge to assist them with their problems. Some of the newer people don't

have that background. They don't know the senior officers, and are blissfully

ignorant. (I8)

Opinions of I8 and I9 differed regarding the organization's commitment to formal

training and appreciation for learning. I8 believed that funding for substantial training of

a high quality was poorly supported, and that management funded only inexpensive short

courses.

Unfortunately so much of that is budget driven. As a result I see little to no

concrete support in that area. We will receive training but usually in areas that

really don't impact all that much on what we do. Because the program training we

get is based upon cost, so therefore you get what you pay for and we don't get

much… very very poor support in that area. (I8)

He explained the complexity of skills for the investigation of fraud.

Fraud incorporates criminal law, commerce, accounting, auditing, you are dealing

with criminology and all the rest of that. I have my professional designation as

most of us do as a certified fraud examiner. And that covers those spectrums. But

the thing is once you have taken the course, which is roughly a year's worth of

studying, and then you take your exams, the support more or less falls off and you

are kind of left to hunt and seek and get what you can and then request coverage

as far as payment is concerned and there is very little of that there.

He gave an example of a worthwhile forensic course that would have been available to all

investigators of the insurance site, offered over 4 days of training, plus access to on line

training, books and materials, for a cost of $10,000. However, the organization would not

approve the funding. Instead, the organization expected investigators to find shorter, less expensive courses. I8 also believed that courses the newer investigators pursued such as interviewing skills, were elementary. In order for courses to qualify for maintenance for the Certified Fraud Examiner designation they did not have to be long, but they had to be relevant, such as social reporting for criminal investigations and intelligence.

In contrast, I9 believed the financial support given for training was sufficient, and which as manager of the program, he approved these for all investigators. Training he expected entry level investigators to have was either completion of the Investigations and Enforcement certificate program offered through the Justice Institute, or a bachelor degree, demonstrating they have, "…gone through that education process and have that mind set and think in a way that is very different from other people who have not done that."

I9 believed some of the new investigators were very curious by nature and wanted to learn new things and new ideas, whereas the older investigators nearer to retirement lacked the same interest in learning. Both I8 and I9 advised the newer recruits were interested in courses on basic skills sets of the job, "interviewing, information gathering, using social media as an investigative tool, different computer systems" (I9). I9 said he supported these types of shorter courses and workshops that would qualify as maintenance education to meet the required 20 hours per year of developmental training to maintain an investigator designation. As manager, I9 hosted internal departmental meetings twice per year that included sponsorship for continuing education training modules.

A number of our investigators are certified as fraud examiners and to keep your

certification you have to have 20 hours of developmental training per year. So

when we have a conference call or when we have our meetings I make sure that

I bring somebody in that will satisfy that need, including an ethics portion of

training. We do a lot of ethics training internally to satisfy that.

As far as learning at the level of the entire organization, I9 perceived a significant

lack of cooperation among departments. Departments worked in silos and managements'

attempts to break the silos down had failed. He speculated that fraud likely existed in

some departments, but the department managers had not wanted the insurance program to

investigate. He believed that the new vice-president overseeing the insurance site might

become proactive and create opportunities for investigations. I9 saw this as a lack of

support overall for organizational learning.

  4.4.3.1 Summary of ROLE scores at the insurance site. Overall ROLE scores

of the two insurance participants reflected their discrepant views of the program. While

I8 believed the program needed to improve on all areas of the ROLE related to learning

and learning from evaluation, I9 believed the program was doing well.

  4.4.4 Organizational cultural context. The insurance site was aware of biases

within an "insurance claim culture" that existed for many of the referrals to their

program, where investigation into disability insurance claimants was requested.

  4.4.4.1 Insurance system culture. The insurance site was aware that referrals

made regarding disability insurance claimants were often predicated on biases that claims

adjusters held about claimants they suspected to be dishonest, or on the adjuster's lack of

information. The investigators strived to maintain a neutral and respectful attitude toward

the clients in these cases, and strived for objective investigations that achieved clarification.

> The vast majority of material we receive to investigate … we establish that the person really is being truthful in which case there is no problem. And if there is a question of misunderstanding we will clarify that and give that information to [the referral source]. (I8)

"We tend to find that a lot of this is really you know a very small problem which is primarily one of misunderstanding" and in many cases when "dealing with someone you create a bias".

> I9 valued that the objective of investigations was to "protect the integrity of the accident fund". This objective could result in the appropriate expenditures, or alternatively could have clarified no eligibility for expenditures.

> There is also investigative work that we do for instance we will go out and do surveillance and the person is very disabled. And we see that as a success, in that there were some questions about the person's credibility or there was some issue we have been able to resolve.

I9 explained how investigations could resolve claims by obtaining missing information.

> From my investigation perspective it may be that the claim owner is able to talk to one person or two people…and they may just leave it at that. Whereas with us we will go out and we will interview the person …and if they say well there were five other people around we are going to want to talk to the five people as well to determine the veracity of what this person has said. It may lead to either a negative decision or a positive decision for the person, but at the very least you

can be assured that now I have six or seven people telling me the same story or I have one person saying this thing and I have six other people telling me something completely different.

Through their awareness of tendencies for bias within an insurance claim culture, the insurance site attempted to maintain a neutral position and gathered information to remove the biases.

4.4.5 Evidence evaluation was grounded in an insurance paradigm. The insurance site conducted investigations to acquire proof of medical impairment and disability to ensure that provision of funding and services were warranted. The insurance paradigm in general has a "strong moralistic element…where it is necessary to clearly differentiate between 'honest' and 'dishonest' claimants" (Schultz et al., 2000, p.276). However, the insurance site of this study did not share the major tenet of the insurance paradigm that, "claimants who anticipate financial benefits…. are likely to be dishonest about their symptoms" (Schultz et al, 2000, p.275). Based on I8's experience in the program for over 20 years he concluded that many referrals for investigation stem from misunderstandings due to lack of accurate information, or biases on the part of the referring claim managers who suspected claimants were being dishonest about their disabilities. The insurance investigation program succeeded in mitigating these biases. I8 estimated that in over 50% of cases investigated information provided by the program confirmed medical impairment and disability and concluded that the person being investigated was being compensated appropriately. I8 and I9 reported valuing the integrity of their work that culminated in elimination of misunderstandings or biases, and protecting the accident fund to ensure that funds were spent appropriately.

Formal evaluation at the insurance site was focused on cost savings and timeliness, as well as ensuring reports were completed to standards set by the program supervisor. This attention to cost savings was consistent with the insurance paradigm insofar as the key determinant of the paradigm in general relates to financial concerns and mitigating risks due to secondary gain of clients. The accounting department of the organization had developed a formula to calculate cost savings resulting from cases managed by the insurance site. This calculation compared costs of claims that had accessed the services of the insurance program, to estimated costs that would have been sustained had the program not accessed the services. I8 described dissonance related to his criticism of the organization's failure to report cost savings achieved by the insurance site in their publications, as a means of using these evaluation findings.

4.4.6 Discussion. Evaluation criteria and standards in place at the insurance site were: timeliness for each type of service to established standards, completion of final reports to acceptable standards established by the program supervisor, cost savings, and satisfaction of potential referring sources.

Evaluation at the insurance site was organized and consistent. On every case an assessment was conducted as to whether timeliness standards had been met, and on every case the program supervisor reviewed the final report to ensure services were completed to his satisfaction before the case could be closed. An accounting formula was used to assess cost savings of cases handled by the program compared to estimated costs on cases had the service not been accessed.

Neither I8 nor I9 believed the organization fully appreciated accomplishments of the program or the potential value of services the program offered. I8 saw the failure to

publish evaluation outcomes of cost savings as overlooking an opportunity to educate internal and external stakeholders on the program's worth. I9 perceived those in power overlooked opportunities to use the program's services for internal investigations, and attributed this to other departments working in silos and not wanting the insurance site involved. There was a lack of confidence expressed by both I8 and I9 regarding the organization's commitment to openness and learning.

The insurance site characterized one aspect of the insurance paradigm as described in the literature, a positivist perspective that objective evidence of impairment and disability is required to verify entitlement to benefits (Schultz et al., 2000). However, the program did not manifest another common aspect of the insurance paradigm described in the literature: that people who anticipate secondary gain are likely to magnify disability. The investigators saw themselves as providers of accurate information to overcome biases and misunderstandings of decision makers within the insurance system, which had resulted from inaccurate or missing information. Based on evaluation of their services I8 estimated that in over 50% of cases, information they provided verified the clients were truthful about their disabilities. They believed their services introduced integrity within the insurance system regarding appropriate allocation of funds.

Evaluation at the insurance site focused mainly on providing accurate information on objective evidence of medical disability, cost savings and the impact of their services to overcome bias in the system. They were less but somewhat concerned with understanding the potential influences of context and diversity among stakeholders.

4.5 Participants' Perceptions of What is Most Important to Evaluate

Participants were asked what they believed to be the most important criteria for evaluation of disability management programs. Table 13 summarizes program objectives and evaluation criteria grouped by paradigm.

Table 13　　　　Program Objectives and Participant' Suggestions of what is Most Important to Evaluate

	Biomedical	Labour	Biopsychosocial	Insurance
Main program objectives	Accurate assessments and useful reports	Accommodating every employee with suitable and meaningful employment and doing so collaboratively with union and management	Counselling, referrals, guidance, career exploration, and support during the earliest stages of acute conditions	Protecting the integrity of the accident fund
What are the most important criteria to evaluate?	Return to work Useful reports Helping the client recover	Return to work Flexibility of the organization in providing work accommodations Early intervention Cost savings	Helping clients manage their disabilities Rapport	Accuracy of information provided for validation of clients disability Appropriate allocation of client support from the accident fund

4.6 Analysis of the ROLE Results

The Readiness for Organizational Learning and Evaluation Inventory (ROLE) was completed by all participants of this study prior to their interviews and discussed during each interview. This was meant to stimulate participants' thinking regarding organizational learning and to elicit participants' perceptions of strengths or weaknesses of evaluation and learning in their program or organization.

ROLE findings for each participant are interpreted by comparing ROLE findings between/among participants from each site and triangulating ROLE findings with interview data. ROLE findings among all nine individual participants are then summarized (see Table14).

4.6.1 Biomedical (BM) site ROLE results. The BM site participants reported on the ROLE by referring to their department/unit, which was a small independent program and not situated within a larger organization. The BM site was the only site of the four included in this study primarily evaluated externally.

The lowest ROLE scores for both BM participants were in the subcategory 'rewards and recognition systems and practices' (2.8 for BM1:1.0 for BM2). The BM site had no formal internal systems in place for recognition of their services, and relied on repeat referrals to indicate their success.

Ratings on the ROLE category Communication of Information were similar for both participants (3.5 for BM1: 3.3 for BM2). Both participants gave low scores on 'availability' of information (3.0 for BM1: 1.7 for BM2), and high scores on 'dissemination' of information (3.8 for BM1: 4.2 for BM2). The lower scores given on

Table 14   Individual Participants' Responses to Readiness for Organizational

Learning and Evaluation Inventory (ROLE*)*

ROLE Dimensions	Participants								
	BM 1	BM 2	L 3	L 4	L 5	BPsy 6	BPsy 7	I 8	I 9
CULTURE	3.9	4.1	4.1	3.3	3.8	3.8	3.6	2.4	3.8
Collaboration & Problem Solving	4.0	4.1	4.3	3.3	3.8	4.1	3.7	2.8	3.8
Risk Taking	3.0	3.8	3.8	3.0	3.6	3.6	3.8	2.4	3.4
Participatory Decision Making	4.2	4.4	4.0	3.5	3.9	3.5	3.5	2.1	4.1
LEADERSHIP	3.9	4.2	4.0	3.3	3.8	3.6	4.0	1.3	4.2
SYSTEMS & STRUCTURES	3.7	3.1	3.8	2.9	3.3	3.9	3.6	1.6	3.9
Open & Accessible Work Environment	4.0	4.5	3.3	3.0	3.0	3.8	3.9	1.5	4.0
Rewards & Recognition Systems & Practices	2.8	1.0	4.0	2.6	3.2	3.8	3.0	1.0	4.0
Relationship of Work to Organizational Goals	4.7	4.7	4.3	3.3	4.0	4.3	4.3	2.7	3.7
COMMUNICATION OF INFORMATION	3.5	3.3	4.1	3.4	3.1	2.5	2.9	2.8	3.6
Availability	3.0	1.7	4.3	3.7	2.3	2.0	2.2	3.0	3.3
Dissemination	3.8	4.2	4.0	3.2	3.6	2.8	3.4	2.6	3.8
TEAMS	-	-	4.1	-	-	3.8	4.3	1.9	4.0
EVALUATION	3.3	3.4	4.1	2.5	3.9	4.6	4.0	2.5	3.9

Note: For the BM site neither participant gave a rating on the ROLE category 'Teams' as their program was small with only 3 or 4 employees including themselves, a receptionist and on occasion a third practitioner who performed assessments at their site. Two of the three Labour site participants left the category 'Teams' blank as well without explanation.

'availability' corresponded to participants' perceptions of a lack of communication between their program and the main referring agency. The higher scores on 'dissemination' confirmed interview data that these participants made conscious efforts to work together and share information.

The BM participants had low ratings on 'Evaluation' (3.3 for BM1: 3.4 for BM2). Scores provided by BM1 and BM2 on this category confirmed interview data regarding their perceived lack of meaningful evaluation by the main referring agency. Evaluative feedback from the referring agency focused on timeliness of services, number of client satisfaction surveys returned, and a single digit client satisfaction rating. The BM participants believed timeliness standards were important to contribute to the delivery of optimum services to clients; however, they were dissatisfied with the limited overall feedback from the referring agency regarding their services. Their preference for evaluation would have included feedback from referring individuals regarding the usefulness of their reports and detailed client satisfaction information regarding their services, so that they could use the information to make improvements.

Under the ROLE category Systems and Structures, on subcategory 'open and accessible work environment' scores were (4.0 for BM1: 4.5 for BM2), and on subcategory 'relationship of work to organizational goals' (4.7 for both). These high ratings coincided with their explanations of strengths in the BM program. Both BM participants had described successfully sharing work space and equipment, and collaborating to ensure services provided met their standards.

These participants regularly sought each other's professional opinions on cases they were managing, especially when they encountered problems, finding opportunities

to learn from each other. Higher ROLE scores on the category Culture (3.9 for BM1: 4.1 for BM2), and the three subcategories 'collaboration and problem solving' (4.0 for BM1: 4.1 for BM2), 'risk taking' (3.8 for BM2), and, 'participatory decision making' (4.2 for BM1: 4.4 for BM2) reflected their reported collaboration. They also gave high ratings on 'Leadership' (3.9 for BM1: 4.2 for BM2).

4.6.2 Labour site ROLE results. Two of the three labour site participants (L3 and L5) referred on the ROLE to their department/unit (disability management program), and the third (L4) referred to the organization within which the department existed. All of L4's ROLE scores were lower than those of L3, and all but two L4 scores were lower than the scores of L5. L4 and L5 gave no ratings on the ROLE category Teams perhaps due to oversight, as much of their work involved teams. I did not ask about this during the interview as the ROLE was submitted at the beginning of the interview and the scores were not reviewed ahead.

All three labour participants` scores were similar on the subcategory 'open and accessible work environment' (3.3 for L3: 3.0 for both L4 and L5), the only category where there was high agreement. This was the only category on which L3 had a score under 3.5 and indicated he believed improvement was required. ROLE items under this subcategory refer to organizational influences including: bureaucratic red tape when trying to do something new or different, open workspaces, and having minimal boundaries between departments facilitating working together. These lower scores confirmed the labour participants' perceptions that their program was required to work within strict organizational structures, including: a collective agreement, formal job descriptions, separately functioning departments, and dealing only with medical or

psychological illnesses not labour relations issues. Additionally, during interviews all three participants had described the labour program as unique within the organization because it was overseen by a joint labour management committee, and operated differently from other organizational divisions because it served employees of the organization rather than clients of the organization.

L5 (the administrator of the Labour program) gave low scores on 'Systems and Structures' (3.3), and its subcategory 'rewards and recognition systems and practices' (3.2); as well as on category 'Communication of Information' (3.1), and its subcategory 'availability' (2.3). These items relate to feedback given to employees regarding their achievements, and informational feedback regarding departmental or organizational performance, mostly from outside sources. These areas may have been of particular interest to L5 as the program administrator, giving her a heightened critical awareness of any shortcomings.

The low score that L5 gave on Communication of Information, subcategory 'availability' (2.3) contrasted with the higher scores given by L3 (4.3) and L4 (3.7). This was the lowest score that L5 gave, whereas it was the only score above 3.5 for L4 and was one of the three highest ratings given by L3. Higher scores indicated this was an area on which the program (in the case of L3) and the organization (in the case of L4) were considered to be functioning well. The low score given by L5 reflected her perception that information available did not meet what was required, perhaps in relation to her unique duties as an administrator. For example, as part of her administrative responsibilities L5 had described working with the organization's human resources

department to develop a data management system that could in the future be used for evaluation organization wide.

High scores were given by L3 and L5 in areas related to a culture of learning, participatory decision making, leadership and goal setting within their department (ranging 3.8-4.3 for L3: 3.6-4.0 for L5). Items under 'Collaboration and problem solving' reflect respect, cooperation, collaboration, constructive problem solving, and willingness to learn and improve practices toward shared success among professionals. Items under 'participatory decision making' reflect employee-managerial openness and cooperation in sharing information to facilitate learning and decision making. 'Relationship of work to organizational goals' reflects how well program goals align with organization goals. Interviews with L3 and L5 confirmed these high ROLE scores. These two participants perceived their program to have open and cooperative communication systems that were inclusive, encouraged feedback from stakeholders within and outside of the program, and integrated labour and management perspectives. L5 believed their program was a model for other organizations.

L3's scores on the ROLE were almost all high, including all six main categories, and seven of the eight subcategories. L5, the administrator of the program, was more discriminating on the ROLE, scoring high on three of the six main ROLE categories, and five of the eight subcategories. Her scores were higher on ROLE areas related to culture, leadership, relationship of work to goals, dissemination of information and evaluation, and lower in areas related to systems and structures and availability of information.

The scores of L4 (who referred to the organization not the program on the ROLE) were almost all low, with eleven out of 13 scores under 3.5. L4 had been a former labour

organizer within the organization, and during her interview reported bringing a unique labour perspective to the program, including knowledge of the collective agreement and the organization. While identifying strongly with labour, during her interview L4 expressed commitment to the joint labour-management administration of the program, and impartiality in serving labour and management stakeholders, and had been constructively critical of organizational constraints impacting the program.

The lowest ROLE score given by L4 was on category Evaluation (2.5). As L4 was referring to the organization not the program, her low score reflected a perceived need for improved evaluation and learning from evaluation on an organizational level. In contrast, L3 and L5 rated Evaluation high (4.1 and 3.9 respectively). During interviews L3 and L5 both described having participated in the development of evaluation systems for the labour program, and having benefitted from the use of evaluation findings, whereas L4 did not report having had similar experiences. L3's and L5's prior involvement in evaluation activities may have contributed to a greater awareness of and appreciation for the role of evaluation in their program.

The only two L4 scores not below 3.5 were under Culture, for 'participatory decision making' (3.5); and Communication of Information, for 'availability' (3.7). ROLE items in these categories refer to employees and managers openly sharing information to make informed decisions, and availability of information on performance from multiple sources. These higher ROLE scores reflected areas L4 saw as strengths within the organization.

L4's lower scores were consistent with her attempts to provide constructive criticism. She identified ways she believed the labour site program and organization were

succeeding and ways they were struggling, and suggested areas she believed were in particular need of improvements. For example, L4 described her perceived lack of organizational commitment to disability management because of a lack of funding to train returning work clients how to use the new computer system, which was a critical job demand. Her pattern of lower ROLE scores reflected her criticism of the organization and perception that improvements were needed in culture, leadership, systems, and dissemination of information to support learning and evaluation.

4.6.3 Biopsychosocial (BPsy) site ROLE results. Both BPsy participants referred to their organization not the program when completing the ROLE. They explained during interviews how all programs within the organization, including the BPsy program, were integral to multi-disciplinary and inter-disciplinary services available simultaneously to clients.

Both BPsy participants gave high ratings on Evaluation (4.6 for BPsy6: 4.0 for BPsy7). Although both confirmed during interviews there was no formal program evaluation within the organization, the ROLE items in this category refer not only to evaluation that is in place, but how evaluation could lead to improvements if it were in place. These higher scores were indicative of the BPsy participants' awareness of ways evaluation could benefit the organization.

These participants gave the lowest scores (all under 3.5) on Communication of Information (2.5 for BPsy6: 2.9 for BPsy7), and its two subcategories 'availability' (2.0 for BPsy6: 2.2 for BPsy7) and 'dissemination' (2.8 for BPsy6: 3.4 for BPsy7). Items refer to the availability of feedback from multiple sources regarding effectiveness of services, an area these participants believed needed improvement.

The low scores given on Communication of Information confirmed BPsy6's perception that there were no organization wide systems of data collection on program services and no formal evaluation of programs. She thought it likely that some of the programs maintained their own data systems or evaluated their program outcomes, but this information was not shared organization wide. As administrator of the BPsy program, BPsy6 maintained a data base of information on services provided to clients, and planned to code the data such as "return to work" or "went to school to retrain", to include in her annual reports. She believed the organization would support the staff if they chose to conduct research or evaluation on their own programs, something BPsy hoped to eventually do.

The other low ROLE score was BPsy7's score (3.0) on 'rewards and recognition systems and practices'. BPsy6, the administrator of the program, rated this higher (3.8). This subcategory refers to organizational recognition given to employee innovation or team learning. BPsy6's experiences as administrator may have provided her opportunities for greater awareness of recognition that staff or programs within the organization were receiving. During her interview she had stated that some programs were considered more essential than others.

The other ROLE scores were above 3.5 and confirmed interview data regarding the organization's strong value of a learning culture, collaborative leadership, and a participatory system of teams. The organization so valued their multi-disciplinary services, that when faced with health budget cuts they had reduced management within the organization down to two or three individuals in order to maintain program services.

4.6.4 Insurance site ROLE results. The Insurance site participants referred to their department/unit when completing the ROLE. Their department was in a large organization that managed insurance claims within one of its divisions. None of the fourteen ROLE category/subcategory scores of the two insurance site participants overlapped.

Communication of Information, subcategory 'availability' of information (3.0 for I8: 3.3 for I9) was the only area of agreement between I8 and I9. Both scores were below 3.5, indicating this was an area both I8 and I9 believed needed improvement.

During his interview, I8 criticized the organization for not publishing information on the cost saving of the insurance program, similar to the way the organization published the accomplishments of other departments. He believed this represented an absence of organizational insight into the possible benefits of public knowledge of the program's outcomes. Additionally, I8 questioned whether the organization appreciated the social and economic accomplishments of the insurance program. I9, however, was critical of the lack of information sharing at an organizational level. He believed directors were not communicating with each other, resulting in departments working in isolation.

On all other areas of the ROLE these two participants' scores contradicted each other, which was consistent with their discrepant views during interviews. On Leadership I8 gave one of his lowest scores (1.3) and I9 gave his highest score (4.2). These discrepant scores reflected differences they expressed regarding program management.

Both insurance participants described a top down management for the insurance program and organization. I9 was responsible for making administrative decisions within the program such as approving training costs. Organizational directors made higher level

decisions that impacted the program from outside. The investigators' contributions were mainly limited to case work, report writing and communicating with other professionals who worked in the claims department. Both I8 and I9 had expressed concern about the hierarchical decision making processes within the organization, and the failure to take advantage of opportunities to learn at an organizational level.

During interviews the two participants expressed different views on managers' decisions related to selection of new investigators and approval of ongoing education for investigators. I8 valued decades of prior policing work experience, whereas I9 valued prior high level academic achievement. I8 and I9 also disagreed on what continuing education was most useful to program investigators. I8 valued contracting with experts to train groups of staff on advanced methods in the field, and I9 (who as administrator of the program approved all training costs) supported basic courses such as interviewing skills to be taken by individuals (interviewing was a skill I8 believed was critical and should have been mastered prior to becoming an investigator within the program).

Scores of the two insurance site participants were also discrepant on Systems and Structures (1.6 for I8: 3.9 for I9), and its subcategories 'open and accessible work environment (1.5 for I8: 4.0 for I9), 'rewards and recognition systems and practices' (1.0 for I8: 4.0 for I9), and 'relationship of work to organizational goals (2.7 for I8: 3.7 for I9). These items describe open work spaces where employees are able to share information, where employees and the program work toward mutual goals, and where recognition is given to employees for contributions they make in meeting common goals. The lowest score given on the ROLE by I8 was under Systems and Structures, on subcategory 'rewards and recognition systems and practices' (1.0). During his interview

I8 criticized the program for not engaging investigators in team work, and criticized the organization for not demonstrating appreciation of innovative work accomplished by the insurance program.

Participants' scores were also discrepant but somewhat closer on Culture (2.4 for I8: 3.8 for I9), and its subcategories 'collaboration and problem solving' (2.8 for I8: 3.8 for I9), 'risk taking' (2.4 for I8: 3.4 for I9) and 'participatory decision making' (2.1 for I8: 4.1 for I9). These items refer to ways a program encourages employee inclusiveness, their sharing of information and ideas, taking risks to be innovative, leading processes of change, feeling no fear of making mistakes, and the degree to which an organization views the capacity of all employees to learn, as the organization's greatest resource. The lower scores of I8 reflected his perspective that the program was not conducive to employee inclusiveness. I9's higher scores indicated he viewed these areas as satisfactory or strengths of the program. During interviews and on the ROLE, I8 was critical of the insurance program's administration and believed investigators' opinions were ignored.

I9's scores were below 3.5 on only two of the fourteen ROLE category/subcategories, and above 3.5 on the other twelve. These two areas I9 believed needed improvement were Communication of Information, subcategory 'availability' (3.3) and Culture, subcategory 'risk taking' (3.4). During interviews I9 was critical of a lack of communication and planning between program personnel and organizational directors outside the program. I9 believed due to territorial attitudes and poor collaboration among directors of the organization, departments were isolated and impenetrable, and he suspected possible corruption may be going undetected as directors protected their departments from scrutiny.

I8's overall low ROLE scores (below 3.5 on all fourteen ROLE categories/subcategories) reflected his criticism mainly directed at administration of the insurance program and of the organization. I8's highest scores related to information sharing: Culture, subcategory 'collaboration and problem solving (2.8), Communication of Information (2.8), and subcategory 'availability' of information (3.0). During the interview, I8 had identified the individual work accomplished by investigators that resulted in learning opportunities for claims staff as a strength of the program. For example, he perceived contributed objective information on client functioning that resulted in referring persons making decisions with reduced bias.

I9's scores were above 3.5 on twelve categories/subcategories. The twelve higher scores reflected his perspective as administrator that the insurance program was operating successfully in relation to learning and learning from evaluation. I9 rated the following categories 4.0 or higher: Culture, subcategory 'participatory decision making'; Systems and Structures, subcategories 'open and accessible work environment' and 'rewards and recognition systems', and Leadership, Teams and Evaluation.

Out of six main ROLE categories, I8 gave his lowest scores on Leadership (1.3), Systems and Structures (1.6) and Teams (1.9). I8 gave somewhat higher scores on the other three: Culture (2.4), Evaluation (2.5) and Communication of Information (2.8). The lowest score that I9 gave was on Communication of Information (3.6) and his other five scores were even higher: Culture (3.8), Systems and Structures (3.9), Evaluation (3.9), Teams (4.0) and Leadership (4.1).

4.6.5 Summary of ROLE results. Six participants referred to their program when completing the ROLE (BM1, BM2, L3, L5, I8 and I9) and three referred to the

organization within which their program was situated (L4, BPsy6 and BPsy7).

Participants who referred to their organization had significant previous work experience at an organizational level, whereas those referring to the program had greater prior work experience at the program level. The program and organizational perspectives contributed to the identification of what was or was not being done at organizational levels to develop learning and learning from evaluation, and what impacts organizations had at the program level.

I8 scored under 3.5 on all categories/subcategories of the ROLE, and L4 scored under 3.5 on all but 2 categories. The ROLE category/subcategory scores of other participants were reasonably high, indicating that conditions for organizational learning and learning from evaluation were in place, particularly at a program level.

Across all nine participants, three ROLE categories/subcategories were rated lowest (under 3.5 by more than half the participants), indicating areas that required improvement in relation to organizational learning and learning from evaluation: Communication of Information (6/9 participants), Communication of Information subcategory 'availability' (7/9 participants), and Systems and Structures subcategory 'rewards and recognition systems and practices' (6/9 participants). Specific items in these three areas identified as needing improvement were: 1) a lack of information from stakeholders that could inform programs about their effectiveness, and 2) not encouraging employees to share information and create opportunities to learn.

Across all participants, five ROLE categories/subcategories were rated highest (3.5 or over by more than half of the participants), indicating areas of strength that could be leveraged in building organizational learning and learning from evaluation. These

were: Culture (7/9 participants), and Culture's subcategories 'participatory decision making' (8/9 participants) and 'collaboration and problem solving' (7/9 participants), Leadership (7/9 participants), and Systems and Structures, subcategory 'relationship of work to organizational goals' (7/9 participants). These items referred to having a culture of respect, inclusiveness, collaborative problem solving and decision making, shared rewards, openness to employees questioning and making innovative contributions, and overall, organizations having the highest regard for employees working together toward common goals.

CHAPTER 5: DISCUSSION

5.1 Introduction

Systematic judgment of the quality of a grounded theory study includes explicit identification of how the study has met the canons and procedures of this methodology. The usual scientific canons including significance, generalizability, reproducibility and verification are redefined for grounded theory, and their positivistic connotations guarded against (Corbin and Strauss, 1990). It is the grounded theorist's responsibility to identify relevant conditions, and determine how individuals are responding to those conditions and the consequences of their actions. Grounded theory is not about accuracy of descriptive units of data but is about conceptual abstraction, explaining patterns of behavior or issues in context (Holton, 2007). "The procedures of grounded theory are designed to develop a well integrated set of concepts that provide a thorough theoretical explanation of social phenomena under study. A grounded theory should explain as well as describe. It may also implicitly give some degree of predictability, but only with regard to specific conditions" (Corbin and Strauss, 1990, p. 5).

Grounded theory begins with open coding to compare and group data and to stimulate analytic questions. This is followed by conceptual coding, a more abstract analysis and development of hypotheses that can be tested in relation to new data and to identify themes, develop context, and explain processes or changes over time (Corbin and Strauss, 2008). Concepts, as the basic unit of data analysis in grounded theory, "earn their way into the theory by *repeatedly* being present…or significantly absent" (Corbin and Strauss, 2009, p.7). The researcher groups conceptual themes into more abstract themes in terms of their properties and dimensions, explaining the conditions under which they

arise, actions or interactions that have occurred in response to conditions, and

consequences produced.

During analysis a core conceptual category often emerges that is central to the

grounded theory, around which the other categories can be related to explain variation in

the conditions, actions and consequences (Corbin and Strauss, 1990). Writing theoretical

memos including illustrations from the start to the end of the study maintains records,

makes hypotheses, generates questions that evolve through analytic processes, and

preserves conceptual detail (Corbin and Strauss, 1990).

Analysis of processes must be built into a grounded theory (Corbin and Strauss,

1990) and can be achieved through different means, such as breaking a phenomenon

down into stages, phases or steps, or, identifying a purposeful action/interaction that

changed with different responses to conditions. Broader conditions that may affect the

phenomenon, such as economic conditions, cultural values, political trends or social

movements, must be analyzed and brought into the grounded theory to show "*specific*

linkages between conditions, actions, and consequences" (p. 12). Bacharach (1989)

suggested a testable theory should state how it is bound in time (applicable at different

times or not) and bound in space (applicable within different types of organizations or

not). Testing hypotheses within a grounded theory one presumes that exact conditions

cannot be replicated. The more abstract the concepts and theory, the more variation that is

uncovered, and the wider the theory's applicability (Corbin and Strauss, 1990), but at the

cost of detail (Bacharach, 1989).

A summary of grounded theory analyses followed in this study is included in

Appendix C.

The following discussion of the findings of this study follows the key steps of grounded theory as summarized by Harry, Sturges, and Klinger (2005). Table 15 summarizes the progression of analyses for this study.

Table 15   Data Analysis Map

1. Open Coding	*80 open codes* See Appendix C
2. Conceptual Coding	*22 conceptual codes* • Accurate information, Adaptation, Bias, Client satisfaction, Collaboration, Communication, Culture, Diversity, Feedback, Formal evaluation Criteria, Funding, Goals, Integrity (in vivo code), Learning, Meaningful evaluation, Performance based model, Qualifications, Rapport, Reports, Self evaluation, Standards, Stigmas.
3. Developing Themes	*11 themes from original 22 conceptual codes* • Adaptation, Bias, Client satisfaction, Collaboration, Communication, Culture, Diversity, Evaluation criteria, Feedback, Learning, Reports. *2 themes elevated from open codes* • Cost savings, Timeliness *3 themes elevated from interview data* • Client functioning, Data management, Return to work.
4. Testing the Themes	*6 conceptual themes* • Collaboration, Communication, Culture, Diversity, Bias. • Adaptation. • Evaluation criteria: Client functioning, Client Satisfaction, Cost savings, Return to work, Timeliness. • Data management. • Feedback, Reports. • Learning.

5. Inter-relating Explanation	*6 inter-relating themes and their abstract relationships explaining disability management evaluation* Multiple sources of disability management evaluation that reflect diverse stakeholder perspectives and values coexist. • Evaluation that does not integrate internal and external stakeholder perspectives risks a potential loss of important information and evaluation that produces narrow insights. The integration of internal and external stakeholder perspectives is accomplished by: *Collaboration, Communication,* and sensitivity to *Culture*, *Diversity* and *Bias.* • Disability management programs are required to adapt to influences from their broader context. *Adaptation* is not necessarily positive for the program or additive for the clients, but can be necessary for program survival. • The primary disability management evaluation criteria are: *Client functioning, Client satisfaction, Cost savings, Return to work and Timeliness.* These criteria are not used uniformly among disability management programs. • Programs and organizations are developing *Data management systems* with increasing use of technology to be used in evaluation and performance management. • *Feedback* and *Reporting* of evaluation findings contribute to *Learning* for stakeholders, programs and organizations. • Evaluation of disability management follows a consumer working logic. Evaluation is primarily concerned with use of findings, and secondarily concerned with pluralistic values of multiple stakeholders.
6. Theory Development	Evaluation that is meaningful requires insight into how impairment environment interactions are being managed by the program.

	The presence or absence of collaboration among stakeholders has the potential to contribute significantly to an explanation of variability in disability management and disability management evaluation.
	Disability management programs that highly value opportunities to learn and to learn from evaluation encourage collaboration and communication among stakeholders during program services and evaluation.
	Understanding how disability management programs are adapting to influences from their broader contexts has the potential to contribute significantly to an explanation of variability in disability management and disability management evaluation.
	Disability management programs that are concerned with understanding contextual influences and integrating multiple stakeholders' interests can leverage new insights to reach their goals.
	There are currently five primary disability management criteria: return to work, cost savings, timeliness of services, client satisfaction, and client functioning. These criteria are not used uniformly among disability management programs.
	Disability management programs and their funding organizations are increasingly using technology to develop new data management systems for use in evaluation.
7. Purpose and Logic of Disability Management Evaluation	Disability management evaluation is primarily concerned with Christie & Alkin's "use" branch of evaluation assuming a social accountability orientation for program improvement, and is secondarily concerned with multiple stakeholders' values.
	Consistent with the emphasis on "use", disability management evaluation primarily adopts a consumer oriented working logic, while some programs also incorporate the logic of connoisseurship and pluralistic evaluation approaches.

Note: The order that grounded theory data was managed moved through the levels from 1 to 6. Christie & Alkin, 2013.

5.2 Open Coding

5.2.1 Line by line coding. A total of 80 open codes were generated from the nine interviews (See Appendix D for a list of the 80 open codes.) These codes were generated using NVivo software to code the nine interview transcripts line by line. An example of open coding of interview data is illustrated in the following example. The open code "participants' evaluations of their program services" was based on excerpts from multiple interviews where participants explained their prior evaluation of their disability management services, including the following excerpts taken from the interview with BM2.

> BM2: You should have ways to know that what you do is effective and also to improve your ability where you are learning or your effectiveness on an ongoing basis. So whether I am an OT here or an OT within a bigger organization it would still be my duty to ensure that what I do is effective. My services are effective in some way or even questioning myself as to what would I judge to be effectiveness in terms of the services I provide.
>
> Researcher: You said that you're aware of the fact that you should always be evaluating your own services. In what way have you evaluated yourself?
>
> BM2: We used to have our own satisfaction questionnaire before [the referring agency] came up with their own…It was for the worker to assess our service. In terms of whether they had been treated in a respectful manner, whether they had learned something during the evaluation that I provided, or through the service that I provided. So there was a series of questions… maybe some of them were not as good as others or not as objective as others, but then once the contracts

started being awarded for services [the referring agency] came up with their own satisfaction questionnaire and they were quite a bit different from the one that we used to give to clients. Nevertheless this is now the main source of information that I have to know if my services are effective.

Another example of the code "participants' evaluations of their program services" is from a document BM2 prepared in advance of our interview. Under a heading "My criteria for success" BM2 listed four criteria she used to evaluate functional capacity evaluations (FCEs), one type of services provided by the BM site. These were: satisfaction questionnaire; debriefing sessions with clients; repeat referrals from referral source; and review of my FCE report – have I clearly answered all questions and provided evidence?"

Open coding included reading and coding each interview transcript individually. As new transcripts were coded, data from previously coded transcripts was reread to confirm whether any of the newer codes had been overlooked or labelled differently during original open coding, with the objective of attributing reliable meanings to the open code labels. One of the outcomes of conducting the interviews myself, then transcribing and open coding every individual transcript and document, was that I gained an overall awareness of what information was included in the data.

5.2.2 Transition from open codes to conceptual coding.

5.2.2.1 Visual display of open codes. In grounded theory the researcher uses memoing and diagramming to examine, analyze and think about data in complex ways, by sorting, analyzing and coding. "When memoing a topic analytically, the researcher

generates a set of categories, contrasts, comparisons, questions, and avenues for further consideration which are more abstract than the original topic" (Lempert, 2007, p. 251). After open coding was completed the 80 open codes that had been generated were each printed on separate slips of paper, and they were organized to create a visual chart (24 x 36 inches) grouping and ordering codes based on commonalities. Groupings of open codes included: descriptions of disability management programs; the ROLE; organizational cultural issues; paradigmatic approaches; and external evaluation of the biomedical site (refer to Appendix E for examples of groupings from the chart).

This visual display contributed greater clarification of open codes that had been created. The visual display of grouping open codes together provided an elementary analysis of how the open codes related to each other. Some groups of codes were straightforward and descriptive, for example describing clients or disability management practices at the sites of this study. Other groups of codes led me to undertake initial levels of more abstract analysis, as I perceived some individual open codes and groups of open codes were foundational to more abstract understandings requiring further analysis.

5.2.2.2 Concrete versus abstract. One of the first analytical realizations I had was to differentiate data that was descriptive or concrete, from data that was abstract to varying degrees. A definition of concrete includes "specific, definite; denoting a material object as opposed to an abstract quality, state or action" (Barber, 1998, p. 294), and the definition of abstract includes "to do with existing in thought rather than matter, or in theory rather that in practice, not tangible or concrete (p. 6).

Examples of concrete were open codes timeliness and cost effectiveness, both measurable evaluation criteria. Examples of open codes I initially identified as more

abstract were: integrity (in vivo code), communication with external stakeholders; referring organization communication with participant program; and multi-disciplinary cross pollination. These open codes were identified as more abstract as they involved a higher level of analysis than linear measurement of outcomes, and had the potential to contribute to a thicker explanation of disability management evaluation at the sites.

When new open codes were identified, I returned to previously coded transcripts to compare data to data, data to codes, and codes to codes. An example was the open code integrity (in vivo code), referring to appropriate allocating of funds, a term originally used by a participant from the insurance site in one of the final two interviews conducted and transcripts coded. I gave reasoned consideration as to whether the code integrity would apply to statements made by BM2 a participant from the biomedical site, who had referred to biases in the reports of other organizations, during the second interview I conducted and transcript I coded. Coding of integrity included obtaining a dictionary meaning (Barber, 1998), and hypothesizing the reference to integrity at the insurance and biomedical sites were both positivist referring to seeking one truth, fairness and ethics.

Initial awareness of the potential for more abstract interpretations of data gained from open code analysis was carried on to the next phase of analysis, conceptual coding of raw data. I anticipated there was potential to examine some open codes further at higher levels of abstraction, as they could reasonably be elevated to conceptual codes, for example open codes integrity or bias. Furthermore, given evidence identified during open code analyses that conceptual findings did exist within the data that could lead to more meaningful explanations than descriptions, I commenced conceptual coding with an

expectation that this higher level of abstract analysis would lead me to the discovery of additional new concepts.

5.3 Defining Conceptual Codes

The original interview transcripts were coded again one at a time on a more abstract level using NVivo software. This involved reading each transcript line by line and attaching conceptual codes to relevant sections of the transcripts. Twenty-two conceptual codes were generated.

Accurate information, Adapting, Bias, Client satisfaction, Collaboration, Communication, Culture, Diversity, Feedback, Formal evaluation, Funding, Goals, Integrity (in vivo code), Learning, Meaningful evaluation, Performance based model, Qualifications, Rapport, Reports, Self evaluation, Standards, Stigmas.

NVivo files facilitate access to all excerpts for each conceptual code, so that original data sources were easily located in the transcripts. Following Bringer, et al's. (2010) recommendation to "search previously coded documents for instances of a newly developed category" (p.254) I returned to the original raw data using constant comparisons rereading each interview transcript repeatedly until I perceived all excerpts representing each conceptual code had been identified, and the code files were saturated. This analysis involved iterative processes where I returned to the NVivo files of excerpts from transcripts pertaining to each of the 22 conceptual codes, comparing conceptual codes to each other and to the original interview transcripts and documents. I entered some memos into conceptual NVivo files describing my reactions to codes.

Paper and electronic documents were also coded a second time conceptually. The documents were coded by hand without NVivo. Given there were so few documents and little information on them to code, computerization that would have been necessary to manage large amounts of document data was not required to assimilate document excerpts into the computer files on transcript conceptual codes.

5.3.1 Grouping open codes into conceptual categories. Grounded theory methods include comparing data to data, data to codes, codes to codes, and codes to concepts. I categorized each of the 80 open codes under one or more of the conceptual codes. Open codes were grouped together when I identified commonalities among them that related to the conceptual code under which they were grouped. Table 16 summarizes the conceptually categorized open codes.

Table 16   Conceptual Categories and Open Codes

Accurate information: Client credibility, Consultants to program, Contributing new evidence, Fraud example, Integrity (in vivo code), Non validity, Surveillance, Validity.
Adaptation: Political limited funding, Referring organization communication with participant program.
Bias: Bias overcome, Impartiality.
Client satisfaction: Client summative evaluation of program (survey), Survey evaluation.
Collaboration: Employer input, Labour joint union management, Multi-disciplinary cross-pollination, Multi-disciplinary early intervention.
Communication: Communications with external stakeholders, Problem solving, Reporting evaluation findings, Sharing personal beliefs.
Culture: Claims managers high stress, Cultural issues, Labour relations roles and issues, Labour site limited information on client disabilities, Problem with increased temporary workers, Psychological disability.

Diversity: Disabled DM counsellors, Gender, Stigmas.
Formal evaluation: Cost effectiveness of program, Evaluation needs assessment, Referring organization criterion for program services, Referring organization formative evaluation of program, Referring organization summative evaluation of program.
Goals: Participants suggested DM evaluation criterion, Setting specific behavioural objectives, Timelines of DM for services programs provide.
Learning: Learning, Participant learning from client feedback surveys, Participant training through NIDMAR return to work coordinator, ROLE, ROLE open and accessible work environment, ROLE area needing improvement, ROLE collaboration and problem solving, ROLE decision making, ROLE strengths from organizational learning.
Meaningful evaluation: Referring organization communication with participant program, Performance based evaluation.
Qualifications: Participant qualifications, Participant program evaluation background.
Reports: Reporting evaluation findings, Weekly statistical summary of services provided.
Self evaluation: Introspection, Social learning copying standards of others (for reports), Validity.
Descriptions of DM sites: Biomedical DM site, Biopsychosocial DM site, Define disability for this organization, Definition of clients for the organization, Definition of services provide by this program, Insurance model site, Labour site joint union management, Labour model site, Non DM services of program, Participant description of program, Participant role in organization, Preventative aspect of DM program, Program clients, Program situation within organization, Referring sources.
Evaluation at the sites: How participant knows program is doing what was intended, Participant evaluation of their program services, Participant formative evaluation of their program services, Participant summative evaluation of their program services, Who is responsible for evaluating program.
Data management technology: Participant opinion new computer system interferes, People Soft DM record keeping software, Technology.

Some of the original conceptual codes were joined (condensed) as commonalities became clear, and new codes were developed as distinctions among data became clear. An example of condensing in this study was when two open codes (Setting specific behavioural objectives, Timelines for providing disability management services) were originally categorized as "Standards", and then were combined with the open code (Participants suggested DM evaluation criterion) under conceptual code "Goals". An example of creating a new conceptual code was "Data management" under which several open codes were grouped (Participant opinion new computer system interferes, PeopleSoft DM record keeping software, Technology). Figure 6 illustrates these examples of condensing and adding conceptual codes.

5.4 Developing Themes

  5.4.1 Selecting conceptual codes relevant to emerging themes. Conceptual codes were further analyzed in relation to each other, to open codes, to the data and to the original research questions of this study through constant comparisons. This analysis revealed ways that concepts related to each other, and ways that concepts grouped together into more abstract themes. This process also resulted in understanding of how some concepts were more important than others in their significance to the data and to meanings participants had been conveying. Themes that began to develop during this stage of analysis included: disability management evaluation criteria, the significance of collaborative communications, recognition of diversity, learning from evaluation, program adaptation to contextual influences, and the role of data management systems.

Figure 6 Condensing and Adding Conceptual Codes

Standards: Setting specific behavioural objective, Timelines for providing disability management services.

Goals: Participants suggested DM evaluation criterion.
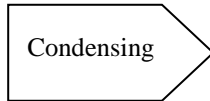
Condensing

Goals: Participants suggested DM evaluation criterion, Setting specific behavioural objective, Timeliness for providing disability management services.

Open codes:

- Participant opinion new computer system interferes
- PeopleSoft DM record keeping software
- Technology

Creating a new conceptual code
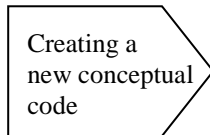
Data Management: Participant opinion new computer system interferes, PeopleSoft DM record keeping software, Technology.

As conceptual categories were being analyzed, 11 of the original 22 conceptual codes were significant to the emerging themes, and the rest were dropped as they were not considered as important. The 11 robust conceptual codes were:

Adaptation, Bias, Client satisfaction, Collaboration, Communication, Culture, Diversity, Evaluation criteria, Feedback, Learning, Reports.

Two open codes were elevated to conceptual codes as their importance to the emerging theme of "evaluation criteria" became evident through constant comparisons of the data. Open code "Cost effectiveness of program" was elevated to conceptual code "Cost savings", and Open code "Timeliness of disability management for services programs provide" was elevated to conceptual code "Timeliness". The importance of these

conceptual codes to disability management evaluation at the sites became evident through

data analysis, and cost savings had also been a primary evaluation criteria in the disability

management evaluation literature. The two elevated codes were:

Cost savings, Timeliness.

Three new conceptual codes were elevated from the original interview data as their

importance to the emerging themes became evident during constant comparisons of data.

These concepts had been significant at the sites of this study, and within the disability

management literature, but had initially been underemphasized during open coding and

conceptual coding of data. The three elevated codes were:

Client functioning, Data management, Return to work.

The new conceptual code "Client functioning" was created to reflect a

characteristic of the emerging theme "evaluation criteria". During interviews participants

described how disability management programs monitored individual client functioning

under varying conditions. To define the new conceptual code "Client functioning" I

utilized the disability literature. The World Health Organization developed an

international system to classify functioning, disability and health (ICF), a conceptual

framework intended to describe and understand the components of functioning, disability

and health (Escorpizo et al., 2011).

> [The ICF system] includes *body structure* and *body functions* at the body level,
>
> and *activities and participation* at the community/society level. Functioning and
>
> disability as embodied in the ICF also consider the influence of contextual factors
>
> such as those related to the person and those related to the environment, on

functioning. Therefore functioning is a result of the interplay between and among

these components" (p. 129).

The components of functioning and disability in the system can be expressed in two

ways: (1) problems or disabilities (impairments, activity limitations or participation

restrictions); and, (2) non-problematic (neutral) aspects of health and functioning, where

"a person's functioning and disability is conceived as a dynamic interaction between

health conditions (diseases, disorders, injuries, traumas, etc.) and contextual factors"

(ICF, 2001, p.8). The ICF classification system considered functional and environmental

interactions, a view of disability management that involved more than biology, requiring

the consideration of contextualization (Smart & Smart, 2006; Loisel, et al., 2001).

5.4.2 Summary of conceptual codes relevant to emerging themes. Sixteen

conceptual codes that originated from the data and those most relevant to explain the

emerging themes were retained for the next phase of analysis, testing the themes.

Adaptation, Bias, Client functioning, Client satisfaction, Collaboration,

Communication, Cost savings, Culture, Data management, Diversity, Evaluation

Criteria, Feedback, Learning, Reports, Return to work, Timeliness.

5.5 Testing the Themes

5.5.1 Reflexivity. Analysis involves inductive moving from data to explanations

of the data; however, researcher interpretations are also involved. As a researcher I have

pre-conceived understandings based on over 20 years experience in the field of disability

management that I bring into data analyses. "Researcher reflexivity works hand-in-hand

with the iterative nature of the research to bring preconceived beliefs into the dialogue,

rather than seeking to omit or ignore them" (Harry, et al., 2005, p.7). In contrast to trying

to control researcher bias as though it would confound the data analyses, I have attempted to recognize how my understandings of the themes may contribute during data analyses. For example, I had been aware of characteristics of paradigms within services at the various sites. During analysis of data I was aware of missing codes important to themes and returned to the interview transcripts and the literature to identify them.

5.5.2 Categorizing the themes. In grounded theory the researcher "treats the various code clusters in a selective fashion, deciding how they relate to each other and what stories they tell" (Harry, et al., 2005, p.5).

5.5.2.1 Themes. The themes that emerged were based on grouping conceptual codes into the five following categories.

1. Five themes were grouped as they represented common disability management *program evaluation criteria* evident in the data of this study, and confirmed as primary criteria in the disability management literature. They were: *Client functioning, Client satisfaction, Cost savings, Return to work*, and *Timeliness*.

2. Five themes were grouped as they reflect the importance of *program context and interactions* within programmatic contexts. They were: *Collaboration, Communication, and sensitivity to Culture, Diversity,* and *Bias*.

3. Two themes were grouped as their common purpose was to contribute to *learning* from evaluation at individual, program and organizational levels and labeled *Feedback and Reporting*.

4. *Data management* was a theme. Participants described how programs or their organizations were in the process of developing data management systems with increasing use of technology for evaluation and/or performance management.

5. *Adaptation* was a theme regarding broad contextual influences on the programs, variations in how programs responded to those influences, and the consequences of their responses.

5.6 Differentiating Evaluation in the Four Paradigms

Table 17 is a visual display of conceptual codes and themes, to assist with differentiation of evaluation at the four sites. The table summarizes the concepts present at each site and within each of the paradigmatic approaches, and whether their presence or absence was perceived as constructive or as contributing to dissonance.

Table 17   Conceptual Themes Evident at the Paradigmatic Sites

	BM	Labour	BPsy	Insurance
PROGRAM EVALUATION CRITERIA				
Client functioning:				
Physical	+	+	+	+
Cognitive		+	+	
Environmental participation		+	+	
Client satisfaction	-	+		+
Cost savings		+		+
Return to work	-	+		
Timeliness	+	+	+	+
PROGRAM CONTEXT INTERACTIONS				
Collaboration	-	+	+	-
Communication	-	+	+	+
Culture	-	+	+	
Diversity		+	+	
Bias	-	+	+	+
LEARNING	+	+	+	+
FEEDBACK AND REPORTING	-	+		+
DATA MANAGEMENT SYSTEMS	+	+	+	+
ADAPTATION	-	+	+	-

Note: (+) denotes some constructive dynamic regarding the theme, (-) denotes some dissonance regarding the theme (due to its presence or absence), and (blank) denotes no conceptual focus noted, in relation to evaluation at the site.

5.6.1 Conceptual themes at the paradigmatic sites. The Insurance and BM paradigms share the assumptions that physical impairments can be objectively measured and symptoms are directly proportionate to pathology. These sites were concerned with clients' physical impairments: the BM site measured physical capacities, whereas the Insurance site investigated consistency in symptoms they considered to be valid evidence of impairment.

Evaluation at these two sites involved measurable program outcomes. At the Insurance site, evaluation criteria were cost savings and timeliness of services, quality of reports and satisfaction of prospective referring individuals. At the BM site, evaluation conducted by the referring agency measured timeliness of services, client satisfaction ratings and number of completed client satisfaction surveys returned, and evaluation conducted by the BM staff measured accuracy of equipment and job demands analyses, client satisfaction and quality of reports.

There was no evidence of these sites incorporating collaboration, or concern with culture and diversity into evaluation. They were concerned with evaluating their programs from the view of the evaluator, and, the Insurance site also evaluated referring sources satisfaction. BM site participants reported experiencing dissonance in relation to the lack of collaboration and communication with the main referring agency that evaluated their program. They did not believe criteria evaluated by the referring agency represented what was most important to them or to the clients. The BM site also perceived biases that some clients held against them due to an insurance culture influence had the potential to negatively influence evaluation findings. The Insurance site valued communication with their referring sources, as they believed that in providing accurate

information to them, biases that otherwise existed due to incorrect or incomplete information were reduced.

The Insurance and BM sites both reported dissonance in relation to contextual influences they faced, the Insurance site mainly due to organizational politics from outside their program, and the BM site due to being unable to interact as they would have preferred with the referring agency.

The Labour and BPsy sites both focused more in their services on interactions between clients' impairments and their contexts, and the multiple stakeholder perspectives that this incorporated. The Labour site focused on multi-stakeholder perspectives encountered at the work place and the BPsy focuses on an interdisciplinary whole person approach for clients within their organization. Both of these approaches took into consideration multiple stakeholder perspectives, and client functioning in context with the client's particular environment, including physical, cognitive, social, political, and economic.

Evaluation at the Labour and BPsy sites was concerned with physical, cognitive and environmental participation criteria. Both were concerned with timeliness of services being available continuously as needed. Both services involved collaboration and communication among stakeholders. Both programs reported positive adaptation to contextual influences on their programs. Neither of these sites reported dissonance from any source outside their program, within or from outside the organization. Evaluation at the BPsy site was informal, but the program was starting to develop data management systems to eventually use in evaluation. The Labour site had well developed evaluation in

place including criteria client satisfaction and return to work, and were in the process of developing a data management system to evaluate cost savings.

All four sites evaluated program outcomes. The Insurance and BM sites that were predominantly concerned with use of program services focused mainly on evaluation of services related to the individual. The Labour and BPsy sites that were concerned with services as they related to the individuals' interactions with their contexts incorporated the valuing of multiple stakeholders' perspectives into services and evaluation.

The evaluation literature consistently stresses the vital importance of attending to context in evaluation (Alkin, 2013; Chelimsky, 2013; Rog, 2012). The importance of evaluation goals varies among different stakeholders including those from within and outside the program, impacting also the use of findings. Programs often overlap with other programs with important consequences, for example, sharing staff. Program intensions may differ from program realities. All programs have political contexts that impact them, both positively and negatively, and must be recognized by evaluators. The analysis of conceptual themes present at the sites of this study emphasizes the importance of these kinds of program context interactions to evaluation of disability management.

5.7 Inter-relating Explanations of Evaluation

5.7.1 Introduction. At the point of conducting interrelating explanations I was reminded of the literature describing this process. In their study Harry, et al., (2005) reported their analysis had been firmly grounded in extensive, triangulated data that they referred to as "explanations" of their topic. Continuing the constant comparative analyses, they called the process "inter-relating the explanations" (p.6) as they sought clarification among the explanations. They noted they were unable to diagram this

process in two dimensions due to its complexity, and reported analysis had involved

going back and forth, between and among explanations. One finding they identified was

that no single explanation could stand alone to measure complex social processes. Corbin

and Strauss (2009) found that during grounded theory analysis a core conceptual category

often emerged around which other categories could be related to explain variation in

conditions, actions and consequences of the phenomenon studied.

5.7.2 Collaboration. Collaboration was a core concept I hypothesized to be

foundational to disability management and its evaluation. Its presence or absence

contributed to an explanation of *conditions, responses* and *consequences* of disability

management and it's evaluation at the sites.

The presence or absence of collaboration was significant to understanding the

conditions at the sites, including: diversity of stakeholder values, degree of inclusiveness

among stakeholders, power discrepancies among stakeholders, contextual (including

political) influences, purposes for evaluation, evaluation questions being asked,

evaluation processes, and intended evaluation uses.

The presence or absence of collaboration was also central to understanding

responses to conditions at the sites, including: which stakeholders' values were

predominant at the sites and why, how stakeholders with different values were

responding to the presence or absence of inclusiveness, or to power discrepancies, how

programs were responding to contextual and political pressures (i.e., for funding,

influencing decision making, perceived risks due to lack of external recognition of

program accomplishments), and what evaluation criteria and standards were in place.

Collaboration and its presence or absence was also valuable to understand the *consequences* of conditions and responses at the sites. These included: potential impacts of evaluation, integration of diverse stakeholders' values into evaluation, and which evaluation findings were actually used by whom.

Another important finding was the role that collaboration played in differentiating types of disability management evaluation criteria. There were three possible sources of evaluation identified within the programs: evaluation conducted internally by program stakeholders (administrators and practitioners), evaluation conducted by external stakeholders (such as funders), and evaluation that involved collaboration between internal and external stakeholders.

Evaluation criteria were either linear and normative or pluralistic. Both types of evaluation were potentially important to multiple stakeholders. Evaluation criteria that were normative included: return to work and costing savings (both goals of disability management) and timeliness of services (having potential impact on both goals). Normative evaluation involved monitoring data or performance management comparing program outcomes to criteria. Normative evaluation did not explain reasons why targeted outcomes had or had not been met.

Evaluation that involved collaboration and integration of multiple stakeholders' values was pluralistic and included criteria: client functioning and client satisfaction (referring to multiple stakeholders). Pluralistic evaluation had the potential to contribute information on co-existing values and motivations, to explain different views of why programs may or may not be achieving goals, and to contribute information on ways

programs could make improvements to maximize the likelihood of reaching goals that were valuable to multiple stakeholders.

"Evaluation is often most effective when it engages multiple internal and external customers and uses a participatory and collaborative approach" (Russ-Eft and Preskill, 2001, p. 413). Possible benefits of involving multiple stakeholders include: enhancing opportunities to understand the evaluand from different perspectives, making judgments about the worth of the evaluand that include different values, maximizing opportunities for more people to learn about and appreciate the evaluand, and including those who can act upon evaluation results. In addition, participatory evaluation "implies that, when doing an evaluation, researchers, facilitators, or professional evaluators collaborate in some way with individuals, groups, or communities who have a decided stake in the program, development project, or other entity being evaluated" (Cousins and Whitmore, 2007, p. 87). This emphasis on collaboration is also evident in the organizational behavior literature, where collaboration is defined as: "a joint endeavor, involving two or more people working together to complete a task. Collaboration includes teamwork – the coordination of efforts of a group of people around a stated purpose. It involves constructive discussion among team members regarding the common workgroup goal" (Weingart and Jehn, 2009, p. 328).

Communication supports collaboration. Communication had originally been coded as a separate conceptual code based on participants' emphases of the role of communication. The consistent presence of communication and collaboration together informed my decision to incorporate the concept communication into the concept collaboration. Communication has been described as an everyday activity but with

effectiveness that can vary greatly (Cai and Fink, 2009). It is multidimensional, where participants simultaneously take multiple roles (sender, receiver, speaker, listener, group member, audience), are aware of verbal and non verbal messages; and, participate in multiple methods at once (in person, written, technological).

Stakeholders within programs collaborated, and stakeholders from within collaborated with stakeholders external to programs. For example, the BM site collaborated with employers during return to work plans:

"We are in frequent contact with the worker to find out how things are going, what difficulties they are having…and there is always unforeseen problems that crop up, whether that's with scheduling or whether that's something they didn't think they were going to have a problem doing and they're finding it physically difficult and that's where we either talk to them or go in [to the workplace and meet with the worker and employer] and say I need you to go back and look at is it a work technique issue, is there some type of equipment that would be helpful for the person to do their job, or is it something not physically suitable at this time" (BM1).

Another example involved Labour site stakeholders communicating with clients and with employers.

"The communication part is key. So much is communication. Has the return to work coordinator maintained regular contact in the planning stages of the employee [being off work]? Were you involved? Another part is making sure the managers feel they are part of it, that they have a say. You don't want them to feel we are dictating. Were the department's needs part of the return" (L3)

Collaboration was bounded in time and in space, during disability management services and during evaluation.

   5.7.2.1 Collaboration and learning as illustrated by ROLE scores. Where collaboration was present, stakeholders had increased opportunities for process learning from participation in evaluation. Collaboration provided opportunities to receive formative feedback enabling them to modify and improve their services. There were also increased opportunities to develop more meaningful evaluation criteria relating to stakeholders' values.

   The value of collaboration and learning to the participants of this study was further analyzed based on the ROLE responses of the nine participants related to collaboration and evaluation. Five categories were identified as most relevant: 1) collaboration and problem solving, 2) participatory decision making, 3) open and accessible work environment including items related to a work space conducive of employees participating together, 4) dissemination of information referring to the availability of necessary information to employees when needed, and 5) evaluation including items that refer to the value of evaluation, or the value of evaluation if it were to be present. Analysis of the nine participants' scores indicated the ratings for each individual were generally consistent across the five areas, with minor exceptions. The exceptions consistently related to perceptions of participants regarding particular disability management programs or organizations. See Table 18 for a summary of participants' scores on these five ROLE categories.

   Seven out of nine participants rated high collaboration based on two ROLE categories: 1) collaboration and problem solving, and 2) participatory decision making.

Table 18   Participants' ROLE Responses on Categories Related to Collaboration and Learning

ROLE Dimensions	Participants								
	BM 1	BM 2	L 3	L 4	L 5	BPsy 6	BPsy 7	I 8	I 9
Collaboration & Problem Solving	4.0	4.1	4.3	3.3	3.8	4.1	3.7	2.8	3.8
Participatory Decision Making	4.2	4.4	4.0	3.5	3.9	3.5	3.5	2.1	4.1
Open & Accessible Work Environment	4.0	4.5	3.3	3.0	3.0	3.8	3.9	1.5	4.0
Dissemination of Information	3.8	4.2	4.0	3.2	3.6	2.8	3.4	2.6	3.8
EVALUATION	3.3	3.4	4.1	2.5	3.9	4.6	4.0	2.5	3.9

The two Labour site participants whose low scores on 3) open and accessible work environment, varied from their higher scores on 1) and 2), both referred to their program on the ROLE. These responses reflected their belief that the Labour program's physical work space was not supportive of high collaboration. The two BPsy participants whose low scores on 4) dissemination of information, varied from their high scores on 1) and 2), both referred to their organization on the ROLE. These responses reflected their opinion information was not readily available within the organization. The two BM participants low whose low scores on 5) evaluation varied from their high scores on 1) and 2) reflected their low opinion of evaluation of their program that was conducted by the external referring agency.

5.7.3 Contextual influences. Evaluation context has been defined as, "the setting within which the evaluand (the program, policy, or product being evaluated) and thus the evaluation are situated. Context is the site, location, environment, or milieu for a given evaluand" (Greene, 2005, p. 83). The complexities of context have been recognized

within evaluation for decades, without a unified conceptualization (Rog, Fitzpatrick and Conner, 2012). Context is important to evaluation because of, "its impact on evaluation plans, methods, implementation, and use, few develop the construct in depth" (Fitzpatrick, 2012, p. 8). Rog (2012) proposed five important components of context to consider: the problem being addressed; the intervention being examined; the broader environment or setting; the evaluation context; and, the decision making context. Rog (2012) suggested seven dimensions to consider within the five components of context analysis: physical, organizational, social, cultural, traditional, political and historical.

Programs continue to develop "in response to changed conditions and new knowledge. Such changes do not mean that what was done before was ineffective; rather, it means that as the world changes, the program must change" (Patton, 1996, p. 134). Adaptation to contextual influences was a main theme among the sites of this study, which according to participant interviews occurred because programs needed to adapt to survive or grow.

All programs face dynamic contextual influences and the sites of this study were no exception. I hypothesized that understanding the contexts within which disability management programs function, and understanding how programs adapt to contextual influences, is essential to an explanation of variability in disability management and disability management evaluation. The five component framework for context analysis recommended by Rog (2012) will be discussed in relation to the sites of this study.

5.7.3.1 The problem being addressed by evaluation. The programs of this study provided services related to a theme of employability and facilitating durable return to work if possible. Services were multidisciplinary. Disability management addresses a

wide range of potential problems including: physical, psychological, social, organizational, political, financial, etc. It focuses on the impairment and its contextualization.

5.7.3.2 The intervention being examined. Disability management involves managing interactions between impairments and their environments to overcome barriers (Smart, 2001). The sites of this study provided services at different stages of disability management. The BPsy program provided counselling and career exploration during early stages of acute conditions, with a primary goal of helping clients learn to manage their disabilities. The BM program conducted functional and job site assessments to develop return to work plans, with a primary goal of providing accurate information that could facilitate return to work. The Labour program implemented return to work plans that accommodated employees, with a primary goal that the work be suitable and meaningful. The Insurance program contributed investigative services to obtain evidence to validate client functioning, with a primary goal of ensuring accident fund expenditures were appropriate.

5.7.3.3 The broader environment or setting. On a wider contextual level the disability management programs were influenced politically by government laws, including: freedom of information; employers' duty to accommodate disabled workers; rights of persons with disabilities to physical access; and collective agreements. The programs each faced making decisions in response to contexts such as which conditions for funding matched their political orientations and were acceptable to them, and which were not. Broader environmental influences within disability management that also

required consideration included "systemic factors arising from health care, compensation and other social systems" (Schultz, 2005, p. 30).

   5.7.3.4 The evaluation context. Program evaluation was conducted by different stakeholder groups: internally by program staff, externally by funders, or through collaboration between internal and external stakeholders. The physical settings of the four programs of this study were not threatened due to political or financial reasons, however, each program was aware that they continually relied on ongoing funding. The programs strived to maintain positive profiles within their organizations and to the public, and to be of value to funders and consumers of their services.

   5.7.3.5 The decision making context. The primary decision makers who used evaluation findings included program funders, administrators and practitioners. It was recognized that at the workplace diverse stakeholders (employers, unions, workers, employee health professionals etc.) could learn from evaluation findings about how to better integrate workers back into the system. Although this would be important to understand, the data had nothing to say about this.

   5.7.4 Data management systems. The four organizations included in this study relied on electronic data management systems for storing information, and were developing data management systems with the use of technology that they planned to access for future monitoring, performance management and evaluation. This emphasis resonates with the views that data management systems can make evaluation more effective through evaluability assessments (Wholey, 2013) and can lead to the use of a broad range of data in evaluation (Chelimsky, 2013).

In general, computer and communication information technologies (hardware and software) used to process, store, retrieve, and transmit information in electronic form are increasingly being advanced and miniaturized, "liberating users from *past limitations* of space and time so that they can use computers anywhere and anytime" (Alavi and Yoo, 2009, p. 597). Organizations are leveraging technological capacities to make improvements "in the support of communication and collaboration processes" (p. 600). Maximum potential is realized within organizations that have a culture supportive of learning: where large scale efficiencies can result from centralized data bases, decision making can be enhanced through electronic communications, virtual work groups can interact, new models of data analysis and of business can be applied, and where digital devices are able to communicate directly with each other. Effective learning organizations are able to transfer knowledge quickly and efficiently throughout the organization (Preskill, 1994).

The data that was being gathered and the electronic data management systems that were being developed at the sites of this study were in preliminary stages. Data was not being maintained with specific evaluation questions in mind. The data was reported to include numbers of incoming cases, case assignments, outcomes and costs, but not specifics explaining program processes. Data that was being gathered at these initial stages was more suitable for performance management or auditing, to compare outcomes to objectives. Although participants hoped that eventually information systems could be useful for evaluation, in their present state they did not explain why performance was the way it was, and therefore, were not useful for problem solving, decision making, or resource allocation.

According to the ROLE results overall, the four sites in this study valued learning and were reported to have leadership that supported learning. However, the programs had limited ability to access important information. With support for learning at the sites and preliminary data management systems being established, there was potential for data management systems to eventually be developed that could be used to complement evaluative investigations and become sources of dissemination of important information.

5.7.5 Primary disability management evaluation criteria. There were five primary evaluation criteria used in disability management evaluation, although all were considered only at one site of this study, the Labour program.

(1) Client functioning: In cases where impairments were physical, all four sites recognized how successful disability management related to functioning of the individual. The Labour and BPsy sites particularly recognized the significance of client functioning and the interactions of impairments and environments.

(2) Client satisfaction: Client satisfaction surveys were administered at three of the four sites, to clients at the BM site, to clients and clients' managers at the Labour site, and to prospective referral sources within the organization at the Insurance site, but results were interpreted and used in very different ways. The client satisfactions survey at the BM site was summarized into one numerical rating. Client satisfaction at the Labour site was analyzed by the Joint Committee and findings provided as feedback to the Labour program coordinators, who modified their services in response. The Insurance site reviewed survey findings to modify their services.

(3) Cost savings: Cost savings was a criterion only for the Insurance site. At the organizational level the Accounting Department estimated financial cost savings resulting from Insurance site services. At the organizational level of the Labour site a PeopleSoft data management system was being developed to eventually be used for evaluation of cost savings and other performance management and evaluation purposes. Minimization of social costs to clients during disability management was implicitly valued by all four sites but not evaluated by any.

(4) Timeliness: Timeliness of services related to established standards was a criterion at the BM and Insurance sites. Both programs tried to meet established standards. The Labour and BPsy sites provided services as needed, their goals being continuous availability, but they did not evaluate this.

(5) Return to work: The Labour site was the only site that formally evaluated return to work outcomes. Within all the sites there was either explicit or implicit a goal of contributing to appropriate vocational outcomes: avoiding the need to go off work due to disability; the current or eventual return to work of clients; or, client adjustment to not being able to return to work following illness or injury.

5.8 Evaluation Theory

5.8.1 Situating disability management evaluation at the sites within Alkin & Christie's evaluation theories.

5.8.1.1 Biomedical site. Evaluation conducted externally by the program's main referring agency involved a *Key Performance Indicator* measuring four criteria: time from referral to first client contact; report turnaround time; client satisfaction; and number of *Client Satisfaction Surveys* returned. Each criterion was measured on a scale

from 1 to 10 including the *Client Satisfaction Survey* a questionnaire that the referring agency reduced to one numerical rating. BM site scores on the four criteria were then compared to scores of all the organizations the referring agency contracted with, presumably to be considered in making decisions of which organizations' contracts to renew. Evaluation conducted by the main referring agency was primarily concerned with the methods side of evaluation use.

Evaluation of the BM program was also conducted internally by the BM program staff including: exit interviews with clients; self evaluation of reports; employer and worker reviews of job demand analysis reports; and self evaluation of measuring apparatus. The main priority of the staff was to ensure usefulness and accuracy of their services and reports. The BM staff considered the opinions of multiple stakeholders where possible. Evaluation conducted by BM staff was primarily concerned with evaluation use.

5.8.1.2 Labour site. Formative evaluation occurred throughout implementation of each return to work plan, where the program coordinator, client and client's manager met regularly to provide feedback, modifying services in response. At the conclusion of services *Client Satisfaction Surveys* were administered to every worker, and *Manager Summary Questionnaires* were administered to every manager. Survey results were analyzed by the Joint Committee overseeing the program, and passed on to the program coordinators. The Labour site was operated by the joint labour-management committee with primary commitment of being of use to both labour and management, assisting every client to stay at work, return to work, or adjust to not being able to work. Evaluation at this site was primarily concerned with the value side of evaluation use.

5.8.1.3 Biopsychosocial site. The BPsy site was not formally evaluated, but was informally evaluated in several ways. The team leader met with each program counsellor to consult on individual cases. Program counsellors were each administered a questionnaire pertaining to organizational standards. Formative feedback on case management was shared among multidisciplinary staff of the organization. Client services included setting and evaluating smart goals. Services at this program and organization emphasized sharing professional and non professional stakeholder perspectives. Evaluation at the BPsy site was primarily concerned with the valuing side of evaluation use.

5.8.1.4 Insurance site. The program supervisor evaluated every case to confirm services were complete and reports met established standards. Criteria and standards were established by the supervisor. Report turnaround time was compared to standards set for each type of service. Cost savings were estimated by the accounting department of the organization, using a formula created to compare actual costs on client claims where the Insurance site services had been accessed, to estimate costs had the services not been accessed. Estimated costs were based on costs for other similar cases where services had not been accessed. The Insurance site administered a *Client Satisfaction Survey* to prospective referral sources from the claims division of the organization. Evaluation at the Insurance site was primarily concerned with evaluation use.

5.8.2 Sources of evidence of evaluation use at the sites. Based on (1) interview evidence provided by participants, and (2) situating evaluation at the four sites according to Christie and Alkin's (2013) evaluation theories, disability management evaluation at the four sites of this study was primarily concerned with the use of evaluation findings,

and secondarily concerned with pluralistic values of multiple stakeholders. Use of evaluation findings emanates from the social accountability root of evaluation, to improve social programs. Valuing emanates from the epistemological root, where objective knowledge is based prescriptively on the opinions of evaluation experts, and subjective knowledge is based on pluralistic values of multiple stakeholders.

Evaluation at all four disability management sites emphasized timeliness of services and client functioning as important evaluation criteria. The BPsy and Labour sites integrated multiple stakeholder values as a secondary emphasis, whereas evaluation at the BM and Insurance sites emphasized evaluators' priorities relating primarily to outcomes.

Evaluation was aimed at improving services. Evaluation at the sites was instrumental, contributing to decisions. Programs almost exclusively responded to evaluation findings by modifying their services trying to meet ideal standards (refer to Tables 9, 10, 11 and 12). Other uses of evaluation findings were justification of program funding and planning future program goals. This evidence demonstrated that disability management programs were responsive to feedback regarding their performance, to improve services. This social accountability emphasis is evidence disability management programs would be expected to embrace increased evaluation and organizational learning to improve social programs. ROLE results supported that programs valued evaluation and learning at the sites.

5.8.3 Types of evaluation use at the sites. Evaluation was used in multiple ways. This included instrumental use, both process use involving using evaluation findings for

formative improvements to programs, and conceptual use involving stakeholders' increased understandings of the potential benefits of evaluation and learning to programs.

There were instances of symbolic use, including imposed use where funders influenced program administrators' decisions, for example when the BPsy program eliminated a counselor position that would have been dependent upon results oriented evaluation, a purpose they did not support for their program. There was also evidence of symbolic mechanical use to meet imposed evaluation requirements, such as when BM program administrators complied with evaluation required by the main referring agency, but found the evaluation meaningless, and only cooperated to secure future referrals.

Primary linear evaluation criteria, return to work, cost savings, and timeliness of services, related primarily to performance management. Return to work and cost savings were goals of disability management, and timeliness had potential to impact success in achieving both goals. These criteria had the potential to be significant to multiple stakeholders, as they were valued by funders, workers and employers. However they did not explain reasons why outcomes were what they were.

Diverse stakeholder values were integrated into evaluation in part through the use of client satisfaction surveys and exit interviews, providing information that did contribute to explanations of how program operation may be impacting outcomes, suggesting means for improvements and use. Operationalization of the complex multidimensional concept client functioning in terms of not only medical impairment, but also abilities and contextual influences, did enable case specific explanations of how an individual could overcome barriers.

Evaluation that was concerned with understanding contextual influences and integrating multiple stakeholder' interests had the potential to leverage new insights to reach program goals, and to improve future programs, enlightenment use.

5.8.4 Working evaluation logic at the sites of this study. To further a theoretical understanding of disability management evaluation, the logic of evaluate inquiry at the four sites is examined. Depending upon the purpose of an evaluation, different procedures are followed giving rise to different types of evidence, evaluative claims and justifications of those claims (Smith 1995). Fournier (1995) proposed that evaluators justify their conclusions and claims by following both a general and a working logic. Evaluators identify a problem, define or operationalize a phenomenon of interest in relation to the problem that is the object of evaluation, identify criteria and kinds of evidence to examine, and justify the claims that are made.

Four main approaches to working logic of evaluation are: connoisseurial (based on qualities identified by an expert); pluralistic (based on values of stakeholders); causal (treatment-outcome relationships); and consumer (properties of a functional product) (Fournier, 1995; 2005).

5.8.4.1 Biomedical site. Elements of the BM paradigm include presumption of a medical condition that is a physical pathology and symptoms directly proportionate to the pathology. Services at the BM site included: objective measurement of client functional capacity, job demands analyses, and development of graduated return to work plans. The BM site was evaluated by the main referring agency and by BM program staff.

The main referring agency followed a consumer working logic. Specification of the problem for the main referring agency was a need for timely services and client

satisfaction information. The phenomenon of interest was timely BM site services and submission of completed *Client Satisfaction Surveys*. Questions posed focused on four criteria on the *Key Performance Indicator*: time from referral to first client contact; report turnaround time; client satisfaction; and number of *Client Satisfaction Surveys* returned. Claims about the value of services were based on: BM site scores compared to: (1) standards on the four criteria of the *Key Performance Indicator;* and, (2) compared to scores of other providers that the referring agency contracted with on the same criteria.

Staff at the BM site conducted evaluation that followed a consumer working logic. Their problem was a need for ongoing new referrals to be able to maintain their business, and the phenomena of interest were accurate services that were timely and of value to referring sources. Evaluative questions included: client satisfaction based exit interviews, report quality; job demands' reports accuracy; and measuring equipment accuracy. The BM site staff claimed high client satisfaction, accuracy and timeliness indicated high value of their services.

Although the evaluation of BM site services followed a consumer approach working logic, the main referring agency and the BM site staff identified different evaluation criteria. Different selection of criteria attested to the main referring agency's focus on measuring outcomes, compared with the BM site staff's focus on more descriptive evaluative perspectives of relevant stakeholders. Neither of these two sources of evaluation criteria alone incorporated a complete understanding of the BM site services, and how the program related to its context. A true understanding of the BM site services would incorporate evaluation criteria important to the diverse stakeholders the program had the potential to impact. Evaluation would then minimize the risk of omitting

critical criteria, and maximize the likelihood of contributing information that was most useful.

      5.8.4.2 Labour site. Elements of the labour paradigm include presumptions that: the impairment is best managed within a workplace context; workers and employers needs can be complementary; employers are responsible to accommodate workers with suitable work; and diagnosis is secondary to matching job demands to functional capacities. Services at the Labour site included assisting workers to stay at work, return to work or adjust to not being able to work due to illnesses or injuries. The labour site was evaluated by a union-management Joint Committee.

      The problem for the Labour site program was a need for workers and management to provide information to the program that would assist the program to modify services that would increase the ability to achieve goals. The phenomenon of interest during return to work programs was iterative communications among the worker, worker's manager and program coordinator. Indicators of services were the *Client Satisfaction Survey,* a *Manager Summary Questionnaire* and data on return to work outcomes. Questions posed were meant to provide: feedback during return to work plans, completion of the worker survey and manager questionnaire, and document return to work outcomes. High satisfaction ratings on surveys and questionnaires and high return to work statistics were evidence of program success.

      Evaluation of Labour site services followed a consumer approach working logic. The Labour site needed information on their services so they could make modifications to maximize successful outcomes and to justify the value of the program.

5.8.4.3 Biopsychosocial site. Elements of the BPsy paradigm include presumption of: an interdisciplinary whole person approach; a conceptual distinction between impairment and disability; and impairment does not reliably predict disability. Services at the BPsy site included multi-disciplinary and inter-disciplinary professional services for adults who had sustained catastrophic injury or illness. The BPsy site was not formally evaluated, but program counsellors received informal feedback on their interventions with individual clients from their team leader and from professionals from other programs within the organization who were working with the same clients.

Evaluation of BPsy site services illustrate multiple working logics and can be explained from perspectives of three approaches: connoisseurial, pluralistic and consumer. Multi-disciplinary professionals contributed specialized expertise to develop and implement rehabilitation interventions for individual clients and evaluate outcomes. The problem from the connoisseurial approach was to develop rehabilitation interventions targeted at individual clients. The phenomenon of interest were treatment interventions developed by experts, and evaluation questions addressed whether the interventions achieved the intended goals, and claims were made about the value of the interventions.

The organization was committed to a pluralistic approach because what multiple stakeholders valued was important, and the assumption that clients who set their own goals would be most motivated to succeed. The problem was maximizing the opportunities for clients to succeed at reaching their rehabilitation goals. The phenomenon was inter-disciplinary services available to clients supporting them to set

and achieve multiple goals. Questions focused on what inter-disciplinary service qualities had the intended impact, and claims related to what service qualities were important.

At the BPsy site individual programs were starting to build data bases that could provide evidence to evaluate service outcomes a strategy typical of a consumer approach. The problem was providing services for clients who experienced accidents or illnesses to successfully manage their resultant impairments. Interests included maintaining rapport, vocational counselling and exploration, referrals and counselling support. Evaluation questions were posed regarding what vocational and inter-disciplinary assistance would best facilitate clients' adjustment and managing of impairments. Claims were made that vocational services available on demand to clients contributed best to their short and long term vocational adjustment and their employability.

5.8.4.4 Insurance site. Elements of the insurance paradigm include presumption that: people who anticipate secondary gain are likely to magnify disability; and objective medical proof of impairment and disability can be proven. Services at the Insurance site included investigations into possible fraud. Services provided on individual insurance cases were evaluated by the program supervisor. Overall cost savings of the program were estimated by the accounting department of the organization.

The problem was protection of the integrity of the accident fund, ensuring funds were allocated appropriately. The phenomenon was investigator interventions to obtain accurate information when miscommunication or lack of complete information resulted in inability to confidently allocate insurance funds appropriately. Questions involved acquiring accurate information on individual insurance client impairment. Claims were made that investigations provided accurate information, accurate information resulted in

reduced biases and appropriate allocation of the accident fund, and overall claim cost savings resulted from the insurance site services.

Evaluation of Insurance site services followed a consumer working logic. The program supervisor evaluated report turnaround time, completion of services and quality of reports in relation to established standards, and cost savings resulting from the program were evaluated by the accounting department of the organization.

5.8.4.5 Summary of working logic in program evaluation at the four sites. Disability management evaluation at the sites predominantly followed a consumer approach to working logic. Consumer working logic has a locus of value based on properties of the products or services provided, questioned whether criteria and standards had been met, and integrated data into a claim of merit or worth. The exception to this was at the BPsy site where there were no formal evaluation of services, and informal evaluation followed multiple working logics: connoisseurial, pluralistic and consumer.

Fournier (1995) stressed the importance of defining the phenomenon because it reflects the locus of values, selection of criteria and what kinds of evidence are sought to justify what warrants and make what claims. This is relevant to developing a theory of disability management evaluation. While evaluation that is meaningful emphasizes program outcomes (return to work, cost savings or timeliness of services), stakeholder values and expert interventions are also important. The potential value of these simultaneous working logics was evidenced at the BPsy site, where if programs were formally evaluated there could have been the potential for evaluation to impact to multiple stakeholders.

5.9 A Tentative Grounded Theory of Disability Management Evaluation

The conceptual groupings identified within DM program evaluation are illustrated in Figure 7, and not surprisingly this figure illustrates common components in most - program evaluation.

Table 19 illustrates how theoretical orientations varied across the paradigmatic sites since there are notable differences. For example, the BM site especially emphasized the usefulness of evaluation based on appropriate data collection and analysis procedures, while the BPsy site emphasized usefulness fostered by the inclusion of multiple stakeholders, including experts and clients. The table also illustrates that only one site, the Labour disability management program, used all five primary evaluation criteria, the Insurance site used four, the BM site focused on client functioning, client satisfaction and timeliness, and the BPsy site focused on client functioning and timeliness only.

Figure 8 summarizes the orientation of DM evaluation by drawing on Alkin & Christie (2004) and Fournier (1995) to illustrate the fundamental nature of DM evaluation theory and its logic. This figure illustrates that DM evaluation is grounded primarily in Alkin & Christie's "use" tree branch and therefore seems to have a social accountability orientation. Consistent with this emphasis on "use" DM evaluation adopts a predominately consumer oriented working logic, although the BPsy site was more complex than the others incorporating expertise and multiple stakeholder views and so drew on the logic of a connoisseurship and pluralist evaluation approaches as well. A consumer approach emphasizes evaluation that focuses on the properties of a functional product—in this case, the product is DM program services and the properties are reflected in the five primary evaluation criteria identified. The combination of this

Figure 7 Conceptual Groupings within Disability Management Evaluation
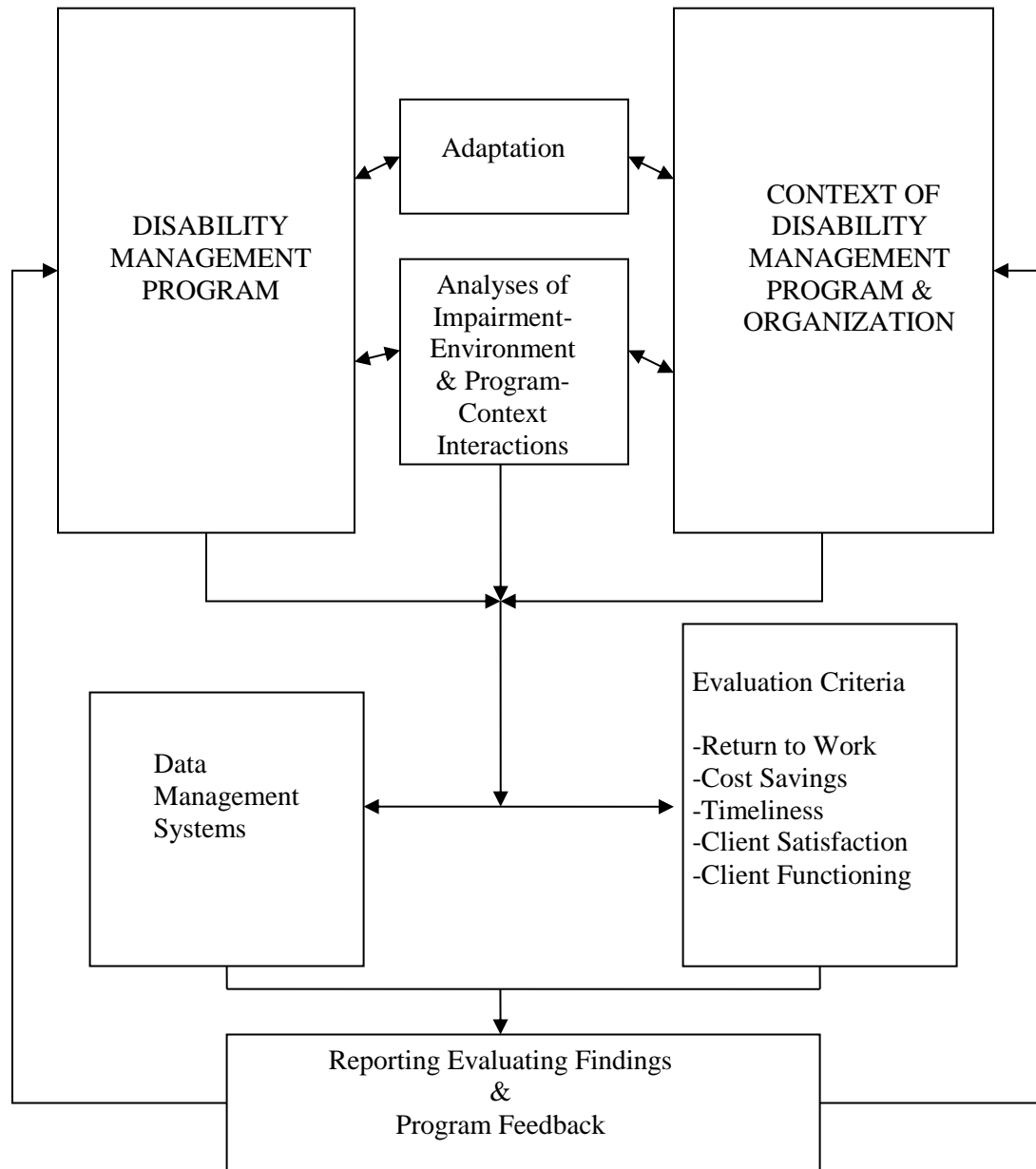
Table 19   Theoretical Orientations for DM Paradigms

Paradigm/Site	Dominant Evaluation Theory	Working Logic Approach	Evaluation Criteria
BM	Use (methods)	consumer	client functioning client satisfaction timeliness
Insurance	Use	consumer	client functioning client satisfaction timeliness cost savings
Labour	Use (valuing)	consumer	client functioning client satisfaction timeliness cost savings return to work
BPsy	Use (valuing)	consumer connoisseurship pluralist	client functioning timeliness

emphasis on use and a consumer orientation orients DM evaluation primarily to program

outcomes reflected in two key common evaluation criteria (client functioning and

timeliness of services). Nonetheless, there is a less prominent but still important

orientation to multiple stakeholder involvement in evaluation that was especially apparent

in the BPsy site evaluation practice.

Figures 7 and 8 and Table 19 present a visual summary of the theory of disability

management evaluation being developed in this study. Figure 8 shows how evaluation is

primarily concerned with use of findings, but secondarily for some programs is also

concerned with multiple stakeholders' values.

This more abstract analysis and interpretation is grounded in the data of this study

that revealed a differentiation between evaluation that measured program outcomes

(where criteria were sometimes critical to limited stakeholders such as funders) versus

evaluation that incorporated the interests of multiple stakeholders resulting in findings with more comprehensive potential for a wider impact among disability management stakeholders.

Figure 8 Disability Management Evaluation Theory and Logic

DM program services ⟷ program adaptation to

organization & context

↓

Evaluation emphasizes **USE**

↓

CONSUMER-oriented evaluation

Outcomes based evaluation   stakeholder values
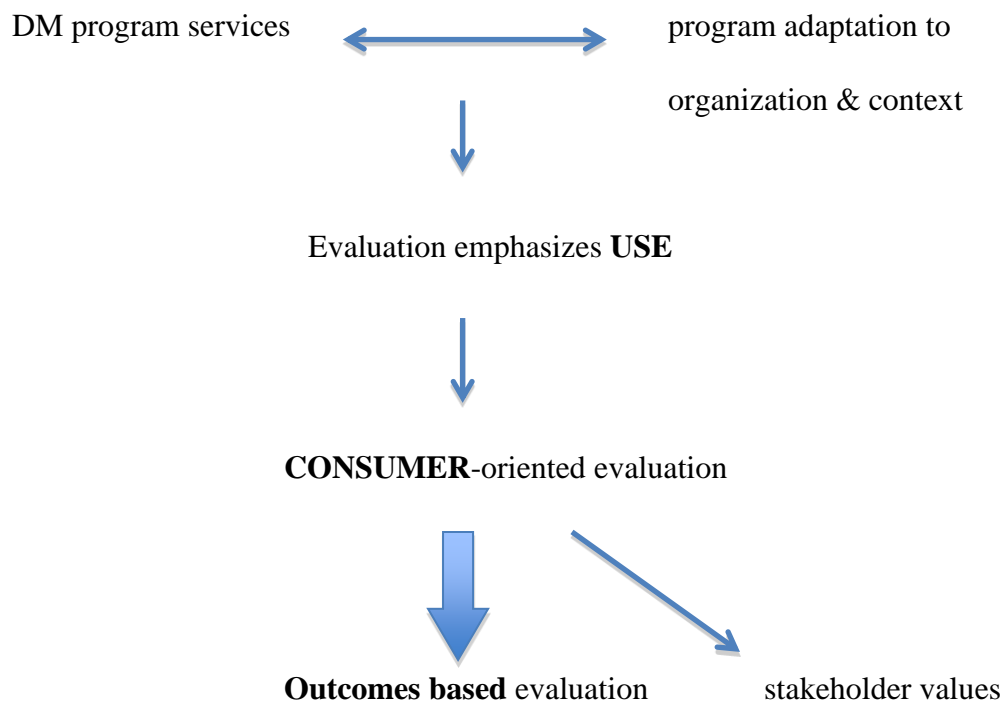
5.10 Answering the Research Questions: A Summary

This study started with an exploration of evaluation in disability management to develop a theory of practice, an area that has not been well documented in the literature. Four research questions guided this study.

5.10.1 What is the extent and nature of evaluation practice within the disability management programs? Three of the four sites of this study formally evaluated their programs but evaluation varied, where some sites focused on program outcomes, including return to work, timeliness and cost savings, as well as client satisfaction. Other sites attended more to contextual interactions and multiple stakeholder perspectives. Evaluation at these latter sites was enriched by more in depth information on programs' adaptation to contextual influences, and how programs managed impairment environment interactions. Evaluation at all of the sites reflected an emphasis on the utilization of evaluation and a consumer or product driven working logic.

5.10.2 How does disability management evaluation practice vary depending on whether the organization is a learning organization? Sites that reported greater collaboration and communication reported high learning within the program, throughout the organization and from evaluation. Programs that reported little collaboration and communication did not consistently report high organizational learning.

5.10.3 How does disability management evaluation reflect diversity and cultural constructions? Disability management evaluation at the sites was primarily concerned with use, grounded in social accountability for learning from evaluation to improve service outcomes. Evaluation was secondarily concerned with multiple stakeholder perspectives. Although the participants described sensitivity to cultural influences at the sites where these influences did exist, the data did not reveal much about diversity, cultural issues or cultural competence in evaluation at the sites.

5.10.4 What evidence is there that disability management evaluation is grounded in a particular paradigm of disability and return to work? This study illustrated that evaluation was grounded in particular paradigms, even though there are overarching similarities. The BM and Insurance sites emphasized program outcomes based on criteria set by the funder, supervisor or staff. Evaluation practice mimicked the assumptions of the DM paradigms and focused on measurable and objective outcomes. The Labour and BPsy sites focused on evaluation that incorporated multiple stakeholders, and included consideration of client context interactions. The Labour site practiced formal evaluation taking into consideration all five primary evaluation criteria, covering the greatest breadth among the sites.

5.11 Analysis of the State of Disability Management Evaluation in Relation to the Evaluation Field

At the onset of this study the limited literature on disability management evaluation indicated evaluation had been primarily normative, focused on summative measures related to economic outcomes such as return to work rates, incidence and duration of absence, lost productivity and cost reduction. These criteria reflected the values of employers and funders. There was a lack of evidence reflective of pluralistic values of multiple stakeholders, and no explanations of why programs were functioning as they were.

This study has shown that evaluation in the field is concerned with a consumer logic and use of findings. These purposes were primarily focused on the use of evaluation findings for program improvement, and limited to performance management. Programs were shown to be secondarily concerned with pluralistic values of multiple stakeholders.

This study identified five primary evaluation criteria that are each likely valued pluralistically.

Situating observed disability management evaluation practices within the greater evaluation field revealed that the questions and problems that disability management evaluation is currently facing are similar to the questions and problems the field of evaluation encountered during its successful evolution, for example expanding away from a mainly scientific approach of cause and effect to more descriptive explanations of programs.

This research revealed a notable lack of concern within disability management evaluation regarding evaluation that prioritizes mainly monitoring and performance management to assess existing program outcomes and accountability. There was little evidence that programs were concerned with other purposes such as empowerment, being transformative, or developing long term policy toward improved disability management programs. The types of values and purposes that would go along with these latter concerns would likely overlap with and be representative of greater appreciation for learning within disability management programs, and perhaps a wider existence of organizational learning. Participants did report valuing evaluation and learning, indicating perhaps developments in those areas related to disability management.

5.12 Contributions of This Study

This exploratory study is a timely and relevant response to the paucity of published literature on disability management evaluation. It offers to the field of disability management a systematic review of evaluation practices that extends what was

previously documented in the literature, with potential to improve future evaluations and facilitate learning.

Research on evaluation practices has the potential to contribute to understanding fundamental evaluation issues, for example through descriptive theory, researching "what evaluation looks like, under different conditions, and what kinds of consequences result from various approaches to evaluation" (Mark, 2008, p. 114). Mark recommends more research on evaluation to stimulate further contributions to an evidence base for evaluation practice, and for classifying, comparing and synthesizing findings that result. Contributing insights into evaluation knowledge, purposes, values, practices, uses and impacts within one industry, disability management, has the potential of raising awareness of applicability in other similar situations, and supplies information that may be applicable to longer term development of best practices.

5.13 Limitations of This Study

I acknowledge that being the sole researcher conducting this study, including doing all the coding myself, was a limitation that had potential to introduce biases. I attempted to counter this possibility by practicing reflexivity, addressing my assumptions, being systematic and maintaining clear records that I returned to repeatedly during analyses. The samples of sites and number of participants were small, and although I did believe that concepts were saturated, doubling each of the numbers would have introduced another interesting dimension of triangulation. Participants were limited to administrators and practitioners from each site, which left out important stakeholders including clients who may have introduced unique insights into evaluation and learning at the sites. However, the participants that were included were primary stakeholders given

they were the practitioners of evaluation. Use of the ROLE did not conform to the

intended purpose, but I believe my use was an appropriate application, and the findings

were important to the analysis. There were limited evaluation documents, and apparently

no evaluation reports available. All instruments used for evaluation were obtained. This

explorative study provided a glance into current evaluation practices within disability

management from a perspective not previously viewed.

5.14 Future Directions

This study identified five primary evaluation criteria, and explained how four of

those criteria could be evaluated in relation to established standards: return to work, cost

savings, timeliness of services, and client satisfaction (this criteria was evaluated through

multiple item surveys and interviews). The fifth evaluation criterion, client functioning,

was more complex to operationalize.

Contextualized client functioning is recognized in the literature and the field as

crucial to disability management (Smart, 2001). Client functioning is no longer

recognized as an impairment or pathology. Its elements include the biomedical,

psychological, social, organizational, political, legal and economic, etc. The World

Health Organization's (WHO, 2001) developed an *International Classification of*

Functioning, Disability and Health that involves a systematic coding to understand health

related states, outcomes and determinations. This classification shifted focus from the

impairment to body functions, abilities and environmental factors.

This study identified how important client functioning was to disability

management, particularly demonstrated at the Labour and BPsy paradigm sites. However,

much has been left to be discovered regarding how client functioning could best be

evaluated. Evaluation has the potential to disaggregate assumptions, and analyze client functioning in context, revealing how links, values, biases, conditions, responses and outcomes operate. Knowing how underrepresented those with impairments are in the labour market and the cost they endure as a result, and being aware of the potential for evaluation to contribute in some way to their empowerment, incites interest, curiosity and hopefully too, has heuristic significance as a suggested future area for research.

5.15 Conclusions

The findings of this study have the potential to stimulate future research on evaluation, enhance evaluation capacity in the field of disability management, and foster appreciation among professional disability management practitioners of the role that program evaluation can play. Little has been previously documented about evaluation practices in disability management. Development of a theoretical framework offers a preliminary road map, where administrators and practitioners of disability management programs have an opportunity to consider current evaluation practices, perhaps identify with one or another of the approaches described, and learn vicariously from the struggles and achievements participants have experienced to make informed decisions of how to best apply resources in conducting evaluations within their own situations.

References

Akabas, S. H., Gates, L. B., & Galvin, D. E. (1992). *Disability Management A Complete System to Reduce Costs, Increase Productivity, Meet Employer Needs, and Ensure Legal Compliance.* New York: AMACOM.

Alkin, M.C. (2013). Comparing evaluation points of view in M.C. Alkin (Ed.) *Evaluation Roots: A Wider Perspective of Theorists' View and Influences,* (p. 3-10). Thousand Oaks, CA: Sage Publications, Inc.

Alkin, M.C. (2013) Context-sensitive evaluation. In M.C. Alkin (Ed.) *Evaluation Roots: A Wider Perspective of Theorists' View and Influences,* (pp. 283-303). Thousand Oaks, CA: Sage Publications, Inc.

Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree in M. C. Alkin (Ed.) *Evaluation Roots: Tracing Theorists Views and Influences,* (pp. 12-65). Thousand Oaks, CA: Sage Publications, Inc.

Alavi, M., & Yoo, Y. (2010). Use information technology for organizational change. In E.A. Locke, (Ed.) *Handbook of Principles of Organizational Behavior (2nd ed.), (*pp. 595-614). Chichester, WS: John Wiley & Sons, Ltd.

American Evaluation Association. (2012). *American Evaluation Association Statement on Cultural Competence in Evaluation.* Retrieved June 4, 2012, from http://www.eval.org/ccstatement.asp

Appleton, J.V. (1995). Analyzing qualitative interview data: addressing issues of validity and reliability. *Journal of Advanced Nursing, 22,* 993-997.

Barber, L. (Ed.). (1998). *The Canadian Oxford Dictionary.* Toronto: Oxford University Press.

Beer, M. (2009). Sustain organizational performance through continuous learning, change and realignment. In E.A. Locke, (Ed.) *Handbook of Principles of Organizational Behavior (2nd ed.), (*pp. 537-555). West Sussex, UK: John Wiley and Sons, Ltd.

Boris, E.T., & Winkler, M.K. (2013). The emergence of performance measurement as a complement to evaluation among U.S. foundations. In P.R. Brandon (Ed.), *New Directions for Evaluation, 137,* 69-80.

Bringer, J.D., Johnston, L.H., & Brackenridge, C.H. (2010). *Using computer-assisted qualitative data analysis software to develop a grounded theory project.* Retrieved October 14, 2010, from http://www.fmx.sagepub.com/.

Bruyere, S. M., & Shrey, D. E. (1991). Disability management in industry: A joint labor-management process. *Rehabilitation Counseling Bulletin, 34*, 227-242.

Bryant, A., & Charmaz, K. (2007). Grounded theory in historical perspective: An epistemological account. In A. Bryant & K. Charmaz (Eds.), *The Sage Handbook of Grounded Theory, (*p. 31-57). London: Sage Publications.

Cai, D.A., & Fink, E.L. (2009). Communicate successfully by seeking balance. In E.A. Locke, (Ed.) *Handbook of Principles of Organizational Behavior (2nd ed.), (*pp. 425-444). Chichester, WS: John Wiley & Sons, Ltd.

Charmaz, K. (2000). Grounded theory objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research (2nd ed.)*, (pp. 509-535). Thousand Oaks, CA: Sage.

Charmaz, K. (2005). Grounded theory in the 21st century: Applications for advancing social justice studies. *The Sage Handbook of Qualitative Research (3rd ed.)*, (pp.

507-535). Thousand Oaks, CA: Sage.

Chelimsky, E. (2013). Evaluation purposes, perspectives, and practice in M.C. Alkin

(Ed.) *Evaluation Roots: A Wider Perspective of Theorists' View and Influences,*

(pp. 267-282). Thousand Oaks, CA: Sage Publications, Inc.

Christie, C.A. (2007). Reported influence of evaluation data on decision makers' actions.

American Journal of Evaluation, 28, 8-25.

Christie, C.A., & Alkin, M.C. (2013). An evaluation theory tree in M.C. Alkin (Ed.)

Evaluation Roots: A Wider Perspective of Theorists' View and Influences, (pp.

11-57). Thousand Oaks, CA: Sage Publications, Inc.

Conner, R.F., Fitzpatrick, J.L., & Rog, D.J. (2012). A first step forward: Context

Assessment. *New Directions for Evaluation*, *135*, 89-105.

Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research 3e*. Thousand Oaks,

CA: Sage Publications, Inc.

Cousins, J.B. & Whitmore, E. (2007). Framing participatory evaluation. In S. Mathison

(Ed.), *New Directions for Evaluation, 114,* 87-105.

Crotty, M. (1998). *The Foundations of Social Research.* Thousand Oaks, CA: Sage

Publications Inc.

Crotty, M. (2003). *The Foundations of Social Research*. Thousand Oaks, CA: Sage

Publications.

Cummings, H., & Paulmer, H. (2010, May). *The Evaluation Framework: An Approach to*

Assessing and Reporting Outcomes and Impacts. Presented at the Canadian

Evaluation Society Annual Conference, Victoria, BC.

Currier, K. F., Chan, F., Berven, N. L., Habeck, R. V., & Taylor, D. W. (2001).

Functions and knowledge domains for disability management practice: A Delphi

study. *Rehabilitation Counseling Bulletin, 44*, (133-143).

Curtis, J., & Scott, L. R. (2004). Integrating disability management into strategic plans.

American Association of Occupational Health Nurses Journal, 52, (298-301).

Dunn, P. (2001). Proprietary rehabilitation: Challenges and opportunities in the new

millennium. *Work, 17*, (135-142).

Dyck, D.E.G. (2002). *Disability Management: Theory, Strategy and Industry Practice

(2ⁿᵈ Ed.).* Markham, ON: Butterworths Canada Ltd.

Dyck, D.E.G. (2009). *Disability Management: Theory, Strategy and Industry Practice

(4ᵗʰ Ed.).* Markham, ON: LexusNexis Canada Inc.

Eakin, J. M., Clarke, J., & MacEachen, E. (2002). *Return to work in small workplaces:

Sociological perspectives on workplace experience with Ontario's early and safe

strategy.* [Electronic version]. Retrieved January 3, 2008 from, Institute for Work

& Health Web site: http:/www.iwh.on.ca/products/images/ESRTW.pdf

Escorpizo, R., Reneman, M.F., Ekholm, J., Fritz, J., Krupa, T., Marnetoft, S-U., Maroun,

C.E., Guzman, J.R., Suzuki, Y., Stucki, G., & Chan, C.H. (2011). A conceptual

definition of vocational rehabilitation based on the ICF: Building a shared global

model. *Journal of Occupational Rehabilitation, 21*, 126-133.

Fitzpatrick, J.L. (2012). An introduction to context and its role in evaluation practice.

New Directions for Evaluation, 135, 7-24.

Fournier, D. M. (1995). Establishing evaluative conclusions: A distinction between

general and working logic. *New Directions for Evaluation, 68,* p. 15-32.

Fournier, D. M. (2005). Evaluation. In S. Mathison (Ed.), *Encyclopedia of Evaluation*,

(pp. 139-140). Thousand Oaks, CA: Sage Publications, Inc.

Fournier, D. M. (2005). Logic of evaluation – working logic. In S. Mathison (Ed.),
Encyclopedia of Evaluation (pp. 238-242). Thousand Oaks, CA: Sage
Publications, Inc.

Franche, R.L., Baril, R., Shaw, M.N., Nicholas, M., & Loisel, P. (2005). Workplace –
based return-to-work interventions: Optimizing the role of stakeholders in
implementation and research. *Journal of Occupational Rehabilitation, 15*, 525-
542.

Franche, R.L., Frank, J., & Krause, N. (2005). Prediction of occupational disability:
Models, factors and outcomes. In I.Z. Schultz and R.J. Gatchel (Eds.), *Handbook
of Complex Occupational Disability Claims, (pp. 93-116). New York: Springer
Science + Business Media, Inc.*

Franche, R.L., & Krause, N. (2005). Readiness for return to work following injury or
illness. In I.Z Schultz and R.J. Gatchel (Eds.), *Handbook of Complex
Occupational Disability Claims*, (pp. 67-91). New York: Springer Science +
Business Media, Inc.

Frankl, V. E. (1963). *Man's Search for Meaning.* New York: Beacon Press.

Galvin, D., King, A., Knuelle, E., & Rushby, M. (2005). *Disability Management
Success: A Global Perspective.* Vancouver, BC: National Institute
of Disability Management and Research.

Gauthier, B., Barrington, G.V., Bozzo, S.L., Chaytor, K., Dignard, A., Lahey, R.,

Malatest, R., McDavid, J.C., Mason, G., Mayne, J., Porteous, N.L., & Roy, S. (2010). The Lay of the land: Evaluation practice in Canada in 2009. *The Canadian Journal of Evaluation, 24*, 1-49.

Glaser, B.G. (2007). Doing formal theory. In A. Bryant and K. Charmaz (Eds.), *The Sage Handbook of Grounded Theory,* (pp. 97-113). Los Angeles, CA: Sage Publications.

Gottlieb, A., Vandergoot, D., & Lutsky, L. (1991). The role of the rehabilitation professional in corporate disability management. *Journal of Rehabilitation*, *57*, (23-28).

Greene, J. (2005). Context. In S. Mathison (Ed.), *Encyclopedia of Evaluation.* Thousand Oaks, CA: Sage.

Guba, G.G., & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. London: Sage Publications.

Hanley-Maxwell, C., Al Hano, I., & Skivington, M. (2007). Qualitative research in rehabilitation counseling. *Rehabilitation Counseling Bulletin, 50*, 99-110.

Harder, H. G. & Scott, L. R. (2005). *Comprehensive Disability Management.* London: Elsevier Science Limited.

Harry, B., Sturges, K.M., & Klinger, J.K. (2005). Mapping the process: An exemplar of process and challenge in grounded theory analysis. *Educational Researcher, 34, 3-13.*

Hatry, H.P. (2013). Sorting the relationships among performance measurement, program evaluation, and performance management. In P.R. Brandon (Ed.), *New Directions for Evaluation, 137,* 19-32.

Holton, J. A. (2007). The coding process and its challenges. In A. Bryant and K. Charmaz

    (Eds.) *The Sage Handbook of Grounded Theory*, (pp. 265- 289). Thousand Oaks,

    CA: Sage Publications.

Hunter, D.E.K. & Nielsen, S.B. (2013). Performance management and evaluation:

    Exploring complementarities. In P.R. Brandon (Ed.), *New Directions for*

    Evaluation, 137, 7-18.

Hurworth, R. (2005). Emergent Design. In S. Mathison (Ed.), *Encyclopedia of*

    Evaluation (pp. 124). Thousand Oaks, CA: Sage Publications, Inc.

Johnson, K., Greenseid, L.O., Toal, S.A., King, J.A., Lawrenz, F., & Volkov, B. (2009).

    Research on evaluation use: A review of the empirical literature from 1986 to

    2005. *American Journal of Evaluation, 30,* 377-410.

Julnes, P.L. (2013). Citizen-driven performance measurement: Opportunities for

    evaluator collaboration in support of the new governance. In P.R. Brandon (Ed.),

    New Directions for Evaluation, 137, 81-92.

Kline, R.B. (2009). *Becoming a Behavioral Science Researcher.* New York: The Guilford

    Press.

Kofman, F., & Senge, P.M. (1993). Communities of commitment: The heart of learning

    organizations. *Organizational Dynamics*, 22, 5-23.

Lahey, R., & Nielsen, S.B. (2013). Rethinking the relationship among monitoring,

    evaluation, and results-based management: Observations from Canada. In P.R.

    Brandon (Ed.), *New Directions for Evaluation, 137*, 45-56.

Lantham, G.P. (2009). Motivate employee performance through goal setting. In

    E.A. Locke, (Ed.) *Handbook of Principles of Organizational Behavior (2nd*

*ed.), (*pp. 161-178). Chichester, WS: John Wiley & Sons, Ltd.

Lemaire, D., & Boyle, R. (1999). Postscript: Evaluation capacity building – a journey

without an end. In R. Boyle and D. Lemaire, (Eds.), *Building Effective Evaluation*

Capacity (pp. 195-198). New Brunswick, NJ: Transaction Publishers.

Lempert, L.B. (2007). Asking questions of the data: Memo writing in the grounded

theory trandition. In A. Bryant and K. Charmaz (Eds.) *The Sage Handbook of*

Grounded Theory, (pp. 245-264). Thousand Oaks, CA: Sage Publications.

Lincoln, Y.S. (2005). Fourth-Generation Evaluation. In S. Mathison (Ed.), *Encyclopedia*

 Of Evaluation, (pp. 161-164). Thousand Oaks, CA: Sage Publications Inc.

Loisel, P., Buchbinder, R., Hazard, R., Keller, R., Scheel, I., van Tulder, M., &

Webster, B. (2005). Prevention of work disability due to musculoskeletal

disorders: The challenge of implementing evidence. *Journal of Occupational*

Rehabilitation, 15, 507-524.

Loisel, P., & Durand, M. (2002). Working with the employer. In I.Z Schultz and

R.J. Gatchel (Eds.), *Handbook of Complex Occupational Disability Claims*,

(pp. 479-488). New York: Springer Science + Business Media, Inc.

Loisel, P., Durand, M. J., Berthelette, D., Vezina, N., Baril, R., Gagnon, R., Lariviere, C.,

and Tremblay, C. (2001). Disability prevention, *Disability Management Health*

Outcomes, 9, 351-360.

Madison, A. (2007). New Directions for Evaluation coverage of cultural issues and

issues of significance to underrepresented groups. *New Directions for Evaluation,*

114, 107-114.

Mark, M.M. (2008). Building a better evidence base for evaluation theory: Beyond

general calls to a framework of types of research on evaluation. In N.L. Smith & P.R. Brandon (Eds.), *Fundamental Issues in Evaluation* (pp. 111-134). New York: Guilford Press.

Mathison, S. (2008). What is the difference between evaluation and research – and why do we care? In N. L. Smith & P. R. Brandon (Eds.), *Fundamental Issues in Evaluation* (pp. 183-198). New York: The Guilford Press.

Mayne, J., Divorski, S., & Lemaire, D. (1999). Locating evaluation: Anchoring evaluation in the executive or the legislature, or both or elsewhere? In R. Boyle and D. Lemaire, (Eds.), *Building Effective Evaluation Capacity* (pp. 23-52).

McMahon, B. T., Danczyk-Hawley, C. E., Reid, C., Flynn, B. S., Habeck, R., Kregel, J., & Owens, P. (2000). The progress of disability benefits. *Journal of Vocational Rehabilitation, 15*, (3-15).

McNulty, J. Symbolic uses of evaluation in the international aid sector: Arguments for critical reflection. *Evidence & Policy, 8,* 495-509.

Mertens, D.M. (2008). Stakeholder representation in culturally complex communities: Insights from the transformative paradigm. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental Issues in Evaluation* (pp. 41-60). New York: The Guilford Press.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis (2nd Ed.).* Thousand Oaks, CA: Sage Publications.

Nagy Hesse-Biber, S. (2007). Teaching grounded theory. In A. Bryant & K. Charmaz (Eds.), *The Sage Handbook of Grounded Theory, (*pp. 311-338). London: Sage Publications.

Nickerson, L. (2000). Quality disability management. *Journal of the Ontario Occupational Health Nurse Association, 18*, 19-21.

Owen, J. M. (2005). Learning organization. In S. Mathison (Ed.), *Encyclopedia of Evaluation* (pp. 225-226). Thousand Oaks, CA: Sage Publications, Inc.

Patton, M.Q. (1996). A world larger than formative and summative. *Evaluation Practice, 17*, 131-144. Publications, Inc.

Patton, M.Q. (2008). *Utilization-Focused Evaluation (4th ed.)*, Thousand Oaks, CA: Sage Publications, Inc.

Pransky, G., Gatchel, R., Linton, S.J., & Loisel, P. (2005). Improving return to work research. *Journal of Occupational Rehabilitation, 15*, 453-457.

Pransky, G. S., Shaw, W. S., Franche, R., & Clarke, A. (2004). Disability prevention and communication among workers, physicians, employers, and insurers – current models and opportunities for improvement. *Disability and Rehabilitation, \26*, (625-634).

Preskill, H. (1994). Evaluation's role in enhancing organizational learning: A model for practice. *Evaluation and Program Planning, 17,* (291-297).

Preskill, H., & Torres, R. T. (1999). *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, CA: Sage Publication s.

Preskill, H., & Torres, R. T. (2000). *The Readiness for Organizational Learning and Evaluation Instrument (ROLE)*. Claremont, CA: Claremont Graduate University.

Purnell, L.D., & Paulanka, B.J. (2003). *Transcultural Health Care A Culturally Competent Approach* (2nd ed.). Philadelphia: F.A. Davis Company.

Rankin, N. (2001). *An Employer's Guide to Disability Management.* Aurora, ON:

Aurora Professional Press.

Richards, L., & Morse, J. M. (2007). *Read Me First for a User's Guide to Qualitative Methods (2nd Ed.).* Thousand Oaks, CA: Sage Publications.

Roessler, R. T., Schriner, K. F., & Fletcher, K. (1991). Partnerships: The bridge from disability to ability management. *Journal of Rehabilitation*, 57, (53-58).

Rog, D.J. (2012). When background becomes foreground: Toward context-sensitive evaluation practice. *New Directions for Evaluation, 135,* 25-40.

Rog, D.J., Fitzpatrick, J.L., & Conner, R.F. (2012). Editors Note. *New Directions for Evaluation, 135,* 1-6.

Rogers, P., Petrosino, A., Huebner, T.A., & Hacsi, T.A. (2000). Program theory evaluation: practice, promise, and problems. *New Directions for Evaluation, 87*, (pp. 5-13).

Rosenthal, D. A., Hursh, N., Lui, J., Ison, R., & Sasson, J. (2007). A survey of current disability management practice: Emerging trends and implications for certification. *Rehabilitation Counseling Bulletin, 50,* (76-86).

Rosenthal, D.A., Hursh, N.C., Lui, J., Zimmermann, W., & Pruett, S.R. (2005). Workplace disability management: Case management implications. In F. Chan, M.J. Leahy and J. L. Saunders (Eds.), *Case Management for Rehabilitation Health Professionals (2nd Ed.),* Vol.1, (p. 331-365). Osage Beach, MO: Aspen Professional Services.

Russ-Eft, D., & Preskill, H. (2001). *Evaluation in Organizations*. Cambridge, MA: Perseus Publishing.

Schmidt, F.L. (2009). Select on intelligence. In E.A. Locke, (Ed.) *Handbook of*

*Principles of Organizational Behavior (2nd ed.), (*pp. 3-17). Chichester, WS: John Wiley & Sons, Ltd.

Schultz, I.Z., Crook, J.M., Berkowitz, J., Meloche, G.R., Milner, R., & Zuberbier, O.A. (2005). Biopsychosocial multivariate predictive model of occupational low back disability. In I.Z. Schultz and R.J. Gatchel (Eds.), *Handbook of Complex Occupational Disability Claims,* (pp.191-202). New York: Springer Scient + Business Media, Inc.

Schultz, I. Z., Joy, P.W., Crook, J., & Fraser, K. (2001). Models of diagnosis and rehabilitation in musculoskeletal pain-related occupational disability. *Journal of Occupational Rehabilitation, 10*, 271-293.

Schultz, I. Z., Stowell, A.W., Feuerstein, M., & Gatchel, R.J. (2007). Models of return to work for musculoskeletal disorders. *Journal of Occupational Rehabilitation, 17*, 327-352.

Schwandt, T.A. (2007). *The Sage Dictionary of Qualitative Inquiry, 3rd* Ed. London: Sage Publications.

Scriven, M. (1991). Beyond formative and summative evaluation. In M.W. McLaughlin and D.C. Phillips (Eds.), *Evaluation and Education: At Quarter Century*, Chicago, Il: University of Chicago Press.

Senge, P.M. (2006). *The Fifth Discipline*. New York: Doubleday.

SenGupta, S., Hopson, R., & Thompson-Robinson, M. (2004). Cultural competence in evaluation: An overview. *New Directions for Evaluation*, 102, 5-19.

Shadish, W.R., Cook, T.D., & Leviton, L.C. (1991). *Foundations of Program Evaluation: Theories of Practice.* Newbury Park, CA: Sage Publications Inc.

Shrey, D. E. (1991). Disability management, occupational bonding & the industrially injured worker. In *Disability Management Interventions for the Industrially Injured Worker (pp. 3-12).* London, ON: London Disability Management Research Group.

Shrey, D. E., & Olsheski, J. A. (1992). Disability management and industry-based work return transition programs. *Physical Medicine and Rehabilitation, 6*, 303-314.

Smart, J.F. (2001). *Disability, Society and the Individual.* Gaithersburg, MD: Aspen.

Smart, J.F. (2005). The promise of the International Classification of Functioning, Disability and Health (ICF). *Rehabilitation Education, 19*, 191-199.

Smart, J.F., & Smart, D.W. (2006). Models of disability: Implications for the counseling profession. *Journal of Counseling & Development, 84,* 29-40.

Smith, N. L. (1995). The influence of societal games on the methodology of evaluative inquiry. *New Directions for Evaluation, 68,* p. 5-14.

Smith, N.L., & Brandon, P.R. (2008). Preface. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental Issues in Evaluation* (pp. vii-xvi). New York: The Guilford Press.

Statistics Canada. (2003). Update on cultural diversity [Electronic version]. *Canadian Social Trends*, Statistic Canada, Catalogue no. 11-008, 19-23.

Strasser, P. B. (2004). Managing transitional work – program foundation. *American Association of Occupational Health Nurses Journal, 52,* (323-326).

Stowell, A.W., & McGeary, D.D. (2005). Musculoskeletal injury. In I.Z Schultz and R.J. Gatchel (Eds.), *Handbook of Complex Occupational Disability Claims*, (pp. 117-139). New York: Springer Science + Business Media, Inc.

Tate, D. G. (1987). The healthy corporation: Disability management and health promotion practices at the workplace. *Journal of Rehabilitation*, 53, (63-66).

Tate, D. G., Habeck, R. V., & Galvin, D. E. (1986). Disability management: Origins, concepts and principles for practice. *Journal of Applied Rehabilitation Counseling, 17*, 5-12.

Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation, 28*, 45-59.

Taylor-Powell, E. (2006). Evaluating food stamp nutrition education: A view from the field of program evaluation. *Journal of Nutrition, Education and Behavior*, *38*, 12-17.

Taylow-Powell, E., & Boyd, H.H. (2008). Evaluation capacity building in complex organizations. In S. Mathison (Ed.), *New Directions for Evaluation, 120*, 55-69.

Ustun, T.B., Kostanjsek, N., Chatterji, S. & Rehm, J. (Eds.). (2010). *Measuring Heath and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0).* Geneva: WHO Press.

Weingart, L.R. & Jehn, K.A. (2009). Manage intra-team conflict through collaboration. In E.A. Locke, (Ed.) *Handbook of Principles of Organizational Behavior (2nd ed.),* p. 327-346. Chichester, WS: John Wiley & Sons, Ltd.

Westmorland, M., & Buys, N. (2002). Disability management in a sample of Australian self-insured companies. *Disability and Rehabilitation, 24*, (746-754).

Wholey, J.S. (1996). Formative and summative evaluation: Related issues in performance measurement, *Evaluation Practice, 17,* 145-149.

Wholey, J. S. (2013). Using evaluation to improve program performance and results in

M.C. Alkin (Ed.) *Evaluation Roots: A Wider Perspective of Theorists' View and Influences,* (p. 261-266). Thousand Oaks, CA: Sage Publications, Inc.

Wiser, C. (1996). Editors note. *New Directions for Evaluation, 71*, 1-5

World Health Organization. (2012). *International Classification of Functioning, Disability and Health.* Geneva: WHO Press.

Young, A.E., Roessler, R.T., Wasiak, R., McPherson, K.M., van Poppel, N.M., Anema, J.R. (2005). A developmental conceptualization of return to work. *Journal of Occupational Rehabilitation, 15*, 557-568.

Young, A.E., Wasiak, R., Roessler, R.T., McPherson, K.M., Anema, J.R., & van Poppel, M. N.M. (2005). Return-to-work outcomes following work disability: Stakeholder motivations, interests and concerns. *Journal of Occupational Rehabilitation, 15, 543-556.*

Appendix A

Department of Educational and Counselling

Psychology, and Special Education

The University of British Columbia

Faculty of Education

2125 Main Mall

Consent Form

The Nature of Program Evaluation in Disability Management

Dear Sir or Madam:

You are invited to participate in a research project which will explore evaluation currently being practiced in the field of disability management. My name is Patricia L. Swenson and I am doctoral candidate at the University of British Columbia, in the Department of Educational and Counselling Psychology, and Special Education. You are being invited to take part in this research as disability management is practiced at your organization.

Purpose of this Research Study

Disability management programs involve multi-disciplinary health, safety and return to work processes, which are proactively applied within organizations to minimize the economic and social costs resulting from time off work due to illness or injury. There is well documented research describing how disability management programs have evolved over the past two decades.

December 21, 2009 version   Page 1 of 4

Little has been published regarding evaluation in the field of disability management. Evaluation is the only objective way of understanding what aspects of a program are working and where improvements are needed. The purpose of this research project will be to improve understanding by developing a conceptual framework regarding the nature of evaluation currently practiced in disability management programs.

Your Participation in this Research Study

As a participant you will be asked to complete a questionnaire (less than 30 minutes). This questionnaire, The Readiness for Organizational Learning from Evaluation Inventory (ROLE), is used to assess the perceptions of personnel about their work environment in relation to learning from evaluation.

You will also participate in a face-to-face interview (approximately 60 minutes). Interview questions will mainly address issues related to evaluation of the disability program you work with, and characteristics of your organization you think relate to learning from evaluation.

I will audio tape the interview and transcribe the audio tape. I will then use a computer software program to analyze the transcript. Your identity will not be revealed as a code name will be given to the audio tapes, transcripts, computer data and on the final report. Tapes and documents will be kept in a secure locked office. Computer files will be saved on a hard drive used only for this research. None of the saved data will be used for any purpose other than this study without your written consent. There are no foreseeable risks.

December 21, 2009 version   Page 2 of 4

If further questions develop from the interview, you may be asked to participate in more than one interview (approximately 60 minutes).

With the approval of your organization, you will be asked for examples of recorded information (e.g. documents, files) that your organization maintains to evaluate their disability management program. General information from these sources may be documented. This information will be managed in the same secure manner as interview data.

Possible benefits of participating in this research include: contributing to a better understanding of how disability management programs are evaluated, and gaining an increased awareness of how your program and your organization may learn from evaluation.

Contact Information

If you have any questions about this study you can contact me by email. This research report will be submitted as a final project for my dissertation study at the University of British Columbia. My supervising professor and the principal investigator for this study is Sandra Mathison, PhD, Professor in the Department of Educational and Counselling Psychology, and Special Education.

If you would like to receive a summary of the research findings please contact me via email or telephone.

If you have any concerns about your treatment or rights as a research subject, you may contact the Research Subject Information Line in the UBC Office of Research Services.

<u>Consent to Participate</u>

Your participation in this study is entirely voluntary, and you may refuse to participate or withdraw from the study at any time. Your signature below indicates that you have read the information in this letter and consent to participate. If you are willing to participate please type your name and date on the space provided, save the document on your computer, and send the saved document to me as an attachment to an email.

I agree to participate in the study.

Name of Participant: _____

Date: _____

Appendix B

The Readiness for
Organizational Learning and Evaluation Instrument (ROLE)[1]

©2000 Hallie Preskill & Rosalie T. Torres

Directions

Before you begin responding to the items, please check one of the two boxes below to indicate whether you will be thinking about the organization as a whole, or your department /unit as the focus for your ratings. Base this decision on the entity with which you are most familiar. For example, if you are part of a large department it probably makes sense to respond in terms of your department. On the other hand, if you are very familiar with the organization as a whole, you can respond in terms of the organization.

❑ I will be thinking about the entire organization.
❑ I will be thinking about my department/unit.

For each of the items below, circle the number that best represents <u>your opinion based on your experiences</u>, and not on how you think other individuals would answer, or your organization's official policy or intent.

	Strongly Disagree				Strongly Agree
	1	2	3	4	5

Culture

Collaboration and Problem Solving

1. Employees respect each other's perspectives and opinions.	1	2	3	4	5
2. Employees ask each other for information about work issues and activities.	1	2	3	4	5
3. Employees continuously look for ways to improve processes, products and services.	1	2	3	4	5
4. Employees are provided opportunities to think about and reflect on their work.	1	2	3	4	5
5. Employees often stop to talk about the pressing work issues we're facing.	1	2	3	4	5
6. When trying to solve problems, employees use a process of working through the problem before identifying solutions.	1	2	3	4	5
7. There is little competition among employees for recognition or rewards.	1	2	3	4	5
8. Employees operate from a spirit of cooperation, rather than competition.	1	2	3	4	5

---

[1] ©2000 Hallie Preskill, Claremont Graduate University & Rosalie T. Torres, Torres Consulting. All rights reserved.

	Strongly Disagree			Strongly Agree	
	1	2	3	4	5
9. Employees tend to work collaboratively with each other.	1	2	3	4	5
10. Employees are more concerned about how their work contributes to the success of the organization than they are about their individual success.	1	2	3	4	5
11. Employees face conflict over work issues in productive ways.	1	2	3	4	5
12. Employees generally view problems or issues as opportunities to learn.	1	2	3	4	5

❑   *Risk Taking*

13. Mistakes made by employees are viewed as opportunities for learning.	1	2	3	4	5
14. Employees continuously ask themselves how they're doing, what they can do better, and what is working.	1	2	3	4	5
15. Employees are willing to take risks in the course of their work.	1	2	3	4	5
16. Employees are committed to being innovative and forward looking.	1	2	3	4	5
17. Employees are confident that mistakes or failures will not affect them negatively.	1	2	3	4	5

❑   *Participatory Decision Making*

18. Employees generally trust their managers or supervisors.	1	2	3	4	5
19. Managers and supervisors view individuals' capacity to learn as the organization's greatest resource.	1	2	3	4	5
20. Employees use data/information to inform their decision-making.	1	2	3	4	5
21. Asking questions and raising issues about work is encouraged.	1	2	3	4	5
22. Employees are not afraid to share their opinions even if those opinions are different from the majority.	1	2	3	4	5
23. I feel safe explaining to others why I think or feel the way I do about an issue.	1	2	3	4	5
24. Employees are encouraged to take the lead in initiating change or in trying to do something different.	1	2	3	4	5
25. Managers and supervisors make decisions after considering the input of those affected.	1	2	3	4	5

243

	Strongly Disagree			Strongly Agree	
	1	2	3	4	5

26. In meetings employees are encouraged to discuss the values and beliefs that underlie their opinions.   1   2   3   4   5

27. Employees are encouraged to offer dissenting opinions and alternative viewpoints.   1   2   3   4   5

Leadership

28. Managers and supervisors admit when they don't know the answer to a question.   1   2   3   4   5

29. Managers and supervisors take on the role of coaching, mentoring and facilitating employees' learning.   1   2   3   4   5

30. Managers and supervisors help employees understand the value of experimentation and the learning that can result from such endeavors.   1   2   3   4   5

31. Managers and supervisors make realistic commitments for employees (e.g., time, resources, workload).   1   2   3   4   5

32. Managers and supervisors understand that employees have different learning styles and learning needs.   1   2   3   4   5

33. Managers and supervisors are more concerned with serving the organization than with seeking personal power or gain.   1   2   3   4   5

34. Managers and supervisors are open to negative feedback from employees.   1   2   3   4   5

35. Managers and supervisors model the importance of learning through their own efforts to learn.   1   2   3   4   5

36. Managers and supervisors believe that our success depends upon learning from daily practices.   1   2   3   4   5

37. Managers and supervisors support the sharing of knowledge and skills among employees.   1   2   3   4   5

38. Managers and supervisors provide the necessary time and support for systemic, long-term change.   1   2   3   4   5

39. Managers and supervisors use data/information to inform their decision-making.   1   2   3   4   5

	Strongly Disagree			Strongly Agree
	1	2	3	4	5

Systems and Structures

❑ *Open and Accessible Work Environment*

40. There is little bureaucratic red tape when trying to do something new or different. 1 2 3 4 5

41. Workspaces are designed to allow for easy and frequent communication with each other. 1 2 3 4 5

42. There are few boundaries between departments/units that keep employees from working together. 1 2 3 4 5

43. Employees are available (i.e., not out of the office or otherwise too busy) to participate in meetings. 1 2 3 4 5

❑ *Rewards and Recognition Systems and Practices*

44. Employees are recognized or rewarded for learning new knowledge and skills. 1 2 3 4 5

45. Employees are recognized or rewarded for helping solve business/organizational problems. 1 2 3 4 5

46. The current reward or appraisal system recognizes, in some way, team learning and performance. 1 2 3 4 5

47. Employees are recognized or rewarded for helping each other learn. 1 2 3 4 5

48. Employees are recognized or rewarded for experimenting with new ideas. 1 2 3 4 5

❑ *Relationship of Work to Organizational Goals*

49. Employees understand how their work relates to the goals or mission of the organization. 1 2 3 4 5

50. Employees' performance goals are clearly aligned with the organization's strategic goals. 1 2 3 4 5

51. Employees meet work deadlines. 1 2 3 4 5

	Strongly Disagree			Strongly Agree
	1 2	3	4 5	

Communication of Information

❑   *Availability*

52. Information is gathered from clients, customers, suppliers or other stakeholders to gauge how well we're doing.   1 2 3 4 5

53. Currently available information tells us what we need to know about the effectiveness of our programs, processes, products, and services.   1 2 3 4 5

54. There are adequate records of past change efforts and what happened as a result.   1 2 3 4 5

❑   *Dissemination*

55. There are existing systems to manage and disseminate information for those who need and can use it.   1 2 3 4 5

56. Employees are cross-trained to perform various job functions.   1 2 3 4 5

57. Employees have access to the information they need to make decisions regarding their work.   1 2 3 4 5

58. Employees use technologies to communicate with one another.   1 2 3 4 5

59. When new information that would be helpful to others is learned or discovered, it gets disseminated to those individuals.   1 2 3 4 5

Teams

60. My department/unit currently operates via (or is transitioning towards) a team-based structure.

❑ Yes, this is true.
❑ No, this is not true.

61. Employees are provided training on how to work as a team member.

❑ Yes, this is true.
❑ No, this is not true.

62. My work is sometimes conducted as part of a working group that is or could be identified as a "team."

❑ Yes, this is true. (Continue with item 63)
❑ No, this is not true. (Go to item 71)

246

	Strongly Disagree			Strongly Agree	
	1	2	3	4	5

Respond to items 63-70 based on your experiences as a team member.

63. When conflict arises among team members, it is resolved effectively.   1   2   3   4   5

64. Team members are open and honest with one another.   1   2   3   4   5

65. Team meetings are well facilitated.   1   2   3   4   5

66. Team meetings address both team processes and work content.   1   2   3   4   5

67. Team meetings strive to include everyone's opinion.   1   2   3   4   5

68. Teams are encouraged to learn from each other and to share their learning with others.   1   2   3   4   5

69. Teams accomplish work they are charged to do.   1   2   3   4   5

70. Teams are an effective way to meet an organization's goals.   1   2   3   4   5

Evaluation

Please use the following definition of *evaluation* when responding to the items below:

> *Evaluation is a process of systematic inquiry to provide information for decision-making about some object – a program, project, process, organization, system, or product. Use of the evaluation results might lead to making refinements to the program or to offering new services or products.*

71. The integration of evaluation activities into our work has enhanced (or would enhance) the quality of decision-making.   1   2   3   4   5

72. It has been (or would be) worthwhile to integrate evaluation activities into our daily work practices.   1   2   3   4   5

73. Managers and supervisors like (or would like) us to evaluate our efforts.   1   2   3   4   5

74. Evaluation helps (or would help) us provide better programs, processes, products and services.   1   2   3   4   5

75. There would be support among employees if we tried to do more (or any) evaluation work.   1   2   3   4   5

76. Doing (more) evaluation would make it easier to convince managers of needed changes.   1   2   3   4   5

	Strongly Disagree				Strongly Agree
	1	2	3	4	5

77. This would be a good time to begin (or renew or intensify) efforts to conduct evaluations.   1 2 3 4 5

78. There are evaluation processes in place that enable employees to review how well changes we make are working.   1 2 3 4 5

Additional Information

79. Which of the following best describes your job category? (Check one.)
- ❑ First-Line Supervisor
- ❑ Middle Manager
- ❑ Senior Manager
- ❑ Administrative
- ❑ Production
- ❑ Sales
- ❑ Non-Managerial Professional
- ❑ Technical
- ❑ Customer Service
- ❑ Other _____

80. Which of the following best describes your organization? (Check one.)
- ❑ Manufacturing
- ❑ Business Services
- ❑ Transportation/Communication/Utilities
- ❑ Health Services
- ❑ Wholesale/Retail Trade
- ❑ Finance/Insurance/Banking
- ❑ Education Services
- ❑ Government (Local, State, Federal)
- ❑ Non-Profit
- ❑ Other _____

81. How long have you worked for this organization? (Check one.)
- ❑ Less than 6 months
- ❑ 6 months – 1 year
- ❑ 1-3 years
- ❑ 4-6 years
- ❑ 7-10 years
- ❑ More than 10 years

248

The Readiness for Organizational Learning and Evaluation Instrument (ROLE)[2]

Purpose

This instrument is designed to help an organization determine its level of readiness for implementing organizational learning and evaluation practices and processes that support it. The instrument's results can be used to:

- Identify the existence of learning organization characteristics
- Diagnose interest in conducting evaluation that facilitates organizational learning
- Identify areas of strength to leverage evaluative inquiry processes
- Identify areas in need of organizational change and development.

In sum, the organization may use the results to focus its efforts on improving or further strengthening areas that will lead to greater individual, team, and organizational learning.

Background and Rationale

In an effort to respond to internal and external demands for growth and success, many organizations have adopted the goal of becoming a learning organization. Organizational learning is "a continuous process of organizational growth and improvement that (a) is integrated with work activities; (b) invokes the alignment of values, attitudes, and perceptions among organizational members; and (c) uses information or feedback about both processes and outcomes to make changes" (Torres, Preskill & Piontek, 1996, p. 2). Evaluation conducted in support of organizational learning provides a means for (a) developing a community of inquirers, (b) harnessing the knowledge capital of its members, and (c) addressing problematic issues that face the organization. It can serve as a catalyst for learning and action on organizational issues (Preskill & Torres, 1999, p. 43). Implementing organizational learning and evaluation efforts, however, is not an easy task. It requires that the organization carefully assess how prepared its structures, policies, procedures, and members are to support organizational learning and evaluation practices.

Description of the Instrument

The items on the instrument reflect the research on organizational learning and evaluation processes and practices. The results from this body of research suggest that an organization must have certain elements of its infrastructure in place if it is to truly support and encourage organizational learning. Research on the use of evaluation findings has also shown that the organization's culture and context significantly influence the extent to which evaluation findings are used to support learning and decision making.

ROLE consists of 78 items grouped into six major dimensions. These include: (a) Culture, (b) Leadership, (c) Systems and Structures, (d) Communication, (e) Teams, and (f) Evaluation.

---

[2] Based on the book, *Evaluative Inquiry for Learning in Organizations*,1999.

Within four of these dimensions are eight subcategories (see Table 1). Three additional questions are included to provide information about the respondent and the organization. As individuals respond to each item, a picture begins to emerge that describes the extent to which organizational learning and evaluation practices and systems are present in the organization. Reliability data for the instrument are shown in Table 1 (see also Preskill, Torres, & Martinez-Papponi, 1999)[3].

Respondents are asked to respond to (a) 75 Likert scale items on a scale of 1 to 5, with 1 meaning "Strongly Disagree," and 5 meaning "Strongly Agree;" (b) three yes/no items; and (c) three multiple choice items. In administering the instrument with organization members, it is important to emphasize that there are no right or wrong answers. What matters most is their opinion based on their experiences. Use of the instrument is most effective when its items are answered honestly and the organization treats individuals' responses confidentially. We recommend that the results for all respondents be aggregated and reported in summary form. The instrument can be administered to single or multiple departments within an organization, or to the entire organization.

<div align="center">Analysis</div>

The instrument data should be entered in a database and mean scores calculated for each dimension and subcategory. The results of this analysis can be displayed on the worksheet shown on page 4.

<div align="center">Interpretation of Results</div>

If a department or organization were to score low in one or more of the dimensions, this would indicate that learning from evaluation might not be supported or allowed to succeed. Likewise, it would indicate that the department or organization isn't prepared to engage in other kinds of organizational learning practices. These kinds of results can help the organization determine where to focus its improvement efforts if its goal is to become a learning organization.

<div align="center">Example</div>

Let's say a training department administered the instrument to its 50 employees. The aggregated results for the six dimensions from the survey are shown below. In interpreting the results, the department's management might conclude that it's leadership, culture, and systems of communication are doing pretty well – at least in terms of supporting organizational learning principles. On the other hand, the unit's systems and structures, its use of teams, and use and/or support of evaluation are less likely to facilitate organizational learning. Based on these results, the department decides to devote further effort to examining the results of the subcategories in the systems and structures dimension (open and accessible work

---

250

environment, rewards and recognition systems and practices, and relationship of work to organizational goals). At the same time they begin looking at ways in which teamwork and evaluation efforts can support organizational goals.

Dimensions with Mean Scores of *3.5 or Above*	Dimensions with Mean Scores *Below 3.5*
Leadership (3.55) Communication (3.90) Culture (3.50)	Systems and Structures (2.60) Teams (3.45) Evaluation (2.95)

References

Preskill, H. & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage Publications.

Preskill, H., Torres, R. T., & Martinez-Papponi, B. (1999). *Assessing an organization's readiness for learning from evaluative inquiry*. Paper presented at the American Evaluation Association annual conference. Orlando, FL, November.

Torres, R. T., Preskill, H. & Piontek, M. (2005). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Second Edition. Thousand Oaks, CA: Sage.

Table 1. Reliability Data for the ROLE Instrument

Dimension/Subcategory	Number of Items	Coefficient alpha
❑ ❑ *Culture* - Collaboration and problem solving - Risk taking - Participatory Decision Making	12 5 10	.88 .85 .89
❑ ❑ *Leadership*	12	.93
❑ ❑ *Systems and Structures* - Open and accessible work environment - Rewards & Recognition System and Practices - Relationship of Work to Organizational Goals	4 5 3	.73 .89 .66
Communication of Information - Availability - Dissemination	3 5	.79 .80
❑ *Teams*	8	.91
❑ *Evaluation*	8	.84
❑ *All Likert Scale Items*	75	.97 (Cronbach's Alpha)

The Readiness for Organizational Learning and Evaluation Instrument (ROLE)
<u>Mean Scores</u>

Culture	Leadership	Systems & Structures	❑ Communic ation	❑ Teams (if answered "yes," to items 60-62) If answered "no", skip this column)	Evaluation
❑ *Collaboration & Problem Solving*	28.	*Open & Accessible Work Environment*	*Availability*	63.	71.
1.	29.	40.	52.	64.	72.
2.	30.	41.	53.	65.	73.
3.	31.	42.	54.	66.	74.
4.	32.	43.	Subtotal Mean	67.	75.
5.	33.	Subtotal Mean		68.	76.
6.	34.	*Rewards & Recog. Systems & Practices*		69.	77.
7.	35.	44.	*Dissemination*	70.	78.
8.	36.	45.	55.		
9.	37.	46.	56.		
10.	38.	47.	57.		
11.	39.	48.	58.		
12.		❑ Subtotal Mean	59.		
❑ Subtotal Mean		*Relationship of Work to Org. Goals*	Subtotal Mean		
❑ *Risk Taking*		49.			
13.		50.			
14.		51.			
15.		❑ Subtotal Mean			
16.					
17.					
❑ Subtotal Mean					
Participatory Decision Making					
18.					
19.					
20.					
21.					
22.					
23.					
24.					
25.					
26.					
27.					
❑ Subtotal Mean					
Total Mean Score	Total Mean Score	Total Mean Score	Total Mean Score	Total Mean Score	Total Mean Score

Grounded Theory Data Analysis

Concepts are the main unit of analysis in the grounded theory method. The following illustrates examples of the grounded theory method followed in this study. One concept, collaboration, is highlighted as an example to demonstrate some of the ways concepts were scrutinized, noting all concepts were similarly analyzed.

<u>Analysis prior to conceptual coding</u>

- I practiced coding with two different data analysis software programs before deciding to use NVivo after confirming its ease for coding and having learned at the 2010 Canadian Evaluation Society annual conference NVivo meets industry standards.
- Each interview was conducted days after initial contact with the participant.
- Each interview was transcribed within days after the interview was conducted.
- Each transcript was coded as soon as possible after transcription.

<u>Open coding</u>

- Transcripts were open coded line by line using NVivo software within days after interviews were transcribed.
- For some open codes memos were entered documenting my thoughts.
- Transcripts were open coded in the order that the interviews were conducted (at the biomedical site, BM1 then BM2).
- Open coding of transcripts and documents involved assigning descriptive codes to all data.
- Transcripts from a given site were first open coded, followed by open coding of documents from that site.
- Open coding of documents was done by hand (without NVivo software). This was mainly because there were few documents. Documents were assigned open codes describing what was being said in respective sections of the data.
- A file of all codes was saved after open coding of transcripts and documents at the site.
- Coding followed the same process at the other three sites in the order that the interviews were conducted, L3, L4 and L5 from the labour site, BPsy6 and BPsy7 from the BPsy site, and then I8 and I9 from the insurance site.
- I returned to the raw data and open codes from prior sites during analyses of all data from subsequent sites, to clarify open coding that existed and to decide whether a new open code or previously created code would be appropriate.

<u>Analysis of open codes</u>

- A file of 80 open codes was saved after open coding of data from all four sites (listed in Appendix C of this document).

- After coding of raw data I created a visual display of all 80 open codes. To do this each open code was printed on a separate slip of paper, and all 80 were organized into groups based on how they related to each other. (Examples of these groupings are listed in Appendix D of this study.)
- This visual display contributed to clarification of open codes, providing an initial understanding of ways codes related to each other.
- Some open codes were straightforward, for example describing clients or describing the disability management programs. I referred to these descriptive codes as concrete.
- Other open codes led me to undertake initial levels of more abstract analyses of what was being said by different participants and at different sites. This led to my initial understanding that some open codes and groups of open codes were more abstract than simple descriptions that I referred to as concrete.
- Awareness of the potential for more abstract interpretations of data and open codes led me to new expectations for conceptual coding. I gained awareness of ways that conceptual coding would be more abstract, and would enable more nuanced analyses and understandings than simple descriptors.
- An example was the open code 'bias overcome' that I understood would require a more conceptualized understanding.

<p align="center">Conceptual coding</p>

Order of conceptual coding

- Conceptual coding of transcripts and documents at each site was conducted in the same order that open coding had been conducted.
- Conceptual coding of interview transcripts from each site was conducted using NVivo software.
- Transcripts from each site were conceptually coded in the order the interviews were conducted, then documents from the same site were coded.
- Conceptual coding of the documents from the same site was conducted by hand (without NVivo software).
- Conceptual coding of data from the labour site followed conceptual coding of data from the BM site, then the BPsy site, and finally the insurance site.

Conceptual coding processes

- Familiarity of data within transcripts and documents that had been gained during the processes of interviewing, transcribing and open coding acted as an advance organizer for conceptual coding.
- Coding of documents involved reading data and identifying sections that pertained to a particular concept and assigning conceptual code names to the section.
- NVivo coding involved reading transcripts line by line to identify sections that pertained to a particular concept, highlighting the section, and assigning a conceptual code name to the section.
- NVivo software enabled selection of prior conceptual code names or creation of new conceptual code names for every excerpt of data selected from a transcript.

- I triangulated information from the multiple sources of data (transcripts, documents and ROLE results) to identify conceptual codes.

<div align="center">

NVivo computer code files
</div>

- NVivo software retained a separate file for each conceptual code that had been created.
- Each NVivo code file included all excerpts selected from all transcripts for that code.

Example of a NVivo computer code file for the concept 'collaboration'

CONCEPT COLLABORATIVE

<Internals\Interviews\I9 Concept> - § 1 reference coded [1.98% Coverage]

Reference 1 - 1.98% Coverage

I9
Yes so what we do is when we have our conference calls if there is something in particular that someone has taken I will ask them if they want to talk for 10 15 20 minutes about what they learned. When we meet biannually then what will happen is if anybody from January to June if anybody has gone to a conference or if anybody has taken any particular courses then what I ask them to do is to put together a 5 to 10 minutes usually a 10 minute presentation for all of the investigators. So they will discuss with them what they learned and what the benefits were as far as work is concerned.

<Internals\Interviews\ BPsy6 Concept> - § 2 references coded [1.10% Coverage]

Reference 1 - 0.53% Coverage

BPsy6
And they have almost have never known there was a VR department in their LTD carrier. So we usually do that, um call the LTD carrier on their behalf and request that they be transferred to the voc rehab department. We always do that right.

Reference 2 - 0.58% Coverage

BPsy6
Yah we used to be silod. And be attached to programs, but we couldn't do that with our small numbers effectively so we went to a consultative kind of approach. Where we now have to have the skill sets for every type of disability. Which has its pros and cons.

<Internals\Interviews\BM2 Concept> - § 1 reference coded [0.92% Coverage]

Reference 1 - 0.92% Coverage

BM2
It is important with [BM1] because often we provide the same service um that we are seen as consistent. We work under the same roof. But he also has to produce something that looks very very similar to me. So often we will talk about you know what we would see a certain situation

what we would do. We are always right next to one another. So of course there is an ongoing consultation between the two of us. Some things I have learned, so we always share.

<Internals\Interviews\BPsy7 Concept> - § 3 references coded [8.21% Coverage]

Reference 1 - 2.36% Coverage

BPsy7
Well you would often have um quite a collaborative model, and I really appreciated that. So if I was working with someone for example and I um…. I was supposed to be working with somebody who had a spinal cord injury for example. But I noticed cognitive issues. I was concerned about the mechanism of injury and the you know the whole presentation of the worker, I could take that person back to the team and suggest in my notes I am thinking there is …so I could suggest okay you know this person might have a cognitive issue, a brain injury that we have missed. Let's have somebody do a neuro psych or let's have the team, the brain injury team look at this guy. Or the neurosurgeon or who ever do a consult to talk about the situation because…Or, similarly if um in a drug and alcohol session somebody was talking about you know their hopelessness and their problems with alcohol and how it related to their the job they were trying to get back to and maybe they should be in a different job, the drug and alcohol guy might say you know you really should be talking with a vocational person about this. Matt is just down the hall, do you want to talk to him? And so we would get cross pollination I called it where we would have that freedom…

Reference 2 - 3.23% Coverage

BPsy7
Yes he was a high quad. So he had the highest cervical injury you can have. Just below the atlas bone. So I think it was C7 or C6 spinal cord injury. And the so all he can do is move his head and use a sip and puff. But he went from vent dependent to out in the community, married his nurse, adopted two kids, started two businesses, became a mentor, joined a bunch or organizations as a Board of Directors member, became a fabulously connected, and public speaker, you know Disney wants to do a movie on his life. The guy is a really amazing guy. And he is the head of peer mentors there in terms of spinal cord injuries. He is one of the best counsellors I have ever met. Peer counsellors. But no official training, his life has been his training. And it was interesting because we would sit around the discharge meeting. We would have the physio, the OT, the doctor, the nurses, the voc rehab, the sexual health, the spiritual care person, the drug and alcohol person if they were required. You know all those people were there. And the different people that provided services interacted. And then we had this peer mentor spinal cord injured fellow who had the least education in the room, sometimes less education than the client. But he had the most to teach in terms of life experience with a spinal cord injury. An amazing fountain of knowledge and wisdom in terms of how to manage spinal cord and how to live with the consequences of the changes. Because his the consequences to his life were so profound. And he had moved through them so well.

Reference 3 - 2.62% Coverage

BPsy7
Yes and I think that often times with the different levels or the different kind of training, because there was lots of cross training, um you know one of the things that was really fantastic I hadn't mentioned, we had an art therapy group. And so the VAMS group was a group of rock and roll

that would come in and play music with folks. And so you would get these guys who have serious brain injuries but they put out an album. You know they actually wrote an album together. That's an awesome therapeutic thing. The therapeutic department the rec department, because there was a whole department just based on rec. And sometimes the rec guys were so gung ho, you know they were extreme cyclists and extreme whatever, I mean we had I remember a spinal cord injured guy who um a great guy, the last day of his in patient program, he had been in the program I think it was six months or a year, and he wanted to do something big and symbolic on the last day. And so he decided to go parachuting. And he was strapped to another guy. And they and so the able bodied jumped out of the plane, and they hadn't quite worked out all the details as well as they could have, and this guy was tied onto this other fellow's chest. I think they were back to chest. And um you know he had just gone through this very expensive program, I think it is $1000 a day in a traction bed in the acute phase.

<Internals\Interviews\I8 Concept> - § 2 references coded [3.21% Coverage]

Reference 1 - 2.76% Coverage

I8
I don't….well if it is, it is individual. And I don't know about that because the majority of the senior investigators you know they are quite able to pick up the phone and contact people. What you have here however, is that within this group there is a percentage of senior level investigators who know what is going on if you will, and how to do things. So they know which individuals to contact that have the information or the intelligence by that I mean knowledge, to assist them with their problems. Some of the newer people don't have that background they don't know the senior officers, uh, and are blissfully ignorant quite honestly.

Researcher
Well the older ones are mostly from policing background.

I8
That is right.

Researcher
And what are the newer ones from?

I8
Some mixture of government employees, maybe one or two from police force but not necessarily the Canadian police force, things like that. So out of the loop so much of this work is based on contacts and you know past experiences and intelligence of what is happening out there.

Researcher
So there isn't too much consistent collaboration among the twelve people.

I8
I would say none, not any more.

Researcher
None, not any more? Did there used to be?

I8
Yes.

Reference 2 - 0.45% Coverage

I8
Well you know again, I mean employees can be asked for opinions and suggestions, but if they are not acted upon pretty soon people stop making suggestions because they feel their opinions are meaningless.

<u>Total conceptual codes created</u>

- A total of 22 conceptual codes were created after conceptual coding of all nine interviews and documents from all four sites.

<u>22 conceptual codes</u>

Accurate information, Adapting, Bias, Client satisfaction, Collaboration, Communication, Culture, Diversity, Feedback, Formal evaluation, Funding, Goals, Integrity (in vivo code), Learning, Meaningful evaluation, Performance based model, Qualifications, Rapport, Reports, Self evaluation, Standards, Stigmas.

<u>Analyses of conceptual codes</u>

- Following conceptual coding I returned to reviewing the transcripts, documents and NVivo conceptual code files multiple times comparing data to data, data to codes, and codes to codes.
- I conducted constant comparisons of data, codes, and the literature
- I analyzed whether data had been missed that should have been included within one of the 22 conceptual codes, or whether additional conceptual codes had been missed and should be created.
- I categorized the 80 open codes under the 22 conceptual codes to compare open codes to conceptual codes.
- I identified conceptual codes that should be collapsed and new conceptual codes that should be added. (Examples of collapsing and adding new conceptual codes was illustrated in a figure in Chapter 5 of this study.)
- I created memos based on the conceptual codes.

- An example of a memo based on the conceptual code 'collaboration':

> (1) Consistency
> (2) To standardize products of service
>   ~with other professionals in the organization
>   ~with other professionals in program
>   ~with learned experience clients
>   ~conference calls and presentations
>   ~with other organizations
> (3) Defer to other specialists
> (4) Knowing when to collaborate versus work alone

- I continued constant comparisons of: conceptual codes to each other, to open codes, to the data, to ROLE results, to the research questions, and to the literature.
- Patterns among codes emerged.
- I hypothesized ways codes could be categorized together in meaningful ways related to disability management evaluation and learning.
- Themes began to emerge from the code groupings.
- This process clarified how some of the 22 concepts were more important than others to understanding disability management evaluation.

<u>Selection of concepts important to the emerging themes</u>

- 11 of the original 22 conceptual codes were considered robust in relation to the emerging themes and were retained:

    Adaptation, Bias, Client satisfaction, Collaboration, Communication, Culture, Diversity, Evaluation criteria, Feedback, Learning, Reports.

- Two open codes were elevated to concepts as their importance to the emerging themes became evident:

    Cost savings, Timeliness

- Three new concepts from the original data were coded as concepts as their importance to the emerging themes became evident:

    Client functioning, Data management, Return to work

Methods to define the concept `client functioning` were considered. Returning to the literature I recommended basing client functioning on the World Health Organization (2011) framework that had been created to describe and understand the components of functioning, disability, and health.

<u>Testing themes</u>

- I identified clusters of codes and how they related to each other and hypothesized what stories they told.
- Five themes emerged by grouping the 16 extant conceptual codes into categories that contributed to understanding disability management evaluation across the sites:
  (1) E*valuation Criteria*: *Client functioning, Client satisfaction, Cost savings, Return to work, Timeliness*
  (2) Program context: *Collaboration, Communication, Culture, Diversity, Bias*
  (3) *Learning* from evaluation: *Feedback, Reporting*
  (4) Data management
  (5) Adaptation
- I created multiple visual displays of the codes and groupings to analyze how they related to each other.
- I consulted a vocational rehabilitation consultant expert who provided feedback on my code groupings, themes, diagrams and hypothesized explanations related to disability management evaluation.
- I created a table displaying `conceptual codes and themes` in relation to `the four disability management paradigms included in this study' (included in Chapter 5 of this document).
  I analyzed the table hypothesizing relationships among the concepts and paradigms. The table illustrated concepts present or absent at each site, and rated whether their presence or absence bad been perceived by the nine participants as having been constructive or as having contributed to dissonance at the site.
- Differing evaluation priorities at the four sites was evident. For example, the BM and Insurance sites prioritized evaluation criterion and standards measurements (i.e. timeliness of reports) whereas the Labour and Biopsychosocial sites tended to integrate pluralistic stakeholders' perspectives gained through collaboration, communication and attention to culture and diversity.
- I compared the themes to results of the ROLE.

<u>Interrelating analysis and theory building</u>
- I hypothesized inter-relating explanations of disability management evaluation based on analyses of themes across the four sites.
- This included consideration of disability management activities, evaluation activities within programs, and interactions between programs and their contexts.

<u>Collaboration</u>
- *Collaboration* was identified as a core theme as its presence or absence was hypothesized as having he potential to contribute significantly to an understanding of variability within disability management evaluation.
- I hypothesized that the concept *Communication* contributed to and should integrated under the concept *Collaboration*.
- *Collaboration* and *learning* were analyzed in light of the ROLE findings.

Context

- Understanding how programs responded to *contextual influences* was identified as a core theme having potential to contribute significantly to an understanding of variability within disability management evaluation.
- Contexts of disability management programs were analyzed according to the five factors suggested by Rog (2012) for context analysis. Contexts of: the problem being addressed; the interventions being examined; the broader environment or setting; the evaluation context; the decision making context.

Evaluation procedures

- Five main evaluation criteria were identified: client functioning, client satisfaction, cost savings, return to work and timeliness.
- Disability management programs were increasingly using technology to develop data management systems that could be used for future research.

Theoretical perspectives

- To further a theoretical understanding of disability management evaluation, evaluation at the four sites was examined with respect to evaluation theories in general, and with respect to the logic of evaluative inquiry.
- Evaluation theory across the four sites mainly emphasized USE of findings, with secondary concern given to the pluralistic VALUES of multiple stakeholders.
- The logic of evaluative inquiry across the four sites mainly followed a CONSUMER approach. The BPsy site also followed connoisseurship and pluralist approaches.

Appendix D

<u>Open Codes</u>

Bias overcome

Biomedical model site

Biopsychosocial model site

Claims managers highest stress

Client credibility

Client summative evaluation of program

Camaraderie among clients

    Communication with external stakeholder

Consultants to programs

Contributing new evidence

Cost effectiveness of program

Cultural issues

Cultural labour relations roles and issues

Cultural psychological disability

Culture political limited funding

Define clients for this organization

Define disability for this organization

Define service provided by this organization

Disabled DM counsellors

Evaluation needs assessment

Evaluative criterion

Fraud example

Gender

How participant knows program is doing what was intended

Impartiality

Insurance model site

Integrity (in vivo code)

Introspection

Labour joint union management

Labour model site

Labour model limited information on disability

Learning

Multidisciplinary cross pollination

Multidisciplinary early intervention

Multidisciplinary risk taking

Non disability management work of programs

Non validity

Obtaining employer input

Participant criterion of program services

Participant description of program

Participant evaluation of their program services

Participant formative evaluation of their program services

Participant learning from client feedback surveys

Participant opinion new computer system interferes

Participant qualifications

Participant suggested disability management evaluation criterion

Participant summative evaluation of their program service

Participant training NIDMAR return to work coordinator

Participant program evaluation background

Participant's role in the organization

Peoplesoft disability management record keeping software

Performance based evaluation

Political limited funding

Preventative aspect of disability management programs

Problem solving

Problem with increased temporary workers

Program clients

Program situation within organization

Referring organization communication with participant program

Referring organization criterion for program services

Referring organization formative evaluation of the program

Referring organization summative evaluation of the program

Referring sources

Reporting evaluation findings

ROLE

ROLE – open and accessible work environment

ROLE – area needing improvement

ROLE – collaboration and problem solving

ROLE – decision making

ROLE – strengths from organizational learning

Setting specific behavioural objectives

Sharing personal beliefs

Social learning copying standards of others for reports

Stigma

Surveillance

Survey

Timeline of disability management for services program provides

Validity

Weekly statistical summary of services provided

Who is responsible for evaluating your program

Appendix E

Groupings of Open Codes from the Visual Chart

Descriptions of the disability management sites

1. Participant description of program
2. Define service provided by this program
3. Define disability for this program
4. Program clients
5. Define clients for this organization
6. Program situation within organization
7. Referring sources
8. Participants role in organization
9. Participants PE background
10. Who is responsible for evaluating your program
11. How participant knows program is doing what was intended
12. Participant evaluation of their program services
13. Participant suggested Dm evaluation criterion
14. Participant formative evaluation of their program service
15. Participant summative evaluation of their program service
16. Client summative evaluation of program
17. Participant learning from client feedback surveys
18. Cost effectiveness of program

The ROLE

Learning
Problem solving
ROLE – open and accessible work environment
ROLE – area needing improvement
ROLE collaboration and problem solving
ROLE decision making
ROLE strengths from organizational learning

External evaluation of the biomedical model site

Referring organization criterion for program service
Referring organization communication with participant program
Referring organization formative evaluation of program
Referring organization summative evaluation of program

Organizational cultural issues

Culture psychological disability
Problem with increased temporary workers
Participant opinion new computer system interferes
Claims managers highest stress
Labour model limited information on disability
Cultural labour relations roles and issues