A CASE STUDY ON THE USE OF DEVELOPMENTAL EVALUATION FOR
INNOVATING: NAVIGATING UNCERTAINTY AND UNPACKING COMPLEXITY

by

CHI YAN LAM

A thesis submitted to the Faculty of Education
in conformity with the requirements for
the degree of Master of Education

Queen’s University
Kingston, Ontario, Canada
December, 2011

Copyright © Chi Yan Lam, 2011
Abstract

Developmental evaluation (Patton, 1994, 2011) is one of the latest approaches to be introduced into evaluation practice. It purports to support the development of social innovation by infusing evaluative thinking through collaboration between program clients and the developmental evaluator (Patton, 2011). In an attempt to build “practical knowledge” (Schwandt, 2008) about this emerging approach, this research seeks to investigate the capacity of developmental evaluation to support innovation.

This thesis reports on a case study of the Assessment Pilot Initiative (API) where developmental evaluation was used to support the development of a novel approach to teacher education. Charged with a vision to innovate their own teaching practices and the learning of teacher candidates, the instructors of the case invited a developmental evaluator onboard in a yearlong collaboration. While the instructors, along with the developmental evaluator, were uncertain about the outcome of the initiative or how best to proceed, this engagement resulted in a novel adaptation of microblogging web technology (Twitter) that came to be piloted with a group of teacher candidates.

This thesis presents an analysis of the development process and the contributions developmental evaluation made in enabling the development of the API. Such analysis is anchored in the records of the program development, and in the perspectives of the program clients and the developmental evaluator. Analyzing the program development records for developmental moments revealed certain trends and patterns that, when triangulated with interview data from program clients and with reflections from the developmental evaluator, provided intricate insights into how the development came about and of the contributions developmental evaluation made in this case.
Development of API proceeded in a highly nonlinear, emergent process through six foci of development. Critical to addressing the uncertainty and complexity that might had otherwise inhibited development, developmental evaluation enabled a data-informed approach that lent a quality of responsiveness to the emergent, evolving nature of the initiative. The developmental evaluator was instrumental in identifying activities that helped make explicit values and assumptions underpinning the initiative and in structuring a learning framework to engage program clients in sense-making. The notion of design emerged from analysis as an important function of developmental evaluation. Implications of the findings are discussed.
Acknowledgements

What a journey this had been! I could not have done this without the support of many people. This work is a celebration of many invaluable opportunities that shaped my thinking as a scholar and as an evaluator.

First, I would like to thank Dr. Lyn Shulha, my thesis supervisor, for her continuing support of my work as I continue to explore the boundaries of evaluation. Her nurturing style of mentorship pushed me to excel academically and intellectually.

I would also like to thank Dr. Richard Reeve, my committee member, and Dr. Ruth Rees, my examiner, for their feedback on my work. Your edits and suggestions were invaluable as I try to wrestle with complexity.

I would also like to thank my clients, “Laurene”, “Albert”, and “Tim”, and the teacher candidates who participated in their pilot program. To my clients: you confronted the disabling feeling of uncertainty as we forged forward in the development. You trusted me, as your developmental evaluator, to lend support to your cause. I hope you found the developmental evaluation useful in furthering your attempts at innovating teaching and learning.

Special thanks to my colleagues in the Assessment and Evaluation Group at Queen’s University for your wisdom, support, and the opportunities to collaborate: K. Luu, C. Doe, L. McEwen, C. Donnelly, J. Woods, A. Chapman, Q. Wong, and E. Katsoulas. Thank you to M. Sayers for her help in transcribing countless hours of recordings. Thank you to the staff in the Grad Studies & Research office for helping me navigate the program requirements. Dr. K. Frederickson, thank you for inspiring me to think educationally. Dr. S. Matrix, thank you for inspiring me to think about innovative
uses of social media. Dr. D. Stockley, thank you for being a wonderful client to my first evaluation.

Finally, I would like to acknowledge the financial support from the Social Science and Humanities Research Council of Canada, Assessment and Evaluation Group, and Queen’s University.

I dedicate this work to my parents and to my supervisor.
# Table of Contents

Abstract...................................................................................................................................................... ii  
Acknowledgements ......................................................................................................................... iv  
Table of Contents.......................................................................................................................... vi  
List of Figures........................................................................................................................................ ix  
List of Tables ......................................................................................................................................... x  
Chapter 1: Introduction .................................................................................................................... 1  
  Introduction to Research ..................................................................................................................... 1  
  Purpose of the Research ....................................................................................................................... 2  
  Research Questions .............................................................................................................................. 3  
  Organization of the Thesis ................................................................................................................... 3  
  A Case for Unconventional Course Planning ............................................................................... 6  
  The Nature of Innovation .................................................................................................................. 9  
Chapter 2: The Developmental Evaluation Context ........................................................................ 12  
  Chapter Overview ............................................................................................................................ 12  
  Contemporary Classroom Assessment ............................................................................................ 12  
  Classroom Assessment Pedagogy ..................................................................................................... 14  
  Classroom Assessment Education at Queen’s University ............................................................ 17  
    Developing a Pilot Initiative in Classroom Assessment ............................................................. 20  
  Conclusion ......................................................................................................................................... 25  
Chapter 3: Theoretical Framework ................................................................................................. 27  
  Chapter Overview ............................................................................................................................ 27  
  Evaluation Theory and Research into Evaluation ....................................................................... 27  
  The Theory of Developmental Evaluation ..................................................................................... 29  
    Social Innovation and Complexity ............................................................................................. 30  
    Developmental Evaluation and Complexity Science ............................................................. 33  
  The Practice of Developmental Evaluation .................................................................................... 33  
  Conclusion and Establishing Lines of Inquiry ............................................................................. 36  
Chapter 4: Method ............................................................................................................................ 38  
  Chapter Overview ............................................................................................................................ 38  
  The Research Context, Revisited ................................................................................................... 38  
  Case Study Research as a Method ................................................................................................. 40  
    Conducting Qualitative Research ......................................................................................... 42  
    Study Participants ..................................................................................................................... 43  
  Aligning Questions to Methods ..................................................................................................... 44  
  Data Sources and Data Collection Strategies ............................................................................. 44  
    Interviews .......................................................................................................................... 46  
    Observational & Other Participant-contributed Program Data ............................................. 46  
    Transcripts of Meetings ......................................................................................................... 47  
    Program Artifacts ............................................................................................................... 48  
    Ethical Considerations ........................................................................................................... 48  
  Data Analysis ................................................................................................................................. 49  
    Stage One: Building the Evidentiary Base ............................................................................ 49  
    Stage Two: Thematic Analysis .............................................................................................. 52  
    Stage Three: Time-Series Analysis ....................................................................................... 55
Stage Four: Coding Interview Data.................................................................................. 57

Enhancing the Trustworthiness of Case Study Research .............................................. 57

Chapter 5: Results of the Study..................................................................................... 60

Chapter Overview .......................................................................................................... 60

Developmental Episode One: Evolving Understanding in using Social Media for
Professional Learning ..................................................................................................... 61
  Situating the episode. ..................................................................................................... 61
  Description of episode................................................................................................. 62
  Significance of episode. ............................................................................................... 70

Developmental Episode Two: Explicating Values through Appreciative Inquiry
for Program Development ............................................................................................. 74
  Situating the episode. ..................................................................................................... 74
  Description of episode................................................................................................. 75
  Significance of episode. ............................................................................................... 81

Developmental Episode Three: Enhancing Collaboration through Structured
Communication............................................................................................................... 83
  Situating the episode. ..................................................................................................... 83
  Description of episode................................................................................................. 84
  Significance of episode. ............................................................................................... 85

Developmental Episode Four: Program Development through the Use of
Evaluative Data............................................................................................................... 87
  Situating the episode. ..................................................................................................... 87
  Description of episode................................................................................................. 88
  Significance of episode. ............................................................................................... 91

Chapter 6: Discussion of the Study ............................................................................. 96

Chapter Overview .......................................................................................................... 96

Revisiting Research Question One: To what extent does Assessment Pilot
Initiative qualify as a developmental evaluation?......................................................... 97
  Preformative development of a potentially broad-impact, scalable innovation....... 97
  Trends in program development. .............................................................................. 100
  Points of divergence. .................................................................................................... 102

Revisiting Research Question Two: What contribution does developmental
evaluation make to enable and promote program development? ............................. 102
  Developmental evaluation as a data-informed process of innovation................. 104
  Implication of adopting a data-informed approach to innovation: Responsiveness in
developmental evaluation......................................................................................... 106
  Consequences to adopting a data-informed approach to innovation: Resolving
  uncertainty...................................................................................................................... 107
  Data-informed developmental evaluative inquiry in action..................................... 109

Revisiting Research Question Three: To what extent does developmental
evaluation address the needs of developers in ways that inform program
development? .............................................................................................................. 113
  Defining stakeholder needs in a developmental evaluation. .................................. 113
  Learning in a developmental evaluation. ................................................................. 114
  Values and valuing as a focus of learning in a developmental evaluation............ 116
  Enacting learning activities in a developmental evaluation................................. 117
Revisiting Research Question Four: What insights, if any, can be drawn from this development about the roles and the responsibilities of the developmental evaluator?

Chapter 7: Conclusion

A Case of Split Identity: Evaluation or Development?
Developmental Evaluations in a Context of Innovation
Design as an Activity and Function unto itself.
Design Thinking and Thinking Evaluatively
Roles of Developmental Evaluator and Knowing Innovation
Role of the Developmental Evaluator as a Facilitator of Learning
Reflections on Conducting Case Study Research
Conclusion: Innovation and Developmental Evaluation as Design Thinking
Implications to the Field of Evaluation
Future Directions
Final Words: A Metaphor for Understanding Developmental Evaluation Use in a Context of Innovation
Aviate, Navigate, Communicate

References

Appendix A Policy Documents in Support of Contemporary Classroom Assessment
Appendix B Letter of Information
Appendix C: Consent form
Appendix D: Interview Guide
Appendix E: Letter of Ethics Clearance
Appendix F: Ethics Certificate
Appendix G: List of Files Reviewed for Analysis
Appendix H: Sample of Document Memo (Doc. 12)
Appendix I: Samples of Assessment Pilot Initiative Document Output (Doc. 107)
Appendix J: Infographic Illustrating a Simplified Model of the Developed Program Model (Doc. 137)
Appendix K: Program Development Timeline
Appendix L: Summary of Results from Appreciative Inquiry Exercise
Appendix M: Design Brief of Assessment Pilot Initiative (Doc. 83)
Appendix N: Theory Spec Sheet (Doc. 94)
Appendix O: Sample of Activity Summary (Doc. 129)
List of Figures

Figure 1. Visualization of the Assessment Pilot Initiative development process........................................................................................................57

Figure 2. Photograph of the output from an appreciative inquiry exercise............................................................................................................79
**List of Tables**

Table 1: Summary of the data sources and analysis used to answer each of the research questions.................................................................45

Table 2: Summary of the different data types and the colour codes assigned to each.................................................................51

Table 3: Descriptions and dates of key developmental episode.......................52

Table 4: Summary of developmental evaluation data collected in the Assessment Pilot Initiative.................................................................104

Table 5: Summary of lessons learned about developmental evaluation from this research.................................................................138
Chapter 1: Introduction

Introduction to Research

This case study is situated at the nexus of teaching and learning, program evaluation, and educational innovations. It describes how a group of educators, each with different backgrounds, interests, values, and substantive skills, worked with a developmental evaluator to overcome teaching and learning constraints and to reimagine how teacher education could transpire. Specifically, it describes how a shared vision of furthering and improving teacher candidates’ learning in assessment pedagogy necessitated a departure from the expected and comfortable. The study describes the processes of adaptation and integration of a novel educational technology and how this was guided by the theories and practice of developmental evaluation. This engagement resulted in a unique approach to the classroom assessment education program at the Faculty of Education, Queen’s University. At its core, this is a case study about an innovation and the extent to which an emerging approach in program evaluation, namely, developmental evaluation, supported the innovation process.

All course work in assessment education is intended to educate teacher candidates in the integrated use of classroom assessment, i.e., assessment of learning, assessment for learning, and assessment as learning, in ways that are both consistent with sound contemporary understandings of learning and reflective of today’s expectations of the profession. For a variety of reasons to be explored in the next chapter, efforts at nurturing teacher candidates’ assessment pedagogy within the required module of PROF 150/155, Classroom Assessment were proving difficult. While learners’ needs have largely been identified through ongoing program evaluations, the two instructors responsible for the
program were not satisfied with the ways the educational program structure, including their own approaches to instruction, were responding to these needs. In July 2010, I was invited to assume the role of a developmental program evaluator to collaborate with these instructors (Laurene, Albert, and Tim) as my clients in developing and implementing an innovative approach to assessment education. This experience came to be referred to as the Assessment Pilot Initiative (API).

**Purpose of the Research**

Developmental evaluation is one of the latest approaches to be introduced into the field. First described by Patton in 1994, it only became codified in 2010 through his groundbreaking 375-pages text *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. As a result of this recent development, there has been a real thirst amongst the evaluation community to understand the use, implications, and consequences of applying developmental evaluation approaches to different contexts. This case study responds to a more general call from the evaluation community for evaluation cases to be systematically analyzed in order to generate “practical knowledge” about different approaches to evaluation (Schwandt, 2008). A case study of a purposefully implemented developmental evaluation is particularly timely given its recent emergence and still tentative acceptance as an evaluation approach.

To address this purpose, this case study describes and assesses the extent to which working with a developmental evaluator (a) contributed to an instructional team’s efforts at creating an innovative module in student assessment and (b) supported the efforts of both instructors and their learners as they explored the instructional potential of social networking technology as a vehicle for developing an assessment pedagogy.
Research Questions

To achieve this purpose, I will address the following questions:

1. To what extent does Assessment Pilot Initiative qualify as a developmental evaluation?
2. What contribution does developmental evaluation make to enable and promote program development?
3. To what extent does developmental evaluation address the needs of the developers in ways that inform program development?
4. What insights, if any, can be drawn from this development about the roles and the responsibilities of the developmental evaluator?

Answering these questions requires that the learning of all those engaged in the process, namely, the instructors, teacher candidates, and the author (as the developmental evaluator of the API) be made explicit.

Organization of the Thesis

The remainder of this thesis reports on the research process and the research findings. Because it is inherently difficult to describe a non-linear and emergent process in a linear way, two strategies are introduced as a way to frame the description of the study. Both provide a visual representation of how the thesis will unfold.

The focus of this research lies in understanding a developmental process as a function of time. To underscore the change process, I use a process that I refer to as book-ending. This process allows me to begin with a description of the initial conditions that gave rise to the project, followed by a description of the project outcomes, before recounting the developmental process. The intention of book-ending is to allow the
The ultimate significance of particular episodes in the development process to reveal themselves within the description, making it easier for the reader to garner the influence of participants and the author in shaping the development.

The second strategy is to have the reader visualize the thesis as the Petrushka dolls. As each doll is opened, a tighter and more refined example of the whole is revealed. Each section of this thesis is written to move from the contextual to the specific and from the theoretical to the practical in order to get to the heart of using developmental evaluation for innovative program development. This metaphor helps to illustrate the notion of unpacking complexity, including the complex nature of the values and relationships that would shape behaviour and decision making at each stage of the project.

The remainder of chapter one describes the two constructs central to this work: the notion of development and the notion of innovation. For the first, it explores the incompatibilities that arise when conventional planning frameworks are put to use in a complex program context and argues for the adoption of a developmental approach to guide development in such contexts. For the second, I discuss the distinction between innovation and improvement, and suggest that innovation requires both a departure from the norm and the comfortable, which necessitates an engagement with the uncertainty inherent in such efforts.

Chapter two begins by describing the theoretical and instructional contexts of the Assessment Pilot Initiative. It tracks the influences of contemporary classroom assessment, how these differ from traditional approaches, and finally relates these influences to the theoretical and practical complexities of teaching classroom assessment. In doing so, I set forth the foundation of the problem confronting the instructors with
whom I collaborated. The chapter concludes with a description of the ongoing efforts of two instructors at the Faculty of Education, Queen’s University, at educating teacher candidates in assessment pedagogy. This description explores briefly the institutional contextual constraints that impinge on assessment learning, and how my involvement at the Faculty as a graduate student evaluator/educator put me in an appropriate position to aid in the development of a novel instructional approach to assessment education.

Chapter three discusses the theoretical framework that informs the evaluation and program development. It presents the emergence of developmental evaluation as an approach to planning using data-informed decision-making and by comparing this approach to more conventional theoretical frameworks of planning and evaluation. Particular attention is given to the promise of developmental evaluation as non-linear response sensitive to complex program contexts.

Chapter four outlines the methods of inquiry used in this research including an argument for conducting this research as a case study, and the rationales for how data were collected and analyzed. It concludes with a specific plan describing the method used to answer each of the research questions.

Chapter five presents the results of the study by recounting four episodes from the development process to illustrate specific features of the development deemed significant by the clients. In particular, I analyze these episodes and features in relation to the theory on developmental evaluation. While some of these features were in keeping with the theory, what emerged from that analysis were the distinct kinds of actions, behaviours, and thinking the developmental evaluator assumed in response to the demand of developing a program from scratch.
Chapter six discusses the findings of the study and attempts to answer the four research questions that frame this study. Analysis of the case data revealed the importance of facilitating learning in a developmental evaluation to enable development. It highlighted the multiplicity of roles that a developmental evaluator adopted to facilitate the process. The data also converged to suggest six different foci of development that could serve as the basis of future development efforts where the objective is innovation.

Finally, the last chapter explores the significance of the findings. The notion of design emerges as a distinct dimension to developmental evaluation that seems to reconcile the roles, behaviours, and thinking required of a developmental evaluator when working in the context of social innovation and program development.

**A Case for Unconventional Course Planning**

This study was conducted in the Faculty of Education, Queen’s University, Kingston, Ontario, Canada. The instructors in this case were deeply interested and heavily invested in improving the teaching and learning conditions in which the learning about classroom assessment took place. A number of concurrent influences operating at various levels of the educational system collectively created a set of constraints that made this learning difficult for many of the teacher candidates. These factors are discussed below. Consequently, however, there was significant motivation on the part of instructors to reflect regularly on the quality of teaching and learning.

Their resulting vision of innovating their practice and the experiences of teacher candidates underpinned the Assessment Pilot Initiative pilot program. Although never made explicit, there was an implicit acknowledgement among the instructors that a departure from the known and comfortable was necessary to innovate the learning.
experience. However, this willingness to embrace the unknown was tempered simultaneously by uncertainty in how they should proceed and teacher candidates would react to such novel experiences. One of the ways in which the instructional team brought some security into their exploration was through the adoption of developmental evaluation as a methodology/framework to undergird the process. The adoption of this approach ultimately led to a journey that spanned over a year, leading to the development of an instructional program prototype that was piloted with teacher candidates.

Indeed, uncertainty about how teacher candidates would respond to the novel features of the proposed learning was not the only unusual aspect of this development. As the process began, the instructors expressed neither agreement nor even an idea about how they might proceed. These two conditions would typically preclude the use of planning frameworks that rely on the following sequence: 1) analyze the knowledge to be learned and the characteristics of the learners participating in the learning, and using this information to arrive at learning objectives; 2) design instruction to match learning objectives; 3) develop learning resources, such as, instructional activities, assessment tasks, and instructional materials; 4) package the activities and materials into a coherent plan, like a tangible syllabus that can be distributed to learners; 5) implement the course, which generally follows a script of teach-practice-test; and finally 6) evaluate the instruction (e.g., Dick, Carey, & Carey, 2004). The only real competitor for this linear process is the increasingly popular, but still sequential alternative labeled backward-design (Wiggins & McTighe, 2005). Backward design probes the planner to foreground the desirable learning outcomes and then work backwards to consider what evidences might be required to demonstrate confidence that the outcomes had been achieved. It
dispenses with starting the planning process with the consideration “what is to be learned”. Following that determination, the planner considers what instructional material and approaches would contribute to this achievement.

These two planning frameworks carry certain conditions and assumptions about effective instructional development that did not necessarily hold true in this initiative. First, using either of these planning processes can require a heavy investment of time. Typically, much time is spent on negotiating both what and how a topic should be taught. Developing instructional material can be equally time consuming. Second, in either backward design or conventional forward design, it is inherently assumed that: (a) the domain and scope of learning can be agreed upon, (b) objectives are definable and can be articulated, and (c) that learning outcomes are definable and can be articulated. As should be obvious by now, at the starting point of this project, it lacked such clarity and certainty (Laurene, interview; Tim, interview). Third, these approaches presuppose a degree of familiarity and proficiency with the genres of activity instructors will use to guide instruction. In other words, conventional course planning has to do with planning known domains of knowledge and predictable patterns of learning. It was obvious that conventional planning frameworks were incompatible with the context in which the instructors were operating.

An additional unconventional feature of the API was the way the program evaluation function was operationalized. Typically, programs are reviewed, appraised, or evaluated after implementation. Indeed, the PROF 150/155 assessment module had been formally evaluated each of the previous three years following the submission of candidates’ grades. This information was then used to adjust the structure, instruction,
and context for the following year. Determining the merit, worth, and significance of a program once it is completed is legitimate for programs that have established a purposeful planning, implementation, and evaluation cycle, but this is not the only place where evaluation can be situated for the purpose of improvement.

One assumption underlying developmental evaluation is that determining the merit or worth of a program at the end of a planning, implementation, and observation cycle is insufficient to guard against the potential of program failure. Given that teacher candidates at this faculty are currently timetabled for only seven hours of instruction in assessment, any radically innovative approach to how teaching and learning might occur can be considered a high stakes activity. Reacting to unpredictable and undesirable effects of instruction only after program completion is too late. Developmental evaluation was selected as a framework for development precisely because it is promoted as a way of buttressing program decisions throughout all phases of the program design and implementation cycle with the use of evaluative data.

**The Nature of Innovation**

A construct central to this work is the notion of innovation. Building on the conceptual work of Patton (2010), I define innovation in this study as denoting both the process and the product in which a novel approach to instruction is purposefully developed to address a perceived need. Michael Quinn Patton, father of developmental evaluation and utilization-focused evaluation, agrees that, when innovating, the process is the product (2010). An implication of this thinking is that innovation can both, and simultaneously, refers to a change in process and in product. The product (i.e., the innovation) that results from engaging in development (i.e., innovating) is co-dependent
on the process taken to innovate; the kinds of thinking, activities, learning, and insights that gets generated during the process of innovating inevitably shape the resulting innovation. Having said this, the focal point for this research is on the process of innovating and the role developmental evaluation played in supporting problem solving and decision making throughout this process.

I differentiate innovation from improvement in that innovations are evolutionary leaps. Innovation necessarily disposes with some aspects of one paradigm (Kuhn, 1970) in exchange for a substantially different approach or way of thinking about a problem. Innovation necessarily unsettles at least one aspect of the conventional way of doing and thinking in exchange for novelty that is presumed to address previous shortcomings of the former paradigm. In contrast, improvement is typically about tweaking or modifying, the thinking or behaviour in such a way that generates slight advantages in process or outcome, but where the predominant modus operandi, i.e. paradigm, remains largely unchanged.

Kuhn, in analyzing how scientific revolutions come about, concluded, "normal science does not aim at novelties of fact or theory and, when successful, finds none" (1970, p. 52). He is suggesting that much of "normal science” would qualify as improvements. Only when we “aim at novelties of fact or theory”, do we succeed in innovating. By making an analogy between normal science and conventional program evaluation, Kuhn’s observation suggests that improvements made to a program may not necessarily address any of the fundamental assumptions underpinning a current programmatic approach. Hence, program evaluations performed for the purpose of
improvement do not necessarily bring to bear the kind of scrutiny necessary for innovation.

Having worked with social innovators for the past two decades, Patton observes that vision often guides innovation (2010). Beyond that level of certainty, the means for problem identification, problem solving, or for identifying outcomes are all far from being explicit or predictable. In spite of a lack of certainty, social innovators recognize a need to act and not stand still, and they embrace the uncertainty and complexity of innovation while accepting the potential and cost of failure. In that regard, Patton suggests that innovators value the value-added exercise of infusing the development process with evaluative data. Developmental evaluation, as will be discussed in chapter three, is built on these theoretical premises to enable innovation. The Assessment Pilot Initiative grew out of an expressed desire by instructors to innovate the teaching and learning experience. Developmental evaluation was implemented to support the efforts of these instructors.
Chapter 2: The Developmental Evaluation Context

Chapter Overview

The purpose of this chapter is to describe the developmental evaluation context in which the initiative took place. It is presented in two parts. In part one, I first introduce the theoretical landscape surrounding contemporary classroom assessment and how that relates to the teaching of classroom assessment at the level of preservice teacher education. Then, I describe the specifics of a classroom assessment education program at Queen’s University and some of the challenges that it is facing. In part two, I describe how a pilot initiative came to be developed in response to the constraints that my clients envisioned to innovate and the circumstances under which a technological approach came to be adopted to inform the development of the initiative.

Contemporary Classroom Assessment

Research produced within the past few decades in classroom assessment, cognition, and learning has contributed to advancing our understanding into the use of classroom assessment in an integrated fashion as part of the learning process. Based on these discoveries came a new understanding of how contemporary classroom assessment could add to student learning over traditional practices. Historically, testing followed teaching (Shepard, 2000a, 2000b). In today’s classrooms, although assessing for making judgments about student achievement is still necessary, teachers are expected to practice an expanded and integrated repertoire of assessment practices; testing for making judgments about achievement remains necessary but is no longer sufficient. Contemporary approaches to classroom assessment focuses on using assessment strategies to support student learning in a more developmental way.
Within the purview of this thesis, I use McMillan, Hellsten, and Klinger’s definition of classroom assessment\(^1\) as the systematic “collection, evaluation, and use of information to help teachers make decisions that improve student learning” (2011, p. 9; emphasis added).

The seminal work of Black and Wiliam (1998) provides an empirical basis to the argument for using this integrated and expanded approach to contemporary classroom assessment. In their work, the authors surveyed the extant body of literature on formative assessment and found strong evidence in linking assessment practices implemented for the purpose of monitoring and promoting student learning to student success. They argued for modern classroom assessment to be conceived of as a cyclical process along with teaching and learning, as opposed to one-off summative testing. Furthermore, Black and Wiliam’s findings converge with those arising from fields like psychology, cognitive psychology, and the learning sciences. These findings are summarized in two excellent

\[\text{-------------}\]

\(^1\) Classroom assessment is not to be confused with other forms of educational assessments. Although accountability assessment or large-scale standardized testing is prevalent in today’s classrooms, they differ in purpose, design, and use from classroom assessment. Appropriate assessment practices require that classroom teachers be cognizant of the purposes, assumptions, and limitations behind different forms of educational assessment for correct interpretations of assessment data.

The *assessment of learning* (AfL) is now only one of three foci in contemporary classroom assessment. Joining it are practices that focus on: 1) the provision of purposeful and meaningful feedback for the purpose of promoting learning, referred to as the *assessment for learning* (A4L) (e.g., Stiggins, 2007), and 2) the development of metacognitive skills in the learners themselves through practices associated with using *assessment as learning* (AaL) (e.g. Earl, 2003, 2006). Data generated through these two newer approaches serve to inform and shape subsequent learning for students and teachers and guide learners in becoming more self-aware and self-regulated as learners. By replacing judgments about achievement with the promotion of learning as the primary purpose of assessment in classrooms (Ontario Ministry of Education, 2010), educational constituencies are shifting the conversations and practices, and thus the paradigm for thinking about assessment.

**Classroom Assessment Pedagogy**

One consequence of this shift is that judgments about student learning will now be embedded within the daily routines of teaching and learning. Heritage (2010), in her report on the assessment reform movement currently taking place in the US, argues that successful assessment practice is no longer fundamentally concerned with the properties of the instruments used or the frequency of testing, but rather with the *process* of conducting assessment (p. 19). Teachers bear the central role of enacting sound assessment in a coordinated effort; it is becoming clearer how the use of grades as a form of punishment or control over students can in fact be detrimental to learning (Shepard,
A capacity to implement sound assessment and, in turn, to use data derived from assessment instruments and processes to make valid inferences and informed decisions on how to promote learning and achievement are ultimately central to a teacher’s assessment pedagogy (cf. Lam, Donnelly, Woods, & Shulha, 2010).

This newer notion of assessment is reflected in professional standards intended to guide policy-making and professional teacher practice (see Appendix A). However, despite formal expectations, research into enacted classroom assessment practices reveals both a gap in implementation and the complexity in embedding assessment for learning principles into daily teaching and learning (e.g., see Bachor & Baer, 2001; Campbell & Evans, 2000; Galluzo, 2005; Graham, 2005; Mertler, 2003, 2009). This has implications, not only for students, but also for aspiring teachers. Anecdotal evidence from teacher candidates during the 2010-2011 in one faculty academic year suggests they have few opportunities to be mentored by supervising teachers who are both informed and experienced in assessment for learning and assessment as learning processes and practices.

For teacher candidates generally, this lack of direct experience is compounded by the fact there continues to be relatively little emphasis on developing a classroom assessment pedagogy during the academic component of teacher professional preparation programs. For instance, in some states in the U.S., competencies in assessment are only required for teacher certification (Stiggins, 1999). In Canada, although most provincial bodies responsible for teacher certification stipulate some degree of assessment competencies in broad terms, assessment education within pre-service teacher education programs can exhibit a discordant array of approaches and varying levels of purposeful
integration into discussions of teaching (DeLuca & McEwen, 2007). These conditions make it difficult for beginning teachers to develop a level of proficiency or articulate how they might implement assessment practices congruent with contemporary assessment policies and practices during their initial practice years.

Given the pivotal role teachers play in student learning, it stands to reason that classroom teachers are the most important stewards in this assessment reform towards sound, contemporary practice. To implement this reform, teachers require 1) a theoretical understanding of assessment theory, 2) a capacity to implement an expanded repertoire of assessment strategies, and lastly, 3) the philosophical understanding and commitment to classroom assessment for student learning and success (cf. DeLuca & Klinger, 2010). Given that assessment for learning has been advocated now for over a decade, it is reasonable to suggest that contexts where this is still not happening actually do reflect the crisis in professional practice that Stiggins (2002) was describing.

Assessment scholars, researchers, teacher educators, and Faculties of Education have a collective moral and practical obligation to respond to this void in sound assessment capacity in the profession by guiding teacher candidates in the development of an informed and sound assessment pedagogy. My research will document a particular response to this call and examine the role that a developmental evaluator can play in supporting the two instructors as they work to guide teacher candidates in the development of an assessment pedagogy that embraces contemporary approaches to assessment practice.
Classroom Assessment Education at Queen’s University

Queen’s University began offering formal instruction in classroom assessment to all 700 of its pre-service teachers in 2005 following an accreditation recommendation from the Ontario College of Teachers (DeLuca, Klinger, Searle, & Shulha, 2010). The Classroom Assessment Module (CAM) is offered as one of four modules within a teaching foundation course aimed at building foundational knowledge and skills necessary for the beginning teaching professional. The CAM offering complements an already-established half-term elective course in assessment. Mounting this module was a huge undertaking given it had to take into account the range of experiences and diversity in teaching backgrounds represented across the 700 candidates required to take the course. It also introduced logistical issues surrounding teaching, learning and assessment that faculty members had not had to consider when working in traditional classrooms with 40 or fewer candidates.

Two instructors co-conduct the CAM to 350 teacher candidates at a time in one-hour lectures. The content of the lectures is modified slightly by certification tracks (primary-junior and intermediate-senior), with approximately equal number in attendance for either section. Although the number of instructional hours varies slightly from year to year, approximately eight hours are allotted each year to introducing teacher candidates to all three forms of contemporary classroom assessment (i.e., assessment of learning, assessment for learning, and assessment as learning). Didactic lecture-style instruction constitutes the predominant method of instruction during lecture hours, with case-problems, group discussions, and most recently, in-class student response systems (clickers) complementing the in-class learning experience. Teacher candidates are
coached in the principles of self-regulated learning in advance of being introduced to the array of module resources.

Recognizing that these instructional activities play a contributory role towards student learning, the module can appropriately be considered as a formal program with goal-aligned activities aimed at achieving an outcome for a group of participants. Conceptualized in this way, program evaluation theory and tools allow for the systematic collection and analysis of evidence to make valid evaluative inferences about the merit, worth, and significance of the instructional efforts.

Information gathered from four, yearly, post-module program reviews conducted by graduate student evaluators led to various infrastructures and support mechanisms being developed to support student learning:

- The use of teaching assistants as an alternate source of expertise in providing course administrative support, and in providing marking of and feedback for student assignments;
- The employment of research assistants to conduct research into assessment education and course development;
- The establishment of a lab as a space where teacher candidates could meet and consult with teaching assistants regarding the course content or to seek technical support; it also houses a resource library of assessment textbooks and literature; and
- The creation and maintenance of an instructional web site as:
  - A repository of lecture PowerPoint files, and,
A virtual environment featuring prior learning assessment scenarios, self-contained learning modules on each of the three assessment approaches, and other problem-based scenarios.

Evaluation findings were also used to identify potential areas of research and development for the following years. As a set, the evaluation findings revealed the complexity and challenges in preparing teacher candidates to practice integrated classroom assessment. These findings revealed two major problems both related to the instructional context for assessment:

1. As a function of dwindling and limited resources, instruction delivered via large-group lecture format was reported to be impersonal and learning was identified as a solitary experience; and,

2. Opportunities and experiences for teacher candidates to actively examine their personal values, beliefs, and assumptions, as these relate to teaching and assessment, were constrained by limited time and space (anchored auditorium seating) and thus less than effective. (Shulha & Klinger, 2010)

An obvious and conventional approach to overcoming these contextual constraints would be to increase instructional hours, strengthen instructor-candidate face-to-face connections, and augment the infrastructures and resources available to the candidates (e.g., more teaching assistant hours to provide more intensive feedback on assignments). These solutions, sadly, are impractical in today’s higher education climate where access to funds is severely limited. Pragmatism dictates that any innovation in assessment education must be feasible within the current and foreseeable institutional climate and constraints. For this reason, the instructors began expressing an interest and
desire into new orientations towards educating teacher candidates in assessment pedagogy. While uncertain as to how or why, they suspected that this orientation might involve the integration of emerging forms of information communication technology. At the same time, they wanted to implement an iterative process of inquiry and action around any initiative so that they could learn how their own ideas might be improved on the fly rather than after the initiative was completed. The 2010-2011 academic year presented such an occasion for the instructors to invite a voluntary group of teacher candidates to try out new learning environments and strategies. My interests in assessment pedagogy, and newer forms of learning using social networking technology, and developmental evaluation positioned me to be a collaborator in the project. In May of 2010, I was invited to join these two instructors, Laurene and Albert, and their lead teaching assistant, Tim, (collectively, referred hereafter to as ‘the clients’) in a developmental evaluation.

**Developing a Pilot Initiative in Classroom Assessment**

The developmental evaluation began with a series of negotiation amongst my clients, with me as their facilitator, into considering the kinds of learning, both among the candidates and among the instructors themselves, that might bring about better learning outcomes. We considered and learned about different kinds of technology that would facilitate the change to we wanted. Specifically, the evaluation explored how the instructors of the required classroom assessment module (CAM) re-conceptualized teaching and learning using collaborative web technology in the design and implementation of a pilot initiative in classroom assessment. Such ventures must themselves be grounded in the realities of institutional constraints (e.g., time, human
resources, funding, hence the notion of feasibility), and be balanced with what might be powerful learning environments technology might offer to the teacher candidates involved. Therefore, in considering what feasibility means in this context, local factors, as well as, broader learning/pedagogical considerations had to be considered. There must be a strong pedagogical argument and rationale for the change.

*Technology-enhanced learning.*

Thinking about integrating web technology into assessment education began through an investigation of the e-learning literature (Lam, Donnelly, Woods, & Shulha, 2010). An appraisal of this literature suggested that e-learning might be able to make a significant contribution to resolving the two problems articulated by the instructors.

As e-learning encompasses a variety of approaches and tools (e.g., Andrews & Haythornthwaite, 2007, pp. 1-52), an approach where educational technologies are used to enhance established pedagogies was most fitting for the current context. In blending what technology could offer for learning with existing theoretically-grounded pedagogies, the pilot initiative adopted this perspective and focused on understanding what technology could offer to the learning of assessment as potential learning environments that could one day be scaled up for use with the entire student population. In other words, the kinds of integration we sought aligned with the notion of technology-enhanced learning (TEL), an approach that has gained tremendous traction in the EU and across the UK. TEL is defined as:

> Any online facility or system that directly supports learning and teaching. This may include a formal VLE [virtual learning environment], an institutional intranet that has a learning and teaching component, a system that has been developed in...
house or a particular suite of specific individual tools. (Browne, Hewitt, Jenkins, Voce, Walker, & Yip, 2010, p. 2)

An explicit focus is placed on learning and how technology could add to that. TEL is understood to be:

… oriented towards the role of technology to enable new types of learning experiences and to enrich existing learning scenarios. To do this successfully, we have to understand not just teaching and learning, but also the context in which the implementation of TEL has to take place. (Laurillard, Oliver, Wasson, & Hoppe, 2009, pp. 289-290)

Adopting the perspective of TEL, the team explored the potentials of Web 2.0 for learning and its impact on modifying the nature of teaching. The term Web 2.0 is associated not only with specific tools, but also with a set of values and “ethos”—ideologies—that define it in comparison to Web 1.0 (O’Reilly, 2005). Because of that, Web 2.0 can be construed as a form of disruptive technology – one that changes and disrupts—established patterns of operation (cf. Christensen, Aaron, & Clark, 2001). Web 2.0’s effects can already be felt and are as far ranging as different industries and sectors begin to experiment with what potential innovations (e.g., Medicine 2.0, Enterprise 2.0, Library 2.0, and of course, Learning 2.0). Due to the relatively new emergence of this technology, the learning implications and consequences arising from the application of Web 2.0 to learning (i.e., Learning 2.0) remain to be explored. It is an area of intense research under the TEL agenda of research (e.g., Selwyn, 2008). An assumption underpinning the initiative is that the disruptive nature of Web 2.0
necessitates a new way of thinking about teaching and learning (i.e., its epistemology and axiology), and hence the need to re-conceptualize teaching in this new paradigm.

As Laurillard et al. (2009) cautioned, paying close attention to the context is essential to TEL implementation. This is yet another reason why developmental evaluation offers a very compatible framework to guide the design and implementation of the pilot initiatives. Developmental evaluation tracks and attempts to make sense of dynamics that emerge under conditions of non-linearity, uncertainty, adaptiveness, and co-evolutionary processes—conditions that tend to typify the introductory phase of innovations (Patton, 2010, p. 7). Therefore, developmental evaluation theory offers a unique theoretical lens through which evaluators and program staff can use to make sense of what is happening.

Conventional evaluation theory provides a coherent set of principles to help examine the context through situational analysis (e.g., Patton, 2008) and program theory (e.g., Rogers, 2008). Developmental evaluation strategies, such as engaging in rapid data collection and sense-making, and in turn, feeding that back into the design and implementation process, support the dynamic and changing nature of the context during implementation.

**The Developed Pilot Program**

The clients, with support from the developmental evaluator, launched a pilot program between January and April of 2011 with 22 teacher candidates. A blended approach was used; teacher candidates were invited to participate in non-formal learning with the clients over six one-hour sessions, while the rest of the program transpired over the Internet. The pilot program invited teacher candidates to articulate instances of their
attempts at practicing classroom assessment and share them with their peers via Twitter (for a more detailed description on the developed program logic, see Appendix J). By nature of the Twitter platform, teacher candidates were exposed to tweets from their peers and were encouraged to respond accordingly in ways they saw fit. Many of the face-to-face program hours were devoted to reducing the hurdles to engaging in participating in this innovative, participatory learning environment by helping teacher candidates learn to use Twitter, as well as, reframing how learning might transpire in such novel environment. As of the end of the pilot, teacher candidates contributed over 300+ tweets.

A series of meetings that took place in the months leading up to the launch helped focus the development to arrive at the program logic. The developmental evaluator supported the clients’ efforts in this development in ways that were informed by developmental evaluation theory, which form the corpus of the analysis. A highly collaborative environment was purposefully created to respond to the emergent context of development. An unconventional feature of the implementation of the program during the pilot was to adopt a developmental evaluation framework that allowed the clients to respond to how teacher candidates were engaging in the program in order to continually develop the program.

Given the focus of this research is on understanding the capacity of developmental evaluation for innovating, this case study focuses not on the product of the development (i.e. the use of Twitter for promoting teaching and learning of assessment pedagogy) but on the process of development. It is intimately concerned with deciphering and unpacking the process of innovating, and assessing the relative contribution a developmental evaluation mindset makes to innovating (Patton, 2011). No assessment of
the merits of the developed program will be made in this case study. The focus will remain on the process of program development.

**Conclusion**

This chapter presented a summary of the evaluation context. It traced the influences of contemporary classroom assessment—assessment of, for, as learning—on assessment education. In particular, the challenges surrounding teaching teacher candidates to practice the three approaches in an integrated and sound fashion, a notion referred to as *assessment pedagogy*, is proving challenging for a variety of reasons, as illuminated through the efforts and experiences of two instructors engaging in assessment education at Queen’s University, Faculty of Education.

A desire, born out of a professional and moral obligation to respond to this significant void in professional capacity among beginning teachers, seeded the efforts to engage in inquiry over conditions that would best enable teachers to learn about classroom assessment that could one day be implemented for the 700 candidates that these instructors are responsible for yearly. For pragmatic reasons, instruction that is coupled with novel uses of technology might best address the learning concerns previously raised and which offer a fit closest to the current context. With the expectation that this new form of instruction could be piloted with a group of voluntary candidates in the year 2010-2011, the instructors raised concerns about a desire to mitigate risks and be aware of unforeseen and unanticipated consequences that comes as a result from engaging in novel teaching and learning. A systematic framework that would allow instructors to make decisions based on available evidence was needed. Consequently,
developmental evaluation was adopted to guide the development of the program and its implementation.
Chapter 3: Theoretical Framework

Chapter Overview

This chapter provides an overview of the theoretical framework that informs the development of the Assessment Pilot Initiative and the current research. Developmental evaluation (DE) is an emergent approach in program evaluation that is intended to support program development. In this chapter, I trace the theoretical and practical foundations of developmental evaluation theory as they are currently discussed by leading theorists. Specifically, I explore how DE is argued to be different from traditional evaluation approaches, both in terms of the purposes and contexts each serve and consequently, how practices differ depending on the approach taken.

This chapter is presented in four parts. First, I situate developmental evaluation within the broader scope of program evaluation and discuss how research on evaluation is generally approached and its significance. Second, I describe the theoretical foundations of developmental evaluation in contrasts to traditional evaluation. Third, I explore how developmental evaluators generally practice DE and where tension arises. Finally, I conclude by outlining lines of inquiry that my thesis research will pursue.

Evaluation Theory and Research into Evaluation

Professionalization of program evaluation flourished in the United States starting in the 1960’s when the federal government started to mandate that all of its funded programs contain an evaluation component. Early evaluators were identified by their expertise in substantive program areas, using research skills that they had acquired from these fields or disciplines. As the importance of evaluation information increased, so did
the status of the profession and the number of practitioners interested in dedicating themselves to this form of inquiry. Shadish and Luellen (2005) recounted in their entry in the *Encyclopedia of Evaluations* that professionalization of the field came about as professional publications, professionals societies and professional code of conduct were developed.

Weiss argued that there are three distinctive attributes that defines the discipline of evaluation today:

- the task – to find out how interventions are functioning, what kinds of effects they are having, and how well they are satisfying the needs and desires of their several constituencies;
- the situation - as part of the world of practice, evaluation should be responsive to the desires and wants of program constituents; and,
- the mission - to help improve the social, economic, and cultural condition of society. (2004, p. 154)

Another important indicator of professionalization is the “accumulation and transmission of a unique transmittable knowledge base” concerning the methodology and practice of the profession (Shadish & Luellen, 2005, p. 186). Today, evaluation theory is of two general types according to Alkin and Christie:

- prescriptive models proposes sets of rules, prescriptions, and prohibitions and guiding frameworks that specify what good or proper evaluation is and how evaluation should be done; such models serve as exemplars; and,
- descriptive models offer a set of statements and generalizations which describe, predict, or explain evaluation activities. (2004, p. 5)
This explanation is useful to illustrate why case study research in evaluation is an important way of contributing empirical knowledge surrounding the utility of evaluation approaches. I will revisit Alkin and Christie’s (2004) distinction at a later point in this chapter.

Evaluation is an applied discipline that seeks to answer socially relevant questions through the application of social science research methods. Instrumental to that end is the use of a flexible framework in approaching the evaluation and the use of appropriate methodologies to produce findings of relevance and importance to those who have a vested interest in the program. An example of that notion is reflected in Patton’s utilization-focused approach to evaluation (2008); it is also reflected in the early distinction made between formative and summative evaluation by Scriven (1967), a distinction that is still very much alive today. These three approaches to evaluation are different in the purposes they serve and the contexts in which they are most appropriate. In sum, approaches to evaluation must fit the questions and the context in which they are proposed. Research into evaluation clarifies that linkage and illuminates how theory can best be translated into practice.

**The Theory of Developmental Evaluation**

Developmental evaluation is an emerging contemporary approach in program evaluation that is popularized through the writing of Patton (1994, 2010), Gamble (2008), and Dozois, Langlois, & Blanchet-Cohen (2010). As a theory, its contribution lies in lending a conceptual framework to support the development of social innovations in contexts of complexity, a context where traditional evaluation is argued to be relatively powerless and incompatible (Patton, 2006).
Proponents of developmental evaluation argue that DE supports the development of social innovations by buttressing program design decisions using evaluation in a data-informed approach (e.g., Patton, 2010; Gamble, 2008). In particular, DE is argued to provide evaluative information in a fashion that is more in tune to the dynamics of social innovation than traditional and more dominant approaches of program evaluation. Patton first espoused this view in 1994, when he reflected on the specific contexts and dynamics in which social innovation occurs are often different to those typical of traditional evaluations.

**Social Innovation and Complexity**

Patton (1994) noticed that social innovators identify “issue[s] or problem[s] and want to explore some potential solutions or interventions” (p. 313). In their attempt to create new knowledge or approaches, innovators attempt to transform the realities experienced by a population for the better (ibid). Social innovations can take “the form of new projects, programs, products, organization changes, policy reforms and system interventions” (Patton, 2010, p.1). Innovation can also be a process that functions as the primary driver of social change (cf. Gamble, 2008). Innovators “eschew clear, specific, and measurable goals up-front because clarity, specificity, and measurability are limiting” (Patton, 1994, p. 312). They work towards a vision, which is couched in values, beliefs, and assumptions that warrant reality testing and validation.

Moreover, solutions to problems are often uncertain and key stakeholders are often at odds about how to proceed in innovative contexts. These properties define complex problems, a distinction applied by Zimmerman to health care, and later popularized through the works of Patton (see: Westley, Zimmerman, & Patton, 2006;
Patton, 2010). Complex problems do not have clear boundaries. This presents significant challenges to problem-solvers and evaluators looking for certainty about what information needs to be included in the inquiry. Patton ties systems thinking into developmental evaluation to help evaluators approach complex problems in evaluation. He explains that:

A complex system is characterized by a large number of interacting and interdependent elements in which there is no central control; self-organizing and emergent behaviour based on sophisticated information processing generate learning, evolution, and development. (Mitchell, 2009, p. 13, as quoted in Patton, 2010, p. 1)

This raises the question of how evaluators are to conduct evaluations in a reality that where there is much fuzziness in the program logic and the intervention.

Gamble (2008) also introduced the notion of wickedness in his description of complex problems. Wicked problems, a construct first published by Rittel and Webber in 1973 when discussing issues in urban planning, and now contemporarily adapted through the works of Conklin (2006), refer to problems that exhibit a set of characteristics that relate to difficulty in defining and bounding.

Horn and Weber (2007) also drew attention to the difficulty of bounding social problems in their use of the construct social messes. Due to interrelatedness to other problems, they observe that “complexity—systems of systems—is among the factors that makes social messes so resistant to analysis and, more importantly, to resolution” (p. 7). This implies that the framing of a complex problem is therefore fluid. The definition of a problem may change even as innovators engage with the problem. The realities that an
innovation targets, and the repertoire of available methods to solving an issue, are likely to be shifting and changing as well. This raises the implication that interventions need to adapt to realities that are constantly evolving in order to address issues optimally and adequately; program officers, likewise, need to exhibit flexibility and agility in their approaches to problems (cf. Patton, 2006). Outcomes will emerge as innovators engage with the messiness of the program (Patton, 2010, p. 5).

It can be inferred from the above argument that the dynamics of innovation do not sit well with the logic typical of conventional approaches to evaluation. According to Gamble:

The very technique that enable evaluation excellence in more static situations — standardization of inputs, consistency of treatment, uniformity of outcomes and clarity of causal linkages—are unhelpful, even harmful, to situations where there is a lot of uncertainty and “moving goalposts”. (2008, p. 14)

In innovative contexts, it can be difficult, if not impossible, to bring about any orderliness when much is unknown and uncertain.

It is generally acknowledged that a fit between the program context and the evaluation approach is crucial to maximize the utilization of evaluation findings (e.g., see Patton’s, 2008). Conventional approaches to evaluation require that the program be relatively stable. Traditional evaluation “aims to control and predict, to bring order to chaos” (Patton, 2010, p. 5). For example, traditional evaluation activities, such as logic modeling or pre-formative assessment, attempt to map out the logical linkages of a program, thereby presenting a static and stable representation of a program to be evaluated. Such order is often lacking in contexts of innovation. Proponents of
developmental evaluation embrace the uncertainty experienced by innovators and look to feed in processes and information in ways that support the visions and decision-making of these clients. Patton explains why innovators appreciate this orientation to evaluation, “[innovators] never expect to conduct a summative evaluation because they don’t expect the program—or world—to hold still long enough for summative review” (1994, p. 313).

**Developmental Evaluation and Complexity Science**

If evaluation is fundamentally about reality-testing—i.e., examining the extent to which aspects of the planned program contribute to the program outcomes—how can evaluation occur in contexts where program planning is ongoing and outcomes are emergent? The most recent response is for evaluators to treat such programs as complex adaptive systems. To describe these programs, evaluators must learn to speak about program dynamics such as non-linearity, emergence, adaptiveness, uncertainty, and co-evolutionary processes. These complexity concepts embrace non-linearity, as Patton argues, and provide an alternate framework in framing program contexts and program interventions (2010). When programs are framed through the lens of complexity concepts, a developmental evaluation approach can take on the following characteristics during evaluation: sensitivity to emergence; attention to unintended effects; use of continuous real-time sensing of perturbations and changes; and feeding back of meaningful information for the purpose of generating adaptive learning (Patton, 2010).

**The Practice of Developmental Evaluation**

Previous evaluation approaches, such as that of participatory, collaborative, and utilization-focused evaluation, have focused attention on the roles the evaluator plays in enabling quality evaluation (Shulha & Cousins, 1997). The types of roles and
responsibilities have grown significantly since the days of summative/formative evaluation when it was most common to place the evaluator in the position of objective external judges of program merit or worth (Yarbrough, Shulha, Hopson, & Carruthers, 2011). Within developmental evaluation, the evaluator’s role is possibly the most multidimensional.

Patton (1994) set the stage for understanding how a developmental evaluator might operate when he began observing that evaluators were bringing to the team “evaluation logic, knowledge about effective programming based on evaluation wisdom, and some methods expertise to help set up monitoring and feedback systems” (p. 6). By 2010, Patton had identified two primary responsibilities for a developmental evaluator. First, the developmental evaluator tracks and documents the nature and results of program development. Situational sensitivity and responsiveness to changing conditions are two skills critical to the sense making of emergent program contexts. Being able to describe both the intended and unintended consequences of a program, as well as the changes in direction and shape of a program model, will allow the individual to carry out a systematic inquiry into the development of a program.

Second, the developmental evaluator facilitates systematic data-based reflections. The evaluator brings to bear his evaluation training in the design, collection, and analysis of program data to the program development. The evaluator may “see patterns that prompt a different way of organizing the data, generating new perspectives and insights for the team” (Gamble, 2008, p. 32). Equally important to the technical aspect of generating high quality data is one’s ability to facilitate reflective sessions. Data are only useful if they are given value. By feeding program data back to different stakeholders,
data can not only inform emergent understanding, but also facilitate adaptive learning. These notions have been adopted and expanded on by those who see themselves evaluating in a program development context.

Gamble (2008) described five functions of a developmental evaluator: 1) accompanying the group to facilitate the progress of development; 2) collecting data that could be used to validate program decisions, approaches or assumptions, or documentary data concerning the innovation process; 3) framing and reporting; 4) developing strategy and aligning programmatic efforts to the strategy; and 5) developing indicators that signal progress or departures/divergence from trodden path in the course of program development.

Similarly, Dozois, Langlois, and Blanchet-Cohen (2010) spoke of four key functions of a developmental evaluator:

1) Orienting the group in the direction of the program, clarifying guiding principles, identifying systems linkages and leverage points, and acting as a strategic coach by agreeing on indicators and progress markers;

2) ‘Watching’ for key developmental moments in relationship formations between and among groups, how group dynamics unfold, and how groups respond to threats and opportunities;

3) Being a partner in the meaning-making process. This involves analyzing and interpreting emergent data

4) Intervening when the group “get stuck, off-track-disheartened, or confused” (p. 49).
Dozois et al. (2010), in comparison to Gamble (2008), focused more on the interpersonal or “soft” skills needed to facilitate quality developmental evaluation. Given the multidimensionality of the development evaluator, successful practice appears to depend both on a capacity to draw on a diversity of skills and versatility in adapting one’s approach most appropriate to an evaluation’s context. Yet to be articulated are the kinds of skills and capacity needed to effectively perform the tasks discussed above.

**Conclusion and Establishing Lines of Inquiry**

Developmental evaluation is an emerging approach where its theoretical and practical foundations are heavily influenced by the seminal texts of Gamble (2008), Patton (1994, 2010), and Dozois, Langlois, and Blanchet-Cohen (2010). These words present a synthesis of the experiences and insights developmental evaluators warrant as important and unique to developmental evaluation. A search within the professional literature reveals no systematic inquiry describing its practice.

Revisiting Alkin and Christine’s distinction (2004) between prescriptive and descriptive models of evaluation, I categorize the theory building work in developmental evaluation as *descriptive* (i.e., offer a set of statements and generalizations which describes, predicts, or explains evaluation activities). Practical knowledge, as summarized by the efforts of Gamble (2008) and Dozois et al. (2010) from current practitioners, has contributed greatly to the development of the theory. In order to bolster the power of the theory, it will be critical to apply the theory across different contexts. A carefully designed case study, like this one, is an important step towards that end.

There remain some points of contention about the legitimacy of developmental evaluation. As Patton pointed out in his entry to the *Encyclopedia of Evaluation* (2005)
on developmental evaluation, a classic debate concerns whether it is appropriate for an 
evaluator to extend his role from rendering judgments to offering advice (p. 186). 
Similarly, is organizational development a legitimate possible evaluation use? To what 
extent does prior knowledge/domain-specific expertise or patterns of effectiveness 
contribute to an evaluation? While answers to these questions will hinge, in part, on the 
values carried by individual evaluators, research on the processes and outcomes of 
developmental evaluation is essential if there is to be informed decision-making on the 
potential power of this approach to evaluative inquiry.

This thesis represents a first step in understanding the capacity of developmental 
evaluation to support educational innovation. This thesis proposes three lines of inquiry: 
1) the contributions developmental evaluation makes in promoting program development 
and innovation; 2) the contributions developmental evaluation makes in addressing 
stakeholder needs in contexts of innovation; and 3) an examination of the roles and 
responsibilities of a developmental evaluator.
Chapter 4: Method

Chapter Overview

This chapter describes the reasoning behind the choices I made concerning the method and methodology I employed to answer my research questions. The first half of this chapter discusses my rationale for using qualitative case study as a method for studying developmental evaluation. I recount how I generated the needed evidence to answer my questions. I also detailed my plan for data collection, with an elaboration on strategies for generating high-quality data that was grounded in appropriate methodological traditions. The second half of this chapter discusses issues related to data analysis.

The Research Context, Revisited

The research capitalized on the knowledge and insights generated, captured as program development records and program evaluation data, among the developers and program participants of the Assessment Pilot Initiative during the development and implementation phases of the development. The Assessment Pilot Initiative convened in early summer of 2010, branched from an on-going agenda of research and course development in assessment education and it concluded at the end of April following the end of the implementation phase.

Prior to the start of my role as an evaluator for the Assessment Pilot Initiative, I was part of a graduate student team that conducted a systematic review of the literature on the use of e-learning for promoting assessment education. Following the completion of that project, and likely, in part, because of this background, the instructors approached me to work with them in a program development project, the Assessment Pilot Initiative.
Even as a beginning evaluator, I recognized the value of using evaluative data in program development, primarily as a way of providing empirical support in informing program decisions. However, for reasons already discussed—uncertainty in program logic at the outset of program development, emergent outcomes, and the high-stakes nature of the context—traditional evaluation did not appear to be an appropriate fit for the context in which I would be working. Soon after learning about developmental evaluation through the works of Patton (2010, 2008) in early September of 2010, I put developmental evaluation into practice with my clients in the Assessment Pilot Initiative. This approach provided a framework for understanding what was happening in terms of program development. The basic premises of developmental evaluation theory also would guide my approach to the evaluation, shaping the roles that I would take on and the behaviours I would exhibit.

This study is a systematic inquiry into the use of developmental evaluation for developing an innovation. This is done through framing the Assessment Pilot Initiative as a case study. Archived meeting transcripts and program artifacts constituted the bulk of the case data. Interviews with the program developers constituted the second pillar of the data sources and allowed for corroboration of findings. By framing this inquiry as a qualitative case study, I seek to answer the research questions as set out in the introductory chapter in the traditions and conventions of qualitative case studies. The remaining part of this chapter is devoted to discussing the specifics of conducting qualitative case studies.
**Case Study Research as a Method**

According to Greene (2007) and Teddlie and Tashakkori (2009), it is good practice for research questions to drive method choices and methodological decisions. To answer the research questions, my study used a qualitative case study as a research method for guiding the data collection and data analysis. In this section, I will first discuss my rationale for choosing case studies as a research strategy (i.e., method) and its appropriateness to the research, before returning to justifying a qualitative approach.

The present research used the case study method to conduct an in-depth study into the phenomenon and context—the development and implementation of a novel approach to teaching and learning classroom assessment in a process guided by developmental evaluation. In following the advice of a noted case study researcher, conducting “case study is not a methodological choice but a choice of what is to be studied” (Stake, 2005, p. 435). What defines a case research is that the primary research goal is to understand the case itself. A case is conceptually defined as a “specific, unique, bounded system” (Stake, 2005, p. 436). This suggests what matters are a system’s activity and its functioning; what matters more are the interactions between sub-systems and with the context, even if no evidence of influence is found. To that end, a case is therefore social, cultural, situational, and contextual (Stake, 2005, p. 452). It is this notion of ‘case’ that I saw a fit to how the present study was already being conceptualized using a developmental evaluation framework; the ‘case’ of this research represented the bounded systems already being examined and evaluated in the developmental evaluation. The research case was, therefore, bounded both by time and by the activities that transpired within the Assessment Pilot Initiative.
As choosing the boundary of a system is inherently arbitrary, what does not fit within the case is an important consideration as well (Campbell, 1979). This study was not primarily concerned, for instance, with the operations of the larger instructional context of the teacher education program at the university. This study was not concerned about the qualities of the developed program that resulted from this development process. I delimited the study to the components, sub-systems, relationships, and dynamics within those that transpired within the Assessment Pilot Initiative. As the purpose of my research was primarily one of learning, Stake suggests that such uses are appropriate and are often referred to as instrumental case study when the researcher hope to get insight into the question by studying it in detail (1995, p. 3).

For this research, I adopted Yin’s definition (2009) of case research to guide my inquiry process. In his two-part definition, he first discusses the issue of scope in case studies:

A case study is an empirical inquiry that

- Investigates a contemporary phenomenon in depth and within its real-life context, especially when
- The boundaries between phenomenon and context are not clearly evident. (p. 18)

In the second part of the definition, Yin addresses the technical aspect of case studies:

The case study inquiry:

- Copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result
• Relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result
• Benefits from the prior development of theoretical propositions to guide data collection and analysis. (p. 18)

Thus, Yin argues that case study research is “an all-encompassing method—covering the logic of design, data collection techniques, and specific approaches to data analysis” (2009, p.18; see also Patton, 2002, p. 447).

**Conducting Qualitative Research**

Adopting an appropriate methodological stance is equally central to conducting high-quality research. Recall that the case study approach does not stipulate any methodological choices or preferences. While it is possible to conduct case studies using a quantitative approach, a qualitative approach is often used (Stake, 1995). To arrive at a methodological decision, I referred to the logic expressed in the research purpose statement for indications. It would appear that a qualitative methodology was most appropriate in this case. Stake (1995) has suggested that qualitative research is particularly appropriate for “understanding the complex interrelationships among all that exists” in a phenomenon (p. 37). This approach of inquiry allows for inquiry into knowledge built through experiential understanding (Stake, 1995, p. 3), as was the case for this study.

Given that the purpose of the research was to describe, understand, and to infer meanings, the theoretical drive (Morse, 1991) of the research was said to be inductive (Newman, Ridenour, Newman, & DeMarco, 2003). In this case, the “method most commonly used will be qualitative and the outcome will be thick description as in
phenomenological or narrative inquiry” (p. 193). Additionally, the methodological aim of qualitative research is said to produce rich, detailed, and concrete descriptions of people and places, hence the notion of thick description (Geertz, 1973). Consistent with case research method, qualitative methodologies allow the reader to understand the phenomenon being studied and to draw interpretations about its meanings and significance. In doing so, the function of qualitative research is to understand the complexity of reality. Inductive logic is used in case studies to infer meaning and significance to aspects of the phenomenon from data collected through multiple sources. Patton outlines the process of constructing case studies as follows:

1) Assemble the raw case data
2) Construct a case record

**Study Participants**

This study involved three participants, who were the clients of the developmental evaluation upon which this study is based. Among the three, two of who were faculty members in a Faculty of Education who had a history of teaching classroom assessment in the aforementioned large-lecture format module as instructors-of-record. They are referred to as Laurene and Albert (pseudonyms) in this thesis. The remaining participant was the lead teaching assistant for the assessment module, who was himself a doctoral student in the same Faculty of Education. He is referred to as Tim (pseudonym) in this thesis. His primary duties included leading a group of teaching assistants to provide instructional assistance to the two faculty members in the delivery of the course. Taken together, the study participants came to the study with a strong personal, scholarly, and
academic interest in the assessment learning of teacher candidates, and by extension, a strong interest in the success of the program and the developmental evaluation. Through interviews, they were to recall and discuss the particular episodes that they deemed as significant to them.

Furthermore, it became apparent during the stage of data analysis that the perspectives of the developmental evaluator needed to be present. In particular, it was important that the behind-the-scene thinking and interpretations of the ongoing development be studied and made transparent for analysis. As a result, I, as the developmental evaluator, became one of the study participants through engaging in document-elicitation reflection and memo-writing to interrogate my practice as the developmental evaluator of the developmental evaluation.

**Aligning Questions to Methods**

The following table (Table 1) presents a summary of the data sources that informed the answering of each of the questions and the analysis that were performed to answer them.

**Data Sources and Data Collection Strategies**

This section elaborates on the process of constructing a case study as introduced above by Patton (2008). I recount the kinds of evidence and data that informed my research and how they were collected. Following that, I discuss the process of data analysis, before concluding with a discussion about the strategies I used to enhance the trustworthiness of this study.
Table 1

*Summary of Data Sources and Analysis Used to Answer each of the Research Questions*

<table>
<thead>
<tr>
<th>Research question</th>
<th>Data sources and collection strategies</th>
<th>Analysis</th>
</tr>
</thead>
</table>
| 1. To what extent does Assessment Pilot Initiative qualify as a developmental evaluation? | Program archival data  
Meeting minutes  
Interview with instructors | Thematic analysis  
Time-series analysis  
Thematic analysis  
Time-series analysis |
| 2. What contribution does developmental evaluation make to enable and promote program development? | Interview with instructors  
Program archival data  
Evaluator’s own reflection  
Observational notes  
Participant-contributed data  
Program outputs  
Participant tweets | Thematic analysis  
Time-series analysis  
Thematic analysis  
Time-series analysis |
| 3. To what extent does developmental evaluation address the needs of developers in ways that inform program development? | Interview with instructors  
Program archival data  
Evaluator’s own reflection  
Observational notes  
Participant-contributed data  
Program outputs  
Participant tweets | Thematic analysis  
Time-series analysis  
Thematic analysis  
Time-series analysis |
| 4. What insights, if any, can be drawn from this development about the roles and the responsibilities of the developmental evaluator? | Interview with instructors  
Program archival data  
Participant-contributed data  
Program outputs  
Participant tweets | Thematic analysis  
Time-series analysis  
Thematic analysis  
Time-series analysis |

I made use of data that had already been collected as part of the developmental evaluation. Data that most appropriately informed my research purpose and that answered my research questions were included. My general strategy was to collect evidence from multiple perspectives, in multiple contexts, and via multiple methods. This strategy was meant to increase the validity and reliability to the claims I make from this study. The following kinds of data were collected and used to inform my research: (a) interview data, (b) observational data and participant-contributed data (c) program development records, and (d) program artifacts.
Collectively, these four data sources formed the two pillars of my evidentiary base. On one hand, the archived program data (#2, #3, #4) provided an objective description of the developmental process and its products. On the other hand, interview data with the program developers provided insights into the complexity of their thought processes, reactions, and decision-making processes.

**Interviews**

The purpose of interviewing is to elicit the perspectives and interpretations of others. As a method of inquiry, it allows the researcher to engage in deep meaningful dialogue that results in an “active” exchange (Holstein & Gubrium, 1995) with the interviewee. This method allowed me to probe and explore the clients’ own constructed realities (Fontana & Frey, 2005, p. 696; Patton, 2002, p. 341).

I conducted three semi-structured interviews with each of the three clients (Laurene, Albert, and Tim) of the developmental evaluation to capture retrospectively their respective experiences and perspectives on participating in a developmental evaluation. These occurred in July of 2011. The semi-structured nature of interviewing provided a degree of comparability across participants, ensuring that relevant ideas were probed, while at the same time, allowing for enough flexibility for me, as the researcher, to pursue lines of inquiry as they arose in the exchange, thereby generating richness. An interview guide (Appendix D) was prepared using themes that emerged following some preliminary analysis of data. The interview questions were designed to aligned with the overall research purpose and with specific enabling research questions, with each interview lasting approximately 30-45 minutes.

**Observational & Other Participant-contributed Program Data**
Patton argues that observational data attunes the researcher to the nuances of how program activities are conducted, supports the inductive orientation to qualitative research, identifies and makes curious of the mundane, and helps to learn about things that participants may be unwilling to talk about in an interview (Patton, 2005, pp. 262-264). “In-class” direct observations were made generally of the activities, people, and their reactions as the instructors implemented the Assessment Pilot Initiative. Planned program activities, unintended emergent practices, behaviours, and interactions were documented. Analysis of this kind of data yielded data on how the program was actualized and communicated to the participants.

The second kind of naturalistic observation involved me as a participant-observer into the activities that transpired over the virtual space. Because Internet access was ubiquitous, teacher candidates could participate online at anytime. A more responsive regiment was adopted in the developmental evaluation to observing and documenting activity over the Internet. As an evaluator, I documented Twitter activity, and responded to participate-initiated queries when such need arose (e.g. responding to inquiries, re-tweeting resources directed at the program account). This kind of virtual observational data captured how the program operated and the conversations that unfolded.

**Transcripts of Meetings**

Transcripts of the meetings I facilitated with the developers were rich sources of data describing the development process. Nine development meetings were held between the months of August 2010 to April 2011. These were audio-recorded, transcribed, and archived as part of the developmental evaluation. Meetings before that period were recorded on paper and subsequently digitized. These transcripts were also critical source
of data as they captured the unfolding conversations between the developers and the developmental evaluator, about the development process.

**Program Artifacts**

The last source of data that included in the dataset was literature and documents generated from within the Assessment Pilot Initiative and the developmental evaluation. They included such documents as: 1) e-mail correspondences between the developers and the developmental evaluator; 2) program literature intended for program participants; and 3) any interim reporting documents produced by the evaluator for the clients (e.g., usage report; observations, etc.). These data provided a behind the scene look at the development process and how they came into being (Patton, 2008, p. 294).

**Ethical Considerations**

This study was conducted under the purview of ethical principles and adhered to established ethical guidelines as implemented by Queen’s University General Research Ethics Board (GREB Ref #: GEDUC-544-11; see Appendix E for ethics clearance letter) and defined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (CIHR, NSERC, & SSHRC, 2010). Participation in this research study was voluntary. Participants were informed of the purpose of the research, their obligations, as well as other ethical considerations through the provision of a letter of information (Appendix B). Potential risks were weighed against potential benefits in determining the ethical merits of this study. The signing of a consent form evidenced informed consent (Appendix C). Participants whose data were included in this study secondary to their original participation in the Assessment Pilot Initiative were informed of this secondary use and were provided with the opportunity to consent in a process that was designed in
accordance with general ethical principles and approved by GREB. Pseudonyms were used to mask the identity of the participants to this study.

Data Analysis

According to Patton and Appelbaum, the “ultimate goal of [a] case study is to uncover patterns, determine meanings, construct conclusions and build theory” (2003, p. 67). A systematic analysis of the collected evidence is necessary to achieve that. The analysis of my data followed a four-stage process, beginning with the construction of the evidentiary base in the first stage of analysis. Second, I engaged in coding and thematic analysis with the program development data. Third, I holistically analyzed the entire data set by conducting a time-series analysis. Following such a plan allowed me to arrive at descriptions concerning both the processes and products of the development process. In the last stage, I conducted a thematic analysis of the interview data conducted with the developers in order to identify the more salient and significant aspects of the development process. I discuss each stage below.

Stage One: Building the Evidentiary Base.

The objective of stage one of my analyses was to construct the evidentiary base of the case. This operation began with an identification of all the relevant documents. I found this digital archeological dig was exhaustive in considering all potential sources of data. Included into the evidentiary base included the data types as described above: (a) audio recordings of meetings, (b) records and transcripts of meetings, (c) e-mail correspondences, (d) outputs of the developed program, such as hand-outs, and (e) development documents, such as design specification, the evaluator’s journal, and program documents that were used internally.
To facilitate interoperability and readability of these files during the latter stages of analysis, all files were converted, or otherwise digitized if they were previously in hard copy, into either Adobe Portable Document Format (*.PDF) format or Microsoft Word (*.DOC) format for text files, and MP3 format for voice files. A total of 147 files were identified to be associated with the development of the Assessment Pilot Initiative. Only substantive documents were included; documents and e-mails that were of a cursory nature that did not add real substance to the analysis were not included.

All files, now in digital formats, were then deposited into a common directory on the researchers’ computers. A date code was appended to all file names to allow for the chronological display of the files. To determine the date code, the date of creation was used if one was recorded in the metadata section of a digital file. In the case of digitized files, the date was approximated by cross-referencing files that were known to have been created around the same time if one was not already recorded originally. Each file was manually prefixed in the form of YYYY_MM_DD (e.g. 2010_11_07), while retaining the original filename. This operation yielded hybridized filenames such as “2010_07_08 - meeting notes - CCF00252011_00010.pdf”. Individual directories were created for each of the months Assessment Pilot Initiative was in operation. Files were relocated from the main directory to folders of the corresponding month. This step allowed files to be displayed chronologically in a way that reduced clutter and enhanced focus.

The files were then color-coded using the label feature of Mac OS. Table 2 describes the color-coding scheme used in this research. This stage resulted in a visual representation of the development process. An initial timeline was then be established to
help establish some preliminary ideas about when different events occurred within the development.

Table 2

*Summary of the Different Data Types and the Colour Codes Assigned to Each*

<table>
<thead>
<tr>
<th>Color</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purple</td>
<td>E-mails</td>
</tr>
<tr>
<td>Blue</td>
<td>Assessment Pilot Initiative documents (internal)</td>
</tr>
<tr>
<td>Green</td>
<td>Assessment Pilot Initiative publications</td>
</tr>
<tr>
<td>Red</td>
<td>Meeting recordings (voice files)</td>
</tr>
<tr>
<td>Orange</td>
<td>Meeting Transcripts and Notes (text files)</td>
</tr>
<tr>
<td>Grey</td>
<td>Design files</td>
</tr>
<tr>
<td>Yellow</td>
<td>Evaluator’s Journals</td>
</tr>
<tr>
<td>(None)</td>
<td>Participant-level data (e.g., survey results)</td>
</tr>
</tbody>
</table>

*Developmental episode.*

These computer files represented periods of the development that I refer to as developmental episodes. Developmental episodes were unique event in time where Initiative-related ideas were communicated, exchanged, or articulated for the purpose of development. Multiple files may constitute a single developmental episode. To determine whether an event qualified as a developmental episode or not, I used the following decision rule:

- A developmental episode must be grounded in data
- It must denote a time period that was sufficiently narrow
- There must be an apparent connection to the development.

This analysis identified a total of 60 developmental episodes from the 147 files analyzed. Table 3 describes key developmental episodes in Assessment Pilot Initiative.
Table 3

*Descriptions and Dates of Key Developmental Episodes*

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010/05/11</td>
<td>• Start of development</td>
</tr>
<tr>
<td>2010/06/03</td>
<td>• Articulation of program design</td>
</tr>
<tr>
<td>2010/07/08</td>
<td>• Group meeting: Assessment Pilot Initiative group meeting</td>
</tr>
<tr>
<td>2010/07/16</td>
<td>• Articulation of program design</td>
</tr>
<tr>
<td>2010/08/05</td>
<td>• Considering the use of e-portfolios</td>
</tr>
<tr>
<td>2010/08/19</td>
<td>• Group meeting: clickers and surfacing of desires and wants</td>
</tr>
<tr>
<td>2010/09/01</td>
<td>• Group meeting: exploring the problematics</td>
</tr>
<tr>
<td>2010/09/27</td>
<td>• <strong>Group meeting: learning Twitter</strong></td>
</tr>
<tr>
<td>2010/12/08</td>
<td>• <strong>Group meeting: appreciative inquiry session</strong></td>
</tr>
<tr>
<td>2010/12/17</td>
<td>• Group meeting: training on the use of Twitter</td>
</tr>
<tr>
<td>2011/01/01</td>
<td>• Articulation of program design: Executive Summary</td>
</tr>
<tr>
<td>2011/01/05</td>
<td>• Articulation of program design: Theory Spec Sheet</td>
</tr>
<tr>
<td>2011/01/06</td>
<td>• Articulation of program design: Conceptual Briefing</td>
</tr>
<tr>
<td>2011/01/06</td>
<td>• Group meeting: Vetting of program design and planning for the first class</td>
</tr>
<tr>
<td>2011/01/07</td>
<td>• Class #1</td>
</tr>
<tr>
<td>2011/01/14</td>
<td>• <strong>Group meeting: planning for the second class</strong></td>
</tr>
<tr>
<td></td>
<td>• Class #2</td>
</tr>
<tr>
<td>2011/01/21</td>
<td>• Group meeting: planning for the third class</td>
</tr>
<tr>
<td></td>
<td>• Class #3</td>
</tr>
<tr>
<td>2011/02/04</td>
<td>• Group meeting: planning for the fifth class</td>
</tr>
<tr>
<td></td>
<td>• Class #5</td>
</tr>
<tr>
<td>2011/02/10</td>
<td>• Group meeting: planning for the sixth class</td>
</tr>
<tr>
<td>2011/03/28</td>
<td>• Group meeting: evaluating the initiative</td>
</tr>
<tr>
<td>2011/04/16</td>
<td>• Class #7 (final)</td>
</tr>
</tbody>
</table>

*Note: Emboldened developmental episodes are those analyzed and described in chapter five.*

**Stage Two: Thematic Analysis.**

In stage two of analysis, my goal was to reduce the complexity of the case into units of analysis in a process known as coding. In order to develop a deep familiarization
of the data and to develop an appreciation for the complexity of the case, I reviewed all of
the audio recordings of the meetings as well as reviewing the related transcripts prior to
engaging in further analysis. As I revisited the development across time, I became
increasingly aware of the concerns and the positions each of the developers held. I
noticed that the audio recordings captured nuances of speech—such as hesitations when
developers grappled with decisions or points of confusion when new ideas were
introduced—that were lost during the transcription process. This was useful in trying to
place myself (as the case researcher) into the roles of the developers.

To facilitate data analysis and management, qualitative data analysis software,
ATLAS.ti 6 was employed to assist with coding and memoing. All files were imported as
primary documents into the software. Because of differences in the way Mac OS handles
filename naming from Windows, I manually verified that all files were imported
correctly.

For each of the documents, I identified and isolated the data into units of meaning
in order to conceptualize and label them. When a new code was being considered, it was
tested against existing ones using deductive logic, in a process described as constant
comparison (Glaser & Strauss, 1967; Strauss & Corbin, 1998, pp. 80-81). Only when a
unit of data was dissimilar to existing codes would a new one be generated.
Developmental moment.

It was shortly after coding the first few documents that I gained an analytical insight. First, it occurred to me that each code represented a developmental moment. It was a moment in time that described an interaction between the evaluator and members of the development team. In that moment, it was possible to identify a unique and specific idea or concept of the development or of the development process that influenced either the developmental process or the development itself. In essence, a developmental moment informed the complexity of the ideas underpinning processes or products. Once identified, developmental moments could be labeled.

To determine if a developmental moment existed or not, I used the following criteria:

- A moment must be grounded in the data;
- There must be an apparent connection to the immediate development; and,
- There must be a plausible *a priori* connection to other preceding or succeeding developmental moments; inclusion of isolated developmental moments were considered on an individual basis

Second, developmental moments could further be classified into three subsets: 1) those representative of the questions or dilemmas that were raised surrounding the development, in which the interaction unpacked some of the underlying complexity; 2) those representative of the ideas or concepts that were raised as a response to those of the first type that informed the program logic; and 3) those representative of the ideas or concepts concerning the development process intended to promote development. To denote this subset classification to the developmental moments, codes were recorded
within ATLAS.ti with a prefix of either CR_, PT_, and DE_ corresponding to each of the three classifications described as above respectively. Analysis of the program data yielded a total of 225 codes, of which 75 were of type 1 developmental moments, 88 of type 2 developmental moments, and 62 of type 3 developmental moments.

This approach to content analysis helped me “remain close” to the substantive content under examination (Kohlbacher, 2005). During this coding process, I engaged in memo-writing as an analytical exercise to think about the underlying concepts; creating code memos also helped me to track the decision logic I used at the time of analysis.

As I revisited the data, I became increasingly aware of my own reactions and emotions as the developmental evaluator to the different developmental episodes. I realized that embedded in the analytic process was an opportunity for me to recall and reflect on the potential meanings and significance of my contributions to the development. This document-elicitation exercise allowed me to capture this reflective data. I recorded my thinking by writing document memos. Appendix H features a sample of my document memos.

**Stage Three: Time-Series Analysis.**

In stage three of my analysis, the goal was to understand how the development came to be developed. Time-Series analysis is an analytic technique used in case study analysis for analyzing events over time. The “essential logic underlying a time-series design is the match between the observed (empirical) trend and either of the following: (a) a theoretically significant trend specified before the onset of the investigation or (b) some rival trend, also specified earlier” (Yin, 2009, p. 145). The technique of conducting time-series analysis originated from experimental and clinical psychology, with the
intention of conducting hypothesis testing against time, which is rooted in the
behaviouralist paradigm. The present research employed no a priori hypothesis, but was
instead sensitive to the idea of emergence, a sensitizing concept from complexity theory.
For this study, I used a specific kind of time-series analysis known as chronicle
construction (Yin, 2009, p. 148), to perform a time-series analysis. The intention was not
to construct causal inferences, or to test against hypothesis, but to create “valuable
descriptive renditions of events” (Yin, 2009, p. 149).

To construct the chronology, I plotted developmental episodes against time.
Developmental moments corresponding to each of the developmental episode were
placed together under the same column. Figure 2 illustrates the finished timeline.
Multiple sheets of gridded flip-chart papers composed the canvas upon which Post-It™
sheets, representing individual developmental episodes were pasted. The x-axis
represented time, and developmental episodes were plotted to scale.

The construction of a timeline as form of data visualization described in detail the
unfolding program development process. It presented elements of the developmental
process for visual inspection. The distribution of developmental moments and the
resulting visual patterns and relationships allowed me to uncover emergent patterns of
practice that might have contributed to program development. It allowed for multi-level
simultaneous analysis on individual elements of the development, as well as holistic
analysis on the development as a whole.
Stage Four: Coding Interview Data.

Coding of the interviews proceeded in a similar fashion to the program documents. The audio recordings of each of the three interviews were first reviewed. They were then transcribed and subsequently imported into ATLAS.ti. Initial coding (Charmaz, 2006), otherwise known as open coding, was performed to break down the data into discrete units of meaning. Using the research questions as sensitizing concepts (Blumer, 1954; Patton, 2002, pp. 278-279), I iteratively analyzed the corresponding evidences to each of my enabling research question. This generated a series of codes, which were then grouped into categories based on relevance to other codes within the category. Categories were collapsed into themes in a similar fashion.

Enhancing the Trustworthiness of Case Study Research

Case study research is increasingly used across many different disciplines and contexts, and is increasingly being recognized as a rigorous research strategy (cf. Stake, 2003). The level of sophistication into designing and conducting rigorous case study
research has been extensively explored in methods textbooks (e.g., Stake, 2003, 2005; Yin, 2009). With a rise in use, strategies have been developed to enhance the trustworthiness of case study research during the data collection and analysis phase (e.g. Lincoln & Guba, 1986, pp. 76-77). Several key strategies have emerged from these discussions of quality: 1) the notion of triangulation; 2) creating a chain of evidence; 3) member checking; and 4) being reflective and reflexive (cf. Patton, 2002, pp. 494-495; Yin, 2009, pp. 114-124, 160-162).

The current case study design allowed for the convergence of evidence as a strategy to enhancing trustworthiness. Because of the iterative nature of the study context, I collected evidence on multiple instances of the phenomenon (i.e., program development) across time, across contexts, and across space. My data collection strategy emphasized the importance of examining all data to determine their relevance and relative contribution (cf. Gillham, 2000). Several analytic strategies were employed to enhance the trustworthiness of the research.

The first strategy that I employed is to triangulate evidences across different data types (e.g., observational data, self-reported data, and archival records), as reflected in Table 3. I also triangulated across the three perspectives comprising of this case: 1) program records, 2) the perspectives of the developers, and 3) the perspective of the developmental evaluator. Member-checking was another well-documented practice that I used to enhance the trustworthiness of my claims (Patton, 2002, p. 560). I also reviewed the completeness and accuracy of the transcribed data. Additionally, I presented some of my initial findings with study participants to assess the degree of agreement in perspectives, such that the reconstructed realities were accurate to that of the participants’
own realities. Triangulating the evidence contributed to the authentic re-construction of the case.

Another strategy that I employed was to be reflective and reflexive as a researcher. As Patton suggests, the researcher is the research instrument in qualitative research (Patton, 2002, p. 14). Thus, the quality of claims arising from analysis as related to criteria like credibility, dependability, and confirmability was dependent upon the interpretive lens that I took as the researcher. Being reflexive, then, was to acknowledge the role of self in the research, and to address the issues of what I knew, and how I knew (Patton, 2002, p. 64). Reflexivity was acknowledged in this study, first, by framing the study through the theoretical lens of evaluation theory and practice; assumptions and values were routinely probed, addressed, and recorded as part of professional practice. In addressing self-reflexivity, I engaged in reflective practices as a mean of learning about my own assumptions, values, and biases. This helped to maintain a separation of the two roles that I assumed during the process of conducting this inquiry.

Another strategy I employed to enhance the quality of the claims was the creation of a chain of evidence—i.e., the notion of audibility. Two separate collections were created following the advice of Yin (2009): 1) the data or evidentiary base; and 2) the report of the investigator, whether in article, report, or book form (pp. 118-119). He argues for constructing a separate database apart from the case study report. In that spirit, I collected and archived all raw data into a digital repository in ATLAS.ti, and maintained a decision trail during the collection and analysis phase of the study by writing memos within the software.
Chapter 5: Results of the Study

Chapter Overview

The purpose of this chapter is to present a detail analysis into the development of the Assessment Pilot Initiative program. It complements the writing in chapter two, where snapshots of the initial conditions and the ending conditions were described, in a presentation style I referred to as book-ending. Having established those reference frames, this chapter, now, delves into the developmental process by highlighting in select episodes the sort of activities, decisions, and struggles undertaken by members of Assessment Pilot Initiative as they engaged in innovating the assessment learning experience.

This chapter builds on the notion that the developmental process is comprised of a series of developmental episodes that build successively upon previous ones. This chapter presents and analyzes in detail four episodes that were deemed significant to this developmental evaluation by the clients. These were chosen on the basis of what the clients self-reported to be pivotal or significant to the development from the interviews.

By presenting these significant developmental episodes, the goal, then, is to illustrate the practice of developmental evaluation and to analyze them for significance. This is done in several ways. The first way in which I am analyzing these episodes is with reference to its intrinsic significance within the episode. Second, I analyze for its practical significance in promoting development amongst my clients. Finally, I analyze it in terms of its theoretical significance, with particular attention paid to the extent to which it is consistent with the theory.
In each of these developmental episodes, I will first provide the context surrounding each episode and situate it temporally within the greater development. Then, I will describe the episode in the way it had transpired. Following that, I step back from the event as the developmental evaluator, and offer my analysis into the event as a researcher. To help underscore this differentiation, I adopt the first-person voice when I recount my thinking and reaction during the developmental evaluation. In later chapters, I revert to the third-person voice when I to analyze and interrogate my own behaviour as a developmental evaluator through the lens of a case study researcher.

Developmental Episode One: Evolving Understanding in using Social Media for Professional Learning

Situating the episode.

My clients (Laurene, Arthur, Tim) and I met on September 27, 2010 in the afternoon. We were four months into the engagement with more than three months to go before launching the pilot program. This is the third of such kinds of group meetings. Although never explicitly agreed upon by the group, this form of meeting was becoming the predominant form of collaboration, a trend that continued throughout the engagement. These meetings were typically collaborative and participatory, with members of the instructional team raising different sorts of issues and discussing them in turn. As the evaluator, I typically assumed responsibility for determining the general thrust of a meeting and assumed the role of the moderator.

Previous meetings with the group have brought about some clarity about some of the desires about what this development could address. By then, there appeared to be a
real desire to explore with innovative uses of technologies to help teacher candidates learn about classroom assessment pedagogy. Although the innovation would be implemented and pilot-tested with a small group, the object of the initiative was to generate transferable lessons and insights about technology use that could be scaled up for large-lecture learning.

Developmentally speaking, previous meetings have yielded little consensus to agreeing on which educational technology to adopt, although there existed the notion that Web 2.0 technologies were to be the focus of this inquiry. Some preliminary conversations have explored the use of e-portfolios (an instance of Web 2.0 technology), but there has yet to be any indication of adoption or forward momentum to formally adopting it.

This meeting followed up on my clients’ interest around experimenting with the microblogging service, Twitter. As the developmental evaluator, I had grown quite interested in this technology, and had seen some interesting implementations of it from other parts of the university. In my assessment around the sorts of issues they would like to learn more about, Twitter struck me as potentially suitable to their needs. As a result, I wanted to introduce and orient my clients to Twitter in order to learn what they perceived to be its benefits and potentials to our context.

**Description of episode.**

On a Monday afternoon, the four of us gathered around a rectangular table that takes up much of the room space. As was typical with these kinds of meetings, the meeting began with some catching-up amongst members of the team. Recognizing the value in these social interactions in potentially strengthening the working relationship
amongst the members, I was conscious not to interrupt these interactions until due time. Tim had brought with him some statistical data results from another project he was working on for Albert to review. When Albert noticed a minor error in Tim’s work, there was general teasing banter amongst my clients. It was Laurene who refocused the group’s attention to the meeting. “Okay, Chi, what are we doing?” she asked (Laurene, para. 21, doc. 48). With that, I assumed the floor.

I began by recapping what had been the focus of the development so far. I structured these introductions to be like briefing reports, with the intention of reminding my clients what some of the achievements had been, what issues and concerns remained, and how I proposed that we invest some time into exploring in order to move forward.

I’ll recap what the conversations had been from last meeting. So, last time, we were talking about how we wanted to… create something innovative to help teacher candidates think about assessment. Laurene, you mentioned that you were interested in modeling the kinds of assessment practices that you’re teaching the candidates. This stemmed from the notion that candidates would often report not seeing the kinds of assessment practices that was being enacted. So what we are trying to do is to explicitly model that kind of behaviour. You also wanted to have a tighter integration between assessment and curriculum classes. You were also becoming more interested in using Twitter as a way of generating reflection. (Chi, para. 22, doc. 48)

Laurene offered a correction to my briefing:
I wouldn’t go as far as saying that because I have no idea how that works, but I am interested in seeing how this 2.0 technology might help that process. (para. 23, doc. 48)

Acknowledging her position, I proposed the following to be the focus of the meeting:

The key messages that I want us to think about are in what ways do we want to use Web 2.0 technologies to explore with the underlying nature of learning that we want to create, that is the learning conditions that we want to create.

I furthered clarified the focus by saying that “it’s not about the technologies that we want to implement or use, but what we want to do with it, and how we can get there” (para. 36, doc. 48).

With that, I wanted the group to consider the notion of technology-enhanced learning (chapter two, this thesis). This engagement was not concerned with designing a technology or a technological platform, but rather with how available technologies could be adapted for use in the context in which we were innovating.

Laurene remarked that she had “forgotten about the business of reflecting and how it was often poorly done ” (para. 39, doc. 48). She added that what she wanted to do was to make sure that if they were doing some reflection that it was meaningful and that it could be used to build understanding as they reflected on their experiences.

Building on these ideas, I offered the following perspective to help my clients conceptualize the possibilities and change that this development could bring about. I argued that up until then, much of what they had done was in creating value and a wealth of information for learning, as evidenced in the development of learning materials and
structural support with TA hours and other efforts in mounting instructional support (doc. 48).

I offered that the developmental question we ought to consider for moving forward was how to generate knowledge and wisdom among the candidates. This was a spin from the notion of expertise that had been brought up in previous meetings. I suggested that problem could partly be addressed using Web 2.0 technologies and its ethos in promoting creation, sharing, and collaboration in a way that might be adapted for learning. I offered the metaphor that what had already been developed were pillars of instruction. Moving forward, the task was to consider how best to connect these pillars and where different people would fit into them. What I was trying to get at was to encourage my clients to consider in what ways our programmatic efforts could serve the needs of our learners. Finally, I proposed that we learn more about how Twitter worked.

As a way of helping my clients learn about Twitter, I introduced them to the notion of a Personal Learning Network, to which Laurene raised the following question to clarify her own learning:

So let me see if I’ve got this straight. A personal learning network is… Is that a network among different technological sites or is that the network that I create through whatever technology I can manage? (para. 43, doc. 48)

In response to her question, I replied, “it’s the latter one” (para. 44, doc. 48). From that, I was sufficiently satisfied that she was grasping the concept. I felt like a teacher trying to gauge the learning of my clients (document memo, doc. 48).
I proceeded to explain that one of the benefits of using Twitter was that a learner could follow and subscribe to the postings of other users, many of whom might be thinkers in their respective fields. That idea raised another thought for Laurene.

Boy, that leads me to something right away, which is something you [referring to Albert] and I talk about a lot, which is how do we, and how do they, establish the authority of those thinkers? How do we make the decision about who has good information or ideas that should be followed up on as opposed to ‘I just thought of this and I don’t have any background to support what I am saying?’ It seems to me if you get the second, what we need is a lot of time to…demonstrate that some of these ideas are not worth holding onto… (para 51, doc. 48)

Albert concurred: “Yeah, my concern when I think of Twitter and blogging is they tend to be uninformed; they are social network sites” (para 52, doc. 48).

Having thought a bit about this very issue, I offered them the distinction of use and misuse. I suggested that the public’s naïve conception of Twitter might be tainted by misuse. I borrowed this use/misuse distinction from the evaluation literature on utilization; an example in of misuse in evaluation refers to the misappropriation of evaluation findings to fuel alternative efforts (e.g., one’s political agenda) for which the original evaluation was not intended. As expected, this distinction resonated with Laurene. Albert agreed as well, and quickly followed up with a query about mechanisms for filtering postings based on legitimacy and authority. Tim offered his insights to this problem:
In Twitter, you can organize lists so you can assemble groups with common people with common interests and form your own areas, if you want to focus on a particular area, so it’s open, but it can also be… (para. 59, doc. 48)

Albert interjected with his assessment: “I understand it can be closed but it can still be very uninformed” (para. 60, doc. 48).

At that point, perhaps sensing that the matter might not be resolved within the means or time available to the group, Laurene suggested that the matter be tabled. In the way that she had suggested, it was suggestive that the matter be kept abreast and be reconsidered amongst other concerns: “… we’ll just dump that out there and you can deal with that as we move along” (para 61, doc. 48). In her use of the pronoun, it seemed the onus was placed implicitly on me, the developmental evaluator, to keep track of these issues and address them when the circumstances arise (document memo, doc. 48).

What follows illustrates how powerful it can be when the developmental evaluator adds novel ideas into the discussion. At some point, I argued that Twitter could be conceived as an asynchronous discussion amongst the learners. Responding to that notion, Albert was concerned that such discussions could easily derail and be of little educational value unless they were moderated. To that end, I agreed with Albert. This analogy sparked Laurene to recall an experience.

That’s neat because, I don’t know about you Albert, but I remember we worked with a teacher really closely who used to do things like a current events session and a topic would come up and he’d say, ‘Okay, we’re going to discuss this’, and he would never moderate it. So, some of the ideas the kids brought to bear were just really wrong. And there were nothing there to actually help kids become
better informed by the discussion. And I think kids used to just get really frustrated at them; it was just kind of like playtime, say anything you want to say, and so we… *that’s what we want to avoid. That’s really what we want to avoid.* (para 67, doc. 48; emphasis added)

As the developmental evaluator, I sensed the emphasis she placed on her interpretation of the experience. What I had taken away from that was a design principle about how our innovation ought to account for. It was this notion that discussions ought to be moderated and made purposeful in a way that facilitates learning.

As a way of demonstrating to them the potential of using Twitter academically, I pulled up the Twitter page for FILM240, a course offered to undergraduate students in the Films and Media department that used a blended learning model involving Twitter. I pointed out different features of the web page, and how students were participating in that environment. In my head, I felt a bit like a car salesperson trying to demonstrate the value of particular features of Twitter. My clients were not particularly concerned with the technical details of Twitter; they were, however, curious about the instructional value of the tweets appearing before them. Referring to a screen full of tweets, Albert pointed at one of them and asked: “What gives me something here that I am supposed to learn?” (para. 153, doc. 48). Laurene inquired about how teaching assistants were being used to support this effort (para. 159, doc. 48).

It can be seen that I was not the sole driver in their learning about Twitter in this episode. As I continued to introduce them to different operations within Twitter and how they could be adapted for our purposes, the clients could be seen taking an active role in constructing their knowledge through different means. For instance, Laurene used
questioning techniques to help her understand how Twitter worked. She would ask
questions to clarify points I had made, and often times, she would ask about what was
appearing before her.

Albert used questioning as a way of building his knowledge, but in a more
nuanced way. His questions were more critical and evaluative. For instance, in my
attempt to explain Twitter by drawing an analogy to discussion boards, Albert asked
bluntly: “So, tell me why Twitter is better as a discussion than a discussion board is”
(para 183, doc. 48). I could not answer with confidence partly because of the novelty of
Twitter; few had researched about it at that point. In another instance, Albert was direct
in counteracting my assessment of Twitter: “I don’t agree with the real-time thing as
much. I think it’s actually slower than having a conversation with a student” (para. 459,
doc. 48). Although I was slightly taken aback by the animosity I had perceived
mistakenly at the time, he revealed its significance in a later interview that I will describe
momentarily below.

The meeting went on exploring the potentials of Twitter. The concerns centered
around understanding the technical features of the platform (e.g. what is a retweet?) and
postulating the educational significance of these actions (e.g. When might a post be
tweeted? What is the significance of that? How might instructors interpret these actions
by the learners?) As the developmental evaluator, I played the role of a teacher in trying
to understand the thinking behind their questions and concerns. Just as I was doing that, I
tried to explore and confirm some of the values and assumptions underpinning their
concerns. The meeting ended abruptly at 2:30 p.m. with Albert having to catch a train to
Toronto. There were no apparent resolutions or actions to be taken arising from this meeting.

**Significance of episode.**

In this meeting, the developmental evaluator assumed a leadership role in introducing his clients to Twitter. By promoting questioning and dialogue throughout the demonstration, the process of participating in the meeting enabled them to clarify their understanding about the technology and draw meaning, relevance, and significance about its potential application to the innovation. What had emerged, as well, concurrent to knowledge construction, was information about the clients’ values, assumptions, and desires that could serve to shape the innovation.

This episode illustrated the dynamic interactions that typified these meetings. It also illustrated the unusual role for the developmental evaluator to facilitate nonformal learning. In the follow-up interviews, two of the three clients recalled specifics from this meeting and attributed special significance to it. Speaking to the value of allotting time for learning, Tim appreciated having spent the time the way we did:

I know we spent some time working out some of the kinks because I know Albert and Laurene are not as technologically savvy as you or me, so it was good that we took some time to familiarize them before they tried these new things with an audience. A few months previous to that first class meeting, it was probably very foreign to them. (p. 4, interview)

This investment of time into learning about the technology contributed to generating buy-in and investment into the development:
And I remember through our conversations we managed to get them to buy into what Twitter was, and perhaps see it as an educational learning tool. In the end, that was one of the important parts of those meetings of getting over the barrier of using technology as a tool as opposed to having it as a learning piece. It was an extra step that needed to be taken. (p. 7, interview)

Laurene independently confirmed the significance of this episode by pointing to the learning it generated for her. In her terms, “[the developmental evaluator] would teach us about the technology, and that was critical because [he] provided us with a context within which a particular aspect of the technology as important to learn”. She reasoned that this process was instrumental to achieving overall goal of the development: “So, we were going to try something, so Albert and I had to learn about that aspect of Twitter” (p. 3, interview). Implicit in her argument was that it might have been more difficult for her to engage in this learning independently. The guidance provided by the developmental evaluator was deemed helpful in promoting her learning. Taken together, their endorsements suggest that this episode is significant on not only a theoretical level, but also it held real-world significance to those who were responsible for the development.

Developmentally speaking, this episode was the first time in the development where a particular solution was explored in detail. Previous meetings had focused on clarifying the intentions of the engagement and on understanding the premises and issues that prompted my clients to innovate. Brief considerations were paid, at one point, to the possibility of using e-portfolios, but that amounted to little momentum. The possibilities of integrating an expanded use of email or trying to use the course management site in
innovative ways were dropped because the clients felt teacher candidates would see these options as simply more of the same experiences already provided by the required course. Neither the clients nor the evaluator had previously invested time to understand the application of Twitter. What appeared evident, though, was that the Twitter remained the focus of the development from this episode onward. In other words, this was a pivotal episode in the development.

The decision to use Twitter was ultimately one made by the clients. They could decide to withdraw their support for Twitter and pursue other technologies. As their developmental evaluator, I was there ultimately to support their efforts, and I would have to respect their wishes. However, their continuing interest in participating in learning about Twitter amounted to an implicit endorsement to keep pursuing it. As the developmental evaluator, I interpreted my clients’ reactions to the information experiences related to Twitter that I brought to the meetings as positive. Having the clients maintain control over how the innovation unfolded was a priority for me as the developmental evaluator, as will be explored below (developmental episode four).

When it came time to pilot the program and the use of Twitter as a teaching/learning tool, meetings were instrumental in developing next steps and quickly adapting to the emergent behaviours, sense-making, and interpretation in response to learners’ behaviours (e.g., Meeting on January 21, doc. 121; Meeting on February 4, doc. 132).

In particular, the meeting where Twitter was endorsed as a teaching/learning tool was illustrative of the learning paradigm that typified this development. As the developmental evaluator, I was deliberate in my practice to foster an environment where
perspectives and concerns about Twitter could be voiced. Having their questions answered likely promoted their learning and the complexity of their thinking about the innovation. It helped identified the values that underpinned the innovation. For instance, in this meeting, Laurene underscored the importance for focused and moderated discussion, which led Albert to think about the importance of engaging in informed conversations by validating speaker’s authority and legitimacy. These were all important elements of the innovation that ought to be surfaced and explicated if the development were to be useful to these instructors, and ultimately to the learners. The assumption here is that my clients’ stances on these issues were informed in part by their past experiences and knowledge about working with teacher candidates. Therefore, their values, assumptions, and perspectives are valid and ought to be taken into consideration in the developmental evaluation. The degree to which these assessments and conjectures were valid could be empirically investigated systematically through evaluation in the formative or summative sense.

For one of my clients, Albert, the question-and-answer discussions that he and I were engaged in, as illustrated above, turned out to be a strategy that promoted his own learning. Recall in the above description where I characterized his involvement as critical and evaluative. At that time, I had misinterpreted this gesture as one of skepticism. It became clear from the interview that he found it helpful:

I can’t speak for others, but for me, I learn by talking things out, so I need to challenge my own thinking, and I need to challenge others’ thinking because it helps me clarify what the other person is saying and it helps me clarify what I am saying. So for me discussions are a really valuable way to challenge thinking and
move [the thinking] along. So I can take a different position when I am having a
discussion and I can start to rationalize or justify the position I have, why it’s
there and it makes me question my own thinking as well. (p. 7, interview)

Engaging my clients in learning about Twitter in a way that encouraged dialogue
contributed to their understanding and complexity about the technology’s potential.

In conclusion, it is often not the case for program evaluators to engage in leading
sessions such as this one. Typically, program evaluators focus strictly on program
evaluation activities, such as the designing the evaluation, engaging stakeholders in
evaluation activities as part of the evaluation process, or the facilitating use of evaluation
data and findings with stakeholders (Patton, 2008). It is rarely the case that program
evaluators engage in facilitating learning on an individual level, as I had done in this case.
I involved my clients in learning about particular facets of the technology, and in
promoting thinking about its adaptation. Important to note was that the learning was not
scripted or structured. I had general ideas that I thought would help my clients in moving
the development forward. This premise led to a learning experience that was highly
adaptive.

Developmental Episode Two: Explicating Values through Appreciative Inquiry for
Program Development

Situating the episode.

This episode describes a meeting my clients and I had that took place on
December 8, 2010 at 10am. With less than one month to go until the launch, I had some
anxiety to see that my clients arrive at some agreement to concerning the development.
For whatever reason, we did not end up meeting in the month of November. Up until this point, the group had met four other times. Despite investing much time into engaging in various activities, many of which were instrumental in furthering the progress, there was little agreement to how this development should proceed. In my assessment, making explicit their desires, wants and wishes would help focus the development.

At the 2010 Annual Meeting of the American Evaluation Society, I learned about the use of Appreciative Inquiry in evaluation (e.g., Preskill & Catsambas, 2006). Appreciative inquiry is “an organization development technique that involves the art and practice of asking questions in ways that strengthen a system’s capacity to apprehend, anticipate and heighten its potential” (Gamble, 2008). Unbeknownst to me then, appreciative inquiry is one the “DE Tools” that Gamble recommends to fellow evaluators for its fit with the developmental context. In what follows is a description of a developmental episode featuring the use of developmental evaluation.

Description of episode.

The meeting commenced with some lively socializing amongst my clients. They were particularly interested in the iPad that I had brought along with me to the meeting. This sparked Albert to share a discovery that he and Laurene had uncovered about the Mac OS X operating system. They had discovered a multi-touch gesture, where a three-finger swipe motion allowed them to navigate forward or backward within a web browser. I was delighted that they were enthusiastic about computer technology. It, at least, meant that they were not technophobes. As the developmental evaluator, I was interested in their attitudes because it could be indicative of their receptivity to experimenting with the technology.
Like the last meeting, I was careful not to stifle social interactions. My belief was that it might encourage cohesiveness between my clients and strengthen the collaboration.

Again, with Laurene’s help, I gained the floor and began with a smile: “Good morning. What I hope to do today is to focus on what this development can be” (para 049). I went on to provide a quick summary of the development progress so far. Recognizing in my mind that the feeling of uncertainty, typical of innovation development, can be disabling, I was intentional in my choice of words to frame it positively. That was important because I considered it as a part of my leadership role to stay positive in an attempt to generate momentum for the development, although admittedly, I was worried about our progress and was anxious move things along. I continued to focus my clients on the task at hand:

As a quick recap of our progress so far, I would label what we have accomplished so far as part of the discovery work. In some of my work, I’ve been referred to what we are trying to build as ‘collaborative social-networked learning’, but what that really is, I don’t think we have any idea what it is or how we can leverage it. I don’t think we have a solid idea, but I think we’re onto something really great, and I think we can definitely make something out of it. As we get closer to the crunch time, let’s get down to business and let’s imagine and describe concretely where we want it go. (para 53, 55; doc. 67)

To my surprise, my introduction triggered Laurene to respond in a way that confirmed my intentions:

Yeah. Can I just say something that I have to tell you? Because I have no idea what this is going to look like, never had an idea what it’s going to be like. The
fact that you are optimistic about it always makes me feel better. No really, like, because, I keep thinking I have no idea how this is going to turn out but, gees… every time we talk to Chi, he goes, you know, this is going to be really great. And I am thinking, okay, I’m ready! (para 56; doc. 67)

What seemed evident was that she was referred to the notion of uncertainty. I went on to acknowledge her feelings:

So I invite you to join me on this journey and I applaud you for, uh, your bravery.

So looking ahead, I think we’ve just finished the discovery phase. Today, we’re going to begin the imagining phase. The goal today is to walk away from this room with some concrete ideas about what are some of things that we want to pursue. (para. 60, doc. 67)

I explained that I would take the data from this session and use it as the basis for building “prototypes” (para. 70, doc. 67).

Before formally starting the exercise, I asked if anyone had anything to add to the agenda or share any other ideas. This was important because it was often the case that there be some issue or an emotion occupying their mind (e.g., January 21, doc. 121). In this case, there was none. I seized this opportunity to consult my clients on a change in the program’s naming that I had noticed in one of the communication forwarded to me. The Assessment Pilot Initiative was originally referred to as the “Assessment Working Group”, but Laurene had called it the “Assessment Pilot Group”. I wanted to see if there was any significance to this change because names often carry meaning behind their literal meaning. In this case, Laurene confirmed that there was not any significance to the labels attached to the initiative (para. 079, doc. 67).
For the appreciative inquiry activity, I laid out a large piece of flipchart paper onto the table and distributed a pad of Post-its, all with a different colour, to each of my participant. Reading off of a list of questions I had prepared before the meeting, I asked my clients, first, to identify some of the things that my clients were already doing well (para. 104, doc. 67). I asked them to write only one idea per sheet. They were then asked to explain what they had written down to the rest of the group. As an example of a response, Albert contributed:

You see, we all focused on slightly different parts, so that’s actually good, because I actually like the teaching aspect of it. So for me, teaching was the first thing I thought we actually do not do poorly of right now. I think we do a pretty good job of introducing some of the broader concepts of classroom assessment so that people have some familiarity with it, and that’s what is not happening in other programs as well. Because of that, the second part is that we are hearing a lot of debate and discussion about assessment amongst faculty and amongst students so we are promoting that debate and discussion. I think one of the things that we are doing pretty well right now is we are separating that formative and summative distinction well. I think we’re doing a pretty good job of giving broad encompassing examples while we are teaching, so they at least have something to hang onto, so they are seeing some relevance to them even if they might be in Phys. Ed. or whatever our examples are. That’s actually a strength and a weakness… (para. 133, doc. 67)
Following that first task, I asked my clients to collaboratively thematize their ideas by physically manipulating the post-it notes and aggregating them by themes. This produced a visual product as seen in Figure 3.

![Figure 2](image)

*Figure 2. Photograph of the end-product from the appreciative inquiry exercise. This photograph illustrates a meeting facilitation technique that makes use of Post-It notes to capture contributions from each of the clients. Working on a large sheet of flipchart paper allowed all participants to observe the big picture, while preserving the inherent relationships among the ideas.*

While my clients were working through that task, I paid attention to the reasoning they were offering to each other to justify their suggestions to the group. Following that, I
invited my clients to speak on their decisions. Four themes emerged as something my clients were already performing well:

1) The instructional focus on assessment *for* learning;
2) The established infrastructure and support already in place for student learning;
3) Attentiveness to student needs; and
4) Promotion of self-assessment and professional self-regulated learning amongst the teacher candidates.

Following that, I directed my clients to consider “what they wish they could do more of” (para 250; doc. 67); I also asked them to thematize their contributions in the way described above. This produced a list consisting of four areas for development:

1) Infrastructure
2) Responsiveness to student learning
3) Self-directed learning
4) Integrating instruction into other programs within the faculty, and not isolating classroom assessment on its own.

In trying to thematize their “wishes”, my clients engaged themselves into thinking evaluative. It became obvious that some of their wishes were more realistic and achievable, while others might have to remain a longer-range goal given the timeframe of this development. In an unplanned move, I invited my clients to come up with “action verbs” that might help describe how one might enact the sort of the change they would like to see:

- Identify
• Track
• Respond
• Monitor
• Motivate
• Combat shared ignorance
• Clarify/demystify (referring to the knowledge base of classroom assessment)
• Expertly facilitate discussions
• Feel connected

Armed with these data (Appendix L), I thanked my clients for their contribution and drew the meeting to a close.

**Significance of episode.**

This episode illustrated a mean of collecting evaluative data through an activity based on appreciative inquiry. By reframing the inquiry into a positive one, the discourse that it generated engaged client to both simultaneously reflect on the strengths of the organization and assess its current capacity, while considering how else the organization could grow in a way consistent with its values and identified opportunities. Generated from this exercise was a condensed articulation of actions, ideals, and values that formed the basis of program development and design (doc. 79, December 17, 2010). This made it possible for me, as the developmental evaluator, to construct a theory of change that makes use of social media Web 2.0 technologies in ways consistent with the values and desires of my clients (see Appendix L).
Not only was the product of this exercise instrumental to the developmental, it is important to underscore the value of the process in engaging clients to collaborate to discuss these issues. At one point in the meeting, Laurene, Albert, and Tim were discussing the notion of “playing with a new infrastructure to help teacher candidates with engaging in self-directed learning” (para 383, doc. 67). Albert observed that while they “won’t get faculty commitment to increase, [it would be possible to] give the students an understanding of it”. This sparked an idea for Laurene:

Yeah that’s what I’d like. I’d like them to see what they’re thinking about in assessment as they move both through the day and maybe even through their practicum in February, you know. I’d like them to have them stop thinking about assessment in a silo and just start thinking about assessment. Think [about assessment] when they are working on their projects and when they are talking to their curriculum instructors. So how do we help them understand assessment in that way and not as a specialized skill set that they don’t have… (para 387, doc. 67; emphasis mine)

At another point in the meeting, I learned about the difficulties my clients experienced five years prior in their implementation of a course web site on WebCT, a course management software. They revealed that the teacher candidates lacked the digital literacy to use the web site and the ways in which they had to mobilize additional resource in an unanticipated move to address their learners’ needs. This newfound knowledge helped me to understand and think about how the development might proactively address this issue.
In summary, this episode illustrated an activity that facilitated systematic collection of data that could then be used to buttress decision-making in a development. In this case, the data collected were used to buttress decisions about the program design. My client placed much significance on this episode. Both Tim and Laurene recalled specifics detailed in their interview. Tim summarized the value of this episode in the following way: “It was a good way to see the big picture of all these little ideas, and I mean, I’m some ideas were used and some were not” (p. 5, interview). In the language of complexity thinking and developmental evaluation, this exercise captured the complexities and constraints within the system (i.e., the program of instruction) in a way that facilitated developmental use.

**Developmental Episode Three: Enhancing Collaboration through Structured Communication**

**Situating the episode.**

Throughout the development, e-mails emerged to be one of two primary modes of communication between my clients and I, with face-to-face group meeting being the other. E-mail was adopted primarily because of its effectiveness in communicating asynchronously with one another and the ease in transmitting large amounts of data at once. From experience, I knew that it was a mode of communication that my clients already used and relied on heavily. All in all, it made sense to continue this trend.

This episode is dated January 05, 2011. It features an e-mail I had sent to my three clients. By then, invitations to participate in the pilot program had just gone out to the teacher candidates. They were asked to register their interest in participating by e-mailing
a designated course e-mail account. Being the developmental evaluator, I had access to the account and I had been monitoring it to gauge receptivity and interest to our invitation.

In response to the sign-ups, this e-mail communication discusses how I went about proposing two ideas with my clients. The first proposal was to inform my clients to the use of a subject-line keyword tag. The second proposal was to propose for collecting baseline data from our program users.

**Description of episode.**

The text of the e-mail reads as follows:

Subject: [Assessment Pilot: decision solicited] Gathering preliminary data on tech use and comfort  
Date: January 5, 2011 3:33:37 AM EST

Good morning:

I am writing as your development consultant. Over the next couple of weeks, I am anticipating that there will be times when I would need to get in touch with you. I will be reporting on what I'm seeing, and feeding that back to you, in order to get your feedback. Any e-mails that are related to our pilot project will be prefixed with the [Assessment Pilot] tag, followed by the action that I require from you. I'll also try to structure my e-mails systematically.

I took a quick peek at the [e-mail] account and was delighted to see a number of replies to our pilot invitation! I think given what we want to do, it would make sense to get some preliminary ideas about their 1) comfort with technology, 2) experience, and 3) what kinds of mobile access they have (smart phones, androids, iPhones, laptops, etc.)

Having this information will allow us to better respond to their needs on Friday.

I recommend that I: 1) reply back and welcome them to the pilot, and 2) gather the above information using a Google Doc form.
How would you like to proceed? Anything else you'd like find out about our participants?

Yours sincerely,
Chi

**Significance of episode.**

This episode described, first, the efforts I took as the developmental evaluator to facilitate collaboration between and among my clients and me. I took into account my understanding of my clients’ busy work schedule, and suspected that structuring the e-mail in a way that would make it obvious the kinds of action needed from them would be beneficial. This sensitivity, in turn, likely helped the clients remain responsive and adaptive to the needs of the evaluator.

That was important because I would act in the place of the instructors. One of the my responsibilities as the developmental evaluator was to observe what was happening in the online community, and redirect my observations to my clients to collaboratively determine what actions was might take. I knew that from the readings of Patton (2010). Furthermore, the clients identified responsiveness as a value that would underpin this development through the appreciative inquiry exercise (see Developmental Episode #2, this chapter). As a result, it seemed to be that the best way to respect these two parameters would be to create a “priority channel” of communication with my clients.

This developmental episode also illustrated how I went about engaging my clients into collecting “baseline” data. It is often the case that the developmental evaluator is responsible for the identification and collection of data to help with data-informed
decision. This notion was discussed briefly above, and will be revisited in the developmental episode to come.

However, what is less typical of a role for a developmental evaluator to play is to draw upon his expertise in a substantive domain to inform the developmental evaluation, a recurring theme of this developmental evaluation. Part of the reason why I became involved in this initiative was that my clients wanted to capitalize on my expertise and knowledge in working in the social media space. In another words, my substantive expertise was sought after by my clients. In bringing with me my expertise, I was able to shape and inform the program development in ways not possible by a developmental evaluator without my expertise. In this instance, I drew on my knowledge to recommend survey items to be included into the questionnaire. Later, I helped facilitate the meaning-making process and helped adapt the pilot program when we discovered that few of the teacher candidates had any real experience in using social media.

Because I was able to draw on my substantive expertise in social media to articulate its potential, my clients were in a position to incorporate my suggestions in order to focus more on adapting the technology for teaching and learning purposes with their students. This contribution is important because the technology was an explicit part of the intervention; any programmatic decisions made without considering the capabilities or the learning curve to using the technology would be incomplete. Working collaboratively with me as the developmental evaluator provided my clients a more complete view of the program and enhanced the validity of the conclusions drawn from the data.
Developmental Episode Four: Program Development through the Use of Evaluative Data

Situating the episode.

The last developmental episodes to be presented in this chapter features another team meeting. This meeting is the eighth in a series of thirteen meetings held during the Assessment Pilot Initiative development. As compared to the September 27, 2010 meeting described in the first developmental episode, the Initiative had entered the implementation phase by now, with a program model finalized on January 3, 2011.

The road to arriving at this program model was long and gradual. Previous meetings had led to decisions being made about the technology to be piloted (September 27, 2010, doc. 48), how the instructors would leverage a blended model (December 17, doc. 79), and there was now a beginning sense to how the teacher candidates might use the technology in a way for learning (January 7, doc. 107; see also, Appendix I). These decisions were made possible partly because of having made explicit the kinds of desires, wants, and wishes the clients would like to see. This visioning process took much time and finally culminated in an exercise using Appreciative Inquiry.

This meeting took place just one week after the pilot program had launched. In the first class session, my clients and I welcomed the teacher candidates to the pilot program and presented to them an overview and rationale for mounting this pilot. A part that was unusual for the evaluator to have facilitated was a training/demonstration session on how to use Twitter. In particular, we wanted to illustrate and model a way that Twitter could be used for learning in a way that builds on the desires, values, and considerations my clients and I discussed extensively in the months prior. Additionally, I was invited to
co-manage the Twitter account for the pilot program. In both of these cases, I was invited
to take on that role because of my facility with social media.

My clients and I met in the morning on a Friday at 9 am with just an hour to spare
between the end of the meeting and the start of the session with the participants. The
purpose of this meeting was to engage the clients in some reflection and sense-making
about the activities of the online community, and to infer any possible actions or
strategies to be put in place with the teacher candidates in the class session that followed.

In fact, because my clients will be meeting with the teacher candidates in two
hours, the most pressing need was to develop an instructional plan. It was not the content
that needed to be developed, but how the instructors would shape the instruction and
learning of the teacher candidates consistent with the goals, values, and desires of this
initiative. It might have been the case that we discovered that very few of the teacher
candidates were adopting the technology; a decision that might have to be made, then,
was to scrap the blended learning piece and resort to face-to-face teaching only.

Part of what my clients had to do in the class session was to check-in with the
teacher candidates about their use and experiences with the technology. This meant that
we, as the instructional team, ought to begin to understand our own behaviours and to
begin to draw some conclusions about Twitter’s utility.

**Description of episode.**

At 9 am on a Friday, the meeting commenced. Laurene began the meeting
with a question about the technology: “How did we get all these people on our list is what
I want to know?” (para. 1, doc. 113). She was remarking on the fact that Twitter users
external to the learning community were being ‘followed’ by the account. This meant that
the posting these users external to our learning community were now appearing on the Twitter feed Laurene was viewing. Tim, being a user offered an explanation: “I think what Chi does is he’ll find some interesting tweets and he’ll re-tweet them” (para. 14, doc. 113). Sensing some confusion in the room, I felt prudent to step in and explain my intentions. It was a purposeful move on my part to subscribe to other accounts that published educational related tweets. The assumption was that the tweets could easily be re-tweeted to our learners. Doing so would enhance the quality of the information available to the network; our teacher candidates could be exposed to tweets that they might otherwise be not aware of. I referred to this practice as “curation” to my clients (para 017, doc. 113), a term used in popular parlance among social media users to refer to the systematic collection and re-posting of tweets from other users to those of one’s subscribers. This came at a cost, as Laurene described in the following way: “I find it really confusing that there are people that are tweeting us that I have no idea who they are” (para 19, doc. 113). Sensing her preference, I proposed that we unsubscribe those users that were not part of our contingency. This was met favorably.

I began the meeting by introducing the two tasks I had prepared to focus the meeting. I explained that I compiled the tweets that our learners sent us, and that we would go through the list and make sense of what was happening. The second task that I proposed to the group was to solicit their “general thoughts and reactions about what was going on” (para 80, doc. 113). I had intentionally left the request open and vague. Laurene responded to the second proposed task right away:

I have a couple. Well, I’ve told you one already, that to clean up the Twitter site so the only people we can see on the account would be our teacher candidates;
that would be really good. But the other thing that I am wondering about, and I’m even wondering whether we have to ask the students this today… When students actually tweet us, it’s like we have got two or three requests for things here. How do we respond to that? You don’t want to take first-come necessarily, so it sounds to me like what has to happen if you’re going to do this as a blended learning, is that we have to try and synthesize at some point everything that’s come in in a particular time period, and the figure out what we have to do to address their needs most of the time. (para. 85 & 87, doc. 113)

Laurene’s remark reminded Albert and Tim one of the teacher candidates’ tweet that had mentioned an instructional method unknown to the team. Reacting to that tweet, Albert proposed that:

So there are two things I think we should talk about today based on that, and one is the language of assessment that is coming out there. So the thing that I can see as being valuable, I think we need to talk about their experience with Twitter and the challenges they found. (para. 105, doc. 113)

Laurene and Tim agreed.

The conversations that followed investigated ways to learn about teacher candidates’ experiences in using Twitter. What emerged out of that conversation was to employ classroom assessment strategies, like exit cards, with the teacher candidates. Framing their experiences with Twitter as a learning experience, the exit card strategy would serve to assess their experiences, as well as, presenting an opportunity for my clients to model the strategy and for the candidates to experience it in action.
In the final part of the meeting, I engaged my clients in interpreting some of the
tweets that I had compiled. In particular, I drew their attention to those that expressed
notions other than classroom assessment. One of the tweets read “if [I] didn’t do this mini
course, [I] would never have done this Twitter thing” (para. 261, doc. 113). Because it
was not clear whether that tweet was an endorsement of the learning experience, I wanted
seek my clients’ opinions. I suggested that “it was really neat” how the experienced
pushed her out of her comfort zone (para. 262, doc. 113); Laurene agreed. Albert saw it
differently: “I wasn’t sure… she said too much tech, too fast. I don’t know if that was…”
(para. 264, doc. 113). I followed up: “like discomfort or anxiety?” (para. 265, doc. 113).
Albert replied: “well it may just be giving up. I wasn’t sure how to take that message so I
think I did respond to her at some point and said, yeah I know exactly how you feel”
(para. 266, doc. 113). In a similar dialogic fashion, we systematically interpreted a few of
the other tweets to get a sense of how the learners were responding to the learning
experience.

Based on the discussion my clients and I had in this meeting, a plan of action
emerged. We would consult the teacher candidates on their learning experience. We
would also provide some ‘teaching’ on classroom assessment based on the ideas that the
teacher candidates had tweeted us.

**Significance of episode.**

This developmental episode was in many ways similar to the first developmental
episode reported. It continued a pattern where I, as the developmental evaluator,
facilitated a check-in with the clients to solicit their feedback in the beginning portion of
the meeting and, later, facilitated a forum to discuss and develop next-steps based on the
available information. During these discussions, clients voiced issues that they were experiencing, reflected on progress, postulated possible actions, and collaboratively participated in the decision-making about next-steps.

What this developmental episode illustrated was the utility in structuring a progress check-in and in providing my clients activity summaries. The value of progress check-in allows the clients to unload ideas and issues they had been wrestling with. More importantly, it allowed me, as the developmental evaluator, to solicit feedback on the program. Activity summaries, in this case, referred to the compilation of tweets prepared by the developmental evaluator for the use of the clients. By compiling the tweets into one document, it provided an at-a-glance access to the data; this was important to reduce the cognitive load of my clients as illustrated in the exchange featured in this developmental episode.

The clients responded favourably to having these summaries. Laurene explained that they were instrumental to the instructional planning and in doing so, lending a degree of adaptiveness to the process:

You did a great job of providing summaries of the discussion flow, so that we could pick up on the issues that were really relevant to them at the moment so that we could talk about them on the Friday (i.e., the day on which the sessions were held). And we actually planned our formal instruction on the Friday. We planned that around the kinds of discussions and needs that came up during the week in the conversations back and forth. (p. 5, interview)

Tim found these summaries useful, but for a slightly different reason. These summaries created a structure and that helped the clients keep their ideas organized:
I think it’s atypical to have someone like you who would track all these different things and put them together into summary sheets for us at the next meeting. Having those ideas summarized for us the next time helped us make clearer connections as we talked further and you had more questions for us, and in the end I think our conversations ended up being more rich and fruitful. (p. 15, interview)

This episode illustrated two of the primary functions of a developmental evaluator, which were to “track the implications and results of decisions and to facilitate ongoing, real-time, data-based decision-making in the developmental process” (Patton, 2010). The developmental evaluator brings to bear information, collected through systematic ways, to the clients that can then be interpreted and given meaning. This allows, then, for the making of evaluative judgments concerning the merit, worth, and significance of the functioning of the program. Because outcomes are emergent in an innovation and hardly predictable, decision-making about next-steps is contingent upon the availability of data. Hence, it is important to track program activities in order to facilitate program adaptation.

Gamble (2008) drew importance to the data-collection function of the developmental evaluator as well. He distinguishes between two types of data in a developmental evaluation. In his terms:

Developmental evaluators are attuned to two streams of data. First, there is information that assists in validating certain decisions, approaches, or assumptions… The second stream of data for developmental evaluation documents the innovation process. (2008, p. 31)
In this episode, the kinds of data collected and presented for use to the clients fell under the first type.

Compared to the first developmental episode featured in this chapter, my role in this session was more passive. My role, in this episode, was to bring in evaluative information to stimulate the conversations, and to ensure that the discussion arrive at some next-steps. I only stepped into the foreground only when needed.

As the developmental evaluator, an assumption I held on to strongly when I facilitated these meetings was that each of my clients held unique concerns that stemmed from their past experiences. They were uniquely qualified and held expertise in certain areas. That formed the basis of my view on collaboration.

Translating that assumption into practice meant that it would be important to facilitate an environment where the weakness of one person is supported and bolstered by the strengths of others. Furthermore, as the evaluator, when we were discussing about issues that I was not uniquely qualified, like teacher education, I was cognizant of how strongly I ought to pursue my convictions. At times when I felt the need to offer a dissenting voice, perhaps because I may had more credible or suitable information that could lead to a better program design, I tried to do so gently. The interaction about “curation” was an example of this.

My assumption was likely informed in part by one of the central tenets of utilization-focused evaluation. Utilization-focused evaluation is an approach to evaluation that “shifts the attention from methods or the object of the evaluation to the intended users of evaluative processes and information, and their intended uses” (Patton, 1994, p. 317; emphasis his). It recognizes that all evaluations are not values free. It
follows that the values that inform an evaluation should be those of the primary users. In
the case of this developmental evaluation, the evaluation was informed by both the values
of my primary users and their expertise as well.

This likely amounted to preserving clients’ control and agency in this
developmental process, something that Laurene valued:

I never felt that I didn’t have control, with Albert, over what we were going to try
and what we were going to do with students. Having said that, I would say you
and Tim were central in providing us with ideas and information that helped us
with those decisions, and sometimes you would recommend things like “why
don’t we try this”. I don’t ever feel that you pushed us to an agenda. I really think
it was more of kind of ‘here is what we have learned, here’s where we can go, and
here’s what the students are saying later on. (p. 5, interview)

This was important insofar as it likely engendered support among the clients to
remain engaged in the development.
Chapter 6: Discussion of the Study

Chapter Overview

Following a discussion of the particulars of the case in the last chapter, this chapter now takes a post-facto look at the development. It revisits the research questions that frames this case study and explores the consequences and implications of developmental episodes in relation to each of the research questions. This chapter is organized into four sections.

The first section determines the extent that Assessment Pilot Initiative qualifies as a developmental evaluation. I do so by comparing this case to one of the five uses of developmental evaluation that Patton (2010) described.

The second section describes how developmental evaluation promoted and supported the development of the Assessment Pilot Initiative. It describes how a developmental evaluation unfolded in a context of innovation in the case. The data-informed model of development is described and its implications and consequences to the development are explored. The second half of this section describes how the data-informed model of development promoted development across the many months of development. What the data documented was the non-linear, iterative dynamic of the development as it cycled between six distinct developmental foci: 1) defining, 2) delineating, 3) collaborating, 4) prototyping, 5) illuminating, and 6) evaluating.

The third section describes the extent that the needs of stakeholders were addressed in ways that informed the program development. It first qualifies what the developers deemed to be the most significant needs of the development. Then, it presents findings from the case that describes its manifestation and how these needs were
addressed. This chapter concludes with an analysis of the roles and responsibilities of the developmental evaluator in Assessment Pilot Initiative.

Revisiting Research Question One: To what extent does Assessment Pilot Initiative qualify as a developmental evaluation?

Up until this point, it has been assumed that this case actually does qualify as an instance of developmental evaluation. The extent to which this premise holds must now be determined. To do this I will revisit the activities and outputs of the Assessment Pilot Initiative, as documented in the program records and reported by the developers. I go on to examine the extent to which the API, as enacted and experienced by the program developers, qualifies as a developmental evaluation. The case findings will be compared to the current literature on developmental evaluation, as introduced in chapter three, to determine the extent of its congruence.

Preformative development of a potentially broad-impact, scalable innovation.

Beyond the basic tenets of developmental evaluation theory as presented in chapter 3 of this thesis, Patton devoted half of his canonical text (2010) to specify and discuss five specific situations that developmental evaluation could serve. In his words:

Beginning with the first chapter, and as the book has unfolded, I’ve been making the case that developmental evaluation is particularly appropriate for but needs to be matched to five different complex situational challenges and developmental purposes. (Patton, 2010, p. 306; emphasis added)
It might be instructive to consult his criteria to determine whether the Assessment Pilot Initiative qualifies as a developmental evaluation. The five different purposes and uses to developmental evaluation are:

1) Ongoing development;
2) Adapting effective principles to a local context;
3) Developing a rapid response;
4) Preformative development of a potentially broad-impact, scalable innovation; and

He described preformative development of a potentially broad-impact, scalable innovation in the following way:

Changing and dynamic systems require innovative solutions and creative new approaches to worsening conditions. Social innovators aspire to major change with broad impact, expecting to engage in disruptive systems change with a new model. But *that new model does not exist and needs to be developed*, reorganizing and exploring new possibilities as old systems show signs of collapse and dysfunction (and may have already fallen apart). (ibid, p. 311; emphasis his)

On that basis it seems fitting that the Assessment Pilot Initiatives be qualified as an instance of preformative development of a potentially broad-impact, scalable innovation. Over an extensive year-long engagement, the developmental evaluator worked closely with Laurene and Albert, and their lead teaching assistant, Tim, to
examine the programmatic context to identify constraints and opportunities inherent in the system to innovate the learning experience.

The instructors became increasingly convinced that change needed to happen both at the pedagogical and programmatic level. Albert was “coming to a belief that [instructors] actually have to teach it or at least create the conditions for [learning classroom assessment]” (Albert, interview, p. 3). Implicit in his statement is the notion that classroom assessment is not something that can be transmitted but one that requires active (re)construction of knowledge and experiences; Laurene explained that the large-lecture format created a distance between the instructors and teacher candidates and that this barrier detracted them from developing an appreciation of the complexity of classroom assessment as it relates to pedagogy (interview).

In response to these deteriorating conditions, a tipping point was reached when they invited me to participate in the development as a developmental evaluator. Overseeing the development process, structured and unstructured activities promoted discussion amongst the group about the nature of the problem with which we were dealing. It was intentional on the evaluator’s part to infuse new knowledge into the conversations to enrich the process, however unstructured they might be. These discussions helped unpack the complexity that underpinned their vision. In fact, my knowledge of social media and web technologies shaped the development in ways not possible otherwise:

What I really appreciated was […] that you do have adeptness and a deftness.

You are quite good at quickly finding things and showing and demonstrating how
they work so that foundational skill set allowed you to be quite adaptable to situation as they came up during discussions. (Albert, interview, p. 11)

The series of meetings resulted in a design concerning the program logic by the midpoint of the Initiative. The program was adapted on the fly as it was implemented with a group of 20 teacher candidates in a pilot. This on-the-fly adaptation resulted in a codification of program logic as exemplified through the enacted program and illustrated in a publication distributed to the teacher candidates (see Appendix J).

**Trends in program development.**

Understanding the program development is best done at the level of the entire program. This is possible by analyzing the developmental moments of the Initiatives. Several patterns and trends become obvious when developmental moments are plotted against time as seen in the chronology (see Appendix K). Together, they strengthen the argument for qualifying the Initiative as a developmental evaluation.

**Observation 1: Development occurred through purposeful interaction.**

Development occurred primarily through interactions between the evaluator and members of the development team. They were purposeful in nature. The evaluator would determine and communicate the focus of these interactions based on what had occurred previously. The kinds of topics discussed within each interaction varied from time to time. It was, however, consistent that each interaction featured an exploration of the problem’s complexity and some responses to them. This phenomenon is illustrated visually in the vertical distribution of the three types of the developmental episode (see Appendix K).
The nature and consequences of these interactions promoted a specific kind of learning, that will be further elaborated and explored in the third section of this chapter under research question three.

Observation 2: Concretization of ideas and thinking.

Referring now to developmental moments across time, there appeared to be a “concretization” of ideas and thinking. The level of sophistication in thinking about the problem of teaching and learning about assessment and the ideas that arose in exploring this problem were initially broad and unfocused (see for example, blue developmental moments and green developmental moments in Appendix K). Towards the end of the development, they became specific and refined. A prime example is how the notion of reflection is leveraged and tailored with Twitter in this context. It was through engaging in the developmental process that my clients come to understand the potential and affordances of the technology in ways that could be used to promote learning.

Observation 3: Intensification of developmental evaluation interventions.

The case experienced a concomitant increase in both the frequency and specificity of developmental evaluation interventions as the development was implemented. This was illustrated visually by the higher density of red developmental moments towards the end of the development (see Appendix K). These interventions were strategic in nature to either further the developers’ thinking about the underlying problem or the API-in-action. This was best illustrated by how communication structures and strategies evolved and adopted. Other developmental evaluation interventions include contributions in the areas of group facilitation, enacting a learning framework, scenario testing and hypothesizing, sensing, and sense-making. These ideas will be elaborated in subsequent sections.
To sum up, the patterns and trends described above provide visual evidence about the developmental nature embedded within the Assessment Pilot Initiative. An assessment into the merits and significance of the resulting program developed from this engagement may help strengthen the argument, but such an examination is beyond the scope of this paper. Suffice it to say, the resulting program represented a significant departure from the conventional thinking and method of teaching and learning typical of the classroom assessment educational programs. It leveraged a cutting-edge social media tool, Twitter, and adapted it for learning in a networked environment.

**Points of divergence.**

On the broadest level, the development appears to have progressed through the following phases: conception, planning, implementation, and evaluation. This convenient view of development is largely fitting with the general thinking about the nature of development. It may be tempting to prematurely describe the API development process as linear. However, a close reading of the developmental moments would indicate a non-linear, non-phasic process.

In what follows, I discuss the kinds of responsibilities, behaviours, and thinking that have emerged as the developmental evaluator responded to the unique demands of innovating in a developmental evaluation.

**Revisiting Research Question Two: What contribution does developmental evaluation make to enable and promote program development?**

Program evaluation as a method of inquiry is only successful insofar if the knowledge generated leads to some perceptible and meaningful impact to the program. One such indicator is to consider the utilization of an evaluation. This premise has long
been recognized by the evaluation field under the approach of utilization-focused evaluation (Patton, 2008). A commonly known anecdote languishes at the wasted efforts (and funds) into producing evaluation reports that sit atop a shelf gathering dust.

Evaluation use is a top concern for many, and one of the top ways to engender evaluation use is to pay attention to the evaluation process. This concept is referred to as process-use in the literature (ibid). There is no reason to expect that this principle would not hold in the context of developmental evaluation: a developmental evaluation is successful insofar as the knowledge it helped generate contributes to the development. By analogy, one of the top ways to promote development is to pay attention to the evaluation process.

Patton argues that one of major responsibilities of a developmental evaluator is the tracking and documentation of the development. Little is said about how it can be leveraged to actively shape program development in a context of innovation. Furthermore, as an evaluator new to the field of developmental evaluation, I cannot help but wonder what role evaluation and evaluation science might serve in the absence of a program? If indeed developmental evaluation is promoted purportedly as a method for enabling program development in complex contexts (Patton, 2006), it stands to reason that the developmental evaluation process contributes in some way to promote and enable program development. Missing in the literature is a discussion of how it is utilized after data is collected.

The present question determines what contributions developmental evaluation makes to enable program development, and to the implementation and short-term outcomes. To do so, it examines the process of the developmental evaluation across time
in relation to understand how program development occurs through the frameworks offered by developmental evaluation.

**Developmental evaluation as a data-informed process of innovation.**

The contributions of developmental evaluation theory in the Assessment Pilot Initiative can best be characterized as creating the conditions for the emergence of a data-informed innovation process. Analysis of the archived records indicated many different kinds of data collected throughout the development. Table 4 summarizes the kinds of data that were collected.

<table>
<thead>
<tr>
<th>Data Form</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>• Collaboration as an approach to learning</td>
</tr>
<tr>
<td></td>
<td>• Explicit alignment between the use of technology to specific learning outcomes</td>
</tr>
<tr>
<td>Assumptions</td>
<td>• Teacher candidates are familiar with social networking tools</td>
</tr>
<tr>
<td>Expectations of the Development</td>
<td>• Pilot the development</td>
</tr>
<tr>
<td>Assessment of knowledge, skills,</td>
<td>• Knowledge in structuring a community of learners</td>
</tr>
<tr>
<td>capacity, and comfort level with</td>
<td>• Ease-of-use with Twitter</td>
</tr>
<tr>
<td>leading change</td>
<td></td>
</tr>
<tr>
<td>Program Activity</td>
<td>• Usage reports</td>
</tr>
<tr>
<td></td>
<td>• Number of tweets</td>
</tr>
<tr>
<td></td>
<td>• Number of active contributors</td>
</tr>
</tbody>
</table>

Across different developmental episodes, the following steps illustrate the inquiry process through which data were generated in the developmental evaluation:

1. A piece of data (e.g., a finding or an observation) is be introduced to the team.
2. The team would collaboratively attempt to interpret and make meaning of it and explore its relevance to the development;

3. A discussion of its applicability (i.e., significance) would ensue; and finally,

4. The identified points of confusion, potential fit, and emergent issues could form the focus of subsequent inquiry (“leads”).

The inquiry process was described to have “injected new ideas for [the clients] to think about” (Laurene, Interview, p. 6). It was through this mean that new ideas would be incorporated into the ongoing exploration about the problem’s complexity. Rather than remaining an empty academic exercise, the concepts or ideas were applied and analyzed against the wishes and wants of the innovation.

The employment of the above inquiry process could potentially identify more issues and questions than could be answered if no filtering mechanisms existed. Given finite time and resources, decisions had to be made in order to deal with the multiplicity of data such process unraveled. What helped to make sense of the data was the employment of evaluative thinking. It worked against the multiplicative effect by producing highly contextualized knowledge concerning the quality of an idea, based on its applicability to the development, and its emergent fit to the complex characteristics of the population and the environment an innovation sought to serve. This thinking was best summarized by a heuristic that I (as the developmental evaluator) adopted in the Initiative: Now that we know what we know, what do we really know? What other information might we want to know in order to move forward?
Implication of adopting a data-informed approach to innovation:

Responsiveness in developmental evaluation.

This data-informed approach to innovation allowed the development process to become and remain responsive in multiple ways. First, responsiveness was experienced at the program level. The program model was adaptive to the emergent needs of program users identified through evaluation. One such example was the addition of a discussion board to be used in conjunction with Twitter. By mid-course, developmental evaluation resurfaced a concern raised previously about the efficacy of Twitter for promoting learning via 140-character tweets. Teacher candidates independently confirmed the suspicion that led to the jointly-made decision to incorporate discussion board technology to complement Twitter.

Analysis of the developmental work of the Initiative indicated that being responsive extended to the ideas that were surfaced, and explicated in the developmental evaluation. One such example could be seen in how the notion of social learning took shape in the development. Recall that the developers were motivated to explore alternative ways to connect the developers with teacher candidates in ways that “helps them see the complexity of classroom assessment in a meaningful way” (Laurene, interview, p. 1). This was first described in terms of using technology to facilitate the “sharing of experiences” (May 24, 2010, doc. 4). The ensuing conversations led me to consider the notion of enculturation (June 03, 2010, doc. 6). Subsequent conversations gave rise to a line of inquiry that considered the literature on computer-supported communities, e-learning and social learning, community of practice/community of learners, and so on. Along the way, the development team also considered about how to
develop in teacher candidates a capacity to reevaluate the validity of the stories they hear from colleagues. They recognized that teachers often swap stories of teaching practices and accept them at face value. This line of inquiry finally culminated in adopting a professional learning community model (e.g. February 10, 2011, doc. 137). Indeed, responsiveness to ideas manifested in multiple ways. Responsiveness also extended from the developers to the ideas proposed and introduced by the developmental evaluator.

With each iteration of the inquiry, the collective knowledge of the group advanced. The sophistication of our collective understanding into the solutions that could be deployed to tackle the issue grew, much like the growth of a tree from a seedling. In a synergistic manner, as our understanding of solutions grew, the complexity of the underlying issues became explicated. The more questions answered, the more questions arose. Evaluative thinking provided the means to identify which route to take among all of the alternatives exposed. The more we investigated about the potential solutions as we engaged with innovating, the more readily able we were to engage and disentangle the complexity.

**Consequences to adopting a data-informed approach to innovation:**

**Resolving uncertainty.**

The consequences of utilizing a data-informed process were that it helped resolve the uncertainty program developers experienced when engaging in a highly complex situation like innovation. While unable to completely rid the process of uncertainty, we, however, confronted it. Analysis revealed that uncertainty underscored not only the initial phases of the development, but indeed, was felt throughout the process. Developers felt uncertain not only about the program under development, but also about how to proceed
next. Laurene described her experiences about the data-informed process in the following way; notice the undertone of uncertainty in this comment:

I really looked forward to our meetings because…typically new ideas would come out of them not only about what we should do with the students but in terms of how I was thinking about teaching and learning and instruction. (Laurene, interview, p. 3)

She explained that “and when we would finish […] we would have a much better idea about what we were going to do next” (Laurene, interview, p. 3).

Related to this notion of resolving uncertainty was the idea that a data-informed process exposes the complexity of the underlying problems, issues, and of the path ahead. Albert talked about this contribution as such: “… it’s helped push my thinking forward, [and] it’s helped me see possibilities I didn’t know were there” (Albert, interview, p. 13).

In a developmental context like the Assessment Pilot Initiative, exposing the complexity of the problem enriches the sophistication of the team’s thinking and capacity to think about the problem.

When data-informed developmental evaluative inquiry is applied to decision-making, the decisions made were about the adaptation of ideas and theories. By adaptation, I am referring to the determination made of which aspects of the data (e.g., idea or observation) fit most with the potential program logic, and consequently, could be retained and “docketed” for later consideration.
Data-informed developmental evaluative inquiry in action.

So far, the discussion described how developmental evaluation gave rise to a data-informed process of innovation. Equally important to consider was the application of this process across time to reveal how different foci of development arose from it.

Captured in the case data was a description of how this data-informed approach enabled the development of the Assessment Pilot Initiative across time. Examination of a specific subset of developmental moments (those in blue, as seen in Appendix K) articulated explicitly the complexity of the ideas underpinning either the process or the product revealed a pattern that described how the Initiative came to be developed. While the Initiative appeared to have followed a linear process of planning-design-implementation, analysis revealed a dynamic that was non-linear and iterative. These developmental moments clustered around six foci of development describing the development process: 1) definition, 2) delineation, 3) collaboration, 4) prototyping, 5) illumination, and finally 6) evaluation. Each is clarified below.

Definition.

Definition, as a foci of the development process, described the acquisition of insights derived from information-seeking activities that sought to uncover and expose the complexity of the underlying ideas and issues underpinning the problem an innovation targets. These insights led to a richer understanding about the complexity of surrounding the promotion of teaching and learning using Web 2.0 tools. The resulting insights provided the “ammunition” (Albert, interview, p. 14) for making evaluative decisions during the design and implementation of the program. By drawing from this rich understanding of the underlying complexities, the developers were in a better
position to shape the program logic and model, as well as respond and adapt appropriately and effectively. Experienced teachers understand where pitfalls are with a concept.

The kinds of defining-related insights clustered around three themes: 1) problem formulation, 2) developing a deeper understanding of the program context, and 3) conceptualizing the program.

First, problem formulation was especially evident during the initial phases of the development. These concerns centered on trying to express and articulate what developers perceived to be the problems and issues they would like to explore within the developmental evaluation. There was much struggle in defining the problem. Was the development about the introduction of software; as in was case with e-portfolio? Was the development more interested in integrating educational technologies into instruction? Was there a model or framework that could be readily implemented, or would we need to design one from scratch?

Second, the developers also struggled with gaining definition into the program context. Some of the discussions centered on learning about the program history (doc. 67), the changes made to program models (doc. 121), its relationship to other programs (doc. 67), the characteristics of the program users (doc. 121, 132), and the place of the innovation relative to other programs. The resulting learning from exploring these issues allowed the development team to anticipate possible pitfalls.

Third, the developers attempted to conceptualize the innovation by drawing upon theories and ideas from the greater literature. Three general domains of knowledge were consulted in the Assessment Pilot Initiative: 1) technology (e.g., e-learning, computer-
supported community, Web 2.0), 2) classroom assessment, and 3) learning communities.

To sum up, definition-related developmental moments sought to define and bring clarity to the problem and issues that underlie the problem.

**Delineation.**

Delineation referred to the distillation of desires, needs, and wants associated with the development. Whereas definition-related developmental moments revealed and exposed the complexity of the problems and issues, the ideas generated under delineation helped with narrowing down the options and focusing the development. Analysis of the Assessment Pilot Initiative highlighted two kinds of delineation: design considerations and design outcomes.

Design considerations described the desirable quality expected of an innovation. Data generated from Appreciative Inquiry helped to elucidate these principles. Developers made such requests as “monitoring students’ changing conceptions over a longer term, supporting informal learning, and having the students see [the developers] as a resource for engagement about assessment and not for completing assignments” (Appendix L). These considerations, once elucidated, could then be addressed and built into a program model.

Design outcomes referred to the explicit articulation of some of the desirable outcomes of an innovation. The design plans were such an example (Appendices M, N).

**Collaboration.**

Collaboration spoke to the negotiation of ideas and decisions that sought to promote collaboration among members of the development team. Developmental moments highlighted such examples as negotiating and adopting structures and processes
that enabled collaboration, such as the use of Appreciative Inquiry as an inquiry method, or the significance of reporting and briefing/debriefing efforts by the developmental evaluator. Collaboration was an especially important aspect of the Assessment Pilot Initiative because of the various strengths, experiences, and perspective each member brought to the team. The assumption was that collaboration would help strengthen the product. According to Albert:

I think the working together helped pull those different strengths and look at some other … to build something bigger than that so it wasn’t just a sum of parts, but it was using our own strengths to look at some of the needs that we had, so I found that really valuable. (interview, p. 4)

This suggests that a major responsibility of the developmental evaluator was to promote collaboration.

**Prototyping.**

Prototyping described the purposeful integration of the accumulated knowledge from a developmental evaluation to produce an implementable program model for the immediate contexts in which development happened. The prototyping (Appendix N) incorporated insights about what had been learned about problems and issues to arrive at a design of the program logic that sought to maximize program outcome/impact based on the best available information and most plausible case.

**Illumination & Evaluation.**

Illumination described the emergence of contextualized knowledge and insights that arose from implementation of the prototyped program. They included emergent unexpected issues, such as teacher candidates perceiving the Twitter interface as busy
and conversations disjointed. Monitoring and evaluation of the ongoing activities provided the mechanism to help identify emergent concerns from the developers and the teacher candidates.

To sum up, developmental evaluation provided a framework of systematic inquiry that in turn enabled a data-informed process to innovation. Various kinds of data were collected and used to shape the innovation. While the former described the process of innovation, analysis into the kinds of concerns that surfaced during the process revealed six general foci of development: definition, delineation, collaboration, prototyping, illumination, and evaluation. The development cycled through these six foci at various points of the development. Once the program was implemented, highly contextualized knowledge and insights about the implemented program, generated through monitoring and evaluation, shaped the ongoing program and enriched the understanding of the complexity underpinning the social problem.

Revisiting Research Question Three: To what extent does developmental evaluation address the needs of developers in ways that inform program development?

Defining stakeholder needs in a developmental evaluation.

The most prominent and pressing need of the developers in the Assessment Pilot Initiative was to resolve the uncertainty related to working in a complex unknown situation that typified the developmental context. Resolving uncertainty, about either the process of innovation or the substantive knowledge needed to innovate, required a continuous input of data into the development process. Developing a deep understanding around the issues and problems the innovation addressed, as well as the capacities to develop such breadth of knowledge were necessary. In light of this need, it stands to
reason that the crux of developmental evaluation, besides enacting a process that moves
the development along, is to promote learning and discovery among the developers and
the establishment of a learning framework. If the framework of developmental
evaluation was the engine that propelled development and innovation forward from a
point of uncertainty to (increased) certainty, learning, and discovery were the fuel that
kept the engine running. This section elaborates and presents evidence in support of this
idea.

**Learning in a developmental evaluation.**

This notion of learning and discovery resonated deeply with the developers. When
asked to describe the development, Albert replied with little hesitation, saying that: “I
guess the best way to describe it [was] a learning experience” (interview, p. 4). This
sentiment echoed the views of the other developers (Laurene, interview; Tim, interview).

Active learning was identified as a major theme of the development. For instance,
the act of learning about how to use Twitter was identified as significant by Laurene
(interview, p. 6). It represented a commitment that required efforts for change. By every
measure, Twittering was a departure from the norm and the expected. The development
provided the developers “a context within which a particular aspect of the technology
was important to learn” (interview, p. 3). Opportunities for learning about the use of
technology and the use of technology for teaching and learning purposes were
deliberately structured into the process. Numerous meetings were devoted solely to
training the developers on using Twitter and the increase in their capacity to tweet with
the students (e.g., Appendix I). The conversations and deliberations focused on the
potential affordances that Twitter might offer.
Active learning also came about as the developers co-operated in the running of the prototyped program. Learning occurred when “[the developmental evaluator] modeled some ways of interacting with students on Twitter” (Laurene, interview, p. 9). In that scenario, a way of communicating with teacher candidates was illustrated to her by the developmental evaluator. When she reported on her reactions during one of the check-in sessions, it engendered learning among other members of the group.

Learning also transpired as a meaning-making activity when developers reflected upon their immediate experiences in engaging with the piloted program. A poignant observation was made during mid-course of the implementation phase that stuck with the developers even after the development had concluded. They noticed “the routines that [they] get involved in here as professors [were] so dependent upon students coming in and going out” (Laurene, interview, p. 7). During the time that teacher candidates went on field placement, the instructors experienced lapses in remembering to participate actively in the communities they themselves enacted and were responsible. Laurene revealed that she felt guilty in “letting go” (ibid). This illustrated the importance of gathering feedback from all stakeholders in a development context. Barriers to participation at the level of program officers could also impeded program success. In fact, Albert learned that “[twitter] was another thing added onto my already busy life, and he just wasn’t ready to manage that. I didn’t know how to manage that” (Albert, interview, p. 19).

Learning also resulted as a consequence to implementing the developed program. The developers experienced transformations in some of their core beliefs. For instance, Laurene, referring to her attempt at teaching with Twitter, “experienced people
interacting around a topic in a way that I hadn’t seen before” (Laurene, I, p. 20). She explained that: “I guess I’m not afraid […] of thinking about social networking as a potential way of communicating with students”. Albert reported a similar insight into seeing the potentials of different Web 2.0 tools.

**Values and valuing as a focus of learning in a developmental evaluation.**

Other significant foci of learning were of the values and assumptions that underpin the decisions made of the development. The act of valuing, i.e., the assignment of worth, is typically relegated to the evaluator in non-developmental evaluations. Data from the Assessment Pilot Initiative suggest that valuing was primarily achieved in collaborative settings. It was through various learning activities (to be discussed) that values were articulated and validated.

The determination of value was not of the conventional dichotomous types (“good/bad”). Rather, value determinations were made based on fit and adaptation. Questions related to fit included: in what ways is this idea fitting to the program context, and to what extent does it inform the problem? Questions related to adaptation included: e.g., what features or aspects of what is being examined can be borrowed and applied? Values remained implicit until surfaced. Developmental evaluation, through its data-informed processes, helped make values explicit. Once explicit, they could inform the decisions behind program design.

Analysis identified values about learning that were made explicit through developmental evaluation that shaped the developed program:

- Teachers as professionals who engage in continuous self-regulated learning
- A capacity for accessing the trustworthiness and validity of information about professional practice
- Helping teachers understand the complexity of assessment practices, as situated and contextual
- Seeing pupils’ experiences through the lens of assessment
- The benefits of engaging in reflection and learning from reflection
- Engaging in student-centered learning, e.g., being adaptive and responsive through “hearing how students think”

Some values were intentionally introduced into the conversations to shape the program design:

- Collaboration and co-creation
- Networked-learning

Some values shaped the conduct of the developmental evaluation:

- Collaboration
- Explicit alignment in the use of technology
- Responsive to teacher candidates’ experiences
- A program design that is adaptive to instructors’ abilities and readiness

To sum up, values were important sources of learning for the developers. Once they were made explicit, they informed the development process and the emerging development.

**Enacting learning activities in a developmental evaluation.**

Developers identified three kinds of activities as particularly useful for informing their developmental needs in learning. These included: 1) the conversations and
facilitated discussions that took place throughout the course of the development, 2) the
demonstration and training sessions, and 3) the summaries and road-mapping of the
process. The following describes each of these learning activities.

*Activities and facilitated discussions.*

Developers spoke highly of the evaluator-led activities and facilitated discussions
in the development for helping them learn and make sense of the development. The
developmental evaluator planned and executed several strategic activities throughout the
evaluations. These activities were intended as data-gathering sessions on their views
and perspectives and were meant to help elucidate the next steps in the development.
These discussions resulted in benefits beyond satisfying the immediate need for
collecting data for development.

The Appreciative Inquiry session on December 8 particularly illustrated this idea.
In the session, developers were asked not to focus on the deficits of the then-current
situation. Various prompts helped make explicit the current capacity and strengths of the
program. Developers were asked to provide input on which aspects of the program they
would like to see grow and how it might be achieved. Sticky notes were used to facilitate
data collection and processing during the activity (see Appendix L). Laurene explained
the learning outcome of the activity this way: “... when we finally got them grouped, um,
lots of things made sense and I think we had some good conversations around it”
(interview, p. 8). Activities helped uncover the “aspirations, doubts, skepticism and
cynicism” (Laurene, interview, p. 9) surrounding the development.

At other times, the evaluator facilitated other activities in the form of
demonstrations and discussions. It was often the case that the developmental evaluator
would “inject” new ideas and thinking into the discussion. Albert recalled that: “when Laurene and I had a question you were very good at the next time bringing some new ideas and new directions to think about” (interview, p. 7).

Conversations and dialogue were desirable because it matched the learning preference of one of the developers. Albert explained that he “learns by talking things out so I need to challenge my own thinking and I need to challenge other’s thinking because […] helps me clarify what they’re saying and it helps me clarify what I’m saying” (Albert, interview, p. 7).

These activities and discussions were not only instrumental to determine next-steps in the development, but they also contributed to generating deep knowledge about the underlying complexities that I argue as necessary for designing an innovation. Laurene remarked on the importance of having conversations with each other during a development process and how that led to articulating underlying values and assumptions. This led to a deeper mutual understanding of what the development could be.

All of those, I think, were really important professional conversations for Albert and I to have with people, with you and Tim, and wouldn’t have happened without that. One of the realities of working with someone for a long period of time is that you learn how to take shortcuts with each other, because we know we’re both so busy, so we know how to […] truncate conversations, we know how to distribute responsibilities to optimize what we can get accomplished in a short period of time, And when you’re doing that a lot, you don’t spend a lot of time just talking about the underlying assumptions and the aspirations and that kind of stuff that you’re carrying with you (Laurene, interview, p. 9).
**Demonstrations and training sessions.**

Another major source of learning was the demonstrations and training sessions planned by the developmental evaluator. Two separate sessions were held with the developers to introduce them to the use of Twitter, and to stimulate their thinking about how Twitter might be used for learning. These sessions utilized the principle of scaffolding to structure the learning. Dialogue was promoted as a way to facilitate learning. The ideas arising from these sessions served as a way to anticipate some of the challenges teacher candidates may experience as they try out Twitter for the first time.

**Summaries and reports.**

The last major source of learning was those related to the use of summaries and evaluation reports throughout the development (e.g., see Appendix O). During the planning/designing phase, summaries were produced of the ongoing development conversations. The developmental evaluator prepared these with ease-of-use in mind; ideas were condensed and then presented as a one-pager in most instances. During the implementation phase, summaries were produced of the ongoing development conversations, as well as of the dialogues transpiring virtually.

This tracking shaped both the development of the program and the implemented program. A specific feature of the implemented program was to leverage ongoing discussions on Twitter to form the foci of class discussion. Summaries were instrumental to that end.

**Roadmapping.**
Roadmapping was an evaluation reporting strategy employed within e-mails during the later phases of the development. A roadmap was organized under the following headings:

- What happened - reviews and summarizes the accomplishments of the week past
- What’s next - proposes some directions to the development
- How next – informs the developers what the developmental evaluator intends to do with the developers at a coming meeting

In some ways, roadmapping could be seen as analogous to sending out agendas. However, it was different in one significant way: the technique of roadmapping was intended to foster a collective understanding of the developmental demands and complexities that might need to be resolved in order to move forward. The summaries allowed the developers to “pick up on the issues that were really relevant to them at the moment” (Laurene, interview, p. 5). Like the chirping of hungry baby birds, these periodic updates reminded the developers of the ongoing nature of program development, and helped them prioritize their work. In that sense, she appreciated the work of the evaluator because he “would not let the importance of this project get buried under a table somewhere” (Laurene, interview, p. 9).

Unsurprisingly, summaries and reports were also instrumental in getting the discussion going in meetings. They served as springboards to conversations about the development. Albert explained that:
[Summaries and reports functioned as] a thinking piece more than anything. So, the handouts I think gave a good summary of the going challenges that we were having in implementing, but I think it was the talk about it afterwards and throughout that really helped clarify it for me and gave me more… ammunition, if you want, for thinking about next steps and where we go from there… but I don’t typically use those other than as talking points at the time when we’re discussing things. (interview, p. 14)

In closing, the need to resolve uncertainty was paramount to promoting the development. Analysis of the Assessment Pilot Initiative development records identified, and later confirmed through interviews, the role of leveraging learning and discovery to resolve uncertainty. This section described three forms of development activities that contributed to learning and discovery: 1) the conversations and facilitated discussions that took place throughout the course of the development, 2) the demonstrations and training sessions, and 3) the summaries and roadmapping of the process.

Revisiting Research Question Four: What insights, if any, can be drawn from this development about the roles and the responsibilities of the developmental evaluator?

Research into program evaluation has typically placed great emphasis on the evaluator as an agent in leading an evaluation from conception through to completion. Evaluators, as experts in the field of evaluation, are typically responsible for all aspects of an evaluation. Recent approaches to evaluation, like participatory evaluation and utilization-focused evaluation, have drawn attention to the roles and responsibilities and
places particular to the evaluator to involve stakeholders in ways that enhance evaluation quality.

Like conventional program evaluation, the literature on developmental evaluation has casted the developmental evaluator in the spotlight. Most lessons learned about effective evaluation practice, and strategies and techniques for involving stakeholders is transferable into the developmental context. Unlike conventional program evaluation, where the development context is complex and the objective is developmental, the roles and responsibilities of the developmental evaluator and their relative emphasis are likely to be different. This question explores the roles and the responsibilities of the developmental evaluator in the Assessment Pilot Initiative.

The developmental evaluator’s role in the Assessment Pilot Initiative was multifaceted as evident by the various roles discussed in the preceding sections. His approach to the developmental evaluation displayed characteristics of fluidity as he responded to the developmental demands of the development. The data identified four major roles and responsibilities to have most contributed to the development process. This finding was gleaned from the self-reported data from interviews and confirmed against the development data through document analysis. In order of increasing importance, the four roles and the corresponding responsibilities of the developmental evaluator were:

1. Project Manager – for managing the progress of development;
2. Facilitator of Learning – for enacting a learning framework and promoting learning opportunities;
3. Evaluator – for lending evaluative thinking expertise and wisdom in understanding social programs and theories of change; and,
4. Innovation Thinker– for sharing the burden in realizing the vision by adding to
team a capacity for thinking about developing innovating.

The following section describes each role in detail.

**Project Manager.**

The developmental evaluator held a substantial role in managing the development progress. Under the umbrella of program manager, he performed various functions. At the simplest level, he assumed a stewardship role in the development activities. He worked to disseminate knowledge and mediate the translation of educational literature to the local context and helped interpret and make sense of observations from activities data. He facilitated the flow of information and knowledge.

He was also the hub of communication. He kept the developers and program participants abreast of all activities via e-mails. He established communication strategies, such as the use of action tags in e-mail subject lines, to help the developers screen out relevant e-mails and prioritize actions. This helped promote a collective understanding surrounding the development, an important precursor for enabling learning.

He was responsible for managing a critical resource: time. The amount of time that could be spent on developing the innovation was limited by other competing demands on the developers’ professional lives. The B.Ed. program also imposed time limitations that needed to be kept at the forefront of the development. As an indication of how precious time was during the development period, two of the developers lamented over the limited time they had with the candidates and wished that they had more time to participate more deeply with the teacher candidates (Albert, interview; Tim, interview).
The timing of the meetings was a deliberately determined by the developmental evaluator by striking balance between the needs of the development and what developers could reasonably contribute.

Lastly, he managed the online community. He was responsible for monitoring the Twitter account and the discussion board space. He monitored the activities over the virtual space and reported on it.

For the work that he assumed under this role, he was credited for not “letting [the project] get buried under the table” (Laurene, interview, p. 9). Under this sense of project management, his responsibilities were (a) to keep the development progress moving forward by balancing the various demands, resources, and expectations among the stakeholders; and (b) to respond reflexively to the various emergent needs that arise during development so to maintain the momentum forward.

**Facilitator of Learning.**

Besides managing the progress, facilitation of learning was found to be a significant role of the developmental evaluator in the Assessment Pilot Initiative, as already discussed in section two of this chapter. Drawing on his training as an evaluator, he facilitated various activities and discussions that promoted learning about the problems underpinning the vision, the skills, and capacity needed to develop the innovation, and the skills and capacity needed to implement the innovation. These learning allowed the developers to think more deeply about the innovation and the program that they were designing and respond in ways that was consistent with their values.
Intentionality was an important feature of the learning. The developmental evaluator sought opportunities to structure learning opportunities. Perhaps more importantly, he structured a learning environment that promoted learning from and about the activities of the implemented program. He monitored and assessed the level of learning that among the developers and that shaped the development process and the developed program.

Evaluator.

Much value was placed on the developmental evaluator to provide evaluation expertise to the group. Although the developers were themselves accomplished evaluators, they found it beneficial to engage fully with the development as a participant (Laurene, interview). However, his role was not entirely straightforward. Evidence suggests that the development called for skills and knowledge that might otherwise be classified as peripheral in strict program evaluation sense. Specifically, the developmental evaluator drew from his knowledge of program theory to help with the innovation.

One important aspect of his roles as an evaluator was to encourage the group to think about what constituted success and what reasonable indicators could be used to infer success. The heuristic used was: is this what we want? Is so, how do we know if we get there? If not, what else do we need to think about, and what else might we consider? Along that line, the evaluator’s understanding of program theory (e.g., logic modeling and constructing program theories of action) was critical conceptualization of the development and in the design of the program design (Appendices M, N).

Innovation Thinker.
The most important role of the developmental evaluator in the Assessment Pilot Initiative were those functions associated with being an ‘innovation thinker’. The developers valued the developmental evaluator’s substantive expertise in using technology. He was able to talk about things that the program staff did not know about (Tim, interview), and “demonstrate ideas, talk about them, look at their potentials and discuss how they had potential” (Albert, interview, p. 13). Having some expertise in innovation and some capacity to think about innovation was perhaps as equally important as bringing some content expertise. To promote the development, the developmental evaluator assessed the instructors’ capacity with Twitter at various points of the development, and inferred their readiness and capacities for change. This led the developmental evaluator to respond in ways that actively shaped the direction of the development. Underneath this notion was the ability to comprehend and visualize change.
Chapter 7: Conclusion

Chapter Overview

This case study reported on the use of developmental evaluation for innovating the learning experiences of teacher candidates in a context of complexity. Neither the means nor ends to achieving the vision were known at the beginning of the engagement. Through continuous engagement and participation in a process informed by developmental evaluation, a program came to be piloted and implemented with a group of teacher candidates.

Captured in the program development data were descriptions of the development progress and its surrounding complexity as the developers engaged in development amidst a context plagued by uncertainty. Also captured in the data were descriptions of how the development had unfolded within a developmental evaluation framework. Follow-up interviews with the developers helped identify episodes in the development that they deemed as significant to the progress.

Analysis of the case data revealed the importance of facilitating learning in a developmental evaluation to enable development. It highlighted the multiplicity of roles and a developmental evaluator adopted to facilitate the process. The data also converged to suggest six different foci of development that could serve as the basis of future development efforts where the object is innovation. These findings were discussed in relation to the literature and they were found to be generally consistent with how developmental evaluation was described.

What has emerged from this case study is the notion of design. In this chapter, I reflect on the possible implication of seeing program development through the lens of
design as a way of reconciling some of the functions not typically associated with
developmental evaluation but was performed in this developmental evaluation in order to
establish a program model. Finally, I consider the implications of my findings to the
evaluation community.

**A Case of Split Identity: Evaluation or Development?**

A lingering issue that I kept returning to throughout data analysis and reporting
was how best to situate this case study. Was this a case of program evaluation or was it a
case of organizational development? The space that this development operated in—the
development of a novel program (i.e., innovation) from scratch to address a social need—
or in Patton’s words, the “preformative development of a potentially scalable program”
(p. 306)—is one that is relatively unexplored in the evaluation literature.

Elements of program evaluation were instrumental to the development of the
innovation studied. Not only did the theory of developmental evaluation provided the
basis of guiding the development process, I’ve argued in the previous chapter that
evaluative thinking undergirded the entire development process. Being able to evaluate
the quality of the emergent ideas and resulting program during implementation and
responding adaptively, but also being able to evaluate the emergent ideas during all
phases of the development process were critical to shaping the program model.

At the same time, the case study highlighted many emergent properties of the
development that do not yet fall under the vernacular of the current conception of
developmental evaluation. For instance, I’ve discussed the role of the developmental
evaluator as a person who (co)develops a program and draws upon substantive skills in a
knowledge domain to guide the innovation process along.
I find relief in knowing that Patton (1994) also grappled with the same issue of placing developmental evaluation almost two decades ago. He recognized that his colleagues might contend with the legitimacy of the developmental evaluation examples he cited as evaluation or not. He wrote: “One reaction I’ve had from colleagues is that these examples aren’t “evaluation” at all but rather organizational development efforts. I won’t quarrel with that.” (Patton, 1994, p. 316). He argued that much of what he did and was perceived to have done with his clients were evaluative in nature. On that basis, he argued that the cases he cited qualified as evaluation and so arose the field of developmental evaluation to describe the unique circumstances, contexts, and demands that development calls for.

In line with Patton’s findings, the findings of this case suggested that much of the development occurred through evaluative means. The data-informed nature of the innovation process was facilitated by developmental evaluation theory. The facilitative and participatory/collaborative approach to developing the program model was informed by utilization-focused and participatory evaluation approaches. Instead of insisting on classifying such case as evaluation proper or not, I find it a more worthwhile exercise to consider the potentials that such evaluation use opens up and its potential implication to the field of evaluation and those it serves. Regardless of the classification assigned this case study, there exists a space in social programming that innovation is called for, and to which the evaluator could respond. I wish to dedicate the remaining of this chapter to this exploration.
Developmental Evaluations in a Context of Innovation

This case study drew attention to the process of developing a program from scratch. It shed light on the processes, concerns, and roles particular to a context as such. Traditionally, this sort of development is referred to planning processes that operate outside of the sphere of evaluation. It is a space where program evaluators do not typically stake a claim. Historically, the power of evaluation was in making judgments about the merit, worth or significance of programs already in operation or completed. As argued in this thesis, such context is not the only one to which evaluators respond. Moreover, evaluators have moved away from making recommendations from evaluation findings in recognition of the rich situated and contextual knowledge that program officers hold that evaluators may not be privy to. Implicit in this view lies a recognition that knowing the quality of a program is not the same as knowing how to innovate a program. Evaluation findings serve to inform the thinking of those responsible for the program; evaluation use becomes paramount to that end. Program evaluators, instead, offer their expertise and services when innovating to facilitate program change.

This case study illustrated how program evaluation supported a creative process in which evaluation played not only a reactive role in actively shaping the program that was, but a generative role in shaping the program to be. The data-informed approach to development enabled innovation.

Design as an Activity and Function unto itself.

Beneath the notion of planning lies the assumption that the act draws upon what is known and available. Whether one is planning for instruction or for a course of medical treatment, options are (mostly) readily apparent and its value or risks already determined,
resulting in the determination of a most-plausible course of action based on some meaningful criteria (e.g., most economical, least risk, maximum gain, etc.). What is uncertain in planning is the outcome. Patients may progress for the worse despite appropriate treatment; an otherwise successful employment assistance program elsewhere may yield little impact on its users at a new site. Evaluation becomes powerful tools for determining the merits, worth, and significance of a program, and in uncovering the unintended consequences that arise.

The case findings suggested a kind of development that was not typical of planning. Development instead occurred in a context of uncertainty. Options concerning which technology to adopt were uncertain. The means to evaluate the options were uncertain. The indicators of success were uncertain (Laurene, interview). The notion of “knowing” became critical in this process. What do we know? How do we know? When do we know we know? Decisions were made about the developed program based on the best available knowledge, an incomplete set of knowledge at best. These characteristics typified the innovation process.

I wish to qualify the kinds of development processes and activities associated with innovation described in this case as design. Design thinking, then, denotes the kinds of knowledge, skills, and capacity to lead change using design such that the output of a design process is innovation. Applying this perspective to developmental evaluation may help to reconcile the unique peculiarities found in this case.

Indeed, the notion of design is not novel and should not be foreign to evaluators or social scientists. Evaluators make decisions about evaluation designs concerning the methodology, timing, reporting mechanisms, etc. Social scientists designs research all the
time. In fact, Patton refers to design in his writing on developmental evaluation (e.g., Patton, 2010, p. 99; 1994, p. 313). However, the notion of design thinking as a problem-solving methodology is novel to the field of evaluation.

Patton and Patrizi (2011) reminded us of the importance to be well-read in the developments of other disciplines. They wrote:

As an interdisciplinary and transdisciplinary field, it is important for evaluation to keep abreast of developments in the larger world. What are leaders talking about? How do they conceptualize the challenges they face and the value-added they bring to their organizations? (p.7)

It is in that spirit that I continue with this discussion.

**Design Thinking and Thinking Evaluatively.**

While the notion of design has come to be associated with the tangibles, e.g., products design or industrial design, and the aesthetics became an end to design, scholars and thought-leaders in businesses have recently drawn attention to the application of design and design thinking as a way of problem-solving and innovating. Design and design thinking is promoted in the business field and is now taught in business school at places like Stanford University and the Rotman School of Management at the University of Toronto. Tim Brown contributed a paper on design thinking to the Harvard Business Review in 2008 that described design thinking as:

a methodology that imbues the full spectrum of innovation activities with a human-centered design ethos. By this I mean that innovation is powered by a thorough understanding, through direct observation, of what people want and
need in their lives and what they like or dislike about the way particular products are made, packaged, marketed, sold, and supported. (p. 8)

He suggested to business leaders who seek innovation to relinquish the norm and the accepted, and start with a deep understanding of people and behavior, and take this insight to shape the products, services, and business models.

Roger Martin, Dean of Rotman School of Management at the University of Toronto, published a book entitled *The Design of Business*, in which he described design thinking as an exercise of validity rather than of reliability. Innovation as reliability-seeking exercises promote change using existing paradigms and thinking, which is often faulty in contexts that desires change. Such change exercises can at best result in improvement. Business decisions made using extensive market research, inferential statistics and various research-derived methods promote reliability-seeking behaviours that often do not produce the leaps that typify innovation. The resulting culture is one that is risk-adverse and “stuck” in a paradigm. True innovation dispenses with preconceived notions and assumptions about the world, and engages in validity-seeking activities in order to design an innovation.

In the literature, design-thinking is posited as a methodology that proceeds through three necessary “spaces” that forms a cyclical process: 1) inspiration, 2) ideation, and 3) implementation (Brown, 2008; Brown & Wyatt, 2010). A starting point in design thinking is the inspiration space, in which “the problem or opportunity that motivates people to search for solution” is given shape through a design brief that describes “a set of mental constraints” that frame the problem. Designers then engage in humanistic observations such as direct observations and field-based interviewing to emphatically
understand human behaviours. Ideation builds upon the insights gained from inspiration and seeks to generate as many ideas and solutions as possible. The third space, implementation, is when “the best ideas generated during ideation are turned into a concrete, fully conceived action plan” (Brown & Wyatt, 2010, p. 36). The idea is prototyped and then “tested, iterated and refined” (Brown & Wyatt, 2010, p. 36).

Returning now to the context of social programing and program evaluation, what are we able to learn from the application of design thinking in business? How might it be translated to a context where the currency of social program is not valued in dollars but complex social needs/change like learning and health?

Under research question two, I described six foci of development that imbued the Assessment Pilot Initiative efforts. These foci were moments in the development where the complexity of the underlying ideas and concepts were unpacked and examined that resulted in some transformation in the groups’ thinking about the problem. To reiterate, these six foci were:

1) Defining;
2) Delineating;
3) Collaborating;
4) Prototyping;
5) Illuminating; and, finally,
6) Evaluating.

These six foci resembled the design thinking “spaces” that Brown described. I contend that is not a coincidence but perhaps point to the innate nature of designing. What the data captured was perhaps innovation as a design exercise. If that was the case, this above pattern lends support to the use design thinking methodology in a social
science context. The six development foci might have captured the design process of an innovation.

The case illustrated the highly dynamic and iterative nature of the process. The case actually began at the stage of delineating with the evaluator working independently and collaboratively with the developers to clarify what it was that they sought to innovate. As their vision became clarified (stage: delineating), the conversation spilled over into informing the understanding of the conundrum (stage: defining)—the underlying complex problem or issue, as a manifestation of a social need that the innovation tackled. In this case, the conundrum was a learning and educational issue in developing assessment capacity among teacher candidates. The design issue then became delineated, which was different from the conundrum. The design issue was to explore how developers could develop meaningful connections with teacher candidates that foster an understanding of the complexity of assessment practices. Necessary to moving the development forward, ideas and decisions surrounding and concerning the exchange of knowledge had to be negotiated (stage: collaboration). A program design, which outlined the relevant theoretical pieces, that informed the implementation and piloting of the Initiative was created based on the insights, desires, and options explored (stage: prototyping). Emergent issues arose, such as the reconsideration of whether Twitter could support deep learning via 140-character tweets, led to adaptive responses in the program, among others, by instating a discussion board to be used in an adjunctive capacity (stage: illumination). This in turn influences the conception and understanding of the underlying conundrum the innovation solves (stage: defining). Developmental evaluation helped
illuminate issues surrounding the program-in-action, as well as the developmental process in the form of the data-informed process (stage: evaluation).

The above description was necessarily a gross simplification of what was previously already labeled as the complex development process, and now, a design thinking process. Rather than being mutually exclusive, the two processes, it would appear to be referring to the same underlying process.

If design thinking is the engine that powers innovation, it bears emphasis that evaluative thinking becomes the fuel that drives innovation. Thinking evaluatively and being evaluative provide for the validity-seeking functions that Martin (2010) described. It helped align the program-in-development with the best-possible conception/understanding of the conundrum to the teacher candidates it served.

I contend that the dynamics of the design process described was unique to this case, but that the process would generally proceed (non-linearly) through those six foci of development in other innovation contexts. This developmental evaluation spent much time on defining the conundrum and delineating the design issue. A second iteration of the Initiative might instead focus more on illuminating and evaluating.

**Defining Design in the Context of Developmental Evaluation.**

To sum up, I wish to offer the following definition of design as a synthesis of its potential significance to the field of developmental evaluation:

Design is the systematic exploration into the complexity of options (in program values, assumptions, output, impact, and technologies) and decision-making processes that results in purposeful decisions about the features and components
of a program-in-development that is informed by the best conception of the complexity surrounding a social need.

Design is dependent on the existence and validity of highly situated and contextualized knowledge about the realities of stakeholders at a site of innovation. The design process fits potential technologies, ideas, and concepts to reconfigure the social realities. This results in the emergence of a program that is adaptive and responsive to the needs of program users.

Roles of Developmental Evaluator and Knowing Innovation

The findings of the case have also drawn attention to the multiplicity of developmental evaluator roles when working on development in an innovation context. I have described the four major roles of the developmental evaluator in the Assessment Pilot Initiative as being: 1) a program manager, 2) a facilitator of learning, 3) an evaluator, and 4) an innovator.

Unlike Patton’s positioning of the evaluator, the evaluator functioned in this case as both the evaluator and the leader of the development. This happened due to necessity, it might be a worthwhile question to ask whether this arrangement of the evaluator as the team leader confers any benefits.

These finding are mostly consistent with that of the literature. Patton (2010) described the developmental evaluator as a *bricoleur*, “a jack-of-all-trades do-it-yourself person who draws on eclectic traditions and integrates diverse approaches to get the job done usefully” (p. 19). Beyond the major roles reported above, the analysis revealed other minor roles of the evaluator: a worker, a teacher, a communicator, an observer, a researcher, and now, a designer. The discovery of these minor roles provides support for
the jack-all-all-trades traits that Patton described of developmental evaluators. In ways similar to the case findings, Dozois, Langlois, and Blanchet-Cohen (2010) drew attention to the developmental evaluator as a strategic thinker, relationship builder, and a one who practices servant leadership. What is revealing from this case study is the emergence of two novel roles of the DE as a designer and an innovator.

The role of the designer-evaluator has been alluded to in the above section; a designer helps structures the innovation process using design thinking, while the evaluator provides the necessary evaluative thinking and skills to help collect and interpret data. The emergent role of the innovator-evaluator is an exciting role for an evaluator to play in this space.

Following Patton’s suggestion of using innovation as a sensitizing concept, the innovator-evaluator’s role might be to “know innovation”—knowing how innovation occurs, and how it can be facilitated—to help the group realize their vision. The innovator-evaluator’s contribution might be to infuse the innovation process with data to enable decision-making. To know what kinds of data might be worthwhile to procure, this requires 1) a sensitivity to the context (e.g., knowing the dynamics of teacher education at a specific site), 2) an understanding to the nature of the issues at hand (e.g., the use of educational technology), and 3) being attune to the role of data and how it might enable development (e.g., the style and preference of the developers; this relates to the negotiated nature of collaboration that gets identified in the design process). To use a systems thinking metaphor, an innovator-evaluator understands the context deeply enough to see how components of a system could be reconfigured in ways to generate
innovation. Part of this may rely upon having content expertise in the contexts that the
developmental evaluator is working in.

This case study in the developmental evaluation of an innovative project in
teaching and learning about assessment education illustrated the benefits of the evaluator
having substantive knowledge in teacher education and educational technologies in
enabling development and innovation. The developers, as reported in the previous
chapter, found it beneficial that the evaluator could draw upon this knowledge to help
inject new ideas into the process. Having his aptitude helped the developers visualize the
potentials. Substantive knowledge about e-learning and educational technologies
contributed to the discussions and shaped the innovation. Relating this innovator-
evaluator role to design thinking, an innovator who “knows innovation” may help design
a learning process for the developers to learn about how the technology. This in turn
allows the developers to visualize the potentials and adapt a technology in ways that
could be used effectively in the local context.

Role of the Developmental Evaluator as a Facilitator of Learning

Another significant role of the evaluator to have played in the case was to be a
facilitator of learning. Previous works in developmental evaluation have drawn
importance to this role. Dozois, Langlois, and Blanchet-Cohen described the
responsibility of a developmental evaluator as having to establish a learning framework
for development (2010). A learning framework “maps the key challenges and
opportunities, identifying 1) what the group needs to pay attention to as they go forward; and 2) what they need to be learning” (p. 31). Applying this notion to the context of the Assessment Pilot Initiative, it would be difficult to map out key challenges and
opportunities with any certainty or specificity due to the highly complex environment. Any attempts at mapping out things-to-watch-for would only be too general to be useful. I am reminded of the saying: You only know what you know, and you do not know what you do not know. The extent of knowing is limited by the one’s current capacity of knowing.

What appears to be helpful, then, as suggested by the case finding, is to facilitate an environment of learning that allows for learning incrementally. Given that developmental evaluation already provides a data-informed approach to development and design thinking provides the direction of learning, the evaluator’s role might be to focus on learning that helps resolve the most immediate roadblock in development. It appears most important to allow for serendipitous learning to emerge; learning environments that promotes discovery, experimentation, and collaboration might be most conducive to this kind of learning. Instead of training the group to pay attention to the learning process as suggested by Dozois, Langlois, and Blanchet-Cohen (2008), it might be more fruitful for the evaluator to take on that role, to sense what has been learned serendipitously, and reframe it to the group and help interpret its meaning. Facilitating opportunities for learning that shapes the developers’ skills, capacity, and cognition helps them to think more deeply about their issues and potential solutions and overcomes their limits of knowing. Put this way, the role of the developmental evaluator might better be labeled as a learning designer. Reframing the developmental evaluator as a learning designer helps to emphasize the strategic and intentional nature of his work in facilitating learning.

In a similar vain, Coffman and Beer (2011) described the use of “strategic learning [as] the use of data and insights from a variety of information-gathering
approaches—including evaluation—to inform decision making about strategy” (p. 1). While their notion of strategic learning does not relate directly to developmental evaluation, the common thread of learning cannot be dismissed but be seen as an emerging focus that evaluators are increasingly concern about.

**Reflections on Conducting Case Study Research**

This experience has taught me much about conducting case study research. I found case study methodology to be a valuable tool in conducting research on evaluation to interrogate the inner intricacies of the decisions and processes of a case. This case highlighted for me the difficulties of conducting *in-situ* research where the context of research prohibited research interventions. As a case study researcher, I had to accept the case as it was in all of its messiness.

This case study utilized two primary sources of perspectives to inform the study: those of the developers and the documented records of program development. Each served a different purpose to the analysis and together they helped strengthen the trustworthiness of the study. It became increasingly apparent during analysis that the voice of the developmental evaluator was missing. Given the importance that program evaluation theories place on the evaluator, it was important that the thinking of the developmental evaluator be articulated.

In this research, I assumed the dual roles of being a case researcher and the developmental evaluator of the case. As I analyzed the data, memories came flooding back as I revisited the various developmental episodes. I recalled specific memories, intentions, and emotions. I had always known the importance of reflexive and reflective practice in evaluation. However, I kept little of those records because I experienced much
difficulty during the developmental evaluation to articulate what I was feeling. But as I revisited the documents, I began to be able to identify and articulate explicitly those experiences. The documents served to elicit and access otherwise trapped knowledge. In that regard, I was fortunate to be able to have that perspective revived.

I may expect some objections to this practice for concerns of reliability. I wish to offer that what may be lost in reliability is gained in validity. Having the researcher conduct a self-study into the contributions of the developmental evaluator allows the researcher to access a perspective in such depth and transparency that would be otherwise be difficult to collect and have its complexity represented using any other research methodologies.

This may be an acceptable practice when one recognizes that the influences of the developmental evaluator is limited insofar as evaluation participants are, in fact, co-participants in decision-making processes. Developmental evaluators can influence, but do not in large part shape the program on their own. Examining the extent that recommendations made by an evaluator are translated into action would make an enlightening study.

This research also reinforced for me the value of being adaptive as a case study researcher. One of my original curiosities was to trace the evolution of the program logic. I wanted to explore how ideas take shape and give rise to subsequent ideas. What I had discovered instead was that the case data lacked any real power to suggest how that could have been done. To be able to do this, a research design might ask the developers to keep individual accounts of the development experience throughout different points in time. But such a request made of the developers need be balanced with concerns of feasibility.
Engaging in case study research is time consuming as one confronts the complexity and multiplicity of data needed to make a case. Analyzing the case data was an all-consuming act; attempting to recreate the complexity of the case for presentation in any coherent manner was a struggle. Nevertheless, case study research remains a powerful indispensable tool for evaluation researcher.

**Conclusion: Innovation and Developmental Evaluation as Design Thinking**

To sum up, this chapter offered to reconcile the unique demands made of developmental evaluation and on the developmental evaluator when working in a context of complexity and where the object of the developmental evaluation is innovation. The notion of design was offered as a lens to understand what was observed in the Assessment Pilot Initiative. The notion of design thinking was introduced and its potential application and implications were explored in relation to enabling innovation using developmental evaluation as a framework.

The concept of design thinking holds great promises and implication for developmental evaluators working in the space of innovation. As a systematic process, it serves a generative function that is unrestrictive and sufficiently structured to complement the evaluative function that developmental evaluation already provides. Developmental evaluators may find value in consulting the literature on design thinking to exploit the creative potential it offers when implemented in a developmental evaluation. Used as a sensitizing concept, design thinking and innovation together provides powerful way of understanding the process of innovation. When design thinking and developmental evaluation are interfaced in an application, developers may stand to
benefit from a rich development experience in ways that help bring an innovation to fruition.
Implications to the Field of Evaluation

This study provided one of the first empirical documentation and examination into the emergent processes and dynamics of a development (the what and how of development) when the theory of developmental evaluation was adopted to undergird and shape a process of innovation. This research was an attempt to bring to bear the insights gained from all those engaged in this developmental evaluation in an attempt to generate “practical knowledge” (Schwandt, 2008), the important of which was discussed in the introductory chapter.

Going in to the developmental evaluation, I had with me a surface understanding of the principles and practice behind developmental evaluation. As a “theory”, it is important to make the descriptive/prescriptive distinction of developmental evaluation, as Alkin and Christie (2004) had of evaluation “theory” generally, as was discussed in chapter three of this thesis. Developmental evaluation offers a set of descriptive ideas, that is, a conceptual framework for approaching evaluations, rather than a set of prescriptions that predetermines a set of steps to be completed or procedures to be followed. It was difficult to appreciate the full depth of the complexity of the issues described in the literature until I had experienced it.

Coming out of the evaluation and the experience of systematically reflecting and analyzing it are instantiations and illustrations of developmental evaluation practice. Table 5 summaries the key lessons learned from this work. Contrary to theory-testing work that typically affirms or rejects aspects of a theory, the findings in this table should be viewed as extending the theory by adding a complimentary dimension of practical knowledge.
### Table 5

**Summary of Key Learning “Going In” to the Evaluation, and Key Insights “Coming Out” of the Evaluation**

<table>
<thead>
<tr>
<th>Key learning that informed the developmental evaluation</th>
<th>Key insights arising from the research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsibilities of a developmental evaluator: Patton, 2010</td>
<td>Undergirding the development of the Assessment Pilot Initiative with developmental evaluation gave rise to a techno-centric approach to teaching and learning using Twitter. This approach was subsequently piloted with a group of teacher candidates (program users).</td>
</tr>
<tr>
<td>1. Tracking and documenting the nature and results of program development</td>
<td>By framing various aspects of the emergent development as data, the developmental evaluator facilitated a mapping of the terrain, clarification of the immediate program context, and made explicit the values and assumptions underpinning the development. Such explication clarified clients’ intentions that were subsequently incorporated into the developed program logic.</td>
</tr>
<tr>
<td>2. Facilitating systematic data-based reflections</td>
<td>Developmental evaluation promoted program development, but also promoted continual program evolution in response to users’ behaviour and feedback.</td>
</tr>
<tr>
<td>3. Frames the developmental context and evaluation in terms of complexity thinking and systems thinking.</td>
<td>The developmental evaluator enacted a learning framework that helped build clients’ capacity in realizing their vision by learning about Twitter and by addressing the uncertainty that accompanied development.</td>
</tr>
<tr>
<td>Gamble, 2008</td>
<td>The developmental evaluator worked closely in collaboration with the clients in support of their efforts at innovating. Engendering trust emerged as central to the client-evaluator relationship.</td>
</tr>
<tr>
<td>1. Accompanying the group to facilitate development</td>
<td>Developmental evaluation gave rise to a non-linear, co-dependent development process that centered around six foci of development:</td>
</tr>
<tr>
<td>2. Collecting data to validate decision and documenting the innovation process</td>
<td>1. Definition</td>
</tr>
<tr>
<td>3. Framing and reporting the data</td>
<td></td>
</tr>
<tr>
<td>4. Developing strategy and aligning programmatic efforts to strategy</td>
<td></td>
</tr>
<tr>
<td>5. Developing indicators that signal progress or divergence</td>
<td></td>
</tr>
<tr>
<td>Dozois, Langlois, and Blanchet-Cohen (2010)</td>
<td></td>
</tr>
<tr>
<td>Five Uses of Developmental Evaluation (Patton, 2010)</td>
<td>Empirical evidence in support for preformative development of a potentially broad-impact, scalable innovation as a use of developmental evaluation</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1. Ongoing development</td>
<td></td>
</tr>
<tr>
<td>2. Adapting effective principles</td>
<td></td>
</tr>
<tr>
<td>3. Developing a rapid response</td>
<td></td>
</tr>
<tr>
<td>4. Preformative development of a potentially broad-impact, scalable innovation</td>
<td></td>
</tr>
<tr>
<td>5. Major systems change and cross-scale developmental evaluation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Roles of a Developmental Evaluator when Innovating (none specified)</th>
<th>• Innovation Thinker</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Project Manager</td>
</tr>
<tr>
<td></td>
<td>• Evaluator</td>
</tr>
<tr>
<td></td>
<td>• Facilitator of Learning</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Legitimacy of the developmental evaluator in bringing to bear substantive knowledge in a developmental evaluation</th>
<th>The developmental evaluator brought to bear his understanding of e-learning and knowledge of social media to recommend technological solutions that might be fitted to the context. The developmental evaluator then engaged the clients to flesh out the suitability and feasibility of the solution. This source of expertise helped shape the development in a co-creation process.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patton, 2004</td>
<td></td>
</tr>
<tr>
<td>• Unclear whether such influences from the developmental evaluator are legitimate; implications unexplored.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tools of Developmental Evaluation (Patton, 2010)</th>
<th>Clients attributed much of the success in development to various features of the enacted developmental evaluation:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamble, 2008</td>
<td>• Structured activities and facilitated discussions to clarify intentions and negotiate next-steps</td>
</tr>
<tr>
<td>• Applicative Inquiry</td>
<td>• Demonstrations and training sessions to learn about specific aspects of the technology</td>
</tr>
<tr>
<td>• Rapid recon</td>
<td>• Summaries and reports that helped the clients stay abreast of the fast-changing program environment. This provided the ammunition to help them interpret and make sense of the evolving context to react accordingly</td>
</tr>
<tr>
<td>• Revised and emergent modeling</td>
<td>• Roadmapping referred to a communication strategy employed by</td>
</tr>
<tr>
<td>• Network mapping</td>
<td></td>
</tr>
<tr>
<td>• Visual language</td>
<td></td>
</tr>
</tbody>
</table>
the developmental evaluator to help focus the evaluation by suggesting a trajectory of the development.

- Appreciative Inquiry exercise was critical in making explicit the values and intentions of the development.

Having participated throughout this arduous journey of leading a developmental evaluation and subsequently researching it, I become increasingly convinced of its potential and worth as one of many evaluation options evaluators can provide. I am reminded of the social mission of program evaluation as an exercise that seeks social betterment for those that a program serves.

By some accounts, we now live in a tumultuous society where the current way of living is viewed to be broken to many. Our health care system is strained; our education systems failing; and the economy remains unstable despite interventions. Pressures are felt at all levels of social endeavours to change to operate in leaner, more efficient, and more effective ways. Innovation is no longer a luxury but a necessity for program officers.

As demonstrated in this case study, uncertainty can be a disabling emotion and an impediment to change. The value in employing a systematic framework like developmental evaluation lies in not only recognizes uncertainty as disabling, but, as demonstrated, it may also serve to address and confront it directly. Through exposure to new ideas, evaluative thinking helps resolve uncertainty and elucidates the path ahead.

As an approach to program evaluation, we are only beginning to discover the potentials of data-informed frameworks like developmental evaluation is enabling social innovations. There is much to be learned and much to be excited about. In this remaining paragraph, I wish to offer some thoughts concerning future directions for research.
Future Directions

As more developmental evaluations are conducted and researched, cross-case analysis could be conducted to further understand the intricacies of doing developmental evaluation and to begin to identify patterns of practice. Several issues will be of interests to developmental evaluator, program users, and development funders. First, I have argued for the developmental evaluator to actively shape the design of a development. What sort of expertise might this require? How might issues of validity play out in this context? How might the development team know that their design decisions are likely to be valid and sound before implementation? Second, how do program evaluation standards translate into a context like developmental evaluation, where it can be difficult to even articulate the object of the evaluation/development? Third, how might theoretical frameworks be integrated into a developmental evaluation? How might it inform the developers’ understanding of the underlying problems and issues in ways that result in a stronger (more valid) innovation? Finally, how might issues of quality in developmental evaluation be evaluated? How might one go about conducting a meta-developmental evaluation? What criteria might one use? Future research in developmental evaluation might address these valuable questions.

Final Words: A Metaphor for Understanding Developmental Evaluation Use in a Context of Innovation

As a teacher, I have long appreciated the explanatory powers of metaphors. In this epilogue, I wish to offer evaluators who are interested in developmental evaluation the metaphor of piloting as a way of understanding the unique demands developmental evaluation make of the developmental evaluator and its developers.
One of the great joys of piloting a plane is to travel through the air at some heights above ground. As a pilot takes off, structures and buildings that were once towering become minuscule. The pilot asserts confidence as he pulls on his control column to continue to climb. Once he reaches cruising altitude, he levels off the plane and adjusts the engine power to achieve the desirable cruising speed. At a height of 8000 feet, a typical cruising altitude for small private planes, such a vantage point affords the pilot breathtaking scenery. To fly towards a destination, the pilot relies on his flight plan, his aeronautical navigation charts, and his weather briefings. Because weather can sometimes be unpredictable, the pilot may unavoidably have to confront inclement weather despite reasonable planning.

To stay safe in inclement weather, the pilot recalls the importance of incremental adjustments; drastic movements could prove disastrous. He continues to navigate, looking outside of his windows for landmarks and matching them to those on his maps. He communicates on the radio with other planes in the air and with control towers to inform them of his presence and intentions. Of course, he continues to aviate and fly his plane ever so carefully. Importantly, the pilot relies on his well-rehearsed routines to help him pilot safely to destination.

**Aviate, Navigate, Communicate.**

There are surprising similarities between piloting and developmental evaluation. Like a pilot, the developmental evaluator invites his clients along on a journey of discovery of the unknown. In both scenarios, there is a beginning and an ending to the journey. The developmental evaluator senses the progress of the development. He seeds ideas and cultivate an environment for learning. The developmental evaluator oversees
the process by lending a birds-eye view to the process and plans a trajectory to the
development. The evaluator continuously gathers data to inform him and his team of the
progress. Like the pilot who relies on his gauges and instruments, developmental
evaluators collect data from a program-in-action from which inferences and decisions
could be made. He asks the following questions with the developers: are we on track, and
how do we know if we are on track or not?

When a developmental evaluator enters a space of heightened uncertainty like
innovation, like the pilot entering incremental weather, he makes small but incremental
decisions and react adaptively to what he senses. He continues to navigate the
development, again in small steps, and react accordingly. He may get sidetracked along
the way, but he relies on the data he’s receiving to help him get back on track. Finally, he
communicates ideas and helps to share lessons learned among the developers and
stakeholders. Despite uncertainty, he relies on data to help propel the innovation towards
fruition.
References

Evaluation roots: Tracing theorists’ views and influences (pp. 153-168).

American Federation of Teachers, National Council on Measurement in Education, &

from http://assessment-reform-group.org/CIE3.PDF

Andrews & C. Haythornthwaite (Eds.), The SAGE handbook of e-learning

classroom assessment practices. Alberta Journal of Educational Research, 47,
244-258.

Education: Principles, Policy, and Practice, 5, 7-74.

3-10.


Technology-enhanced learning: Principles and products (pp. 289-306).
Netherlands: Springer.


Appendix A

Policy Documents in Support of Contemporary Classroom Assessment

The following is a subset of policy documents that makes reference to the contemporary notion of classroom assessment as it is being described here. AFT, NCME, & NEA, 1990; Heritage, 2010; FAST SCASS, 2006; ARG, 2002; Joint Advisory Committee for *Principles for Fair Student Assessment Practices for Education in Canada*, 1993.

Appendix B

Letter of Information

Letter of Information

for Instructional Team Members

Reconceptualizing Teaching and Learning through the use of Web 2.0 Technologies in a Process Guided by Developmental Evaluation

You are invited to participate in a research study being conducted by Chi Yan Lam, under the supervision of Dr. Lyn Shulha, in the Faculty of Education, at Queen’s University in Kingston, Ontario. This study has been granted clearance according to the recommended principles of Canadian ethics guidelines, and Queen's policies.

The purposes of this study are to (a) understand how instructors actively reconceptualize teaching and learning when collaborative Web 2.0 technologies are introduced as a context for learning about assessment and (b) to assess the contribution that developmental evaluation can make to this re-thinking process. As a member of the instructional team who has contributed to the new learning approach, I am interested in the knowledge and insight you have gained as the process had unfolded, as well as, your thoughts on how assessment education can be better designed.

You are invited to 1) participate in an interview, and 2), authorize the inclusion of your contribution (e.g. meeting minutes, worksheets, diagrams) that were archived as part of program development into this research study for secondary analysis. The interview will require 60 minutes of your time, although I will not cut you off if we are actively engaged in conversation. There are no known physical, psychological, economic, or social risks associated with this study.

Your participation is voluntary, and to any extent you are comfortable with. You should not feel obliged to answer any material that you find objectionable or that makes you feel uncomfortable. You are able to withdraw at any time with no adverse personal effect. If you withdraw, you can request removal of all or part of your data; if you so request, all such data will be destroyed.

We will keep your responses confidential to the extent possible and ensure that you cannot be identified in any publication emerging from the study. The interview will be recorded and transcribed to facilitate later analysis. Only members of the research team will have access to the data. Following transcription, you will have the opportunity to review the transcript for accuracy if you wish to do so. You may at that time elect to remove any part of the transcript.

The results of this research may also be published in professional journals or presented at scientific conferences, but any such presentations will report only aggregated findings, which in some instances will be illustrated by short quotes that will be carefully selected so as not to breach individual confidentiality. In addition, a web site will be created that will highlight research findings. Should you be interested, a copy of the findings will be sent to you when the study is completed. In accordance with Queen’s policy, data will be retained for a minimum of five years and may be retained indefinitely for the purposes of secondary analysis. Data will be destroyed after a minimum of five years if it is not retained indefinitely. If data are used for secondary analysis it will contain no identifying information.
Any questions about study participation may be directed to Chi Yan Lam chi.lam@queensu.ca; or my supervisor, Dr. Lyn Shulha (613-533-6000 X 75016) lyn.shulha@queensu.ca. Any ethical concerns about the study may be directed to the Chair of the General Research Ethics Board at 613-533-6081 or chair.GREB@queensu.ca.

Thank you.

Yours sincerely,

Chi Yan Lam M.Ed. Candidate
Queen’s University
Appendix C: Consent form

Consent Form
Reconceptualizing Teaching and Learning through the use of Web 2.0 Technologies
in a Process Guided by Developmental Evaluation

Name (please print clearly): ________________________________________

1. I have read the Letter of Information and have had any questions answered to my satisfaction.
2. I understand that I will be participating in the study called “Reconceptualizing Teaching and Learning through the use of Web 2.0 Technologies in a Process Guided by Developmental Evaluation”. The purposes of this study are to (a) understand how instructors actively reconceptualize teaching and learning when collaborative Web 2.0 technologies are introduced as a context for learning about assessment and (b) to assess the contribution that developmental evaluation can make to this re-thinking process.
3. I understand that this means that I will be allowing the researcher to include into the study for analysis my academic contributions stemming from my participation in this pilot program. This includes microblog posts (i.e. “tweets”) and discussion board submissions.
4. I understand the interviews will be audiotaped and transcribed.
5. I understand that my participation in this study is voluntary and that I may withdraw at any time without any adverse personal effect. If I withdraw, I understand that I can request removal of all or part of my data; if I so request, all such data will be destroyed. I understand that if I withdraw, there will be no effect on my standing in school if I am a student. I understand that every effort will be made to maintain the confidentiality of the data now and in the future to the extent possible. Only members of the research team will have access to the data. The results will be published in professional journals, presented at scientific conferences, or disseminated through other academic means (e.g. research findings web site), but any such presentations will be of general findings, which in some instances will be illustrated by short quotes that will be carefully selected so as not to breach individual confidentiality. I understand my confidentiality will be protected to the extent possible. I understand that I will be sent a copy of the results if I wish.
6. Any questions about study participation may be directed Chi Yan Lam chi.lam@queensu.ca; or my supervisor, Dr. Lyn Shulha (613-533-6000 X 75016) lyn.shulha@queensu.ca. Any ethical concerns about the study may be directed to the Chair of the General Research Ethics Board at 613-533-6081 or chair.GREB@queensu.ca

7. Please sign one copy of this Consent Form and return it to Chi Yan Lam, the researcher. Retain the second copy for your records.

I have read the above statements and freely consent to participate in this research:

Signature: _____________________________________ Date: _____________________

If you would like to receive a copy of the study’s results, please provide your e-mail address here:

_________________________________________
Appendix D: Interview Guide

Interview Protocol

Consent

Hello, ______. Thank you for agreeing to participate in this interview. As you know, my name is Chi Lam and I have been working with you since August of 2010 as a resource person in developmental evaluation. It is now time for me to step back from this development work and assess what can be learned from our experiences. The overall goal of the interview is to understand the different experiences, you and (two of Don, Lyn and King) had in developing the new approach to assessment education and in addition, your impressions of and reactions to the processes and products of our work together as a team. To help you with this process I have put together a one-page chronological summary of our work together and with teacher candidates. I have also prepared some questions to help me stay focused on my goal. Feel free not to answer any question that makes you uncomfortable or you feel is inappropriate. You should know that any comments that you might wish to make about my behaviours or contributions will be taken in the spirit of learning as much as possible about what it means to be engaged in program development and developmental evaluation. Before we begin, I have prepared a letter of information and a consent form. Please take a minute to review the documents, and please sign the consent form if you agree to participate in this study.

Provide LOI to participant
Provide Consent Form to participant
Sign CF, provide duplicate to participant

Define scope

Today, I would like us to focus primarily on the development process that shaped the new program. I will also ask you questions about your impressions of the program that resulted from our efforts (i.e., the developed program itself); and any insights that may have resulted from your involvement in this project.
A. By development process, I mean the meetings, activities, and communications we had with each other and our students over 12 months. As well, I’d like you to identify what you feel were critical episodes or decisions in the development process and hear you talk about their significance.
B. By the developed program I mean the program that we implemented, between January and April, 2011.
C. By asking you to consider insights I’m wondering if you think any differently about teaching and learning in assessment since experiencing this project.
<table>
<thead>
<tr>
<th>#</th>
<th>Purpose</th>
<th>A. Development process</th>
<th>B. Developed Program</th>
<th>C. Insights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Can you tell me what your motivation was for committing to a new approach to teaching assessment education?</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>When you review the summary of our meetings, activities and communications how would you describe the work we did together and with students?</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>How would you describe our decision-making processes as a team?</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>In looking at the whole process we went through, what are some of the more important or significant moments and why were they significant to you? (Note to self . . . allow time for reflection)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>How would you describe my contributions to our efforts, if any?</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>What were the consequences of my contributions?</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>How would you describe your contributions to development process?</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>What contributions did you see other members making to the process?</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>How did this pilot initiative compare with other courses you have prepared and implemented?</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>10</td>
<td>Were the process or products associated with this initiative typical of what you have experienced in the past? If so, how and if not, what made this different?</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
I’m wondering if you think any differently about teaching and learning in assessment since experiencing this project?

**Introductory Question**

Question 1
Can you tell me what your motivation was for committing to a new approach to teaching assessment education?

**Question on process**

Question 2
When you review the summary of our meetings, activities and communications how would you describe the work we did together and with students?

Question 3
How would you describe our decision-making processes as a team?

**Critical Periods**

Question 4
In looking at the whole process we went through, what are some of the more important or significant moments and why were they significant to you? (Note to self . . . allow time for reflection)

**Developmental Evaluator**

Question 5
How would you describe my contributions to our efforts, if any?

> As the pilot initiative unfolded, I purposefully adopted the role of a developmental evaluator.

Question 6
What were the consequences of my contributions?

**Role of Participants in a Program Development Initiative**

Question 7
How would you describe your contributions to development process?

Question 8
What contributions did you see other members making to the process?

**Questions on Insight**

Question 9
How did this pilot initiative compare with other courses you have prepared and implemented?
Question 10
Were the process or products associated with this initiative typical of what you have experienced in the past? If so, how and if not, what made this different?

Impact
Question 11
I’m wondering if you think any differently about teaching and learning in assessment since experiencing this project?

Closing
You’ve been very helpful.

Are there any other thoughts or feelings you would like to share with me concerning your experience of participating in a developmental evaluation or about the teaching classroom assessment with technology?
Appendix E: Letter of Ethics Clearance

May 2, 2011

Chi Y. Lam
Master’s Student
Faculty of Education
Duncan McArthur Hall
Queen’s University
511 Union Street
Kingston, On K7M 5R7

Dear Chi Lam:

GREB Ref #: GEDUC-554-11
Title: “Reconceptualizing Teaching and Learning Through the Use of Web 2.0 Technologies in a Process Guided By Developmental Evaluation”

The General Research Ethics Board (GREB), by means of a delegated board review, has cleared your proposal entitled “Reconceptualizing Teaching and Learning Through the Use of Web 2.0 Technologies in a Process Guided By Developmental Evaluation” for ethical compliance with the Tri-Council Guidelines (TCPS) and Queen’s ethics policies. In accordance with the Tri-Council Guidelines (article D.1.6) and Senate Terms of Reference (article G), your project has been cleared for one year. At the end of each year, the GREB will ask if your project has been completed and if not, what changes have occurred or will occur in the next year.

You are reminded of your obligation to advise the GREB, with a copy to your unit REB, if applicable, of any adverse event(s) that occur during this one year period (details available on webpage http://www.queensu.ca/ors/researchethics/GeneralREB/forms.html – Adverse Event Report Form). An adverse event includes, but is not limited to, a complaint, a change or an unexpected event that alters the level of risk for the researcher or participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GREB within 48 hours.

You are also reminded that all changes that might affect human participants must be cleared by the GREB. For example you must report changes in study procedures or implementations of new aspects into the study procedures on the Ethics Change Form that can be found at http://www.queensu.ca/ors/researchethics/GeneralREB/forms.html - Research Ethics Change Form. These changes must be sent to the Ethics Coordinator, Gail Irving, at the Office of Research Services or irvingg@queensu.ca prior to implementation. Mrs. Irving will forward your request for protocol changes to the appropriate GREB reviewers and/or the GREB Chair.

On behalf of the General Research Ethics Board, I wish you continued success in your research.

Yours sincerely,

Joan Stevenson, PhD
Professor and Chair
General Research Ethics Board

c.c.: Dr. Lyn Shulha, Faculty Supervisor
Dr. Lesly Wade-Woolley, Chair, Unit REB
E-REB: c/o Graduate Studies and Bureau of Research, Attn.: Celina Caswell

JS (g)
Certificate of Completion

This document certifies that

Lam, Chi Yan

has completed the Queen's University online
Course in Human Research Participant Protection (CHRPP).

Date of Issue: April 13, 2011
## Appendix G: List of Files Reviewed for Analysis

<table>
<thead>
<tr>
<th>Document Number</th>
<th>Name</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>P 1</td>
<td>2010_05_11 - thesis topic.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 2</td>
<td>2010_05_2X CCF00252011_00000.pdf</td>
<td>Yellow – Evaluation Journal – Thinking notes</td>
</tr>
<tr>
<td>P 3</td>
<td>2010_05_2X CCF00252011_00001.pdf</td>
<td>Yellow – Evaluation Journal – Thinking notes</td>
</tr>
<tr>
<td>P 4</td>
<td>2010_05_2X CCF00252011_00003.pdf</td>
<td>Yellow – Evaluation Journal – Thinking notes</td>
</tr>
<tr>
<td>P 5</td>
<td>2010_06_03 - a.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 6</td>
<td>2010_06_03 THESIS.docx</td>
<td>Grey – Thesis</td>
</tr>
<tr>
<td>P 7</td>
<td>2010_06_23 - Re- Statement of Intent for Chi Yan Lam's Thesis Form.pdf</td>
<td>Purple – emails</td>
</tr>
<tr>
<td>P 8</td>
<td>2010_06_24 - CCF00252011_00004.pdf</td>
<td>Orange – Mtg Transcripts / Mtg Notes</td>
</tr>
<tr>
<td>P 9</td>
<td>2010_07_05 - CCF00252011_00006.pdf</td>
<td>Yellow – Evaluation Journal – Thinking notes</td>
</tr>
<tr>
<td>P 10</td>
<td>2010_07_06 - CCF00252011_00008.pdf</td>
<td>Yellow – Evaluation Journal – Thinking notes</td>
</tr>
<tr>
<td>P 11</td>
<td>2010_07_06 - thinking notes CCF00252011_00007.pdf</td>
<td>Yellow – Evaluation Journal – Thinking notes</td>
</tr>
<tr>
<td>P 12</td>
<td>2010_07_08 - meeting notes - CCF00252011_00009.pdf</td>
<td>Orange – Mtg Transcripts / Mtg Notes</td>
</tr>
<tr>
<td>P 13</td>
<td>2010_07_08 - meeting notes - CCF00252011_00010.pdf</td>
<td>Orange – Mtg Transcripts / Mtg Notes</td>
</tr>
<tr>
<td>P 14</td>
<td>2010_07_08 - meeting notes - CCF00252011_00011.pdf</td>
<td>Orange – Mtg Transcripts / Mtg Notes</td>
</tr>
<tr>
<td>P 15</td>
<td>2010_07_08 - Meeting notes.docx</td>
<td>Orange – Mtg Transcripts / Mtg Notes</td>
</tr>
<tr>
<td>P 16</td>
<td>2010_07_08 - Writeboard- Minutes to July 8th Mtg.pdf</td>
<td>Orange – Mtg Transcripts / Mtg Notes</td>
</tr>
<tr>
<td>P 17</td>
<td>2010_07_08 m-learning.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 18</td>
<td>2010_07_16 - thesis - overaching research question.docx</td>
<td>Grey – Thesis</td>
</tr>
<tr>
<td>P 20</td>
<td>2010_07_16 Thesis chapter 1.docx</td>
<td>Grey – Thesis</td>
</tr>
<tr>
<td>P 21</td>
<td>2010_07_21 Interesting.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 23</td>
<td>2010_07_23 thesis progress update.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 26</td>
<td>2010_08_05 - e-portfolios.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 27</td>
<td>2010_08_05 - Re- e-portfolios.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 28</td>
<td>2010_08_10 - Writeboard- The PLAN - Part One - The Case for Change.pdf</td>
<td>Green – Output/Publications</td>
</tr>
<tr>
<td>P 29</td>
<td>2010_08_10 - Writeboard- The PLAN - Part Three - The Plan.pdf</td>
<td>Green – Output/Publications</td>
</tr>
<tr>
<td>P 30</td>
<td>2010_08_10 - Writeboard- The PLAN - Part Two - Design Principles.pdf</td>
<td>Green – Output/Publications</td>
</tr>
<tr>
<td>P 31</td>
<td>2010_08_11 AWG Design -The Plan.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>P 32</td>
<td>2010_08_14 [name redacted].pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>Page</td>
<td>Date</td>
<td>File Name</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>33</td>
<td>2010_08_19</td>
<td>CCF00252011_00013.pdf</td>
</tr>
<tr>
<td>34</td>
<td>2010_08_19</td>
<td>first half - CCF00252011_00014.pdf</td>
</tr>
<tr>
<td>35</td>
<td>2010_08_19 1_06 PM.docx</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>2010_08_19 1_06 PM.mp3</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>2010_08_19</td>
<td>meeting arrangement for Sept 1.pdf</td>
</tr>
<tr>
<td>38</td>
<td>2010_08_21 Update and Mtg Confirmation.pdf</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>2010_08_31</td>
<td>Re- Update and Mtg Confirmation.pdf</td>
</tr>
<tr>
<td>40</td>
<td>2010_09_01 Team Meeting.docx</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>2010_09_13 Assessment Working Group Design Mtg.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>42</td>
<td>2010_09_13 Meeting.pdf</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>2010_09_14 - Patton Book.pdf</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>2010_09_17 - CCF00252011_00014 copy.pdf</td>
<td>Orange – Mtg transcripts / Mtg Files</td>
</tr>
<tr>
<td>45</td>
<td>2010_09_17_2 - CCF00252011_00015.pdf</td>
<td>Orange – Mtg transcripts / Mtg Files</td>
</tr>
<tr>
<td>46</td>
<td>2010_09_24 scheduling next mtg.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>47</td>
<td>2010_09_26 - next design meeting - demo.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>48</td>
<td>2010_09_27 - mtg transcript 1_43 PM.docx</td>
<td>Orange – Mtg transcripts / Mtg Files</td>
</tr>
<tr>
<td>49</td>
<td>2010_10_13AWG design meeting.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>50</td>
<td>2010_11_01 - debrief on clicker sesh.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>51</td>
<td>2010_11_01 - talk from Sir Ken .pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>52</td>
<td>2010_11_17 - great news!.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>53</td>
<td>2010_11_20 - tweet re formative assessment.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>59</td>
<td>2010_11_25 - draft 2.docx</td>
<td>Grey - Thesis</td>
</tr>
<tr>
<td>60</td>
<td>2010_11_28 progress update.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>61</td>
<td>2010_12_01 Design Meeting for AWG.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>62</td>
<td>2010_12_01 doodle.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>63</td>
<td>2010_12_02 - Doodle+AWG+Design+Meeting.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>64</td>
<td>2010_12_08 - Meeting notes post its_1.jpg</td>
<td>Orange – Mtg transcripts / Mtg Files</td>
</tr>
<tr>
<td>65</td>
<td>2010_12_08 - Meeting notes post its_2.jpg</td>
<td>Orange – Mtg transcripts / Mtg Files</td>
</tr>
<tr>
<td>66</td>
<td>2010_12_08 - Design Meeting for Classroom Assessment Pilot Group.docx</td>
<td>Blue – Initiative documents</td>
</tr>
<tr>
<td>67</td>
<td>2010_12_08 - mtg transcript - 10_22 AM.docx</td>
<td>Orange – Mtg transcripts / Mtg Files</td>
</tr>
<tr>
<td>68</td>
<td>2010_12_08 10_22 AM.mp3</td>
<td>Red – Mtg Voice Files</td>
</tr>
<tr>
<td>69</td>
<td>2010_12_09 Classroom Assessment Pilot Group -- Design.pdf</td>
<td>Purple – Emails</td>
</tr>
<tr>
<td>70</td>
<td>2010_12_11 Research Background.docx</td>
<td>Grey - Thesis</td>
</tr>
<tr>
<td>Page</td>
<td>Date</td>
<td>Document Name</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>P 71</td>
<td>2010_12_14</td>
<td>bibliography.docx</td>
</tr>
<tr>
<td>P 73</td>
<td>2010_12_15</td>
<td>proposal - cover.docx</td>
</tr>
<tr>
<td>P 74</td>
<td>2010_12_15</td>
<td>proposal - body_final.docx</td>
</tr>
<tr>
<td>P 76</td>
<td>2010_12_17</td>
<td>CCF00252011_00016.pdf</td>
</tr>
<tr>
<td>P 77</td>
<td>2010_12_17</td>
<td>CCF00252011_00017.pdf</td>
</tr>
<tr>
<td>P 78</td>
<td>2010_12_17</td>
<td>Do Well.docx</td>
</tr>
<tr>
<td>P 79</td>
<td>2010_12_17</td>
<td>mtg transcript.docx</td>
</tr>
<tr>
<td>P 80</td>
<td>2010_12_17</td>
<td>Twitter 101.docx</td>
</tr>
<tr>
<td>P 81</td>
<td>2010_12_17</td>
<td>Re- Meet tomorrow at 230.pdf</td>
</tr>
<tr>
<td>P 82</td>
<td>2010_12_17</td>
<td>2011-03-24.mp3</td>
</tr>
<tr>
<td>P 83</td>
<td>2011_01_01</td>
<td>Executive Summary.docx</td>
</tr>
<tr>
<td>P 84</td>
<td>2011_01_01</td>
<td>Invitation to Participate.docx</td>
</tr>
<tr>
<td>P 85</td>
<td>2011_01_01</td>
<td>Mail message.pdf</td>
</tr>
<tr>
<td>P 86</td>
<td>2011_01_02</td>
<td>CSCL Script.docx</td>
</tr>
<tr>
<td>P 87</td>
<td>2011_01_02</td>
<td>sample recruitment script.pdf</td>
</tr>
<tr>
<td>P 88</td>
<td>2011_01_03</td>
<td>Design Meeting.pdf</td>
</tr>
<tr>
<td>P 89</td>
<td>2011_01_04</td>
<td>Design Meeting - schedule - reminder.pdf</td>
</tr>
<tr>
<td>P 90</td>
<td>2011_01_04</td>
<td>Design Meeting confirmed.pdf</td>
</tr>
<tr>
<td>P 91</td>
<td>2011_01_05</td>
<td>[Assessment Pilot meeting notification] Confirming tomorrow at 1PM.pdf</td>
</tr>
<tr>
<td>P 92</td>
<td>2011_01_05</td>
<td>Assessment Pilot vet the first document, familiarize with the second).pdf</td>
</tr>
<tr>
<td>P 93</td>
<td>2011_01_05</td>
<td>reschedule mtg.pdf</td>
</tr>
<tr>
<td>P 94</td>
<td>2011_01_05</td>
<td>theory spec sheet.docx</td>
</tr>
<tr>
<td>P 95</td>
<td>2011_01_05</td>
<td>theory spec sheet.pdf</td>
</tr>
<tr>
<td>P 96</td>
<td>2011_01_05</td>
<td>[Assessment Pilot decision solicited] Gathering preliminary data on tech use and comfort.pdf</td>
</tr>
<tr>
<td>P 97</td>
<td>2011_01_05</td>
<td>[Assessment Pilot notification] Questionnaire has gone out.pdf</td>
</tr>
<tr>
<td>P 98</td>
<td>2011_01_06</td>
<td>Conceptual Briefings.docx</td>
</tr>
<tr>
<td>P 99</td>
<td>2011_01_06</td>
<td>COP.docx</td>
</tr>
<tr>
<td>P 100</td>
<td>2011_01_06</td>
<td>mtg transcript 1_14 PM.docx</td>
</tr>
<tr>
<td>P 101</td>
<td>2011_01_06</td>
<td>1_14 PM.mp3</td>
</tr>
<tr>
<td>P 102</td>
<td>2011_01_07</td>
<td>Assessment Mini-Course- Getting to Know You Questionnaire.xls</td>
</tr>
<tr>
<td>P 103</td>
<td>2011_01_07</td>
<td>betterknowyou.docx</td>
</tr>
<tr>
<td>P 104</td>
<td>2011_01_07</td>
<td>CCF00252011_00018.pdf</td>
</tr>
<tr>
<td>P 105</td>
<td>2011_01_07</td>
<td>Edit form - [ Queens U Assessment Mini-Course 10 Questions to Get to Know You Better ] - Google Docs.pdf</td>
</tr>
<tr>
<td>Page</td>
<td>Date</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>P 106</td>
<td>2011_01_07</td>
<td>Fwd- FYI - OTF technology conference.pdf</td>
</tr>
<tr>
<td>P 107</td>
<td>2011_01_07</td>
<td>Twitter 101 for teachers.docx</td>
</tr>
<tr>
<td>P 108</td>
<td>2011_01_12</td>
<td>[Assessment Pilot schedule a meeting] Consume, review, reflect, and do..pdf</td>
</tr>
<tr>
<td>P 109</td>
<td>2011_01_14</td>
<td>Jan14_1.MP3</td>
</tr>
<tr>
<td>P 110</td>
<td>2011_01_14</td>
<td>Jan14_2.MP3</td>
</tr>
<tr>
<td>P 111</td>
<td>2011_01_14</td>
<td>Jan14_3.MP3</td>
</tr>
<tr>
<td>P 112</td>
<td>2011_01_14</td>
<td>mtg transcript_3.docx</td>
</tr>
<tr>
<td>P 113</td>
<td>2011_01_14</td>
<td>mtg transcript_1.docx</td>
</tr>
<tr>
<td>P 114</td>
<td>2011_01_14</td>
<td>mtg transcript_2_2DF5.docx</td>
</tr>
<tr>
<td>P 115</td>
<td>2011_01_14</td>
<td>mtgnotes - CCF00252011_00019.pdf</td>
</tr>
<tr>
<td>P 116</td>
<td>2011_01_18</td>
<td>Re- meeting recordings.pdf</td>
</tr>
<tr>
<td>P 117</td>
<td>2011_01_14</td>
<td>CCF00252011_00020.pdf</td>
</tr>
<tr>
<td>P 118</td>
<td>2011_01_21</td>
<td>field CCF00252011_00022.pdf</td>
</tr>
<tr>
<td>P 119</td>
<td>2011_01_21</td>
<td>field notesCCF00252011_00021.pdf</td>
</tr>
<tr>
<td>P 120</td>
<td>2011_01_21</td>
<td>Jan21.MP3</td>
</tr>
<tr>
<td>P 121</td>
<td>2011_01_21</td>
<td>mtg transcript.docx</td>
</tr>
<tr>
<td>P 122</td>
<td>2011_01_31</td>
<td>CCF00252011_00023.pdf</td>
</tr>
<tr>
<td>P 123</td>
<td>2011_01_31</td>
<td>CCF00252011_00024.pdf</td>
</tr>
<tr>
<td>P 124</td>
<td>2011_01_31</td>
<td>CCF00252011_00025.pdf</td>
</tr>
<tr>
<td>P 125</td>
<td>2011_01_31</td>
<td>Re- Find upcoming events and notes from recent AEGGRAPE meeting.pdf</td>
</tr>
<tr>
<td>P 126</td>
<td>2011_02_01</td>
<td>CCF00252011_00028.pdf</td>
</tr>
<tr>
<td>P 127</td>
<td>2011_02_01</td>
<td>field notesCCF00252011_00026.pdf</td>
</tr>
<tr>
<td>P 128</td>
<td>2011_02_01</td>
<td>PLC - CCF00252011_00027.pdf</td>
</tr>
<tr>
<td>P 129</td>
<td>2011_02_03</td>
<td>January 21 – February 2.pdf</td>
</tr>
<tr>
<td>P 130</td>
<td>2011_02_03</td>
<td>[Assessment Pilot scheduling meeting] Friday AM.pdf</td>
</tr>
<tr>
<td>P 131</td>
<td>2011_02_04</td>
<td>Q posed to class (classnotes).docx</td>
</tr>
<tr>
<td>P 132</td>
<td>2011_02_04</td>
<td>mtg transcript.docx</td>
</tr>
<tr>
<td>P 133</td>
<td>2011_02_04</td>
<td></td>
</tr>
<tr>
<td>P 134</td>
<td>2011_02_09</td>
<td>[Assessment Pilot Last Class Preparation].pdf</td>
</tr>
<tr>
<td>P 135</td>
<td>2011_02_09</td>
<td>Re- [Assessment Pilot- Last Class Preparation].pdf</td>
</tr>
<tr>
<td>P 136</td>
<td>2011_02_09</td>
<td>Fwd- participant list for.pdf</td>
</tr>
<tr>
<td>P 137</td>
<td>2011_02_10</td>
<td>twitter use infographic.docx</td>
</tr>
<tr>
<td>P 138</td>
<td>2011_02_11</td>
<td>mtg transcript.docx</td>
</tr>
<tr>
<td>P 139</td>
<td>2011_02_11</td>
<td></td>
</tr>
<tr>
<td>P 140</td>
<td>2011_03_21</td>
<td>[Assessment Pilot Interim Mtg to plan evaluation].pdf</td>
</tr>
<tr>
<td>P 141</td>
<td>2011_03_28</td>
<td>6-44-57 PM.jpg</td>
</tr>
<tr>
<td>Page</td>
<td>Title</td>
<td>File Name</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>P 142</td>
<td>2011_03_28 2_12 PM.mp3</td>
<td>Red – Mtg Voice Files</td>
</tr>
<tr>
<td>P 143</td>
<td>2011_04_15 - debrief focus group recording - R09_0002.MP3</td>
<td>Red – Mtg Voice Files</td>
</tr>
<tr>
<td>P 144</td>
<td>2011_04_16 - Assessment Mini Course Agenda Apr 15.docx</td>
<td>Blue – Initiative documents</td>
</tr>
<tr>
<td>P 145</td>
<td>Thesis Interview Tim.docx</td>
<td>Interview</td>
</tr>
<tr>
<td>P 146</td>
<td>Thesis Interview Laurene.docx</td>
<td>Interview</td>
</tr>
<tr>
<td>P 147</td>
<td>Thesis Interview, Albert.docx</td>
<td>Interview</td>
</tr>
</tbody>
</table>
July 2, 2011

What:
Meeting notes for July 8th meeting
In attendance were Laurene, Albert, Tim, and I

Observation:
Captured in these notes were:
- Description of the program context, and the struggles the instructors have been having with their current way of teaching classroom assessment (i.e. barriers to learning)
- Outlined, also, are the parameters of the program for the upcoming year
- Laurene described some of the concurrent changes that are happening with regard to the larger program
- In this meeting, the clients also raised questions that they’d like to see explored.

Significance:
- First design-team meeting with the group
- There is a lot of negotiating and creation of shared-meaning taking place
- there is a whole lot of listening to the clients talk about their issue that they’d like to explore  (*I am reminded of the business about practicing servant leadership*)

What do I remember from then:
- it was difficult to wrap my head around some of the things that the clients wanted to do
  --- partly because they were subject experts (i.e. they already know much more)
  --- and because I perceived myself as lacking the authority/credibility to speak up (at least for the beginning).
Appendix I: Samples of Assessment Pilot Initiative Document Output (Doc. 107)

QU Assessment Mini-Course: Twitter 101

Join us on this exciting adventure of exploring the use of Twitter as a way to continue our conversations beyond our class times.

Let’s get you set up.

Step 0] Get an account.

- Go to www.twitter.com
- Click on: “Sign Up”

Let’s explore!

Step 1] Login to your account. Immediately after you login, you’re presented with a “timeline”. This is where all the tweets the people you follow will appear. (Just like your newsfeed in Facebook). At first, you won’t have any, until you start following people.

Step 2] Post your first tweet!

- Here is your prompt: “Teaching excites me because...”

Follow us and others!

Step 3] “Follow” others.

- Follow us, the official account for our mini-course! @QUAssessment
- You could also follow @QueensU , @OSTF @etfonews

There are two ways of following someone.
1) You can visit their twitter page, www.twitter.com/username (e.g. www.twitter.com/queensu), and click FOLLOW
2) Type in their username into the search box, click on the “people” tab, and click FOLLOW.

Tweet @us!

Step 4] Send a tweet at us! In your tweet, add the “@” symbol before the username. Here is a prompt.

“@QUAssessment I’d like to learn more about...”

Tweet at us, and we’ll reply back! Make sure you do, as we’ll be constructing class list for next week so other can find you! See tweets that mention you via the @Mention tab.

How else can you access Twitter?

2] App Store (Mac OS X)
3] iPhone / iPod Touch / iPad - App Store > Search: Twitter
5] Blackberry - App World > Search: Twitter > Download
Appendix J: Infographic Illustrating a Simplified Model of the Developed Program Model (Doc. 137)

@QUAssessment Mini Course at a Glance

1. Think
During your practicum...
- Whenever you have an assessment question, idea, observation, or advice...
- Whenever you are struggling with assessment...
- Whenever you have an “ah-ha” light bulb assessment moment...
- Or you simply want to bounce assessment ideas off of one another...

2. Tweet
Tweet about it!
(we all get to learn about it)

3. Share
Share, Collaborate, Comment
- Use D2L discussion board if you need more space to elaborate. Tweet about your new post, so we all get to check it out.
- Don’t forget to log on to Twitter to see what others are up to.
- Lyn and I will be on Twitter during your Prac, listening and responding to your tweets.
- Most important, tweet @ each other! Jump in anytime; don’t be shy!
Appendix K: Program Development Timeline

The photograph presented above depicts a visualization of the development process. The relative distribution of the colored flags on this visualization is a matter of interest. The significance of which are described and explored in chapter 6 of this thesis. A horizontal black line running through the figure represents the timeline. Developmental episodes (above the timeline) and the respective developmental moments (below the timeline) are plotted to scale against the timeline.
## Appendix L: Summary of Results from Appreciative Inquiry Exercise

<table>
<thead>
<tr>
<th>Do Well</th>
<th>Wish List</th>
<th>Wish list, elaborated</th>
<th>Action Verbs</th>
</tr>
</thead>
</table>
| Our focus on "Assessment FOR Learning" | Infrastructure | • More contact time  
• Change nature of contact  
• Monitor students' changing conceptions long term  
• Hear how students think  
• Provide time for greater student interaction  
• Address how to questions as a way of learning concepts  
• Build personal connections with students | Identify  
Track  
Respond  
Monitor  
Motivate  
Combat shared ignorance  
Clarify/demystify (relating to concept base)  
Informed, skill of interaction  
Expert facilitation  
Feel connected |
| Established infrastructure and support for student learning | Responsiveness to student learning | | |
| Attentive to student needs | Self-directed learning | | |
| Promote self-assessment / professional self-regulated learning among the students | Integrating, not isolating | | |

182
Appendix M: Design Brief of Assessment Pilot Initiative (Doc. 83)

**Executive Summary**

of discussion from the
Discovery and Dream Phases

Assessment Pilot Group

*This document summarizes our intent, mission, and objectives with regard to the Assessment Pilot Group. It lends a 'big-picture' view of what we hope to do. This document provides the empirical evidence upon which our future instructional decisions and actions are based upon; we will try to map our actions on to our intents.*

**Mission**

Building upon existing assessment learning infrastructure and resources, the pilot seeks ways to:

- Better prepare preservice teachers for assessment challenges facing them in today’s classroom

To overcome some teaching constraints that we believe are currently hindering the quality of learning, the pilot seeks to:

- Understand the process of integrating technology-enhanced learning and to evaluate its impact for learning classroom assessment in a B.Ed program.

The goal of this pilot is to produce some evidence of student learning and experience at the institutional level to support the feasibility of scaling up a TEL approach to assessment learning.

**Objectives**

*Program*

- Try out alternate pedagogical strategies
- Investigate the role of technology in facilitating assessment learning
- Understand the nature of teaching and learning with Web 2.0
- Investigate how we can support collaborative learning in a large-group setting

*Program Learning Objectives (PLObj)*

With the pilot, we want to:

1. Monitor students’ changing conceptions over a long term
2. Support informal learning
3. Hear how students think
4. Address how to questions as a way of learning concepts
5. Support informal learning
6. Work more at integrating our work/examples with other faculties
7. Integration with other courses
8. Combat shared ignorance


C2. Reflect on their own assessment learning journey

If we are successful in acting on our objectives, we hope to:

- Prepare better assessment-informed teacher candidates
- Build professional connections with students
- [Have the students] see us as a resource for engagement about assessment (not completing assignments)
- Experience “non-traditional” assessment

How this document was compiled

I’ve first selected items from the wish list that falls within the purview of this current project. Selected items must be relevant to the stakeholders involved within the current project, be of a reasonable time frame, and be specific and actionable. For instance, “increasing faculty commitment to understanding the role of assessment in the classroom” did not make the cut.
Appendix N: Theory Spec Sheet (Doc. 94)

**Theory Spec Sheet**

Our project operates at the intersections of:

- Teaching and Learning
- Assessment Pedagogy
- Assessment Learning - Teacher Development
- Classroom assessment
- Web 2.0 / TEL

*Classroom assessment as what we teach (the knowledge domain)*
*Assessment learning (+TEL) as how we teach*
*Web 2.0 and TEL as mediating the learning process*

**using a**

**Community of Learner Model**

*Informed by:*

- Professional learning
  - Life-long learning
  - Informal learning
- Mobile Learning
- TEL
  - Web 2.0

*facilitated using:*

- Collaborative learning
- Transformative learning
- Reflective learning

**points us to using:**

- CSCL Scripts
- Assessment Dilemma

<table>
<thead>
<tr>
<th>Track</th>
<th>Respond</th>
<th>Monitor</th>
<th>Motivate</th>
<th>Combat shared ignorance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connect</td>
<td>Clarify / Demystify</td>
<td>Expertly facilitate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix O: Sample of Activity Summary (Doc. 129)

QU Assessment Pilot
January 21 – February 2, 2011

<table>
<thead>
<tr>
<th>Total Users</th>
<th>Tweets in Timeline</th>
<th>Unique Threads</th>
<th>Unique Contributors</th>
<th>Thread Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>70+ Tweets</td>
<td>3</td>
<td>9</td>
<td>1-8 replies</td>
</tr>
</tbody>
</table>

What's Trending?

1. **Doing projects at home is not allowed; @StudentA argues that it limits research**
   - Laurene: depends on what home looks like
   - StudentB: Agree, classroom has too much distraction
   - Student C: it levels the playing field
   - StudentD: if students care enough, they will put in the time regardless
   - StudentE: depends if they are able to or if other things get in the way
     - StudentD: in assessment yes, grading no. Principle of fairness
   - StudentA: Now teachers refuse to mark work done outside of the classroom
     - StudentC: What’s the rationale behind that?
   - Laurene: When is working at home appropriate?

2. **How do you get students to use your AFL/AAL to craft short term/long term goals?**
   - Laurene: scaffolding handouts; see if it gives clues

3. **Teachers grading parents (news article)**
   - StudentD: grading is a lazy solution to a serious problem
     - StudentF: agree; probably won’t get parents involved; leads to finger pointing
     - StudentD agrees back; students get turned off when they are given a bad grade; parents too.
1) Review Progress
2) Evaluate our Efforts
3) Plan Ahead

1) A framework for inquiry:

1a) Technology: i) some concerns about 140 characters preventing deep conversations from happening
   ii) the interface makes it difficult to follow threads

   What does that mean?

1b) Community: i) There's lively activity (with some threads commanding a lengthy debate); ii) Participation is largely polarized.

   What does that mean?

1c) Classroom Assessment: 3 new topics; (7 topics on the docket; 2 addressed last class).

   What does that mean?

2) Evaluating our progress
- As we begin to cement our efforts in building our community, an important evaluand is the notion of “value”.
- How do we add value to the lives of our teacher candidates?

- Value is the new buzzword in social media; generating value and active engagement is key to stake-holder engagement.

3) Planning Ahead

| To what extent are we value-adding to TC’s learning when they engage in the technology-enhanced learning experience? |
| How else can we add value to their learning? |