

SELF-ASSESSMENT PROCESSES OF PROFESSIONAL DEVELOPERS

A Dissertation Submitted to the College of  
Graduate and Postdoctoral Studies  
In Partial Fulfillment of the Requirements  
For the Degree of Doctor of Philosophy  
In the Department of Curriculum Studies  
University of Saskatchewan  
Saskatoon

Copyright Wendy James, April, 2018. All Rights Reserved

By

WENDY JAMES

## PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Curriculum Studies  
28 Campus Drive  
University of Saskatchewan  
Saskatoon, Saskatchewan S7N 0X1  
Canada

OR

Dean  
College of Graduate and Postdoctoral Studies  
University of Saskatchewan  
116 Thorvaldson Building, 110 Science Place  
Saskatoon, Saskatchewan S7N 5C9  
Canada

## ABSTRACT

Professional Learning is a mechanism designed to support K-12 educators in improving their professional lives, knowledge, and instructional practice. It is supported by school divisions because of a belief that it will help teachers keep abreast of current research and improve the overall quality of instruction over time. However, most professional learning occurs without formal evaluation to determine its results.

The purpose of this study is to develop theory about the self-assessment processes of the people who lead professional learning. The main objective of the research was answer the question: How do professional developers describe and demonstrate the process of either self-assessment or internal evaluation? The sub-questions included the following: 1) How do professional developers decide which strategies to use to determine the impact of their professional learning? 2) What influences the perpetuation of evaluation (or lack thereof)? 3) How is the praxis of assessment or evaluation of professional learning influenced by a professional developer's experience?

The study employed a constructivist grounded theory approach. Data were collected via interviews, observations, and extant documents from participants with a variety of roles leading professional learning.

The professional developers in the study engaged in a variety of self-assessment processes focused primarily on teachers' reception of the professional learning and the effectiveness of the facilitation process employed during professional learning sessions. They also shared extensively about barriers to the implementation of the learning in the classroom and took specific actions designed to respond to information they found in self-assessment. The self-

assessment processes were largely self-initiated based on personal questions (wondering). Wondering was followed by typically informal methods of finding out and attribution, then responded to through action or inaction, depending on the professional developers' interpretations of the results.

Implications for theory and practice include a need for greater data literacy and support for internal evaluation and self-assessment conducted by professional developers.

Recommendations include the development of essential evidence-based theory about facilitation processes and directions for further research.

## ACKNOWLEDGEMENTS

I would like to thank Dr. Jay Wilson, my advisor, for his great questions, positive nature, and openness. Jay has been a partner and mentor at every stage of this doctorate.

My committee members and external, Dr. Jim Grier, Dr. Tom Guskey, Dr. Dirk Morrison, Dr. Heather Kanuka, and Dr. Rick Schwier, have been flexible with my process and asked me thought-provoking questions that have improved my scholarship and deepened my thinking. Dr. Gale Parchoma provided such useful support in a close read of my first three chapters.

I would also like to extend a thank-you to Saskatoon Public Schools for supporting me in my desire to lead the best possible professional learning. I am privileged to work for a school division that has supported me financially through two post graduate degrees. Mr. Dave Derksen, Dr. Scott Tunison, and Dr. Kerry Alcorn have supported me in my academic journey outside of work and been my allies in becoming a better thinker and writer.

My study participants, who generously offered to describe their thinking and experiences, are both essential to this project, and to a person, deeply dedicated to high quality teaching and learning. I would like to offer particular thanks to Ms. Terry Johanson, the participant in the pre-study. The quality of Terry's critical thinking inevitably improves the thinking of those around her (I am no exception), and her self-reflection was very helpful in framing this study.

On a personal level, I would to thank my family, particularly Mike, Jodi, and Teela, for helping a dyslexic complete something of this length with relatively few errors, and for valuing and supporting me. All my debate children, biological and otherwise, continue to help me be a better thinker and speaker, and challenge me every day. I have to keep learning to keep up with them, and this study is a small part of their ongoing legacy to me.

## TABLE OF CONTENTS

PERMISSION TO USE .....	i
ABSTRACT .....	ii
ACKNOWLEDGEMENTS .....	iv
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
CHAPTER 1 .....	1
Introduction .....	1
Overview of the study .....	1
Need for the study .....	2
Context .....	4
Gap in the literature .....	8
Purpose of the study .....	9
Assumptions/delimitations .....	9
Positioning as a researcher .....	9
Choice of source materials .....	10
Assumptions about teachers and the profession .....	11
Definitions of terms .....	11
Structure of the Dissertation .....	14
CHAPTER 2 .....	15
Literature review .....	15
The logic of professional learning .....	16

What is effective professional learning? .....	19
Barriers to change in professional learning .....	23
What is good evidence of the impact of professional learning?.....	25
The problem of a lack of good evidence .....	28
Conducting internal evaluation in school divisions.....	32
The impact of the professional developer .....	41
CHAPTER 3.....	46
Research questions .....	46
Significance .....	47
Role of the pilot study .....	49
Narrowing the study design.....	51
Assessing “close” alternative methodologies .....	54
Study design .....	57
Data collection.....	59
Saturation.....	63
Data analysis.....	64
Strategies for increasing the quality of the study .....	67
Ethics .....	71
Participants .....	71
Addressing my perspectives .....	73
CHAPTER 4.....	75

Introduction .....	75
Participant’s descriptions of causal and intervening conditions.....	75
Systemic conditions.....	76
Expectation for evaluation.....	77
Theories of professional learning .....	82
Purpose of professional learning .....	83
Process of professional learning.....	90
Learning to evaluate .....	95
Resources.....	98
Evaluation and negativity .....	101
“Drinking the Kool-Aid” and professionalism.....	104
Personal causal conditions.....	110
Identity as a learner .....	111
Identity as a teacher .....	114
Relationship with professional learning .....	117
Identity as a professional developer or leader .....	119
Core phenomenon.....	124
Wonders.....	125
Extending wonders .....	129
Something to wonder with.....	132
Process of data collection: “finding out”.....	133



Planned assessment .....	133
Gathering information about a wonder.....	134
Attributing and sense-making through the facilitation lens .....	138
Likely interpretation .....	143
Action .....	144
Conclusion.....	151
CHAPTER 5.....	153
Introduction .....	153
Theory of professional developers’ self-assessment processes.....	155
Wondering.....	155
Theory of professional developers’ self-assessment processes.....	158
Finding out.....	160
Interpretation: attribution and sensemaking .....	163
Action .....	165
Implications .....	167
Significance of the leaps in professional learning logic .....	169
Common ground for self-assessment and evaluation.....	173
Gathering data about teachers’ use of new knowledge and skills.....	177
Data Literacy .....	178
Reducing barriers.....	179
Comparing self-assessment and program evaluation standards .....	182

Recommendations .....	183
Research guidelines to help professional developers .....	184
Making the implicit logic of professional learning explicit .....	185
Developing strong goals .....	189
Need for a systematic approach.....	191
Responding to issues of insider evaluation .....	193
Increasing observation.....	194
Conclusion.....	195
CHAPTER SIX .....	196
Theory summary.....	196
Conditions that influenced self-assessment or evaluation processes .....	197
Process for self-assessment .....	198
Theory.....	198
Further research .....	200
Implications of processes like coaching on self-assessment and evaluation.....	200
Study of “accuracy” of self-assessments .....	200
Reflections.....	201
Timing .....	201
Change to personal theory .....	202
Conclusion.....	203
References .....	204

APPENDIX A: .....	215
MAIN THEMES OF RESEARCH IN PROFESSIONAL LEARNING EVALUATION .....	215
APPENDIX B:.....	216
MIND MAP OF POTENTIAL METHODOLOGIES .....	216
DESCRIPTION OF STUDY.....	217
APPENDIX D: .....	218
INTERVIEW AND TALK ALOUD QUESTIONS .....	218
PARTICIPANT CONSENT FORM .....	220
APPENDIX F:.....	224
PARTICIPANT CONSENT FORM .....	224

## LIST OF TABLES

Table 5.1 – Desimone’s logic of professional learning as lived in professional developers’ self- assessment.....	171
Table 5.2 – Planning to use leading and lagging measures .....	187

## LIST OF FIGURES

Figure 3.1 – Decision-making factors in selecting grounded theory .....	57
Figure 4.1 – Systemic and personal conditions .....	77
Figure 5.1 – Wondering as a trigger for self-assessment.....	159
Figure 5.2 – Finding out by collecting data .....	161
Figure 5.3 – Interpreting data.....	164
Figure 5.4 – Taking action .....	166

# **CHAPTER 1**

## **Introduction**

The desire to transform educational practice flows through decades of research literature, pooling in the ideas of thinkers such as Dewey and Vygotsky. Many transformational goals, such as current interests in decolonizing curriculum or digital citizenship, swirl with older goals like constructivism or critical thinking, trapped together in an eddy of educational inertia. In each case, teachers' beliefs and instructional practices create an impermeable layer of established practice that prevents transformation unless the teacher decides, as a professional, to make different instructional and relational choices. Professional developers play a critical role in a teacher's decision to improve instruction, as they tend to be the intermediaries between new theory and current practice. The facilitators of professional learning make many decisions that influence the nature of teacher learning, and ultimately, how teacher instructional practice impacts student learning and engagement. Understanding the facilitators' thinking processes and how they are influenced is important in dredging for the transformation of teaching practice.

### **Overview of the study**

This study focused on understanding the processes that people who lead professional learning, including central office staff, coaches, administrators, and consultants and other professional learning leaders, use in order to self-assess the impact of the professional learning they lead. Using a grounded theory approach, the study utilized interviews and observations to develop theory to describe how professional developers articulate and demonstrate the process of self-assessment. The self-described processes were compared to research literature and implications of the similarities and differences were discussed.

## **Need for the study**

In K-12 education in Canada and the United States, professional learning is often financially supported by employers with the intent that participating in professional learning will result in a change of teacher practice (Fullan, 2007; Reeves 2010). A school division might want change because of a changing student population, like an influx of English as an Additional Language learners. New information available in research, like research about the value of a specific instructional approach, could also be a reason an employer might want a change. Similarly, new curriculum, or belief that some classrooms have inadequate instruction might also cause an employing school division to fund professional learning for a specific group or subject. However, regardless of content focus, the common denominator in funding professional learning is a belief that change is desirable, and professional learning is the main mechanism for sparking that change. To that end, school divisions, educational authorities like a Ministry of Education, and other interested groups, are focused on providing professional learning constructed in a way that is likely to support teachers in changed instructional practice that improves student learning. The series of assumed causal links provides underpinning for much of the criticism that professional learning is unlikely to be effective in changing student results (Fullan, 2007; Reeves 2010).

For the last decade and a half, professional developers have changed the structure and the substance of the professional learning as researchers debated the characteristics of effective professional learning. In Canada, professional learning prior to this time was typically structured as presentations (Broad & Evans, 2006) where teachers listened to information from experts. The logic underlying these so-called “sit and get” sessions of the 20<sup>th</sup> century was that if teachers

had a plethora of new information, teacher practice would change in a desirable way (Reeves, 2010). As widely-cited research claimed that a small fraction of what was learned in these sessions was ever implemented in classrooms (Joyce & Showers, 2002; Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009), more collaborative or job embedded models became the focus of theory about “best-practice” in professional learning (Broad & Evans, 2006; Darling-Hammond et al., 2009; Timperley, Wilson, Barrar, & Fang, 2007). A variety of new forms of professional development became the focus of experimentation and research, including collaborative inquiry groups, coaching models, and data teams. The debate about what structures and elements are essential for effective professional learning is ongoing, as research attempts to clarify the value of various approaches to teacher satisfaction and student learning.

In the 21<sup>st</sup> century, short professional development sessions in Canada and the United States have been augmented with other structures such as: multi-day workshops, professional learning communities, data teams, coaching, lesson study, and other interactive models (Broad & Evans, 2006; Timperley, Wilson, Barrar, & Fung, 2008; Darling-Hammond et al., 2009; Hirsh & Killion, 2009). The newer models are described as more engaging or effective in practitioner literature because they are directly connected to teacher’s classroom work, provide extended time to learn, and assume teachers as agents of change (Broad & Evans, 2006; Hirsh & Killion, 2009). The transition to increased experimentation with newer forms has been a response to the perceived implementation gaps left by older professional development sessions (Learning Forward, 2011).

Changes in teacher practice rarely transpire without effective professional learning (Guskey, 2000), so the perceived relationship between transformation of instructional practice



and effective professional learning has become a catalyst for change. Newer standards for what makes “effective professional learning”, like the *Learning Forward Standards* (2011), have become accepted frameworks for what professional learning should be and are often cited in professional periodicals, sometimes as a de facto requirement for publication in specific journals (<https://learningforward.org/publications/jsd/writers-guidelines> ). However, the change in professional learning practice towards these newer standards has yet to yield wide-ranging results that defend the whole scale change in professional learning practice, such as substantial changes in teacher pedagogy, transformational change in school divisions, or significant differences in student results. Research continues to find limited effects of the professional practice of teachers, leading to extended debates in the research about the efficacy of the reformation and probes into the causes of the potential gaps (Timperley et al., 2008).

## **Context**

In Saskatchewan, professional development for teachers is primarily the responsibility of their respective school divisions. While there is a formal set of expectations for school divisions described in the provincial Continuous Improvement Framework and strategic planning process, there is no form of provincial testing used to assess teacher instruction or student achievement in core curricula each year. At the time of the writing of this dissertation, one measure is required for extended time (*Our School*, a standardized national measure of student reported engagement) and variety of diagnostic measures are used in grades three and under (<https://www.saskatchewan.ca/residents/education-and-learning/prek-12-education-early-learning-and-schools/student-assessment>), predominantly to assess literacy skills. Mathematics and writing assessments in three grade levels were introduced in the fall of 2017.

While divisions remain responsible for the quality of instruction and learning, there are no specific consequences for failure or success on a common measure tied to funding, a school's reputation, or a school's enrollment. Continuous improvement remains an ideal without the specific consequences for success or failure that one might find in other jurisdictions in North America. Professional developers have the freedom and responsibility to assess the changes to student results with little provincial mandate or oversight. The freedom from high stakes testing has the additional consequence of rendering growth and change in teacher practice free from any form self-assessment or evaluation.

Provincially, the main group responsible for assessing the impact of professional learning is the people in school divisions who provide that development. Often, the professional developers are consultants, coordinators, instructional coaches, or superintendents based out of central offices. They may also be school-based teacher leaders or administrators. They are commonly tasked with introducing new research or ideas, or growing teacher expertise in areas like instruction, curriculum, or assessment. Like classroom teachers, they plan for what they will teach, and then lead the learning. Unlike classroom teachers they have no formal responsibility for conducting classroom assessment or evaluation, and rarely assess their personal impact on teacher learners with any regularity (Reeves, 2010).

The professional developers' lack of self-evaluation represents a gap that two decades of attempts to transform the quality and efficacy of professional learning have failed to fill. Guskey (2000) describes the most commonly used framework for the assessment of professional learning, which includes evaluations at five critical levels: "participants' reactions, participants' learning, organizational support and change, participants' use of new knowledge and skills, and

student learning outcomes” (p.82). Self-assessment of impact differs from formal program evaluation, but even extended program evaluation conducted by a professional cannot easily isolate professional development as the element that caused the change in student achievement (Guskey, 2000). However, data-based decision-making requires that professional developers be able identify if the professional learning is changing teacher practice and influencing student results with some accuracy. Newer forms of professional learning from the last two decades emphasize collaborative, job-embedded learning over time, which is even more complicated for professional developers to informally assess; the newer forms of professional learning are often offered by those embedded in schools rather than solely from central office or in conferences (Broad & Evans, 2006; Learning Forward, 2011; Timperley, Wilson, Barrar, & Fung, 2008). The changing nature of professional learning has made it both more difficult to gather data on centrally lead initiatives and hard to measure change over time with few constants.

The context in which potential evaluation may occur is also complex because of the nature of the support offered to professional developers. Studies focus on teacher perception or standardized test results for students, and causal links between newer forms of professional learning and impacts are not attempted or unclear (Desimone, 2009). In addition, many studies are written in academic rather than professional language and published in journals educators rarely access (Tunison, 2016). As researchers have noted these complexities, they are beginning to offer more plain language guides that are readily available online, like Blitz & Schulman’s (2016) summary of potential evaluation tools for professional learning communities. However, even these tools require an understanding of logic models of program evaluation, access to

research, and a specific understanding of the exact goal of the research. Despite being plain language, the summary in the example below is not simple to follow:

Step 1: Consult appendix B (the PLC logic model) to determine what information you want and for what purpose (planning, implementation, or evaluation).

Step 2: Depending on your goal (planning, implementation, or evaluation), use the table in appendix C to identify the key tasks you are interested in (column 1) and select the key indicators you are most interested in measuring (column 2). Finally, note the class or classes of measurement instruments most relevant to measuring the indicators of interest (column 3).

Step 3: Click on the hyperlink of a specific instrument in appendix D's table of contents to retrieve the information about the instrument, determine whether it meets your needs, and find out how to obtain a copy. (Blitz & Schulman, 2016, p.3)

Information about the impacts of professional learning and how to evaluate them is a complex problem for professional developers, school divisions, and other stakeholder groups.

Professional developers, of all stakeholders, have the best access to the decision-making processes around professional learning. As the main leaders of professional learning in the field, they know the reasons why specific strategies and processes are used or omitted, and they have the greatest incentive to change professional learning based on their own results. While many decisions are made independently of professional developers' practice, they are tasked with considering those factors when deciding about how to structure the professional learning processes designed to change instructional practice. Understanding the thinking of professional developers provides critical, if not wholly sufficient, information about how the professional

learning is composed and evaluated. While there has been little focus on the process professional developers use to make determinations about the impact of professional learning, their conceptualization of the process and their underlying assumptions can inform the field about evaluating professional learning in praxis. This study uses the voices of professional developers to better understand the processes of, and conditions of, professional developers' self-evaluation.

### **Gap in the literature**

Professional developers are the central decision makers in professional learning in K-12 education. They provide materials, plan processes, convene and resource groups, and determine what the learning focus will be and for how long, within the constraints of budgets. Most factors that influence the quality of professional learning are influenced or directly controlled by professional developers, absent of ongoing guidance or interest from senior administrators, stakeholder groups, researchers, or the provincial Ministry of Education (Johanson, personal communication, Oct. 6, 2015). Professional developers are often the sole stakeholders making decisions about professional learning, including choosing how to self-assess or evaluate it. Theory about the decision-making of the professional developer is essential in understanding factors that impact evaluation, and the subsequent changes to professional learning, but such theory is largely absent in the research literature. Understanding the thinking processes and evaluation practices of professional developers can help determine why professional learning evaluation is less common and less effective. It can also provide guidance for actions designed to make professional learning evaluation more common and useful in K-12 school divisions.

## **Purpose of the study**

Most practitioners do not have the time to evaluate professional learning, and when they do, they focus on the organization and content of the learning experience (Hirsh, 2013). This study attempts to understand the process professional developers use to determine the impact of their work with teachers.

The main research question of this study is, “How do professional developers describe and demonstrate the process of self-assessment?” The sub questions will include:

1. How do professional developers decide which strategies to use to determine the impact of their professional learning?
2. What influences the perpetuation of evaluation methods (or lack thereof)?
3. How is the praxis of professional learning influenced by the professional developer’s experience?

## **Assumptions/delimitations**

### **Positioning as a researcher**

I am an insider researcher (Bonner & Tolhurst, 2002) in professional development theory, and began my formal research in 2002, while still a teacher. My school division granted me a year’s leave in the 2005-06 academic year, when I started my master’s program and completed my thesis, a narrative inquiry entitled *Pedagogy and Practice: The effective use of computers through the eyes of professional developers*. In 2006, I returned to the field and progressed from working as a coach and teacher support person, to a teacher-leader, a consultant, and finally, a coordinator of curriculum and instruction. My time in academia has always been brief, bracketed by professional practice. My focus is praxis and my lens is pragmatic. My experience

with professional learning has lead me to a rich conceptualization of the structures, context, processes, and impacts of professional learning in the provincial context, and I have a strong interest in the transformation of education. The combination of pragmatic thought processes and transformational interests focuses my thinking on understanding the mechanisms needed to support professional learning in creating change in instructional practice.

Insider researchers closely study the group they belong to (Bonner & Tolhurst, 2002). They may have greater access, understanding of the culture, and the ability to act naturally and maintain intimacy with participants (Bonner & Tolhurst). I position myself as a co-investigator and co-learner, particularly as I am striving to generate theory that resonates with my participants as an accurate framing of their thinking processes. Like other insider researchers, I need to be wary of making assumptions based on prior knowledge, confronting myself in the data I am collecting, and responding to ethical issues associated with insider status. A detailed analysis of potential issues and steps to address them is found in chapter three.

### **Choice of source materials**

Like many jurisdictions, Saskatchewan experiences a substantive gap between the informational available to researchers and academics and the information available to practitioners (Tunison, 2016). The theory-practice gap results in classroom practice that may not be research supported, and research that may not seem relevant to teachers and students. In this dissertation, I have made the deliberate decision to consult both academic journals and texts, and the journals and books commonly circulating among practitioners. While it may be valid to critique the practitioner-focused materials as less rigorous and less current, the academic work may have been seen as esoteric to some teachers. The focus of the study is the thinking of

professional developers who attempt to read both types of materials, and it is my goal to understand their thinking. I have used both types of materials, and where possible, found common ground between them.

### **Assumptions about teachers and the profession**

As a teacher, I assume that teachers wish to help students and support their learning. I also believe teachers are professionals and should make judgement calls about their instructional process based on the students in front of them and their contexts. I think teachers have an obligation to teach required curriculum and that our ongoing growth as instructors is key to student success. I see professional learning as one vehicle for helping teachers meet the needs of students. I reject pre-packed instructional scripts and other interventions as unresponsive and disengaging, leading me to view certain forms of professional learning with skepticism.

### **Definitions of terms**

Terms like “professional learning” and “professional development” are sometimes used without clear distinctions, making it difficult to know what is being referred to in the literature and professional materials. Traditionally, “professional development” referred globally to training offered to teachers in events or activities like workshops, lectures, post-secondary schooling, and conferences (Villegas-Reimers, 2003). Now the term is often used synonymously with the more preferred phrase “professional learning”, which describes a much wider variety of both formal and informal activities that can affect teacher knowledge or classroom practice, including personal processes like reflection (Borko, 2004; Desimone, 2009; Guskey, 2000; Ingvarson, Meiers, & Beavis, 2005). Increasingly, research recognizes that the most effective



opportunities to learn are as close to teachers' practice as possible (Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009; Reeves, 2010; Villegas-Reimers, 2003) and researchers describe a wide range of activities from formal learning to hallway conversations and personal reflections as professional learning (Desimone, 2009; Guskey, 2000). Professional learning in the formal sense is typically facilitated or lead by professional developers, who sometimes have formal roles in school divisions or districts as consultants, coordinators, administrators, or coaches. Other professional learning occurs informally and has no appointed leader, or is led by teachers themselves. The broad understanding of what might be professional learning, and the wide range of people who lead it, makes the already challenging problem of evaluating the impact of these activities profoundly complex (Guskey, 2000; Reeves 2010).

In the last decade and a half, professional learning has gone from episodic training to something much more embedded and essential. In addition to the widening of the definition of how and why praxis develops, professional learning is now viewed as critical to the improvement of student learning because it is the precursor to enhanced praxis. While it is openly acknowledged that most professional learning in a teacher's career is never evaluated, evaluation is essential when we want the best possible professional learning to help improve student results (Guskey, 2000; Hirsh, 2013). Innovation in education, driven by a changing world, necessitates growth and change in teaching practices (Earl & Timperley, 2015; Schwille, Dembélé, & Schubert, 2007). However, substantial changes to teacher practice require thoughtful, well-executed professional learning as a precursor (Guskey, 2000). Links between professional learning and student learning (Darling-Hammond et al., 2009; Guskey, 2000; Hattie, 2009; Marzano, 2007) and general dissatisfaction with the quality of professional learning

among teachers have resulted in substantive change in professional learning practices (Broad & Evans, 2006; Darling-Hammond et al., 2009; Reeves, 2010). However, so many changes have occurred simultaneously there are no fixed variables, and both researchers and practitioners have struggled to establish which actions, forms, content, or elements are essential for high impact professional learning processes (Boyko, 2004; Desimone, 2009; Earl & Timperley, 2015; Taylor, Getty, Kowalski, Wilson, Carlson, & Scotter, 2015). Professional learning practice has changed dramatically to be theoretically more likely to enhance teacher instruction, but it has now become even more difficult to measure and increasingly diffused.

To discuss the evaluation of professional learning, it is essential to agree on a clear definition of the distinction between general assessment of professional learning and evaluation of professional learning. Guskey defines evaluation as “the systematic investigation of merit or worth” (2000, p.41) and describes it as being purposeful and thoughtful. He depicts an evaluation process as collecting relevant, valid information to rendering judgement about the professional learning. Guskey distinguishes between merit and worth in an evaluation of professional learning. Questions of worth are about the value to the professional learning in the context of our goals and needs, and they help us think about big issues like alignment, needs assessment, and relative benefits. Merit, however, is about the excellence of a program in meeting its own goals; merit considers the program’s success. Educators typically need to consider both merit and worth when thinking about professional learning, and the process requires informed judgement.

Issues with evaluation of professional learning begin with the word “evaluation”, because it is misconceived as judgement of teachers, rather than judgement of professional learning.

Nearly all the professional developers in the study have used the word “assessment” instead, and stated it is acceptable to gather general information about teachers but not to render judgment. While assessment measures can be very helpful, particularly as formative assessment that shapes professional learning in progress, they are not a substitute for evaluation, which is essential to help appraise impact. Evaluating professional learning must be imbued with systematic collection of data about teacher practice in a way typically characterized as innately threatening. Understanding professional learning evaluation processes in school divisions will depend on distinguishing between professional developers’ generalized assessments of professional learning and the systematic investigation of merit or worth.

### **Structure of the Dissertation**

This first chapter describes the context for the study, terminology, and the lens of the researcher. Chapter two is a review of related literature around professional learning practice, evaluation processes, and thinking processes, and an overview of current practice. Chapter three discusses the process of the research, including the selection of the methodology, the process of ethics approval, and attempts to increase the quality of the study. Chapter four is an overview of the data collected, while chapter five discusses the development of the theory and its implications.

## CHAPTER 2

### Literature review

“Too many schools remain captivated by staff development experiences that were discredited in previous centuries. Indeed, the critical thinking in the Lyceum of 325 BCE exceeds that of the purveyors of contradictory and evidence-free staff development” (Reeves, 2010, p.3). Effective processes for professional learning are at the heart of any substantive, useful change in how teachers teach and students learn (Guskey, 2000). However, knowing the most robust ways to conduct and evaluate professional learning is problematic. Current practices of professional development are often distinct from what research considers good professional learning (Darling-Hammond et al., 2009) and may miss any form of assessment or evaluation that is likely to improve them (Reeves, 2010). Professional learning is ubiquitous as the main way to support the evolution of teacher instructional practice, but the understanding of its impacts is scarce:

PD is required by virtually every work contract in the country, and teachers participate in PD every year. Foundations and federal agencies spend large sums on the design and implementation of PD programs. Yet despite widespread agreement about its importance, there is little consensus about how PD works, that is, about what happens in PD, how it fosters teacher learning, and how it is expected to alter teacher practice. The actual form and substance of PD programs is tremendously various, raising questions about why something so various is uniformly assumed to be a good thing. (Kennedy, 2016, p.945)

Professional learning is routinely offered, but often devoid of the logic found in formal program design and evaluation, resulting in substantial investments of time and resources without clarify

of impact or return. An uneven mix of the self-assessments of the people who lead professional learning, school division strategic planning or data-based decision-making, and rare formal evaluation provide a patchwork of attempts to understand if investment in professional learning is paying dividends.

Formal program evaluation, the general forms of assessment attempted in data-based decision-making, or the self-assessment of professional developers have two key elements in common. First, they share a common purpose of understanding the merit or worth of a professional learning process. Second, the ways of considering professional learning are all parts of a continuum that mirrors a basic process in a classroom that utilizes desired outcomes, learning experiences, and assessing impact or evaluating results. Like teachers, professional developers must effectively self-assess the value of what learning they support; improvements in the outcomes of professional learning require those with the greatest agency have the tools to appraise and enrich those learning opportunities.

### **The logic of professional learning**

When program evaluators assess the value of an activity, they start by trying to understand its purpose and internal logic (Owens & Rogers, 1999). Referred to as a logic model, the process has four key elements: the inputs, the goals, the actions, and the projected impacts in the immediate, intermediate, and long terms. Logic models are critical because they help to make the tacit logic and assumptions of a process explicit, and make it possible to render the judgements that distinguish evaluation (Hill, Beisiegel, & Jacob, 2013; Owens & Rogers, 1999; Worthen, Saunders, & Fitzpatrick, 1997). Professional learning in K-12 education rests on underlying assumptions about the circumstances under which learning occurs and the existence

of a specific sequence that theoretically occurs. First, the teachers' understanding improves as a result of the professional learning experience. Next, teachers improve practice in the way suggested in the experience, resulting in specific improvements to student learning, engagement, or some other learning related target.

The logic of professional learning faces challenges both in its implementation and in its measurement. Critics argue that professional learning is not focused on implementing change, but is actually antithetical to change in practice, because it is a discussion of general theory, poorly delivered. Fullan (2007) contends professional learning lessens the push for change by focusing educators away from addressing major underlying causes, and towards brief activities that tinker at the surface and do not have results. Because of the way activities are offered, preparing for change may be the best case scenario for many professional learning experiences (Cole, 2004). Instead of condemning professional learning outright, much of the opposing thinking highlights the way professional learning is delivered as the reason it is ineffective (Birman, Desimone, Porter, & Garet, 2000; Garet, Porter, Desimone, Birman, & Yoon, 2001; Darling-Hammond et al., 2009, Reeves, 2010). Practitioners and researchers call for investments in professional learning to provide a more substantive impact on student results, indicating funding may be another challenge.

Even if professional learning was delivered exceptionally skillfully, however, the assumptions that underlie the logic model are potentially problematic. The assumptions establish beliefs about education that can be challenged, rendering the causal thinking behind professional learning questionable. Desimone (2009) describes the logic:

1. Teachers experience effective professional development.

2. The professional development increases teachers' knowledge and skills and/or changes their attitudes and beliefs.
3. Teachers use their new knowledge and skills, attitudes, and beliefs to improve the content of their instruction or their approach to pedagogy, or both.
4. The instructional changes foster increased student learning.

This model allows testing both a theory of teacher change (e.g., that professional development alters teacher knowledge, beliefs, or practice) and a theory of instruction (e.g., that changed practice influences student achievement), both of which are necessary to complete our understanding of how professional development works. (p. 184-185)

The four assumptions seem to be at the heart of most forms of professional learning, regardless of structure. If the goal is to offer professional learning that improves over time, there is one more assumption, namely that the education sector has the resources and understanding to make the causal link between professional learning and improvements using identified measures.

The questionable nature of the links in the logic of professional learning makes the decision-making processes of professional developers challenging. If any of the assumptions is incorrect, then it is not likely that professional learning will positively affect student results. While many teachers might want to grow their practice over time, it is unlikely they all want to change their current practice in the way professional developers intend (Kennedy 2016). In addition, adult learning and change is a complex and difficult process that cannot assume all teachers wish to grow and change their practice (Cuban, 1993; Fullan, 2007; Kennedy, 2016; Mezirow, 1991). Similarly, it is difficult for educators like professional developers and teachers to determine where growth and change is needed and there are many resources designed to help

educators be more accurate in assessing need (Dufour, Dufour, Eaker, & Many, 2010; Hume, 2009; Peery, 2011). If change is substantively resisted or efforts are misdirected, it is unlikely that professional learning will result in improvements for students.

The next assumptions, namely the success of methods of professional learning and that the growth in teacher practice will create the desired change in student results, are both hotly contested within research (Birman et al., 2000; Garet, et al., 2001; Borko, 2004; Darling Hammond & Richardson, 2009; Hirsh 2013; Kennedy, 2016; Reeves, 2010). Since effective methods and change in teacher practice would both be need to be true to have the desired impact, it is vital that the education sector be able to determine what activities are most likely to lead to improved student results. However, the lack of information about what professional developers can do to garner results is frustrating. While research has clearly established that no significant changes in education occur without effective professional learning (Guskey, 2000), and pressure to create change is mounting in education systems world-wide (Earl & Timperley, 2015), effective, frequent assessment of professional learning remains elusive.

### **What is effective professional learning?**

Of the four assumptions, the assumption that professional learning methods are effective has been the most broadly challenged in recent professional writing, and the push for more effective professional learning has primarily focused on the transformation of the process (Broad & Evan, 2006; Darling-Hammond et al., 2009; Elmore & Burney, 1997; Fullan 2007). There have been substantial changes to the philosophy and basic structures of professional learning in the last 10 to 15 years. The “sit and get” models (Fullan, 2007) that dominated the 20<sup>th</sup> century teacher training were based on a transmission-oriented process of thinking that teachers would



hear about how they could be better, understand it well, and want to make the change. Then the same teacher would make the change without supports, and it would go well despite time constraints, lived realities, and a wide variety of competing pressures. While sit and get sessions are helpful in sharing information or changing technical procedures like learning new software, oft-cited summaries of current research suggest this form of professional learning alone rarely results in improved student achievement (Broad & Evan, 2006; Darling-Hammond et al., 2009). Common referenced critiques (Darling-Hammond et al., 2009; Elmore & Burney, 1997; Fullan 2007) focus on teachers' descriptions of professional learning as boring, irrelevant, unlikely to cause change, too brief, and consistently under-resourced.

Although meta-analysis reveal that many teachers in North America are still experiencing mostly older forms (Broad & Evan, 2006; Darling-Hammond et al., 2009), newer forms are dominating research, and are being experimented with in divisions throughout Saskatchewan (Johanson, personal communication, Oct. 6, 2015). As early as the 1990s, theorists and academics were discussing the need for a shift and debating what elements might be essential in professional learning that would influence student results. The often referenced Elmore and Burney (1997) article argued there was clear evidence of what worked at the end of the century, but a more general taking up of the ideas in both research and the K-12 sector required a more extended timeframe (Hill, H. C., Beisiegel, M., & Jacob, R., 2013):

We know a good deal about the characteristics of successful professional development: it focuses on concrete classroom applications of general ideas; it exposes teachers to actual practice rather than to descriptions of practice; it involves opportunities for observation, critique, and reflection; it involves opportunities for group support and collaboration; and it

involves deliberate evaluation and feedback by skilled practitioners with expertise about good teaching. (Elmore and Burney, 1997, p.2)

Garet et al. (2001) and Desimone (2009) also articulated there is common agreement about what is good praxis in professional learning, and their work is followed by a decade of different researchers articulating common principles of good professional development using different frames and terms, but with similar ideas (Blank, 2013; Hill et al., 2013), including:

1. Good professional learning is driven by student need and rooted in teacher practice (Garet et al. 2001; Elmore and Burney, 1997; Learning Forward, 2011; Timperley et al., 2007; Reeves, 2010).
2. Teachers are active, ongoing participants rather than passive listeners (Blank 2013; Broad & Evans, 2006; Darling-Hammond and Richardson, 2009; Dufour et al., 2010; Elmore and Burney, 1997; Garet et al. 2001; Kennedy, 2016; Learning Forward, 2011; Timperley et al., 2007; Reeves, 2010).
3. Professional learning is as directly related to the classroom as possible and designed to address a specific student need (Blank, 2013; Broad & Evans, 2006; Darling Hammond et al., 2009; Dufour et al., 2010; Elmore and Burney, 1997; Learning Forward, 2011; Reeves, 2010).
4. Professional learning has a sustained focus over time, although the amount of time or duration is a point of debate. Blank (2013), Broad & Evans (2006); Darling Hammond et al. (2009) and Garet et al. 2001 focus on time, while Guskey (2009) and Timperley et al. (2007) argue effective use of time is at least as important as amount of time.

5. Collaboration with other teachers is an essential element (Blank 2013; Broad & Evans, 2006; Darling Hammond et al., 2009; Garet et al. 2001; Kennedy, 2016; Learning Forward, 2011; Timperley et al. 2007).

Other elements remain supported by some thinkers (Garet, et al., 2001; Hawley & Valli, 1999; Ingvarson et al., 2005), like content focus. Kennedy (2016) notes an exclusive content focus is associated with less effect on student learning. An attempt to reach common agreement is found in standards like the professional learning standards offered by Learning Forward (2011).

While a sense of best practice in professional learning seems to congeal into consensus, a sense of why those forms are important and how they should be offered is also developing. A particular challenge is that while newer forms are associated with higher satisfaction among teachers, they have been difficult to directly link to improved student performance (Hill et al., 2013). In addition, some researchers argue that newer forms may appear to be more effective simply because they are allocated more time than traditional forms, and then automatically have more of the key features like teacher interaction (Garet et al., 2001). Hill et al. (2013, p.486) contend outright there is no evidence that newer forms are superior:

Through studies conducted over the past two decades, scholars have identified program design elements thought to maximize teacher learning, including a strong content focus, inquiry-oriented learning approaches, collaborative participation, and coherence with school curricula and policies (e.g., Cohen & Hill, 2001; Garet et al., 2001; Penuel, Fishman, Yamaguchi, & Gallagher, 2007). Agreement about this list had reached a level such that many in the field felt comfortable characterizing support for the list as a

consensus (Desimone, 2009; Penuel et al., 2007; Russell, Kleiman, Carey, & Douglas, 2009). Yet disappointing results from recent rigorous studies of programs containing some or all of these features have turned this consensus on its head (Arens et al., 2012; Bos et al., 2012; Garet et al., 2008; Garet et al., 2011; Santagata, Kersting, Givven, & Stigler, 2011).

A consensus, possibly without proof, is further complicated by the theory-practice gap in professional learning practices; much of the professional learning teachers experience does not resemble the experiences described by the reform movement (Broad & Evans, 2006; Darling Hammond et al., 2009). Absent clear evidence about what professional developers can do to support teachers in improving student results, professional learning fails to discharge the duty of school improvement. Even if professional developers can overcome the barriers to changing and evaluating their own professional learning facilitation, they risk of replacing old, ineffective practices with barren, novel processes. Professional learning process are significant because they are highly associated with the likely impact of the professional learning (Kennedy, 2016), but also because of what professional learning seeks to do: “Any new idea offered by PD requires not merely adoption but also *abandonment* of a prior approach” (Kennedy, 2016, p.984). Trying to cause abandonment of previous practices is not merely related to understanding of new approaches, it is influenced at every turn by a wide variety of types of barriers that a professional developer must consider.

### **Barriers to change in professional learning**

Barriers to professional learning evaluation include assumptions of professional learning, the amount of change, and the complexity of sharing information between researchers and practitioners. Each element is an impediment to improvement in professional learning and its

assessment. However, many obstacles are established within the structures in field, and each is substantial.

Resourcing from central offices is one of many reasons that professional learning does not meet the basic requirements to cause change, let alone improvements in student results. Newer forms require much more lead time and planning, which many divisions do not have, potentially requiring more than double current funding levels (Garet et al., 2001). Evaluation is challenging as “central office leaders struggle to find the time to capture, think about, and respond to data” (Honig, Copland, Rainey, Lorton, & Newton, 2010, p. 2), let alone making effective professional learning decisions based on data or monitor results. Both professional developers and school-based leaders of professional development struggle with the skills to lead new forms of professional learning (Stein, Smith, & Silver, 1999).

Some elements of effective professional learning are actively resisted by teachers although the majority of teachers characterize traditional professional learning as not being useful (Darling-Hammond et al., 2009). For example, “teachers generally avoid using the most effective means for promoting professional learning within the school - classroom observation, feedback, and lesson study” (Cole, 2004, p. 2). Cole (2004), Fullan (2007), and Reeves (2010) also contended teachers may not have intentions to create change, or may describe the need to change and grow as resting with others, like parents or students. In addition, many school leaders must take more responsibility for establishing a professional learning culture within the school (Cole, 2004; Fullan, 2007; Reeves, 2010), as professional learning is often an event that happens with no intention of any change in teacher practice (Fullan, 2007). The combination of

division-level barriers and teacher barriers make leading and evaluating professional learning designed to improve instruction fraught with obstruction.

### **What is good evidence of the impact of professional learning?**

Program evaluation has its roots in education. The *Encyclopedia of Evaluation* describes program evaluation broadly, encompassing all potential elements:

Evaluation is an applied inquiry process for collecting and synthesizing evidence that culminates in conclusions about the state of affairs, merit, worth, significance, or quality of a program, product, person, proposal, or plan. Conclusions made in evaluations encompass both an empirical aspect (that something is the case) and a normative aspect (judgement about the value of something). It is the value feature that distinguishes evaluation from other types of inquiry, such as basic scientific research, clinical epidemiology, investigative journalism, public polling. (Fournier, 2005)

Although there are many ways to study the impacts of professional learning in schools, any form of evaluation is rare (Hirsh, 2013; Reeves, 2010; Sherman, Tibbetts, & Condelli, 1997). There is strong agreement in the research that evaluation of professional learning from within divisions or districts is focused almost entirely on participant reaction to professional learning (Guskey, 2000; Ingvarson et al., 2005; Broad & Evans, 2006; Desimone, 2009; Desimone 2011; Hirsh, 2013; Mark, 2014). This focus is problematic because while teacher engagement in learning may be a precursor to change, it is not a sufficient measure of professional learning leading to improved student results (Guskey, 2000; Hirsh 2013). Professional learning structures that are genuinely accountable require more complete evaluation.

Historically, forms of program evaluations are enmeshed with calls for sector-wide accountability and the push for continuous improvement in education. School inspections, Horace Mann's (1865) work, or Smith and Tyler's *Eight Years Study* (1942) are all examples of an early desire to understand the impact of educational practice. Attempts to deeply understand the impacts of social sciences interventions in the 1950s and 1960s culminating with *National Defense of Education Act* in 1958 in the United States, helped to shape program evaluation as we understand it today (Worthen et al., 1997). Like its modern relative, the now repealed and critiqued *No Child Left Behind Act* (2001), the foundational *National Defense of Education Act* (1958) was responsible mandating forms of program evaluation in the United States and the growth of specialists and codification. Each major piece of legislation resulted in the assessment of changes in student results, and an attempt to determine what school districts could do to increase the pace of positive change. Improved learning and improved instruction were summarily linked, and a discussion of methods of teacher and school improvement ensued.

Unlike the United States or some parts of Canada, Saskatchewan has few forms of standardized assessments that might generate external accountability and support extended academic study or program evaluation. In the last decade, provincial *Assessment for Learning* and other similar provincial measures of student achievement in core subject areas have been replaced with *Our School*, a national student perception survey focused on engagement. Different measures of early literacy rates have also been introduced as formative assessments. In addition, Saskatchewan participates in the Organization for Economic Co-operation and Development's (OECD) *Program for International Student Assessment* (PISA), but the test is not specific to local curriculum and student performance cannot be directly tied to causal factors

in specific instruction. Failure to change student results has no specific consequences and does not required local action. Because of a lack of external accountability, formal program evaluation in Saskatchewan is relatively rare. Regardless of one's philosophical position on large programs of standardized assessments, there is no current provincial mechanism for program assessment of province-wide professional learning to improve the quality of instruction or student results.

The provincial Ministry of Education and its system partners are currently in a strategic planning process to attempt to establish agreement about, and actions related to, targets related to student achievement, engagement, literacy, and graduation rates

([http://publications.gov.sk.ca/documents/11/85636-](http://publications.gov.sk.ca/documents/11/85636-Education%20Strategic%20Sector%20Plan%20Matrix%202014-2020.pdf)

[Education%20Strategic%20Sector%20Plan%20Matrix%202014-2020.pdf](http://publications.gov.sk.ca/documents/11/85636-Education%20Strategic%20Sector%20Plan%20Matrix%202014-2020.pdf) ). As schools are pressed to improve credits earned, student retention, literacy and other agreed upon measures, the attention to factors that might influence results is increasing. However, as Reeves (2010) notes, a growing push for accountability has bypassed professional learning:

The fundamental thesis of effective accountability systems is that our zeal for holding 9-year-old children accountable should be matched by enthusiasm for assessing the performance of adults in the system. In the decade since Guskey implored schools to take a more rigorous approach to assessing professional learning, however, accountability is least in evidence in staff development seminars. (p. 2)

The lack of school division focus on evaluation of professional learning has a myriad of causes as complex as the process of evaluation itself, and the problem is rooted in the culture, process, and knowledge base of the divisions themselves. However, the evaluation problem has



not been solved by either academics or program evaluators for equally complex reasons, although the three factions, academics, program evaluators, and practitioners, are all observing separately that evaluating professional learning is essential in improving professional learning, and therefore, student results (Earl & Timperley, 2015; Desimone, 2011; Guskey, 2000; Hirsh, 2013; Reeves 2010).

### **The problem of a lack of good evidence**

While Sherman, Tibbets, and Codelli argued in 1997 that the days of unfocused and unexamined professional development have passed, the practice persists to this day. Guskey's work (2000) in academia, then Reeves' (2010) and Hirsh's (2013) writing targeted at the profession a decade later, all contend that school divisions struggle to gather good evidence for many deeply-rooted reasons. Desimone (2011) noted that "for decades, studies of professional development focused mainly on teacher satisfaction, attitude change, or commitment to innovation, rather than professional development's results or the processes that make it work" (p. 68). There are three key reasons change has not been forthcoming despite unprecedented accountability pressures.

The most significant problem is one of access to teachers and understanding of why evaluation is important. School division personnel do not understand the need for evaluation of professional learning (Hirsh, 2013) and even when they do, "administrators and evaluators lack the skills, experience, and ability to effectively evaluate the true impact of professional development" (Mark, 2014, p. 1648). Second, as professional learning becomes more school-based and less centrally led, professional developers are no longer conducting all of the professional learning (Johanson, personal communication, Oct. 6, 2015). Much of the

professional learning being offered is led by teachers and administrators with little expertise or experience in even planning professional learning. In addition, professional learning “is home-grown; it arises from district or local developers’ needs and interests, has a relatively short shelf-life, and proceeds with little or no formal evaluation” (Hill et al., 2013, p. 476).

The other two main issues, expertise and resources, are both related to gathering data, turning it to evidence, and rendering sound judgments. Hirsh (2013) noted that isolating the impact of professional learning “from other factors that promote school and individual improvement is challenging” (p. 12) and most school systems and schools have neither the research capacity nor time to link any activity to any specific outcome. It is not essential that divisions be able to establish the causality that a researcher might determine with a formal study (Guskey, 2000), but the reasonable likelihood of a link is even too complex for many divisions to establish. As school districts shift to more data-based decision-making, the final main problem of gathering and using data has been revealed. Most lack infrastructure or resources to use the data they gather (Hirsh, 2013) and they do not know how to determine what data they would need to assess if professional learning is effective because of a lack of guidance from researchers (Taylor et al., 2015). Even if division leaders knew which programs are worth offering to impact learning, they do not have the capacity to gather and use data about teacher practice, which is essential to seeing a relationship between professional learning and student results (Hirsh, 2013; Sherman, Tibbetts, & Condelli, 1997).

The data gathered by academics and researchers should provide clear guidance to help divisions wrestle with the problem of high impact professional learning, but it largely does not (Desimone, 2009). The common problems of transforming educational research to praxis

abound: lack of access to academic research in school divisions, academic writing is disconnected from daily practice, and the possibility that form and language of academic writing may be inaccessible to teachers (Tunison, 2016). However, the problems are compounded because research is failing to provide clear guidance because of its foci; research continues to focus on evaluating specific programs, rather than large scale comparisons of forms and process that can be generalized for planning purposed professional learning (Desimone, 2009). The questions that divisions currently must answer to perform effective internal evaluation require substantive comparisons, not small non-generalizable studies (Mark, 2014; Taylor et al., 2015). The logical links between professional learning strategies and changes in teacher knowledge, classroom practices, and student outcomes require large-scale studies with the capacity to test these relationships in many circumstances (Desimone, 2009; Ingvarson, et al., 2005). Research studies have been predominantly naturalistic and descriptive studies using ethnographic or case-study methods (Desimone, 2009), providing a plethora of theory about how teachers are experiencing professional learning, potentially leading school divisions to think that asking teachers about their experiences is the best link that can be made.

What divisions really need are studies that ask questions about impacts, trends, and connections (Desimone, 2009) and that report their results in a way educators can understand, allowing school divisions to make good decisions about how to offer professional learning that is likely to be effective. The need for clear evidence is pressing based on the pace of reform (Taylor et al., 2015) and compounded by the studied pace of research efforts. Most school division initiatives are over before research could study them, and the pace of academic research can mean professional learning practice has moved on before research even provides clear

guidance (Desimone, 2009). In addition, school divisions may block access to students and teachers for privacy concerns, ethical considerations, or a general sense that the research will not provide value to the division. Access, timeliness, and cooperation for mutual benefit remain major barriers for researchers and school divisions.

There are encouraging steps to link evaluation and newer forms of professional learning. One such example is related to professional learning communities, which are collaborative structures enjoying popularity in Canada and the United States (Dufour et al., 2010). Summary documents like Blitz and Schulman's (2016) aggregation of potential measures are helpful in many ways. The document starts by explaining the need for a logic model in evaluation, and describes one for professional learning communities that is sufficiently general and related to the common elements. Next, it connects specific elements of professional learning to what professional developers might need to know about them to determine if they are effective. Finally, it links to research studies assessed as part of a meta-analysis that might provide examples of evaluative metrics that school divisions might access. Although the access to the studies in journals may still be a problem, Blitz and Schulman's summary is publicly available online and written in professional, non-academic language, reducing both obstacles to access. Summaries of evaluation practices in plain language hold the potential to connect research and practitioners through the process of knowledge brokering (Levin, 2011; Sudsawad, 2007; Weiss, 1979). While relatively rare in education, theory regarding the practice in health is substantial, and may be referred to as knowledge brokering, knowledge translation, knowledge transfer, or knowledge mobilization, depending on its characteristics. Research can move beyond

conceptual and symbolic use to instrumental use if research is deliberately and specifically translated into easy to use formats (Sudsawad, 2007).

Like research and division practice, formal program evaluation has proved inadequate to meet the majority of the need. The evidence gathered by program evaluators is better designed to provide useful information about impact, but is not focused enough on design to support professional developers in decision-making (Borko, 2004; Hill et al., 2013). The problem is essentially that single program research studies assess a package of professional development, not its specific features that professional developers need to make decisions (Wayne, Yoon, Zhu, Cronen, & Garet, 2008). If the program seems positive based on teacher comments, student results, or other measures, it is hard to discern what factor or factors were causal, and the same issue exists if the outcome is negligible or negative. A second problem is that most professional development is locally developed and implemented; the life cycle of any approach or program is relatively short (Hirsh, 2013; Killion, 2013) and there is little access for evaluators. With a shifting policy agenda, professional development is frequently designed and redesigned to meet teachers' and districts' needs, and program evaluations are too infrequent or longitudinal to provide critical information (Hill et al., 2013). Finally, many forms of program evaluation have a poor record of being employed to shape practice, because they are written in inaccessible forms, not given to those who could most use them, or are do not contain practical, easy to implement solutions (Patton, 2008).

### **Conducting internal evaluation in school divisions**

Increased focus on interactive, utilization evaluations (Patton, 2008) and a comprehensive focus on measuring the impact of forms of students rather than assessing specific programs

(Desimone, 2009) can yield only some of the clarity that school divisions need. While they can make judgements and conclusions about professional learning impact more accessible and generalizable, they cannot address the systematic problems with evaluating professional learning within school divisions. Since causal links may be difficult for divisions to find (Hirsh, 2013), championing and normalizing the process of internal evaluation is essential if schools and divisions hope to offer professional learning that impacts student results.

Models for practical evaluations as a part of professional learning practice in school divisions have their roots in Kirkpatrick's (1959) model. They look at what new understanding educators develop, what they do in the classroom, and what impacts their actions have on students (Desimone, 2009). Guskey (2000) proposed the most commonly referenced set of stages for evaluation of professional learning:

1. Educators' reactions
2. Educators' understanding of the new ideas and skills
3. The reduction of barriers
4. Educators' attempts to use the new practice
5. The impact of the new practice on students. (p. 82)

Guskey's model absorbed the most common current measure, that of participant response to professional learning, and makes the logical links through understanding, to use, to impact. He articulated one concept unusual in other frameworks – the reduction of substantial barriers to change built into the system. The component of barrier reduction helps to address the problems raised in the literature about the lack of change in school systems. Guskey (2011) also directly

addressed the problem of the causal link by describing the level of evidence that strikes the best balance between utility and accuracy within school divisions:

To obtain proof—by which I mean to show professional development uniquely and alone leads to improvements in student learning—is very difficult. It requires a level of experimental rigor that is hard and often impossible to obtain in practical school settings. But most policymakers, legislators, and school leaders are not asking for ironclad proof. What they want is evidence that things are getting better. (p. 13)

Guskey's levels of professional learning evaluation require school divisions to have a complete understanding of the change in teacher behavior they aspire to and how it connects to both student learning and organizational change. Broad and Evan's (2006) review of Canadian professional development practice reinforces the value of a model like Guskey's, and they articulate the need to connect to program goals and standards that Hirsh (2013) echoed. Clarity of goal is a precursor to conducting evaluation, and can be challenging as professional developers may struggle to connect processes they are using with goals (Ingvarson et al., 2005). Reeves (2010) added greater complexity to Guskey's third element of barriers by suggesting specific evidence of barrier reduction. In addition to substantiation of better student results and improved teacher practice, he noted that evidence of improved leadership decisions and system-level learning should also be gathered.

Wiggins and McTighe's (2005) commonly used model of unit planning asks teachers to start with what they want their learners to be able to do, and then ask themselves what they will accept as evidence that learning has occurred. Professional developers, similarly, need to articulate what new skills and understanding they hope teachers will acquire (Hirsh, 2013), and

what evidence would confirm that teachers are using the new learning with sufficient skill that student results are improving. The problem of acceptable evidence of teacher skills is rooted in the difficulties of collecting data about both teacher practice and student impact. Broad and Evans (2006) articulate the need for many measures like observation, portfolios, behavioral measures, and assessment data.

One of the most debated measures of teacher practice is self-reporting. It has a long history largely because teacher survey data is much easier to gather than alternatives like observation, and provides more efficient and quantitative source of data than the theming of interviews. It is also the most cost effective method (Desimone, 2011). The struggle to find the time and opportunity to collect primary evidence in observation is often theoretically avoided by self-reporting on surveys (Desimone, 2009; Ingvarson, Meiers, & Beavis, 2005). While teachers' reports about their perceptions of the value of professional learning programs are often frank, their self-reports of use, and especially their assessment of the impact of student learning, are less reliable (Ingvarson et al., 2005). Borko et al. (2008) state that "self-report data have obvious limitations" (p. 418), but Desimone (2009) notes it is not clear what the limitations are in many studies. Professional developers seeking teacher assessment of transfer of the professional development to the classroom can likely trust teacher assessment of the perceived usefulness of the professional learning and can ask questions that will establish teacher understanding of the skills or concepts taught. However, if professional developers wish to accurately understand how fully the skills transfer or with what frequency, they will need to triangulate educator self-reporting with observations.



Observation is a useful way to gather data that creates a sense of causality about the impact of professional learning and it also helps establish the quality of implementation occurring. Since a direct link to impacts on student learning is the most important element in determining impact and it varies by teacher, documentation of that link occurs primarily at the classroom level. For professional developers, observational data of student learning can be gathered in student products, observations, and conversations, which are the same tools used by teachers themselves as they conduct classroom assessments. Reeves suggested observation is essential because it examines one classroom and student at a time, rather than an aggregate, providing a more direct line to the decisions of teacher and leaders (2010). Lawless and Pellegrino (2007) reinforced the value of a unified professional development evaluation plan coupled with direct observation. Foreshadowing the impact of the newer forms of professional learning on professional developers, Stein, Smith, and Silver (1999) noted the lack of transfer to the classroom is best seen through observation, and that observation will tell the professional developer if the professional learning even has the potential to impact students:

A fundamental discovery that awaits all professional developers who have worked primarily with teachers on a short-term basis removed from classroom settings is the lack of correspondence between what one "teaches" and what is "learned." Professional developers or teacher educators who haven't followed their students into the classroom may believe that transferring knowledge and skills from coursework or workshop activities into the classroom is relatively straightforward. (p. 265)

While engaging in regular observation is essential for professional developer's evaluation of the professional learning they led, the factors described next make that observation problematic.

Observation is more time consuming than survey, and has its own validity issues depending on the bias of the observer. However, observation may fail before it is even tried. Teachers are least willing to engage in forms of professional learning that require observation (Cole, 2004). Despite the value of instructional coaching as professional learning process, observation creates accountability pressures and may provoke a sense in teachers that they are inadequate (Knight, 2014). However, only observations “enable coaches and principals to determine whether teachers are using reform practices perfunctorily or effectively” (Desimone, 2011, p. 71). Observation in classrooms can take many forms and may include formal observation, walk through, or learning walk. Since effective use is essential to create significant impact on student learning, observation is an essential form of measurement in professional developer self-evaluation because if the new approach is tried, observational data can help the professional developer see how well the new instructional process is understood and utilized. Knight argues that using video may be the disruptive innovation that allows observation to be both autonomous and a mechanism for accountability (2014) and suggests a plethora of methods of observation that respect professionalism but allow for accurate assessment and growth.

Properly conducted observation is complex and requires specific processes and expertise that professional developers are unlikely to have training in (Knight, 2014; Stein, Smith, & Silver, 1999). While Saskatchewan professional developers are learning about coaching stances and other key tools, there is still much to learn to use observation well (Johanson, personal communication, Oct. 6, 2015). In addition, school divisions must recognize the value of observation, and commit substantial resources to learning how to observe and how to coach: “Properly conducted observation can provide comprehensive, objective measures of what occurs

in professional development and resulting classroom instruction. But observation is burdensome and expensive” (Desimone, 2009, p. 190). Video is helpful in recording simultaneous small group interactions one observer could not catch (Borko et al., 2008) and in providing the opportunity for feedback (Knight, 2014). Video also allows teachers to see their own classrooms from the third person perspective, and provides rich information for self-assessment which adds to a general culture of reflecting about professional learning. Observational data may be essential for evaluation of professional learning, but it is also helpful as a tool for professional learning and reflection. Capturing and using video observation is a key and emerging tool for establishing causal links in professional development evaluations (Desimone, 2009; Knight, 2014).

An examination of the purpose of the professional learning evaluation in divisions yields another key method often overlooked. Because professional developers conduct internal evaluations focused on determining if current processes are working and how they can be improved, some developers are using an informal form of what Owens and Rogers (1999) call interactive instruction. In this form of evaluation, formative process elements are useful barometers that support formative evaluation and stakeholders help with data collection and direction. Earl and Timperley (2015) contended that formative evaluations strengthen and help to form professional development by examining delivery, quality of implementation, and context. While outcome evaluations, which investigate if the professional learning caused the intended outcomes and impact evaluations, which assesses the overall effects, are typically considered summative (Trochim, 2006), in the hands of professional developers their purpose is often formative as they shape an evolving practice. Towards that end, formative assessment

should enjoy the same lauded place in professional learning it enjoys in the research about instructional practice. Formative measures provide useful information about teacher understanding, reaction, and the potential for implementation that Saskatchewan professional developers can use to evaluate and shape their practice (Johanson, personal communication, Oct. 6, 2015).

Reeves (2010) contended that formative evaluation is not merely a nice addition to the evaluation of professional learning, but is actually a way of looking at professional growth that is essential to ensure a sense of agency and growth in educational institutions. Like Fullan, Rincón-Gallardo, and Hargreaves (2015) and Knight (2007, 2014), Reeves saw the engagement of teachers as coinvestigators as critical to success and effective professional learning. Reeves paralleled what his *assessment process* (as opposed to an *evaluation process*) with the change in what educators are attempting to do around student assessment. He highlighted four key elements:

1. Plan for what you will accept as evidence before you start, and deliberately collect it before, during, and after professional learning.
2. Focus on assessment for learning, rather than evaluation, and consider teachers, building administrators, system administrators, policy makers, and parents as sources of data and ideas.
3. Create preflight checklists for teachers, principals, and system administrators before action occurs so the goal and processes are clear to everyone.
4. Use effective formative feedback often. (Reeves, 2010, p. 57).

Newer forms of professional learning assume teachers will start from data, determine what their students need them to know, learn new strategies, then test those strategies in the classroom while collecting data to see if they make a difference. Those processes, when used, normalize self-assessment process and make them cultural norms rather than threatening behavior. Formative types of evaluation are naturally embedded in many newer forms of professional learning and just need to be systematically gathered to improve the quality of the professional learning.

Frameworks for internal evaluations in school divisions can yield enough information to determine if professional learning is affecting teacher practice and, when paired with the most appropriate measures of student growth given the goal, can provide the information school divisions need to make educated decisions about professional learning. Evidence of student growth must be more than just test scores; school divisions must hold themselves accountable for deeper learning and engagement for all students (Fullan, Rincón-Gallardo, & Hargreaves, 2015). Schools and divisions must consider philosophical and structural changes, methods for conducting high quality professional learning, and methods for gathering, interpreting, and using data and evidence about the impacts of the professional learning they fund. If this complex process is formative assessment for growth rather than summative measurement of success, it is more likely to drive effective professional learning and increase autonomy and accountability (Knight, 2014). Many school divisions in the province would engage in an evidence-based decision-making model they already request from their teachers, increasing internal consistency and credibility.

## **The impact of the professional developer**

As significant as the system, the form of professional learning, and culture of self-assessment are, they are always couched in the context of each professional developer attempting to conduct an internal evaluation. While insider evaluation needs support of division leadership, it will be problematic if assigned as extra work on top of an existing workload (Owens & Rogers, 1999). Since it is rare for educational organizations to employ trained program evaluators (Owens & Rogers), professional developers need substantial training and support. Few professional developers can purposefully select from the new forms of professional learning, feel confident in leading something efficacious, or know how to assess the effectiveness of that work (Stein, Smith, & Silver, 1999). Decades of accepted practice assumed professional developers would lead something episodic with no actual expectation that the training would even impact a teacher's classroom (Reeves, 2010; Stein, Smith, & Silver, 1999), ensuring that evaluating the impact of professional learning was highly improbable. Ingvarson, Meiers, and Beavis (2005) described how professional developers struggle to even distinguish specific strategies and purpose:

Designers of professional development activities select from a wide range of strategies to promote professional learning. They often describe the strategies they have chosen in ways that are not necessarily helpful as measures for research purposes. They may use terms such as, 'hands on', 'action research', 'workshops', 'training sessions', and 'case methods'. What these terms actually mean in terms of teacher learning processes is not always clear. To make the research task even more complex, designers often say they use a large number of these strategies in the same activity. So we found it difficult to gain

useful measures of actual teacher learning processes by asking program designers about the strategies that characterized their activities. (p. 7)

If professional developers struggle to define their own practice such that experienced researchers can build “useful measures”, facilitators of professional learning could not design their own evaluations without learning more about the connection between forms of professional learning, purpose, and ways to measure. Professional developers need a foundation in program evaluation.

Professional developers also struggle as internal evaluators to challenge organizational assumptions freely and must address their own personal assumptions developed over years of praxis (Owens & Rogers, 1999). The way they use data, make sense of their experiences, and how they attribute ideas are all factors that affect their interpretations of the value of professional learning. While there is little research on the thinking of professional developers themselves, the push for data-based professional learning is yielding information about how teachers engage in using data to decide what to do.

Teachers work through four phases as a part of a professional learning process like a teacher learning group or data team. They gain access to the data, relate the data to their experience and understanding to make it something they can act on, reframe it as knowledge, and then engage in a response (Bertrand & Marsh, 2015; Mandinach, Honey, Light, & Brunner, 2008; Marsh, Pane, & Hamilton, 2006). Research about these professional learning processes predominantly considers how the response improves instructional practice (Bertrand & Marsh, 2015). The link to improved practice will also be a main focus of understanding how

professional developers think about self-assessment or internal evaluation, but the process of thinking through the data is worth closer examination.

When teachers plan to take action on something, they need to consider its cause. The process of determining cause is referred to as attribution theory, and it has three main characteristics (Seifert, 2004; Weiner, 2010). First, teachers think about the locus of control for something, for example when a student is not reading. Sometimes they place the locus on control internally (with themselves as teachers) or externally (with student, family, the school system). More complex and pervasive problems, such as Indigenous student achievement in Saskatchewan, result in teachers attributing the cause outside themselves; this misattribution occurs even when all other stakeholders, like student, parents, and administrators, perceive teachers as powerful in causing and addressing the issue (Steeves, Berryman, Carr-Stewart, Kovach, Laliberté, Meyer, Merasty, Sloboda, & Stelmach, 2014). For professional developers, attribution is of particularly significant because of their agency in the evaluation process. Johanson (personal communication, Oct. 6, 2015) notes that professional developers may see other teachers as resistant to the need for change or as inadequate instructors, leading the facilitators to offer professional learning that blames, and prevents the facilitator from considering how the learning process contributed to the likelihood of a teacher understanding or using ideas.

The next two characteristics, stability and controllability, are equally significant for professional developers. A person must also see a cause as enduring (stability) to consider it worthy of action, and controllable for any action to have any potential for effect. How a person attributes elements directly influences both perception and behavior (Seifert, 2004; Weiner,



2010), including persisting in overtime and in the face of failure, general motivation, and intensity (Bertrand & Marsh, 2015; Dweck & Leggett, 1988; Nicholls, 1984). Understanding the process of attribution for professional developers is critical, as their thinking process is the main basis for either self-assessment or internal evaluation when divisions support the assessment of professional learning.

Bertrand and Marsh (2015) also identified a second key process, sensemaking, as significant in how teachers, and by extension professional developers, make sense of data. Sensemaking is a process we engage in to help us choose what to pay attention to. Our experiences and existing knowledge help us filter all possible stimuli and focus our attention on a select group of them (2015). Professional developers receive little formal training in leading professional learning, and develop their knowledge through experience, leading to filters that are experiential rather than research-based. Their decisions and actions, therefore, are profoundly ingrained based on their conceptualizations of how teachers learn (Stein, Smith, & Silver, 1999). The process of sensemaking leaves an indelible mark on how professional developers self-assess the impact of the professional learning they lead.

One of the perplexing parts of studying the impact of new forms of professional learning is discovering there is little information about the thinking or actions of professional developers themselves. Professional learning is driven by the expectations and offerings of school divisions. Within those divisions, professional developers are the only stakeholders with specific responsibility for planning and leading professional learning. As new realities demand an understanding of the impact of the professional learning, it is professional developers who are asked to determine impact; only professional developers can make changes based on the findings

and judgements of any form of assessment (Stein, Smith, & Silver, 1999). Even in distributed professional learning models, professional developers typically “train the trainers”, making them the key to shifting any practice in professional development or evaluation. All of the variables of professional learning, including form, expectation of impact, duration, purpose, and delivery, are in flux and influenced by professional developers. Kennedy’s 2016 review of the impact of professional learning on teaching noted: “Many of the more effective programs reviewed here were offered by individuals or groups who had long histories of working with teachers, were very familiar with teachers and the problems they face, and based their programs on their own personal experience and expertise” (p. 973). She noted that “there is little discussion in the literature about the nature of PD expertise, how PD providers are selected, how they prepare their work, or how their efficacy is assessed. These topics need to become part of our discussion as we generate and test of PD theories in action” (p. 973). A focus on professional developers’ sensemaking, processes, and practice can yield critical information in the search to understand how to increase the frequency and quality of evaluations of professional learning. The process of self-assessment, as practiced by professional developers, is a crucial clue in finding ways to better determine the impact of professional learning, and ensure that it is connected to student results.

## **CHAPTER 3**

### **Research questions**

The study was designed to develop theory about the process professional developers use to self-assess or internally evaluate the impact of the professional learning they lead; the study describes the conditions under which professional learning evaluation actions, intentions, and processes emerge or are muted. The study of participant-identified causal factors, intervening conditions, strategies, context, and consequences can provide essential information to help educational stakeholders make effective decisions that make effective evaluation, and therefore improvements to professional learning. Because professional developers are the main actors with lived-experiences connecting leading professional development and their own self-assessment or internal evaluation processes, they are best positioned to describe experiences and their interpretations. My main research question asked: How do professional developers describe and demonstrate the process of either self-assessment or internal evaluation? My sub-questions included:

1. How do professional developers decide which strategies to use to determine the impact of their professional learning?
2. What influences the perpetuation of evaluation (or lack thereof)?
3. How is the praxis of the assessment or evaluation of professional learning influenced by the professional developer's experience?

## Significance

Teacher learning endeavors fuel improvements in students' learning, but teachers continue to indicate that they are dissatisfied with their learning opportunities (Darling-Hammond et al., 2009). Commonly cited research points to 50 hours of professional learning (Darling-Hammond et al., 2009) before substantive changes in teacher instructional practice, so professional developers must strive to make professional learning effective. In Saskatchewan, resources devoted to professional learning are shrinking as the education sector grows (Johanson, personal communication, Oct. 6, 2015), and recent provincial budget cuts are further reducing resources. A better understanding of the factors at play, specifically the decision-making process of those responsible for professional learning, is a much needed piece of the puzzle.

Understanding the thinking and actions of professional developers offers an opportunity to influence the quality of professional learning as it is developed and implemented, because it provides much needed information about how the main actors in the assessment or evaluation of professional learning conceptualize and value determining impact. Constructing theory regarding their thinking also offers the opportunity to alter the quality of professional learning midstream as some self-assessment processes are formative and could occur as the professional learning is being led. Like formative assessment in classrooms, formative assessment in professional learning offers the opportunity to re-teach and clarify based on participants' responses (William, 2011), improving the quality of teacher learning, and potentially, transfer and later use in the classroom. Researchers have argued that using feedback to alter instruction has a high effect size on student learning (Black & William, 1998; Hattie, 2008), and comparable effect could be felt in adult learning. Although it is a formal extended process, rather than a regular part of any learning activity, formative evaluation and classroom formative assessment uses

similar logic about refining the quality of an instructional program. If attribution and decision-making processes of professional developers were better understood, researchers and program evaluators would have more information about what parts of evaluation and self-assessment processes are routinely understood, and which parts could be better developed and more fully supported to create the regular, effective analysis about the impact of professional learning that writers like Guskey, Reeves, and Hirsh are calling for in education. For example, with regular, accurate feedback, professional developers could alter professional learning processes so they were more likely to result in teacher understanding of key concepts and teachers would be more likely to leverage what they have acquired in classroom settings. But because it is not clear what professional developers do for formative assessment and why, researchers are missing a much needed piece of information about why professional development remains largely evidence-free.

Understanding professional developers' conceptualizations of how professional learning is best evaluated will also help to explain why professional learning evaluation practices evolve the way they do in the field. Current theory provides little insight about why certain elements of professional learning evaluation are embedded in the culture of professional learning, while others are not. While researchers critique the quality of professional learning evaluation and offer frameworks for conducting such evaluation (Desimone, 2009; Guskey 2000), they rarely work directly with those who lead the vast majority of professional learning. The nature of the theory/practice gap, beyond general concerns about the dissemination of research, remains largely unclear. Understanding the reasoning and actions of professional developers provides the possibility of targeting proposed reforms more directly. Recommended changes could be tailored to organizational cultures, lived realities, and the beliefs and understandings of those whose job it is to lead professional learning. When internal evaluation or self-assessment

become innate for professional developers, using evidence in considering the merit or worth of professional learning has the possibility of become a regular practice. Without professional developers understanding and valuing the processes of gathering evidence about the impact of the learning they lead, professional learning cannot be sculpted to create change in teacher practice or student results.

### **Role of the pilot study**

To frame the main study, a pilot study was conducted in late 2015 and early 2016. The first recruiting was purposeful and a selected provincial leader in professional learning, Terry Johanson. Johanson is the director of the Saskatchewan Professional Development Unit of the Saskatchewan Teacher's Federation, and has been responsible for leading the provincial facilitator cadre. She was the participant who helped to construct and revise the main study through a pilot. The pilot study was framed specifically to help refine the methodology and research questions of the main study, and yielded both themes for exploration in the main study and refinement of the research questions and methods. The study also combined elements of narrative inquiry and case study to determine the main issues that could be the focus of the main study.

The pilot study generated themes and insight for a longer-term, more extensive process of contributing to new theory that described and explained the process professional developers use to assess the impact of the professional learning they lead. In the main study, participants came from stakeholder organizations in the education sector with specific investments in the professional learning of teachers. The bulk of participants were from provincial school divisions, as school divisions have the responsibility for funding and providing for professional learning in the current provincial framework. In addition, input from participants from other

stakeholder groups like the Saskatchewan Professional Development Unit (a part of the Saskatchewan Teachers Federation) and the Ministry of Education enriched the resulting theory. Since Johanson has extensive experience with professional learning in many different divisions, she was an ideal partner in developing sufficient background for the main study.

Johanson and I discussed each stage of data collection as it was completed, and considered implications for the main study. She was a research partner, suggesting ways to refine interview questions and ways to use images and documents as comparison to statements in the interview. Johanson made several observations about the provincial professional learning context based on her broad experience, and was very articulate about her personal and framework theories. For example, I gained significant insight in last half of the talk-aloud method, where Johanson described her personal attribution theory related to professional learning. The nature of her thinking process was not a part of my original literature review, but became significant enough in the theming to add additional foci related to sensemaking, data analysis, and attributions (Bertrand & Marsh, 2015). The attribution process became one of the foci during interviews in the main study because of its significance within the pilot study.

The process of collecting and theming the data in the pilot also informed this study in several ways. Triangulation of the artifacts, field notes, and transcription strengthened both the themes and Johanson's reflection on her own practice. The storied experiences she shared manifested in the observations of the professional learning as it was planned and delivered. In addition, the most useful insights were gathered during the talk aloud, which allowed Johanson to see her own process as an observer. The pilot study yielded significant themes, like the pivotal nature of framework theory absent accountability, that were valuable directions for inquiry in this study.

Theming Johanson's statements in parallel with the field notes helped to provide a check and balance for the impact of the researcher's perspective in this study. Because I have worked as a professional developer for nearly a decade, I sometimes assigned significance based on my preconceptions. Because of the influence of the pilot study, the data gathered in this study were analyzed on two levels: the manifest level, which was a descriptive account of the data where line by line open coding focuses on meaning in the words, and a second level, a more inductive one (Charmaz, 2014), which was used to generate initial theory that further shaped this study. The method of analysis and the variety and depth of data also helped to address the complication of the researcher's perceptions. The pilot was very valuable in designing the data collection process for the main study.

### **Narrowing the study design**

From the outset, I was interested in a process of inductive analysis, and I knew there was a wide range of literature that could inform my thinking about methodology. Creswell (1998) uses comparisons of five major methodologies to help researchers with qualitative questions clarify their approaches and select the most appropriate methodologies. While narrative inquiry, phenomenology, grounded theory, ethnography, and case study may all use interviewing as a primary method of data collection, the processes used in describing, classifying, and interpreting data in each methodology can help to reveal an approach best suited to a specific question. I read at minimum a major theorist in each area to generate a basic understanding of the content and processes of each methodology, and a familiarity with the specific approaches or traditions, including phenomenology (van Manen, 1990), narrative analysis (Leiblich, 1998), and grounded theory (Strauss & Corbin, 1990; 1994). I excluded case study as a methodological approach early on because I did not want to work with a tightly bounded group. As I become more



interested in program evaluation, I looked at Patton (2002; 2008) and an increasing interest in current trends in grounded theory brought me to Charmaz (2014).

I used four of Creswell's points of comparison (1998) to help select my potential methodology in the winter of 2016. First, I considered what I wanted to understand and my positionality, and wrote many iterations of my research questions. I knew purpose should be the controlling force in selecting methodology (Patton, 2002), and my decisions about design, collection, analysis, and reporting should extend from my purpose. My main objective was to understand the process professional developers use for self-assessment, so the iterations of my research questions clarified my thinking and intent. I categorized those questions by the methodologies designed to address them, and considered which of the types of questions were most compelling and most in need of study. Next, I looked at the types of data analysis and representation processes in each methodology as Creswell describes them, and considered which of the methodological approaches were best suited to the type of thinking I wanted to do and to my emerging research problem. In my increasingly metacognitive process, I made two key mind maps, one of the main themes in the research about professional learning evaluation (Appendix A), and one about the prospective methodology and my positionality (Appendix B). Finally, I read studies in the various writing styles associated with each methodology, and wrote short sample passages in each style, to help me consider how the findings might be shared, utilized, and represented.

Having considered each element, I was leaning strongly towards grounded theory, when deeper reading of methodological theory caused me to question my thinking. While Creswell (1998) placed grounded theory as pulling from qualitative and quantitative traditions, he does not give it a positivist stance. Glaser and Holton (2004), however, located grounded theory in the

positivist tradition, and argued that a mix with qualitative traditions only burden theorists with excessive description. They argued those who see grounded theory as qualitative have the benefits of neither methodological stance because their attempted methodology has incompatible assumptions and processes. Glaser (1992) perceived qualitative theory as becoming more popular due to a mislabeling as grounded theory, and saw misconstrued grounded theory as plagued by qualitative problems of rigor, types of data, the participant's voice, rich description, accuracy, and constructivism. He defined grounded theory as a set of interwoven, inductively generated hypotheses. Glaser contended that the process, which has a series of iterative steps to check the accuracy of a systematic conceptualization, has the value of being both very structured and exceptionally versatile. He argued the constructivist lens erodes the power of grounded theory, and Glaser positioned the researcher outside the research context, and contended the researcher needs as little preconception as possible. Glaser's characterization of grounded theory meant I could not select it as a methodology. I position myself as an insider, and I know my prior understanding must influence every element of the study. Declaring my insider stance seems the most useful research perspective, and is not positivist.

Additional reading, however, helped me find a way of conceptualizing grounded theory which is current and addressed my concerns about the use of grounded theory for the study. Strauss and Corbin (1994), whose work laid the foundation for Charmaz's (2014) constructivist grounded theory, were categorical in their own positioning, that "truth is enacted" (p. 279) and there is no exact reality to discover through grounded theory. They have strayed from what Glaser and Holton (2004) called classical grounded theory purposefully, and directly critiqued the idea that a researcher can be most accurate if she has little information or personal stake. In addition, Charmaz argued that if everything contributes to data as Glaser contends, then seeking

rich description or a participant's voice enriches, rather than erodes, the quality of end theory. Charmaz's (1995a, 1995b, 2000) and Strauss and Corbin's (1990, 1994) conceptualization of grounded theory as constructivist are echoed by other contemporaries (Annells, 1996, 1997; Charmaz & Mitchell, 1996; Clarke, 2005; Mills, Bonner, & Francis, 2006), and gave me the sense that grounded theory, which was so attractive because of its theorizing and constant comparative method, could still be an appropriate approach for the study.

### **Assessing “close” alternative methodologies**

Creswell (1998) characterizes narrative approaches as describing through story, often chronologically, while locating participants' epiphanies and interpreting larger meaning in context. In recognizing the processes of professional learning evaluation, narrative might offer a strong sense of experiences of individual professional developers and key events in their conceptualization processes or development. Questions of character, setting, key conflicts, chronology, and epiphanies would dominate the description of the participants' lived experiences (Czarniawska, 2004), and likely provide excellent, non-generalizable, insight about influences on the process of specific participants. In particular, Chase's (2005) conceptualization of socially situated narrative, and how individuals are enabled or restrained due to social resources, has good connections to some of the professional learning literature on coaching (Knight, 2007). What narrative would not provide, however, is the development of a general framework or matrix for sensemaking of the process that could drive decision-making about professional learning evaluation. The process of storying is not easily overlaid with program evaluation processes in a comparative, and an attempt to do so oversteps from a narrative paradigm. Narrative seeks to understand and respect participant voice, and to avoid superimposing theorizing of the researcher where possible. In addition, narrative focuses on individuals

(Polkinghorne, 1995), and it would not be valid to decontextualize participants' stories in pursuit of theory. Narrative is not ideally suited to the research question of this study.

Phenomenology offers rich, thick, descriptions of lived experiences, particularly how participants make meaning of essential elements. It finds the essence or nature of something (van Manen, 1990). Phenomenology could be very useful in understanding the meaning making and attribution processes of professional developers as they parse data and plan for evaluation. However, it is best at uncovering the essence of shared experiences, and is not focused on solving problems except those identified and speculated about by participants. Since the pilot study indicated participants may not have thought in depth about all stages of professional learning evaluation, phenomenology may not yield rich information. Regardless, it is not designed to provide utilizable theory to help address the problem of evaluation. In addition, phenomenology seeks to describe the essence as a philosopher, not explain or analyze (Moustakas, 1994). Since I wanted to explain the process professional developers use to assess professional learning, phenomenology is antithetical to my purpose, despite providing valuable, related information.

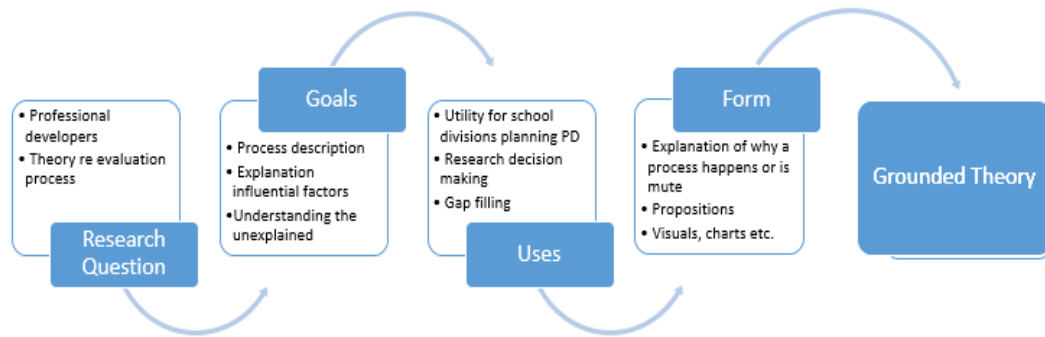
Ethnography is a good tool for understanding processes like the contextual evaluation of professional learning. Ethnography, even more than case study or narrative, brings a variety of lenses to help understand the common culture and process of a group, as described by forms like critical or realist ethnography. An ethnographic approach would help to establish common ways the professional developers viewed things, like concepts of learning journeys or self-assessment, and ways in which the professional developers commonly behave. Results are represented in narrative form, but also in sketches, tables, and other forms that could provide useful insight for

those planning professional learning. The focus on organizational structure and culture would help clarify both barriers to, and the impetus for, professional learning evaluation.

One of the two reasons I did not select an ethnographic approach is my background. I position myself as a professional developer, and I have led or participated in a wide variety of activities designed to help explain and grow organizational practice around professional learning in several divisions and provincially. While an ethnographic, or even autoethnographic, approach would have merit in understanding culture or process stemming from culture, I was looking to extend my learning professionally in a newer vein for me personally. In addition, ethnography is best designed to address questions of a social and cultural nature, and my question, while influenced by these factors, was more directly looking at the thinking and processes of individuals. My interest was generating theory describing a thinking and self-assessment processes, rather than the social circumstances or culture of professional learning.

In weighing grounded theory against other close methodologies, I used four key factors in my final decision once I determined its positionality was compatible. I considered the research question, the goals of methodology, the potential uses of the research, and the writing style and voice commonly used by researchers in that methodology (see Figure 3.1). I selected these elements because they highlighted key differences in the methodologies. I found there were many areas of similarity, for example using interview and observation for data collection, or creating and organizing files for data management. For each area, I articulated the key elements for myself, then compared them to the main tenets of methodology and the theory underlying it.

Figure 3.1 – Decision-making factors in selecting grounded theory



### Study design

The purpose of my proposed grounded theory approach was to develop theory regarding the influences on the process of self-assessment or internal evaluation used by professional developers in the K-12 education sector in Saskatchewan. Two common current approaches to grounded theory are Strauss and Corbin’s (1990, 1994) systematic approach and Charmaz’s (2005b) constructivist approach. Both use a mix of interviews, observation, and documents in a constant comparative method to develop theory. A systematic approach is focused on developing a core category or theory using a specific, prescribed process. A constructivist approach uses many of the same analytical approaches, but is more focused on openly acknowledging the researcher’s views in the conceptualization of resultant theory. I have used a systematic approach to compare each participant’s description of interpreting data to the other participants, for example, then looked at the resulting theory and asked myself how my lens has influenced my own sensemaking processes, trying to describe that explicitly as a part of later part of chapter four.

There were three primary sources of data used to develop the theoretical framework. The first was a conversational interview process designed to help participants articulate their current professional learning practice and to describe the assessment or evaluation process they use. The second source of data was the researcher's observation of a professional learning or planning process selected by each participant. The process is an example of the participant's engagement with a part of the assessment or evaluation process, such as: goal setting, planning for professional learning and assessment, leading professional learning and engaging in formative assessment, data gathering or analysis, or rendering judgement. If the participant was demonstrating and should not be interrupted in process to engage in reflection, a third process called a talk-aloud protocol was used to provide the opportunity for metacognitive commentary. A participant-selected part of a process (like leading professional learning) was videotaped, and shared with the participant for reflection. In the think aloud protocol, the participants described their decision-making and thought processes about the evaluation. Extant documents referred to and utilized by the participant for evaluation, or lead professional learning, were also analyzed and compared to interview content. Each participant was invited to participate in both the initial interview and the think aloud, so the opportunity to discuss, demonstrate, and engage in metacognitive reflection about the evaluation process was provided. Participants also had the opportunity to participate in a member check of the resulting emergent theory and suggested alternative explanations or additional information, which two participants chose to do. Rephrasing and summary strategies routines were employed after the first interview to compare the interviewer's grasp of the participant's thinking with the participant's personal understanding.

## **Data collection**

The 12 participants were K-12 educators from Saskatchewan. In most cases, these individuals were employed by school divisions, but 3 participants in the study represented the Ministry of Education or the Saskatchewan Professional Development Unit (part of the Saskatchewan Teacher's Federation). Two additional participants left their positions during the study and their data was excluded from theorizing although it had been partially collected and analyzed. Three others agreed to participate but had not started data collection when budget cuts returned them to the classroom. Of the 12, half had leadership roles with supervision components, like superintendent, coordinator, or administrator. The remaining participants were consultant type roles or teacher coaches. The school division participants came from four different divisions, all within the Saskatchewan provincial education system. Three participants were male and nine were female, a ratio relatively typical of those working in professional development roles in Central Offices. All participants were former teachers, and all except one held other professional learning leadership positions before their current roles.

Typical case sampling (Patton, 2002) was used initially to recruit female and male educational professionals who have, as a part of their professional responsibilities, substantive responsibility for the professional learning of teachers and educational leaders. While such samples are illustrative, not definitive, they provided a solid basis for an initial understanding of the process. The initial sample was predominantly generated by school divisions themselves. The divisions were approached for permission to complete the study in the division (see Appendix C), then a contact person with responsibilities for professional learning was asked to help with recruitment. The initial sample included school-based administrators and central office staff from rural and urban school divisions. A third, separate, school division gave permission to



participants but recruited no volunteers. The participants were unknown to the researcher or were casual professional contacts. The more open sampling process followed the open coding process used initially in grounded theory (Charmaz, 2014; Strauss and Corbin, 1990).

Once transcription, coding, and memoing were underway, theoretical sampling was necessary. In grounded theory, theoretical sampling is used to provide challenge and depth to emerging theory (Glaser, 1978) and it is directly connected to both the coding process and the constant comparative process (Strauss and Corbin, 1990, 1994). There is a direct link between the stage of the coding and theorizing, and the need for the move to a different method of sampling:

Strauss and Corbin (1990) connected specific theoretical sampling strategies to the three types of coding. They suggest that open coding requires open sampling in which data are gathered to uncover as many relevant categories as possible. Open sampling can be done purposefully, by choosing sites, persons, or documents deliberately to gain the maximum amount of data needed to unearth potential categories and their dimensions; systematically, by moving from one person to another on a list to uncover subtle differences; and fortuitously, by gathering data during field observations that were unexpected but are seemingly relevant to category development. Axial coding requires relational, or variational sampling, in which data are gathered to uncover and validate the relationships among categories that have been discovered. This type of sampling can also be done purposefully or systematically. Selective coding calls for discriminate sampling, in which data are gathered to verify the emerging theory and to further develop categories that have not been well saturated. (Draucker, Martsof, Ross, & Rusk, 2007, p. 1138)

As categories emerged, the process of axial coding and theoretical sampling began. The composition of the theoretical sample depends on information garnered from initial participants and occurs in service to theory, so it should not be established prior to the study (Hood, 2007). After the eight initial participants, four additional participants were recruited to deepen the emergent categories established in the initial sample. They were asked to respond to emergent theory generated by the initial sample (see Appendix D) but recruited differently via conversation or email to deepen early categorization. Finally, sharing of the foci of the research by a provincial organization also led to four volunteer participants who were asked questions as a part of the theoretical sample because they were a part of group explicitly taught to deliver and assess professional learning.

The second part of the data collection with each person was focused on participants' metacognitive assessment of their processes. They self-selected a process they believed could demonstrate their thinking. A variety of processes for data collection comprise a grounded theory approach, as the researcher seeks different sources and directions in an attempt to confirm or extend theory (Strauss and Corbin, 1990). In this case, participants were told the purpose of the observation, which was to see how they determined if the professional development they led was effective, and asked to select a process that would allow the researcher to see their thinking. The collection of field notes and extant documents, and memoing, continued throughout the second stage of observation. Photographs of professional learning processes designed to yield visual data were also captured depending on the focus described by the participant. Participant-selected observations included planning meetings, debriefing processes, leading professional learning, examining data, and examining materials produced by the attendees of professional

development sessions led by the participant. In one case, the first and second meeting occurred concurrently due to time constraints.

Finally, some of the professional developers engaged in a talk-aloud about recordings of their processes or data use to elicit commentary to make thinking processes explicit. The study employed Polkinghorne's (2007) data gathering structure of three meetings, including: a semi-structured first interview, observation, and a talk aloud protocol to debrief participant thinking. The structure was consistent in most cases, although in some, the second and third meeting were combined because stopping to reflect during the process was easy. Probing questions (see Appendix D) to support the elaboration of the evaluation process and the participant's thinking were used in the final meeting. In addition, extant documents proffered by the professional developers and the materials used by professional developers were gathered and used at this stage to support the participant's elaboration of his or her thinking. Where the participant chose to look at already collected data from a professional learning activity, observational questions and a talk-aloud protocol were combined. Where possible towards the end of the theoretical sample, emergent theory was connected to participant's metacognitive process through direct questions following an interviewer summary.

Complexities of budget and a small participant pool affected the sampling process in the second year. Three participants initially offered to participate but withdrew when their employment changed due to provincial budgetary constraints. One also retired before signing the final transcript release, so her data was withdrawn from the study. While the study achieved diversity in perspectives from administrators and central office employees, the perspectives of school-based teacher leaders, like instructional coaches are under-represented, as many of them were returned to the classroom following budget reductions. Also, one participant completed the

initial interview in one school year, then was hired into a professional development position that created indirect conflict when she became employed in a school the researcher supported. Her initial interview was included, but subsequent interviews were not completed.

### **Saturation**

One of the complexities of grounded theory is the concept of saturation. Grounded theorists describe a sample size in pursuit of saturation, so no purposeful set amount of participants is pre-selected (Bowen, 2008; Charmaz, 2014; Mason, 2010). After the first interviews, coding revealed organizational influences on self-assessment processes, so additional perspectives from members of organizations other than school divisions were added to the purposeful sample and two additional school divisions (one rural, one urban) were added to the sample group to challenge and deepen the categorization. Similarly, except for the two administrators in the initial sample, all participants had the same role and described that role as limiting the chance to observe. Additional roles, like teacher leaders and administrator, were added to the sample to see if it altered initial categories and emergent theory. Purposeful sampling was used to extend the sample and refine emergent theory.

In grounded theory, using theoretical sampling improves analysis. Charmaz (2014) describes theoretical sampling as directly related to categorization, because it is used to check the relationship between categories, determine variation, delineate what belongs in each category, and ensure conjecture is grounded in data. The goal is to make an “analysis of mundane experiences in a field setting become more analytic, abstract and potentially generalizable” (p. 212-13). At some point, more data no longer yields new theory or insight; the same elements are repeating over and over in the analysis process. This point is called saturation. Glaser notes that seeing similar stories is not saturation. Rather, he clarifies that when the comparison of concepts

is no longer yielding new “properties of the pattern” in conceptual density, it means there is theoretical completeness (2001, p. 191).

Despite describing what saturation may mean, there are still questions about how many participants are required. Glaser (2001) contends the size can be small. Mason (2010) suggests a skilled interviewer may have sufficient data in ten interviews. Bowen’s (2008) clarification of purpose helps to underscore that sampling is not designed to achieve generalizability or representativeness. The goal is saturation of theory, and sample size is determined in relation to consistency in service of theory (Bowen, 2008). Morse, who writes often of the subject of saturation, also relates theoretical sampling and saturation as she considers the characteristics of saturation (2011).

### **Data analysis**

Grounded theory might be best described “as a constellation of methods” (Charmaz, 2014, p. 14). One of the distinctive elements of grounded theory is that analysis and data collection are iterative and concurrent. In 1967, Glaser and Strauss provided clear guidelines for the collection and analysis process designed to legitimize qualitative research as credible and rigorous. Over time, Strauss and Corbin diverged from Glaser, and two constructions of grounded theory developed; Charmaz’s constructivist grounded theory follows in Strauss and Corbin’s tradition, which acknowledges subjectivity (Charmaz, 2014). Emergent theoretical categories are constructed through early coding and memoing, and early theory is deliberately challenged by selection of additional participants (Charmaz, 2014; Strauss and Corbin, 1990). The methodology of this study does not seek to evaluate, from a positivist stance, the value of the professional learning conducted by the study participants. Rather, it seeks to uncover the multiple realities that they describe as their experiences with assessment or evaluation. How

they construct and act on their understandings of assessment or evaluation is critical information about the process of professional learning evaluation. Similarly, I interpret participants' meanings and actions, and they interpret mine as the researcher. Relativity and subjectivity are addressed by acknowledging the researcher's role in construction and interpretation during analysis (Charmaz, 2014).

Charmaz articulates specific criteria for grounded theory that distinguish it from general inductive qualitative strategies. She describes processes or actions as criteria for grounded theory, many of which interweave collection, analysis, and theorizing:

1. Conduct data collection and analysis simultaneously in an iterative process
2. Analyze actions and process rather than themes or structures
3. Use comparative methods
4. Draw on data (e.g. narratives and descriptions) in service of developing new conceptual categories
5. Develop inductive abstract analytical categories through systematic data analysis
6. Emphasize theory construction rather than description or application of current theories
7. Engage in theoretical sampling
8. Search for variation in studied categories or processes
9. Pursue developing a category rather than covering a specific empirical topic. (2014, p. 15)

Charmaz argued that many researchers engage in actions one to five, but not all nine, specifically failing to adequately address sufficient development of theory. Theory must be an analytic process, not merely a description.

The basic process analysis is the mainstay of theory development, and I used many elements of the process in the study. While the development of the research question and some recruitment occur prior to analysis, the analysis and data collection were simultaneous. The initial coding process was a line by line analysis, followed by focused coding and categorization following each interview. Qualitative software was used for the line by line coding, and words and phrases were grouped in the software to generate emerging categories. These were compared to memos, documents, and observations to create a logic model for self-assessment or internal evaluation process of each participant, and a diagram of their processes. Memo writing was used for multiple purposes, but especially to identify early theory and incomplete understanding during comparative processes or to summarize analytical thinking at the end of observation. For example, participants provided extensive description of using facilitation and conversation to gather information, but used generalized statements to describe data analysis processes. Abductive reasoning was used to consider a variety of possible explanations for the generalizations, and they were documented for further questions in the theoretical sample interviews.

Comparison was a mainstay of the analysis processes to see how the comparative illuminated theoretical categories. The process of appraisal related and contrasted participants with other participants and themselves, transcript coding with observational coding, transcripts with documents, and codes to categories or early theory. The theoretical categories, generated from the coding and memo writing process, led to theoretical sampling to test the emerging theory, which began the comparative cycle again at recruitment. The iterative nature of the process continued until saturation, when no new precepts were being generated by data collection, and theory building began.

One of the complexities of grounded theory is that the answer to even the most basic questions about methods may be “it depends.” For example, grounded theorists emphasize theoretical usefulness of interview data rather than on meticulous accuracy (Charmaz, 2014). Interviews are semi-structured and the nature of the questions or elements for follow-up change as theory emerges. In the study, the interviews were open-ended and conversational, and shifted to mutual conversations about prospective theory with increasing regularity. Appendix D contains examples of earlier and later interview questions. As the theory became increasingly detailed and simplified, participants were asked directly about elements of the self-assessment process that needed greater clarification or invited to explain about the reasoning behind their choices.

One of the key elements of grounded theory is coding. The line by line coding in the initial phase provided the analytical structure for the focused coding. Axial codes described by Strauss and Corbin (1990) were not the focus, and the analysis processes favored the more emergent and theoretical codes that distinguish constructivist grounded theory. Codes were collapsed into categories in the coding program, and sample versions of the theory were generated using diagrams. In addition, quotations from the research were grouped by topic and used to challenge the categories. The differences between research and the emergent categories was used to generate topics for refining thinking and further explicate the categories.

### **Strategies for increasing the quality of the study**

As a relatively novice academic researcher, one of my pressing concerns was the development of a high quality study. There seemed to be little agreement on what would make a study more reliable or valid, particularly across different ontologies. There are a plethora of qualitative terms used to describe what makes a better, more believable study including, but not



limited to: trustworthiness, validity, authenticity, reliability, objectivity, credibility, transferability, and external or internal validation. Whittemore, Chase, and Mandle (2001) have four primary criteria in their synthesis helpful in describing the worth or merit of a study; the primary ones, in particular, helped to construct this study. They define credibility in terms of an accurate interpretation of participants' meaning, so this study confirmed meaning as a part of the third interview with participants. Their second criterion is authenticity, which seems to speak to a variety of voices. The second criterion relates to the type of sampling used, but also relates to diversity of role, gender, level of power, or other factors. This study attempted to address this through theoretical sampling as the study progressed. The participants provided different voices or lens to critique the theory about professional developers' assessment and evaluation processes. The next criteria is critical appraisal, which is essentially an analytical consideration of all aspects of the research. Both the memoing and comparative process and the feedback from others was used to think critically about the study's design. This section, and others, are explicit, vivid, and designed to provide the opportunity for genuine critique of the process the study used to provide the opportunity for the reader to engage in critical appraisal. The final primary criterion, integrity, is also related to these concepts because it focuses on being self-critical. Sharing thinking with peers, seeking alternative interpretations, and trying to explicitly record my assumption and logic were all used to address this final element. Detailing criticism of my own study, like my descriptions of specific issues in participant recruitment, help to provide detailed information for the reader that support critical thinking about the study.

Creswell (1998) suggests eight ways that qualitative researchers try to increase the validity of their work, then specifically recommends two this study utilizes. The first is member checking, which Lincoln and Guba (1985) see as the most critical. Utilizing the methods

suggested by Charmaz (2014), my theorizing will be directly shared with participants in a member check, and participants will be asked to judge the accuracy given their experiences. In the third interview, this process has begun with strategies like summarizing, rephrasing, and comparing to emergent theory. The second technique used in this study was attempting to provide sufficient quantity and quality of description to allow the readers to determine for themselves if the ideas represented legitimate theory.

I also looked to methodology-specific frameworks for high quality study, and used Strauss and Corbin's (1990) questions to shape both the study and the framing of theory. During chapter three (and the upcoming chapter four), description of specifics provides the reader with information about the process of the research that reflects Strauss and Corbin's questions:

Criteria 1: How was the original sample selected? On what grounds?

Criteria 2: What major categories emerged?

Criteria 3: What were the events, incidents, actions and so on (as indicators) that pointed to some of these major categories?

Criteria 4: On the basis of what categories did theoretical sampling proceed? Guide data collection? Was it representative of the categories?

Criteria 5: What are some of the hypothesis pertaining to the conceptual relations (that is, among categories), and on what grounds were they formulated and tested?

Criteria 6: Were there incidences where hypothesis did not hold up against what was actually seen? How were these discrepancies accounted for? How did they affect the hypotheses?

Criteria 7: How and why was the core category selected (sudden, gradual, difficult, easy)? On what grounds? (p. 253)

Charmaz's (2014) *Constructing Grounded Theory* has a plethora of reflective questions designed to provoke thinking, improvement, and critical thinking at every stage of a grounded theory study. She also specifically addresses criteria for quality in grounded theory studies. Two of her four categories, credibility and resonance, ask similar sub-questions to other theorists, including both Strauss and Corbin and Creswell, but her other two, particularly as seen through the sub-questions, are very focused on constructing grounded theory. Originality, for example, has four sub-questions:

- Are your categories fresh? Do they offer new insights?
- Does your analysis provide a new conceptual rendering of the data?
- What is the social and theoretical significance of this work?
- How does your grounded theory challenge, extend, or refine current ideas, concepts, and practice? (p. 337)

The sub-questions were helpful for me in several key ways. First, they highlighted the need for solid grounded theory to add new concepts and insights to existing theory. Second, they provided a framework for self-assessment of significance that memos in the study were built around. Like her questions on originality, Charmaz's questions on usability help distinguish grounded theory as a methodology, and are actually reminiscent of Patton's (2002) case for utilization-oriented program evaluation. Charmaz's usefulness questions are:

1. Does your analysis offer interpretations that people can use in their everyday worlds?
2. Do your analytic categories suggest any generic processes?
3. If so, can you examine these generic processes for tacit implications?
4. Can the analysis spark further research in other substantive areas?

5. How does your work contribute to knowledge? How does it contribute to making a better world? (p. 338)

The usefulness questions helped frame both responses to early comprehensive questions and were the most useful in the development and articulation of the theory framework. As theory was developed, each category was considered for what it offered for description of process, and what the implications of those processes were for supporting future evaluate processes.

Facilitating professional learning involves many complex, heuristic decisions. Designing research studies using a specific methodology seems to require a similar, quirky cocktail of a deep understanding of what you are doing and why, and a double dose of deep analytical thinking. There are two areas in specific, ethics and positionality, where advice from experts and valuable guiding questions were used to shape and refine this study.

### **Ethics**

An ethics application with the Research Ethics Board at the University of Saskatchewan was completed and approved in the winter of 2016, and renewed in the winter of 2017. This section describes the contents of the initial submission.

### **Participants**

Since the early selection process began with contacting school boards, there were limits to anticipated confidentiality. Participants chose whether they would like to be anonymous, but given the small community provincially, there were also limits to confidentiality based on context. Four participants choose a pseudonym. Because observing professional learning was a part of the proposed collection process and the researcher observed people working with division data, it was necessary to have approval from divisions and other organizations that participants

are drawn from to ensure ethical data practice where institutional professional learning was the focus. Asking for permission to interview meant organizations were aware of who was participating and gave approval. Sometimes the recruiting organization was not a school division, and in that case, the organization offering the professional learning did the recruiting and the employing school division was not contacted.

Participants gave written consent prior to enrolling in the study (Appendix E) and were informed of their right to withdraw before the results are submitted. All had the opportunity to review data, including transcripts and video (Appendix F), and seven participants added observations or changed the recorded data, particularly transcripts. Participants were verbally asked for continuing consent at each meeting. Because the pilot study indicated some professionals might want their ideas attributed to them, participants chose between a pseudonym and personal identification. Three changed that decision part way through the data collection process. Names and identifying statements were removed during the transcription process if a pseudonym was selected or others were identified. Identifying details were changed, for example Victory school would have been changed to x school.

Many educators who lead learning for teachers have working relationships with each other, including relationships with me. Those potential participants that I supervised or might wield power over were excluded from the study. Co-workers who expressed interest were included. Given the safeguards, the project was deemed minimal risk by the Research Ethics Board.

Participants stated that the research provided an opportunity for self-reflection, and that the process was valuable for them. At the conclusion of the study, participants were given a copy of the study summary and will be invited to participate in various forums for sharing of the results if they chose to be identified. Some participating school divisions request results as a part of

their approval process. I distributed a summary of results with participating divisions and anticipate sharing more widely at conferences and potentially through publication in professional journals.

### **Addressing my perspectives**

I have my own perspectives about the processes professional developers use or choose not use to evaluate the professional development they lead. Even the belief that professional developers have significant influence represents a conceptualization of professional learning, as does a belief in the need for professional learning. Some might believe we need professional learning because of deficits in the education system, some because of a growth in what we know through research, and some value lifelong learning and a growth mindset. My personal experiences and worldview influence what I might see as a reason we need professional learning, and as a corollary, influence what I think about when I consider how to evaluate impacts. Perceptions and tacit assumptions, like my confidence in the potential value of professional learning in improving education, can bias elements in data collection and analysis. In addition, as an insider, I have relationships with some of professional developers in my home province. Those relationships and the compact nature of the community may have prevented participants from honest disclosure.

I took several steps to minimize, but not eliminate, the factors that might bias my findings, to increase the credibility of the study. First, I consulted a provincial expert in professional learning as my participant in pilot study, and reframed both my data collection and analysis processes because of her advice. Next, I asked participating school divisions to select potential participants in my initial group for me, removing myself from the initial selection process, but potentially enabling divisions to influence the types of voices I heard. Finally, I selected

participants for diversity of role and perspective, including provincial officials, teacher union employees, and school division staff from both rural and urban divisions to provide breadth and depth of perspective. To address my own disciplinary perspectives about professional development evaluation as a practitioner and researcher, I drew from both psychology and program evaluation to address different frames for how we gather and interpret data to form judgments as professionals. During the process of memoing, I overtly described my assumptions and logic, and worked through a process of elaborating and questioning them. Finally, I sought and used the helpful advice of my advisor in questioning assumptions in the research literature and my writing and thinking. Constant memoing and the development and use of logic models also required me to make the tacit assumptions of the field explicit and helped me to analyze rather than presume or believe.

## **CHAPTER 4**

### **Introduction**

The purpose of a grounded theory study is to deeply understand a process through the eyes of a specific group of people that engage in that process, and generate theory to describe the process or processes. Chapter four uses the words of professional developers to delineate and compare their descriptions of their processes of self-assessment. The professional developers in this study framed the discussion of their self-assessment process through conditionality. Their descriptions were rich in contextual details related to the cultures of their workplaces and the teaching profession, and they frequently described their decision-making processes as contingent on those details. The context-related preconditions shaped their thinking and they frequently attributed them, either overtly or indirectly, as causal conditions that started or prevented self-assessment or intervening conditions that affected personal decision-making processes. Broadly, the precursor factors that influenced them fell into two main categories, systemic and personal. When describing attribution processes, participants often referenced multiple causal and intervening conditions in passing as determinants in their thinking processes. Generally, the described systemic conditions increased the barriers to assessment and evaluation, while the personal conditions decreased them and made assessment and evaluation more likely.

#### **Participant's descriptions of causal and intervening conditions**

In grounded theory, understanding the circumstances that give birth to and shape a process is an essential part of the understanding the overall theory. This chapter begins with the systemic factors and moves to the personal ones to establish the environment the theory relates to. The

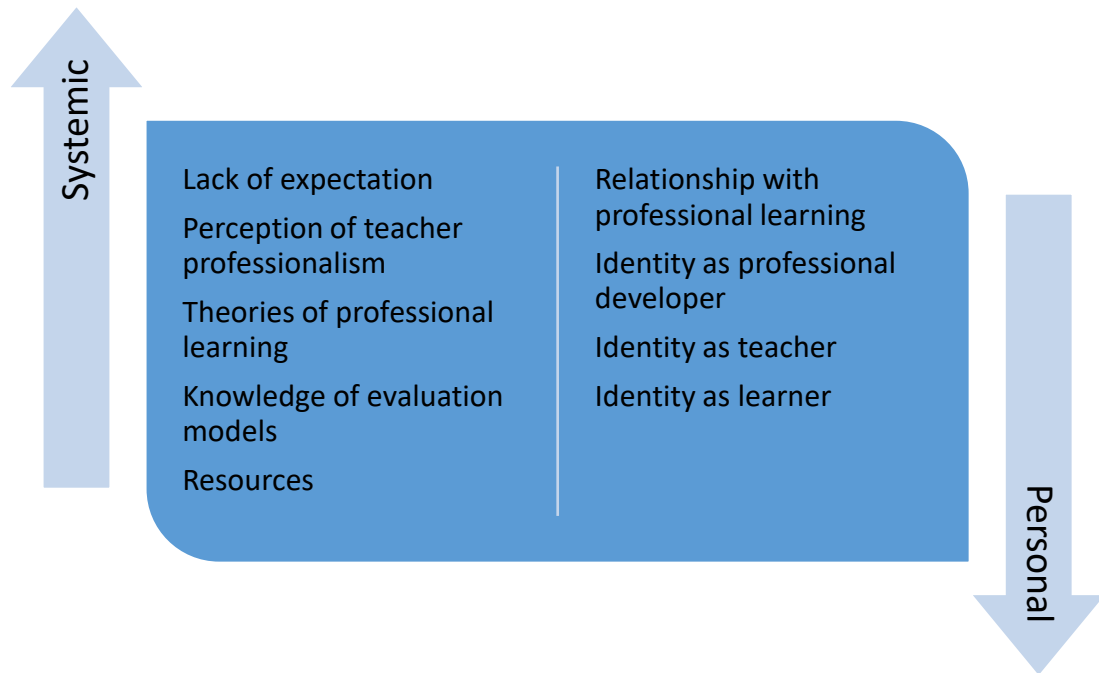


next part of the chapter weaves together the general thinking of the participants in the stages of the process. Elaboration of the theory and its implications occurs in chapter five.

### **Systemic conditions**

The elements described as systemic conditions in the study were factors that professional developers ascribed as outside of their control, either partially or significantly. Participants often responded to questions about their processes with initial descriptions of the circumstances they had to work within (see systemic conditions in Figure 4.1), indicating they believed an understanding of those circumstances was important in explaining their thinking or their actions. In addition to describing circumstances initially, the participants routinely added additional conditions when they explained their choices (see personal conditions, in Figure 4.1). For example, many participants tangentially referred to reduced resources for professional learning as part of a new reality and described ways in which they tried to make the best of circumstances they regarded as less than ideal. While some of the elements in the systemic category, like the professional developer's theory of professional learning, could be described as personal, they have been referred to as systemic because the professional developers described them as how professional learning is (when participants described elements of professional development as choices, they were categorized as personal). Broadly, systemic conditions included the expectations for evaluation of professional learning, the cultural understanding of professionalism, theories about professional learning, knowledge of evaluation models, and resources. The conditions are elements the participants described as critical external circumstances that affected their decision-making about leading and assessing professional learning. The systemic conditions include expectations for evaluation, theories of professional learning, knowledge of evaluation models, perception of teacher professionalism, and resources.

Figure 4.1 – Systemic and personal conditions



### **Expectation for evaluation**

Expectation for evaluation is the first of the conditions the professional developers described, but mostly because they were articulating they rarely experienced it. Reeves (2010) describes “evidence-free staff development” (p. 3) as pervasive. For something to be systematically supported, it should be a part of discourse within the workplace. Spontaneous references to expectations for evaluative practices might provide clues about systemic expectations or lack of them, but participants in the initial sample were also directly asked about potential expectations: “Who wants to know the impact of the professional learning you lead?” Given that professional learning is the main mechanism for creating change to teacher practice,

an education sector focused on the attainment of the sector's strategic goals would be likely to have different stakeholders who need information about the progress of professional learning. However, only one of the participants in the study described multiple stakeholders who were interested in the impact of professional learning, indicating a gap between goals and perceived need for evaluation.

Most commonly, participants took time to think about who might be interested in the impact of the professional learning, or they characterized the question about who wanted to know about the impact of professional learning as thought provoking, indicating it was not something they thought a lot about. Unlike other areas of focus in the interview process, prompts or additional questions were frequently needed to focus or extend thinking. With those prompts, some participants generated a person or group of people beyond themselves:

Jennifer: It's probably true, actually, no one asks I think.

Julie: No, we have one person who asks.

Jennifer: X does.

Julie: Yes, X asks.

Jennifer: And more, out of, how did it go?

Interviewer: And X is someone else you work with?

Jennifer: She's our supervisor of instruction, yeah, but the person we directly report to—there's an assumption that what we're doing is the right thing to do, I guess. Which is a very nice thing (pause).

Interviewer: Flattering? (laughs).

Jennifer: It is flattering, and I would hope I'm hired for that, right, that they trust our judgements.

When characterizing the lack of checking as a sign of trust, Jennifer stated:

And I do believe that our team really trusts our judgements *very* highly, which is a nice compliment. It would be nice to have someone say, “How did that make a difference?” Even these questions are challenging me today actually to think “Huh, I don’t know” it’s an interesting thought. I would say no one really.

Similarly, Juanita ascribed a lack of monitoring her impact as partially an indication that her superiors trusted the quality of her work.

Some participants, like Maxwell, wondered if their work as professional developers, or the work of professional developers they supervised, was assessed for impact at all:

I truly wondered, do they want to know what we are doing? I think about the two consultants [I supervise], I don’t know that anyone has ever asked or checked about the work they are doing and the impact it is having. I hear stories. People say to them, “I hear you are doing good work.” But I don’t know that they could ever...They would only want to know something if something wasn’t going well.

Sherry was unsure who wanted to know about her impact in her work leading professional learning provincially or in her school division.

Sherry: (pause) Who wants to know?

Interviewer: I know you want to know but –

Sherry: - about my impact, or the impact of the event?

Interviewer: You can tell me about both. The thing I really want to know is, if you lead an accreditation workshop, who wants to know, and how do they check, that it actually helped teachers?

Sherry: I have no idea.

Interviewer: Okay (laughs). So... no one checks, in your experience?

Sherry: Often participants are asked to complete an exit survey evaluating the PD [Professional Development]. I don't know if and how they check what participants think of me as facilitator, or how they know if the PD has affected practice in desired ways. I know I want to be personally evaluated, and to see that feedback.

For many participants, the question of who was or might be interested in the outcome of the work created a pause. The lack of ready answer might be attributable to a culture of "evidence-free staff development" as Reeve's suggests, or it might be connected to a lack of structural supports, a lack of information about gathering and using data, time constraints, or other issues.

Some participants did describe people who were interested in the results of professional learning. As an administrator, Pat saw a pattern in people asking about the impact of professional learning in two of the divisions he recently worked in:

Administrators themselves [want to know about results], when we share our learning journeys, which we do in x division. We have an opportunity to share our learning journeys at the end of the year in terms of what we have learned in our buildings and where we have gone. That, for me, was really informative to hear what other people are doing, just like we did in x other division when we did our end of the year story of what was learned, right?

Pat characterized "the board office" as wanting to "know how we are doing and have feedback in terms of the professional development that they are offering." He also mentioned students and parents as stakeholders who want to know how things are going at the school and that the provincial Ministry of Education was interested in understanding overall results: "My lens is a little bit different, having been a superintendent as well, but I think the Ministry is certainly

wanting to see what the results are. They are not into micromanaging what we are doing in professional development, but the outcomes are certainly of interest to the Ministry.”

Participants did not typically describe overt connections between the professional learning they were leading and a systematic attempt to monitor the impact of that learning on teacher classroom behavior by members of the education sector. Except for Pat, who described five ways that professional learning impacts are systemically discussed, most participants described few examples of systemic expectations for monitoring the impact of professional learning. Several participants referred to mandated data collection about student reading, but that was typically not paired with assessments of the professional learning, except when two administrators described learning walk processes connected to reading scores. Gwyneth and Pat, the two administrators, both described visiting classrooms to see the practices discussed in the professional learning being used, and Gwyneth gave an extended description of a process designed to help teachers visit classrooms without students and take pictures in the environment to look for evidence of specific classroom environment and instructional practices. Given that participants described a wide variety of data collection processes (like pre-assessments, concensograms, survey, feedback forms, observation in professional learning, and follow-up conversations) throughout the interview processes, and many described a personal interest in knowing they were making a difference, there seems to be some relationship between how educational systems are connecting data and professional learning internally for decision-making purposes. However, the relationship was not described as direct or explicit. Even when the study participants described mandated professional learning and mandated data collection as focused on the same target, no participants described a systematic internal or external process for assessing the impact of professional learning through to its impact on teacher practice and

student results. Administrators, who described themselves as responsible for the collective actions of a school, were the most likely to describe ongoing sources of data and actions related to interpretation of that data. Most forms of data that participants described using in decision-making were collected as a result of personal interest in the impact of a professional learning process and occurred as a part of facilitation, like Amy's observation of participants' notes on the large chart or Jennifer's description of listening in to conversations during facilitation. The lack of expectation for evaluation reduced the likelihood of it occurring and made its focus and process a personal decision of the professional developers.

### **Theories of professional learning**

The professional developers situated their comments about how they assess their own impact in their understanding of what effective professional learning is or is not, making is a condition of the theory, like the lack expectation for evaluation. The study participants engaged in theorizing about professional learning that supported their assumptions about what was likely to be effective and what issues to monitor. Sometimes they specifically referenced adult learning theories, specific theorists, or goals they had been given when speaking about the precursors of their processes of self-assessment. When the participants treated the theories of professional learning as facts that dictated planning or professional learning, they were included in the systemic categorization in this study. When participants attributed their thinking to personal epistemology or pedagogy, they were categorized as personal. In both systemic and personal thinking processes, participants made links between what they saw as essential to good professional learning and what they focused their self-assessments on.

## **Purpose of professional learning**

The third of the systemic conditions is the study participant's conceptualization of the purpose of professional learning. The process of program evaluation begins with understanding the purpose of the program being evaluated, and the participants' self-assessment was also focused on goals or inquiry questions. However, for nine participants, the goals tended to include both a specified focus articulated as a general idea rather than a measurable or time-bound goal. For all participants, the main goal of particular professional learning was typically accompanied by goals the participant extended to many professional learning events because the participant saw that goal as significant. When participants were asked what they were hoping to influence or change by leading professional learning, they all articulated that having professional learning as a mechanism related to specific topics were preset conditions of their work as assigned by their employers, for example, improving teacher's understanding of curricular outcomes. Once the given goal of professional learning and the content focus were articulated, all the participants described additional goals as systemically necessary for success of the professional learning in the long term, but personally decided. For example, Dean articulated that the professional learning he leads often has a content goal, like learning to understand the new science outcomes or developing shared resources for colleagues. The goals related to learning about renewed curriculum are the goals of his employment as a curricular consultant. However, he also described a more far-reaching purpose related to the on-going leadership of the instruction of science:

I hate the phrase "building capacity." I just hate it, but that is a big part of what I have to do. I look at that original Science 10 leadership team of 14 that we had from 2005-2007. Over half are consultants already. A bunch of them have said, "Well, it is because you



brought me into this and showed me I can do a whole bunch more things in education other than just being a teacher.” Obviously we need lots of people to be teachers, but we need people to lead in some of these other roles, too. I really see that as a huge part of the role.

The additional goals participants like Dean described were sometimes a focus of evaluation and sometimes articulated a higher calling. Participants referred to specific purposes that were content-related and required by an employer, like Pat or Julie’s focus on improving reading instruction and Sherry’s goal to help teachers understand the English curriculum. However, these goals were typically also paired with other goals, just as they were for Dean. Tracey talked about building effective leaders who could support her division’s vision of a successful classroom, and both Pat and Maxwell talked about building common beliefs to make respective work on reading assessment and culturally competent instruction successful. Gwyneth wanted her teachers to develop strong classroom environments focused on reading and cultural responsiveness as suggested by her division, but she also wanted her teachers to be effective observers and inquirers who see interdisciplinary intersections. The professional developers described their additional goals as personally determined but essential to the success of the employer-required goal.

Holding a variety of different goals is significant for several reasons. Many participants articulated more goals than one evaluation could undertake, making the evaluation either less likely happen for each goal, and/or less effective. Including additional goals that are far reaching goals designed to change or develop systems also adds greater complexity in trying to evaluate impact, as different purposes are best assessed using different processes (Patton, 2008). Having a few, clearly articulated goals makes evaluation easier, but no participants described just a few easy to measure goals.

Determining the goal was partially within the professional developer's locus of control. The professional developers described themselves as empowered to add the additional goals, meaning the complexity was self-imposed, but it was built on experiences that taught the professional developers that the additional goal was necessary. If the professional learning requires the complexity to achieve its purpose, but the complexity also makes its evaluation profoundly difficult, then systemic evaluation is at once more necessary and less likely to be effective, posing a dilemma for the professional developers, even as they begin a process of professional learning. Given that the complexity of adult learning and the nature of trying to plan for it (universally described as challenging by the professional developers), determining a purpose that can be effectively evaluated presented a significant challenge.

Determining purposes or goals was further complicated by professional developers' epistemology and pedagogy. Amy articulated that her purpose depends on her participants, which is similar to what Gwyneth and Jennifer described. They all related the professional learning to the division goals, but described learners as having distinct needs and capacities that dictated both goals and processes, as Amy did:

It is different for different people, and I think that is key to the whole thing. This is about people and this is about who are we working with. For some people, I would be hoping to influence their willingness to try something new, their willingness to change. For other people, I would be hoping to influence their willingness to share. Some people don't need me for the learning part of it, they are very naturally learning, and growing, and doing all kinds of awesome things. I am looking for ways to encourage those people to share that with others. And actually, those two kinds of people are linked. If the second group of

people would share more, it would have more influence on the first group of people than I ever could.

Amy's description reflects her belief in socially-constructed learning, which she described extensively throughout the interview processes. Like other participants, she explained "differentiating" professional learning because it was "essential to differentiate", making it less likely that a process is easy to evaluate because a greater variety of variables must be included. Amy's process of instruction and goals were conditional on her learner's attributes, because she believed the purpose of professional learning was to meet her teachers' learners' individual needs. Her conceptualization of the purpose and the resulting pedagogy made any process of evaluation more complex.

Juanita's description of the overall goal of all professional learning was focused specifically on students as "clients", and she described professional learning as trying to "impact learning positively", making her description the closest to the ultimate purpose articulated in professional learning theory. But even as she was describing the goal, she immediately connected it with difficulties in trying to determine if "you have actually helped student learning":

Lots of times it is really challenging because student achievement is impacted by so many different factors. The one thing that is a possible indicator that something has been successful... is over time we see that student learning is improving and maintaining itself. That means that whatever has been shared has been implemented. It also means you will see impact if you have a group that has received the professional development and a group that has not, because then you can see a change from one group to another and it eliminates that some of those factors are simply time-oriented. Something that everybody may have

been exposed to, or they may all have been part of. It is very difficult to separate, and I think sometimes what we do is we end up looking at perception as well, so qualitative data gathering in addition to the student achievement.

Interviewer: Can you give me an example of what kinds of things you might look at for perception or implementation? Because you are saying ‘sometimes I can see that it is there based on behavior’. How would you know that?

Juanita: One example might be survey that you would give to teachers. Not just following the professional development, because that is immediate feedback, but over time, two or three months later. If you have planned your professional development such that you can bring them back together again, you can talk about, “Okay, how did this change your practice? What changes did you see? Which of those changes were positive in terms of teaching, which of those changes were positive in terms of learning and which may have had negative impact, again for teaching and for learning?” So that perceptual survey is one piece. One thing we can also go to – and it depends on the question you ask, so sometimes you can pull things out of other student data such as the *Tell Them From Me [Our School]* in terms of overall learning experience. And if you see learning experiences over time, one would hope that some changes have occurred at the school level to lead to positive learning. You can kind of look backwards and see, what happened in that school? Oh, this is something that everyone attended, so it may not be a direct question, but sometimes you can link it back.

Juanita explicitly identified the problems she experiences when trying to link professional learning in general to goals for student learning, and how hard it is to isolate specific factors as a part of her process. When asked later in the interview what she might change about how school

divisions determine their impact, Juanita identified what she sees a critical, system-wide issue in the four school divisions she has been employed in by the province:

If I could change one thing, it would be to have a common expectation that professional development and training – all the professional development, be assessed to determine its impact immediately and then later after some time has passed. And, use some consistent questions so that we can actually see if it is impacting and having change over time.

Juanita described a progression of thinking about each of logical links of professional learning, from participants liking it, to their understanding increasing, and their classroom practice changing. She linked all of the steps to the ultimate purpose, then articulated the complexity of knowing causality with all the variables. While Juanita articulated strong personal attachment to the research-articulated purpose of professional learning, she also stated serious reservations about the frequency and effectiveness of evaluation of professional learning.

Gwyneth linked her goals for the professional learning and an iterative process of finding out what was working that took the school she was leading three years. She was focused on teachers having a compelling vision and observables that linked classroom environment and instruction when thinking about reading, math, and cultural responsiveness. Because the vision was complex and wove different complex ideas together, she described an extended process for helping staff understand the goals deeply. Gwyneth noted this was essential, and that leaders, teachers, and parents often make the mistake of “stealing the learning.” She described her personal constructivist vision for goals in professional learning as a layered process:

We spend a little time co-constructing the beginning of ... initiatives and then we really forget to delve in and to come back. And so I think the work that I'm most proud of doing, there's been a vision of ourselves as learners, as an adult a vision of our students learning,

the vision of the community as a whole , that is layered. A vision is layered. And it's really, really complex when you start to piece it out. And I think that the other thing I would enjoy or at least articulate that really had an impact and that worked well, was helping staff articulate their vision. And developing a common language, a common understanding as we're moving forward, so you're really talking that language.... I can't do the work for people. It's not what a great teacher, a great leader, a parent does, doesn't take away the learning of others. They help construct opportunities for other people to see the learning. Right? That's really, really important.

Gwyneth's goal setting process was complex, and included the nature of her staff, her identity as a teacher, learner, and leader, and the division initiatives. Gwyneth tried to lead process that helped staff look for commonalities. However, Gwyneth described the complexity of interleaving all the different goals as increasing the time, energy, and difficulty of knowing the professional learning made a difference. Part of the way she tackled the process was to ensuring learning was visible and expressly linked to the goals by placing celebrations of specific children's progress towards the goal in a highly visible place:

Looking at all the learning, the professional development that we have taken as a staff, lets me draw a straight line from our A3 [strategic plan], to my CIT [collaborative inquiry teams teachers are a part of] to my PLP [personal learning plan written by a teacher], to that child on the wall. There is my impact. I specifically chose a very, very, visible part of the school, where they were front and center. It would be hard to go to the bathroom, go to the office, or go to copy without seeing the learners. It became so visible that we changed it two or three times a year. The parents loved it, they would come and see it. It took the

very simple concept of labeling and visualizing our evidence along the way to that next level, then it took it to a community level.

Gwyneth made overt links between the goals, the professional learning, and the student progress literally visible to keep the staff focused on achieving the goals and understanding how all their efforts worked towards them.

Each of the systemic conditions, as articulated by each professional developer, influenced both the potential success and likelihood of any formal evaluation. First, the professional developers had multiple goals, and only some were designed to be specific or measurable. Second, many professional developers did not link the professional learning they led to goals for changing student learning, and articulated they did not have the opportunity or mechanism to make a direct line to formal practice. Finally, those who articulated thinking deeply and consistently about their goals and evaluation, like Michelle, Gwyneth, Pat, or Juanita, described difficulties in isolating their impact from other factors and making any form of causal link. Each issue with evaluating the professional learning based on purpose presents a significant, system-wide barrier to professional developers' self-assessment or alternative external evaluations.

### **Process of professional learning**

Like the purposes professional developers articulate, the process used by professional developers also has systemic implications for the likelihood and success of evaluation processes. While professional developers did not connect expectations for evaluation or their purpose with routine self-assessment, they made overt ties frequently between self-selected processes for facilitation and self-assessment. While they characterized facilitation practices as external, because they referenced research and how they had been taught, they articulated substantial

personal agency in selecting facilitation processes. They also connected those processes explicitly to information they gathered to self-assess.

Michelle stated her process for constructing a sequence of professional learning activities she learned as a part of her current job. She starts with what she calls “the dream,” which is vision for five years out, and supports teachers in discussing “what you’d need to get it and what you need to do differently to get there.” Michelle described an example where grade three teachers “might know about math stations but can’t make it work.” She noted the “knowing/doing gap” as significant in helping her determine where to focus the content of the professional learning. Michelle described how “when a teacher can speak to the how and competently knows the part, but still doesn’t do it, it shows me there are pieces missing or maybe it works never because the kind of teacher I am.” She stated that often “solid pedagogy and understanding are not the issue”, knowing how to make something work is. The goal for the professional learning is co-determined with the teachers, and Michelle wants it to be clear. Michelle described the goal as the foundation for the selection of processes for professional learning.

In each example she described, Michelle linked her formative assessment of the learning process with her theory of professional learning as a part of a specific process. Since she believed that a well-articulated “dream” provides motivation and direction, she assesses the mind map of the dream for its “depth” and complexity, and looks to it for potential topics. Because she has experiences that tell her the “knowing/doing gap” is often a source of the problem, she listens to facilitated conversations to see if teacher statements show lack of understanding conceptually or trouble making something work in the classroom. Then she listens to see what potential causes for difficulty enacting in the classroom might be, to see if she can target the



professional learning at resolving it. Michelle constructs the process with the teachers, so much of her attempt to determine her impact early in a process is directed at pre-assessment and formative assessment. Michelle articulated that before she learned her current process, she was offering professional learning, but was much less clear on the process she could use or how to determine her impact.

Jennifer described adult learning as constructivist, saying a discussion process was helpful because, “it honors their professionalism, and it allows them to share their own expertise, question their own learning or their own practice in their classroom, and get advice or ideas from colleagues. I think that’s very critical.” Like Julie and Michelle, she described a process of self-assessment that focused on listening to discussion among teachers during professional learning. As funding for professional learning was reduced and more professional learning occurred online, she described a reduction in information about the learning process, and she responded by providing a plethora of choices so adults could construct based on their own interests. However, both Jennifer and Julie indicated that providing many choices delivered online makes it harder to assess your personal impact. They used administrator informal comments to assess their impact, assuming the lack of questions and concerns likely meant things were going well during the second year. The process for professional learning dictated both of their ability to assess their impact, although they referred to a much more extensive consultation process as they first moved to the new model. Jennifer, Julie, and Michelle made explicit links between what they had teachers do and what they believed about effective professional learning. They also described deliberately constructing processes that gave them conversational data and observational data related to their success in helping teachers understand or reduce barriers. Each also had clear assumptions underlying the decision-making they were doing.

Sherry also described framing professional learning processes partially around an understanding of adult learners: “Often we talk about who is the adult learner? What do we do with the adult learner, quite specially?” She noted that many people offer constructivist professional learning based on a belief about how adults learn best: “One of the assumptions is that adult learners need to be free to construct their own meaning. And that adult learners working in a group to do so will self-correct.” Since Sherry was not sure adults always had the information needed to self-correct or generate deep learning, much of her evidence gathering focused on the conceptual understanding of big ideas, like the curricular Broad Areas of Learning, and listening for teachers’ questions and wonders in an inquiry process. Sherry built her processes based on her beliefs and experience about the processes most likely to help teachers, and then determined her impact by observing teachers in those processes and seeking feedback from them.

Like Sherry, Juanita listened and observed during and after processes to determine her impact. Juanita noted that in addition to constructing learning, transmission can also be an effective process because educators are used to learning by listening: “It is how they want to learn. I think that many of us grew up as students and young teachers in an environment where we were given information. So we were very good at doing school. Many of us as professionals are good at doing school. We are good at sitting and receiving information.” When transmission is the goal of the process, Juanita uses successful transmission as an indicator of success. In processes with distributive leadership groups, she listened for how information is understood, internalized, and acted on. In describing her thinking process while facilitating, she most frequently attributed her decision-making process to judgements about the level of understanding and use of information. She referred to specific words her participants used that she saw as

indicating comprehension or barriers, then described how she framed her responses to the teacher learning. Like the other participants, she build the self-assessment into the facilitation process because she said she is deeply committed to being a good facilitator, and needs the information to improve.

Gwyneth's process of professional learning and process for assessment were inextricably linked. She described a process of "gradual release" where participants co-constructed criteria, looked at many examples and compared them to the criteria and thought deeply about what each criterion meant for them as adults and for their students. Gwyneth paired early conversational data from professional learning with what she saw on learning walks, and invited her staff to do the same during professional learning. Self-assessment occurred for Gwyneth and for her teachers at all five levels of Guskey's (2000) levels of evaluation, but not initially. Over the three year process, Gwyneth described facilitation processes purposefully constructed to have staff encounter their own data at various levels and make their own decisions related to that data. The most "important data" was always connected back to the goals of the professional learning, and processed Gwyneth and her data team, then by her teachers themselves. While time consuming, Gwyneth's process of cross-pollinating professional learning and data collection for extended time offers a potential way to address teachers owning their own learning but being responsible to student results.

Like Juanita, each professional developer framed the most useful data through the lens of the purpose of the professional learning, and used facilitation processes that yielded what Gwyneth called "visible learning" to determine if the learning was effective. In most examples used by participants, the self-assessment was focused on the success of the process in helping participants acquire new knowledge or skills or helping reduce barriers that were preventing the

learning from being used. Sometimes data about what people attending the professional learning actually did with the knowledge or how it ultimately affected students was gathered, but participants universally expressed that information was hard to obtain. The sole exception was Gwyneth, who described data about teacher actions and student results as a way of thinking about professional learning that was essential to the visible learning process. All of the professional developers, including Gwyneth and Sherry, described being cautious in gathering such data for a wide variety of reasons, most notably teacher professionalism and implied judgement. However, Gwyneth and Sherry described their caution in terms of a careful walking along beside teachers and using gradual release, rather than an automatic limiting factor.

### **Learning to evaluate**

Trying to understand the process of evaluation expressed by the professional developers requires an understanding of their cognitive processes and experiences. The professional developers did not typically describe establishing processes for planning evaluation or engaging sensemaking processes using the language of program evaluation or formal research, except for Michelle's use of the concept of a logic model and Terry's descriptions of elements of evaluation in the pre-study. Many said they had no formal training or had read little about evaluation. Three referenced research in professional learning evaluation, and all who did received formal training in facilitation from one provincial body. Michelle described her evaluation thinking when describing the backward planning of the professional learning process:

So that process is the brainstorming process, and we use Guskey's four levels or five levels of professional development but flipped around. So we start by brainstorming what we hope we would see from kids in the room or what we hope students would experience after a teacher had come to our PD, and then work backwards.

Like Michelle, participants universally linked evaluation and processes for leading professional learning. More than half of the participants spontaneously volunteered information about training in facilitation, or professional learning about facilitation, when asked directly about their learning about evaluation. The same participants did not reference specific learning about evaluation. Juanita summarized an identity as a facilitator that many participants expressed: “I got this complex skill set from facilitation workshops, but mostly I love facilitating and really care about it. I want my work to be good. I want their work [workshop participants] to be good and I want people to tell me if I am doing something wrong.” The desire to help teachers learn and be effective as a facilitator or leader was the main motivation for evaluation that participants expressed. Understanding the influence of their circumstances and identities provides helpful clues in their process of sensemaking and attribution, particularly in the absence of formal evaluation training.

Julie noted that she had no training about evaluation of professional learning or connecting it to facilitation. She said, there was “never any training. I think in my case, its common sense. Yes, I find myself using some of the strategies or techniques I used as a teacher. Even some of the ones I used when I taught grade one core French.” She described observing professional learning to learn the skills she needed: “So I went to a PD session with X, and I was thinking, ‘What was so great about her session?’ So thinking back – ‘Oh right! This, this, this. How do I bring that into what we’re trying to plan?’” Jennifer characterized it as “beg, borrow, and steal” with a laugh, and Julie called it “mentoring.” Julie compared learning facilitation skills to learning to teach:

We always do that as teachers, I think. Especially at the beginning of my career, I always took bits from different teachers that I thought, “Oh, that probably would work for me, for

my personality—and I think my students would respond to that.” Most of my learning I learned in the classroom, on the job, from colleagues, and from my co-op teachers when I was interning. Very little of what I did at university I was able to transfer and use. Some, but not as much as I thought.

Julie described connecting to strategies because she found them personally exciting, then wanted to use them. Both Jennifer and Julie emphasized the value of working with others in polishing their craft as facilitators, because you get to observe what others do well and learn how to do it. They also described times when formal training on specific resources was provided and helped them with the content, but they emphasized the ability to work collaboratively. Jennifer said, “I really feel for me, the team approach in the last two years has really helped me feel more confident in the professional development that I’m delivering.”

Other participants echoed the value of observing other professional developers. Maxwell described learning facilitation strategies from watching others, and Gwyneth described the value of watching expert facilitators, like Karen Hume, or other administrators as mentors. Michelle, Sherry, and Amy described working with peers to plan and lead facilitation as shaping their thinking and promoting reflections. As members of a group of facilitators in the province receiving specific training and practice in facilitation (Michelle called it being “indoctrinated into the SPDU method of planning”), Michelle, Sherry, and Amy described the training and collaboration often when describing their processes in learning to facilitate well. They also all referred specifically to using facilitation strategies to gain information they needed for responses. All study participants, even Dean who had the least opportunity to facilitate, described the value of trying to anticipate potential needs, plan learning opportunities with others, and debrief the process. In most cases this cooperative learning process was only focused on evaluation in areas

where it might inform the facilitation, but it provides useful clues to about potential directions for learning more about evaluation of professional learning. The professional developers in the study expressed a strong desire to be effective but little exposure to formal training in how to evaluate if they were effective compared to formal and informal learning about facilitation. However, they used multiple tools to try to determine their own impact anyway.

## **Resources**

This study occurred at a time of financial constraint in K-12 education in Saskatchewan. Over the course of the study, the pool of potential participants was reduced as school divisions returned coaches, consultants, and coordinators to the classroom to stretch core operating grants from the provincial government. In addition, study participants described careful calculations about how to use fewer funds to support teacher learning. Reduced time for learning was a concern for many participants. Participants characterized resources as important, and given the provincial context, consistently referred to reductions in resource levels or outright insufficient resources.

A lack of resources was described by all participants, but they articulated different issues with resources. Julie described insufficient time: “I would like for us, in the context of what we’re dealing with, to have more PD days for teachers to receive training because the reality is that we’re very limited in what we can provide.” Amy referred to time constraints and the time available, noting “there isn’t the meaningful time.” She stated professional learning is often slotted into after school times when teachers are tired or during “report card time”, and concluded it was unlikely to be effective given those time periods, so there was little point in assessing it. She referred to the value of “extended time with the participants” and “processing time” that are hard to get for the professional learning she leads within her school, but more

available when she facilitated externally. Amy's description of specific types of time for good professional learning aligns with Learning Forward Standards (2011) and other newer standards, although she did not refer to them.

Maxwell noted his goal of changing beliefs often requires time he "doesn't have." He describes genuine change as taking "time and resources" and stated follow up as critical, but missing given the time and resources he has been allocated. Both Jennifer and Julie described time as their greatest barrier to success. Juanita was specific in her calculations:

For example in our system, just the way we have it set up, for internal PD we are allowed 20 substitutes per day. That is 4000 days of teacher release every year, if you are there every day. Take off the days where there is no school, because there are only 197 days, and those days that don't need teacher substitutes, because they are non-student days, take off two days before and after every major holiday, because teachers take personal days because it is part of their LINC, their contract, and take off one day on either side of a short holiday [long weekend] because again they take extended leave and you are down to 3,500 days of professional development or less. We give 1000 of those days to our schools based on a formula. We give those days as Professional Learning for Schools, and we don't want to touch those because they use those to move forward internally. So that leaves us with 2,500 days as a system. It is very hard to do anything with 2,500 days as a system. And last year, because of budget, we cut it to 1,300. 1,300 days of subs across a system. It sounds like a big number.

Interviewer: I know it is not.

Juanita: When you have 15,000 staff, it is less than a day per person of PD, and you don't distribute it that way. Some people get more, others get none. That, for example, is why I



need to really scrutinize the PD that we offer. Because I need to offer what is going to have the greatest impact, have the greatest collective impact, need to make sure it is sustainable, so it can't be a one time, it should really be one or two times that we gather, and it has to be extremely targeted. It can't just be it would be nice if you came. If you are coming you are required to be here, and I am not bringing anybody that doesn't really need it. That is really limiting! Everyone deserves the opportunity to grow professionally, and everyone should be growing professionally. It is often those who don't think they need to that are the ones who need it the most. So that makes me scrutinize professional development. I go into meetings, and I have to fight, fight, fight for support for professional development. It is not people don't think it is important, but the resources just aren't there. Further to that, we have cut all of our instructional consultants next year. We will not have one. We will have coordinators, and my supervisor role remains, but we have over time built our consultants up to about 24. We are down to zero next year.

The perception of a lack of time and/or resources was consistent and universally negative for participants, who expressed the frustration Juanita did. Non-school division participants, like Dean, spoke of times when they had more opportunity to work with teachers or more resources due to more staff, and note that assessment of their own impact was more likely. Participants also all articulated a lack of control of the amount of time or resources allocated overall, although some could choose the internal allocation of what they perceived as limited resources. Almost all participants suggested ways to get more time or commented on the short-sighted nature of government cuts. While time was reported as a fundamental systemic barrier to getting to the evaluation process or believing the professional learning has sufficient resources to impact teachers, participants' descriptions of decision-making about evaluation diverged.

There was a noteworthy difference in participants' perceptions of the need to evaluate in the face of scarce resources. Many felt there was little point in evaluating since they felt they knew they were offering professional learning too limited to have the desired effect. Others, like Sherry, Gwyneth, Michelle, and Juanita described a greater need to assess their personal impact if resources were limited. Those same participants gave more examples of how they tried to assess their impact and all stated that the best evidence of their personal impact was observed changes in teacher behavior. They also all suggested student data was needed to determine if the professional learning was effective and expressed the difficulty of trying to determine causality. Since they generally articulated, as Juanita did, that they "thought a lot" about how to make a difference, further research about the implication of personal agency or exposure to ideas about evaluation in the sensemaking process would be helpful. Three participants came from different employers and different roles (learning leaders, supervisor and professional developer for the teacher's union), but expressed common thoughts about the need to know your own impact to feel effective. The mix of systemic and personal factors in their thinking about scarcity may have significance in understanding the way professional developers construct a process of self-assessment.

### **Evaluation and negativity**

Like a lack of resources, the idea that evaluation is negative formed a barrier to self-assessment for some of the professional developers in the study. While they expressed interest in understanding their impact of teachers and students, some participants also expressed concern regarding perceived criticism. Maxwell noted that he sometimes needs confront negativity about the goals or likelihood of success of professional learning he is leading:

I have had many people say to me, ‘Who is going to be to blame if we don’t reach the targets for graduation in 2020?’ I have said to those people, ‘Blame me. If it helps you move this work forward and continue the work, then blame me. I will take the blame. But don’t ever use that as an excuse to stop doing this work, because that is what you are looking for. You are looking for a reason to not do this.’

Maxwell noted the problem came up often, but the goal was so important to him personally that he continued to lead the professional learning although he felt response to the goal was sometimes negative.

Like Maxwell, Dean noted taking criticism is a part of facilitating. He noted that he often asks teachers to help facilitate, but he needs to defend them from the criticisms of their peers: “I jump in, I guess it is part of my structure, whenever the questioning gets to the point. Whether it is a question my fellow facilitators shouldn’t answer – I don’t think anybody should take any flack for Ministry decisions. That is my job.” Dean notes the criticism is not always about the curriculum, which he wrote and supports teachers to adopt. He says teachers routinely critique many things about professional learning, and he argued developing an immunity to criticism while still caring about your affect is a critical challenge for anyone leading professional learning: “I think that is a huge role that whoever facilitates PD has, to know when to step in and take the crap. Whether it is about a meal, a room, something in the curriculum, whatever. It is not always about the curriculum.”

Gwyneth spoke about negatives only twice, and always when comparing them to positives. She described negative information as useful data that was not personal but yielded valuable intellectual data about potential issues to be addressed: “It was important for people to be able to label and visualize what was moving us forward in both positive and negative ways. It

was very important to see what was working well that we would continue to do well and replicate, and what was a barrier, and we just needed to remove it.” Because Gwyneth’s process was always collaborative with her staff and she had extended time to develop relationships, she was able contextualize “negatives” as opportunities for action, not personal indictments. She described focusing on the intellectual implications as a part of her personality.

Juanita noted negativity has a real impact on her because of her personality: “I am a pleaser and I want to do well on everything. I reflect on everything all the time. I often think a lot and get emotional at the end of the meeting.” When thinking aloud about a video of her facilitation, she described how she tries to help teachers who are facilitating try to deal with negative feedback by framing the comment as a “common challenge” or trying “to relate the challenges I am observing to solutions they have already generated to focus on their skills and solutions.” Juanita was distressed when she watched video and wanted herself to respond differently: “In video, I realized I should have affirmed her when she identified the problem and justified. I just moved on. I wish I had affirmed and offered problem solving.” She also expressed concern about having to tell anyone they might need to do something differently, as she was concerned about her personal position, wanted to be a part of the team, and did not want to harm another teacher:

I struggled to give the negative feedback. I want to be a team player and it is a public concern I am raising with in scope teacher leaders. I don’t want to undermine the teacher with the principal. A principal’s voice is the most important in our system. I am concerned and want to share it in a way that is appropriate.

Juanita's concerns echoed the concerns other participants expressed when discussing the potential of evaluation leading to criticism of teachers or a loss of teacher autonomy. Those concerns about negativity were often linked to a desire to respect teacher professionalism.

### **“Drinking the Kool-Aid” and professionalism**

When the professional developers in this study were asked to describe how they determine their impact on teacher practice or student results, they were always careful to couch their descriptions in a belief in teacher professionalism and choice. Over and over again, they described providing what Isabel characterized as options for teachers; then they noted teachers could choose whether to use the new information. Amy contrasted the level of professional discretion her adult learners have with her work as a high school teacher, where attempting the learning was not optional. As she discussed it, she expressed concern about infringing on teacher professional responsibility and choice. While the participants universally stated they wanted to know what they offered to teachers influenced the classroom and students, they did not want to be seen as evaluating teacher practice, even when their roles required it, as Tracey's, Gwyneth's, Amy's, or Pat's roles did. Each professional developer contextualized personal self-assessment processes as a careful process that needed to respect teacher professionalism. Dean referred to it as drinking the Kool-Aid - a process whereby you facilitate professional learning and believe that if teachers tell you it is useful, then it worked: “It is so weird. In my world we facilitate a lot, we deliver a lot, and then we don't get that feedback very much. You have to somehow *believe* (participant's emphasis), drink the Kool-Aid and *believe*.”

Jennifer described the influence of role when she characterized her learners as equals in a co-learning process: “They are professionals too, I mean I'm really at the same level as all the people that I'm working with, I just have a role where I'm doing professional development with

them.” She expressed concern about seeming to evaluate peers. Michelle also emphasized the professional learning process as learning for everyone including her, and a co-construction with peers. Michelle specifically described instances where her employer might have a contract with a school division that wants specific changes in teacher practice. She said, “It is my job to empower teachers” and fixing is “not the goal. I don’t fix or report on teachers.” Without exception, participants explicitly reinforced a belief in a teacher’s choices, which they described as at odds with evaluation. Gwyneth explain that she was substituting the word *assessment* for *evaluation* in responding to interview questions specifically because evaluation has a connotation of judgement or supervision. Although as a principal, she has a teacher supervision role, like all the other participants she distanced professional learning from supervision and avoided the use of the term evaluation. Gwyneth described the intersection of evaluation and supervision as problematic.

The professional developers in the study who also have the role of supervisor were all equally clear in describing professional learning as both optional and expected in a way that causes cognitive dissonance. Amy stated: “I guess adult learning is always invitational to a certain degree. You can’t force that...It is not invitational that you learn. It is invitational how and when you learn.” She acknowledged the complexity of her approach to adult learning: “There is that professional respect that you will make some meaning of this, you will engage with it in your own way in your own time. There are challenges with that. Does that always happen? No. I guess I feel like I have to start from that premise. I have to start from that premise that you are a professional and you will make that happen.” She also described a belief that: “We have this collective responsibility to learn together, not only for your individual professional growth, but we have responsibility to further the program of this school and school

division, and we have a responsibility to do that together.” As a vice-principal and professional developer she concluded: “It really does come from a place that I want to respect you as a professional. I am sending the message that this is what we do at x school and this is what lifelong learning looks like. But I am just putting it out there, and you are going to do with it what you will. The onus for your professional learning is not on me. It is on you.” Amy’s characterization may reflect thinking about a key difference in adult learning.

Other participants used similar characterizations that reduced any perceived authority to enforce or judge learning of teachers, like Tracey’s position as superintendent, which she characterized as “supporting adults in their learning journey.” Pat, as principal, as also positioned himself as an equal in a professional learning process: “I have been blessed to work with good professionals. When we work with teachers who are learners, it allows us to understand and work together, learn alongside.” Peer relationships in adult learning were consistently articulated when participants were asked about the process of self-assessment. They problematized the process of self-assessment through its potential to be seen as rendering judgement on teachers themselves.

Jennifer explained the conundrum faced by the professional developer when the concept of evaluation is introduced, because evaluation may be akin to holding teachers accountable for specific results: “I think accountability would also be part of that too. We can do as much PD as we want, but the actual taking it back and implementing it into your classroom is optional. The expectations around that aren’t always clear or...monitored. Enforced.” Jennifer described professional learning as a supportive offering and stated trusting your impact was a part of the process.

The sole person to trouble the concept of professionalism was Sherry, a teacher-coach participant in the study. She described the complexity of the problem of nurturing adult learning, characterizing professional learners as “more like orchids than cactus”. Sherry argued teachers “need to feel that what they do in their classrooms matters” and “believe that their organizations are investing in them”, which she described as potentially at odds with leaving teachers alone to make their own best decisions. Sherry argued professionalism and responsibility are actually mutual not antithetical, and that as instructional coach, she needs to avoid fearing overt discussion of teacher practice:

We have to get over our fear. We have to increase our capacity for difficult conversations. For example, I have to be able to talk about outcomes and assessment, for example, including assessments designed to give students easy and ready access to multiple outcomes all at once. But if I am hesitant to say, outcomes, because I might be pushing someone faster than they want to go, it doesn't matter how genius the PD opportunities are in the division. It is saying to everybody, “It doesn't matter if you change. It doesn't matter if you grow.” And I have said earlier, that is more deeply problematic because it can seem to imply, “It doesn't matter. You don't matter. What you do, not really that important. Whether you are a mediocre teacher or a great teacher, not really that important.” But people are giving their lives to this, their whole lives to this, and I think that we want to be truly excellent.

Like other participants, Sherry repeatedly articulated caring about teachers and valuing their judgement, but she also articulated a concept of responsibility that both Amy and Tracey expressed. All three participants described wrestling with the issues of teacher responsibility and



professionalism while attempting to come to terms with their own self-assessment of the impact of the professional learning they lead.

Gwyneth was the sole participant who did not refer to professionalism once in the interview. Her perspective was closest to Sherry's; Sherry believed teachers needed support in tackling hard issues and leaving them to make judgement calls about what to learn was a mistake, but she characterized it as a belief in the power of all teachers. Gwyneth credited the crystallization of this thinking to one of the more effective instances of professional learning she experienced as a leader:

One of those would have been the community education action plan as it once was known. It has taken different incarnations throughout the years. In its early years, I was very, very fortunate to be a part of that when Lyn Sharratt was here. My own predilection and my own biases, the research that I was looking at [were addressed]. Think of her core belief: 'All teachers can teach to high standards given the right supports.' I had already been doing that work. I had been thinking about that already but you put it in one sentence for me. She really helped me see the work. That was really informative.

Gwyneth assumed that everyone on her staff, including her, would be learning what students needed them to know. She built the choice elements of professionalism into being part of a team planning for action together, but she did not leave it up to her professional learners to decide to learn. Gwyneth described the belief that if the professional learning was effective, teaching would change, but she articulated change in instruction as an ongoing, supported process, not an independent one, and she checked forward progress on her goals for her school for over three years.

As a part of the theoretical sample, when Gwyneth was asked why she was substituting the word assessment for the word evaluation, she came the closest to discussing anything as negative, even as she reinforced her focus as positive. Her thinking was not directly about professionalism, but rather, was focused on evaluation as an anathema to a “learning journey”, a phrase Terry, Tracey, and Amy also used:

Why the preference for using assessment over evaluation? I think I prefer using assessment because for me it is more organic. Assessment for me evokes images of cycles of learning and perpetual growth. Assessment in my mind, is typically appreciative in its orientation. I use formative assessment practices to monitor growth of my students and teacher and use it to monitor my own leadership practice. Evaluation on the other end, is for me more pejorative in nature. It could be the influence of constructivism on my educational practice but evaluation does not evoke a journey for me, but rather an end. In order to respond to the complexities and demands of educational leadership, assessment rather than evaluation has been more helpful in deepening the rigor of my work.

Leadership and learning are a never-ending story.

My lens as professional developer was noticeably different than my participants in thinking about professionalism. I knew concerns about respecting teacher autonomy were common, but did not see them as mutually exclusive with evaluation of the degree to which teachers implement an instructional practice they are learning about. Initially, participants were not asked about the concept of professionalism as a result. However, how much evaluation of professional learning might infringe on teacher professional discretion was raised tangentially or directly by each participant in the initial sample without the participants being asked about a link, indicating that the two ideas have a link in the minds of the study participants. When overtly asked as a

part of the theoretical sample, participants confirmed the perceived connection. Most participants in both sample groups also expressed that the goal of professional learning (to change teacher practice, and therefore, student results) must be viewed through the lens of teacher choice and autonomy. However, there was also a strong cognitive dissonance, as professional developers articulated that they were often suggesting new instructional or assessment practices specifically because they believed them to be superior enough that it was worth teacher's time to make the change. In that sense, professional developers described various forms of "drinking the Kool-Aid", or believing that if the professional development was sufficiently persuasive and effective, it would make a difference.

Like all systemic causal conditions, a desire to protect teacher professional discretion was a concept that the professional developers described as influencing their process of evaluation. They mentioned it when making decisions about goals, choosing process for professional learning, and determining what good evidence of success was. The systemic causal conditions were factors that professional developers perceived as (at least partially) affecting decisions they made about assessing their own impact. The conditions provide a useful window into elements that could be a focus for further research or action designed to increase the frequency or effectiveness of evaluating professional learning. Since professional developers have the greatest access to planning and self-assessment, the systemic causal conditions they commonly articulate provide useful information about potential systemic barriers to the process of evaluation of the professional learning.

### **Personal causal conditions**

In the pre-study, Terry framed her motivation to lead professional learning and her desire to know it was helpful through her identity. She drew on two key areas, formal thinking about

program evaluation and her personal experiences. Terry theorized that few professional developers were likely to have formal training, and that the lack of theory might make personal experience even more significant when understanding professional developers' self-assessments. During the interviews and observations with her in the pre-study, she referenced a personal love of learning, how much she loved professional learning as a teacher, and how she planned and assessed as a teacher. Terry's interest in relating her evaluation practice to her personal identity and experiences helped to shape questions for the participants in the main study. Initial sample participants were asked specifically about their own professional development and teaching experiences to provide them with the opportunity to connect personal elements with their evaluation practice.

### **Identity as a learner**

The participants typically characterized themselves as learners. Some described it as a professional state, and a few described it as a key part of their personal identity. Julie described herself as using professional learning regardless of its quality because she loves to learn:

I'm always interested in learning—that has never been an issue for me. We were talking about that just before you arrived. I've always been of the frame of mind that I'd go to a PD session when I was a teacher and I'd always find something to take back with me, even if it might not have been the perfect fit for my situation or my context.

Like all the study participants, she described learning and changing as a key part of her identity. Julie's love of professional learning is not universal, and she described circumstances in professional learning she led where her participants did not share her love of professional growth. There may be interesting implications if professional developers who love learning plan

what might work for them for a wider audience, and then need to determine if the professional learning is working for everyone.

Juanita described childhood experiences and playing classroom as early examples of a learning disposition: “My mom was a teacher-librarian and I loved her. I saw what she did. I wanted to be her. She was a teacher-librarian and I had my own little library. We lived on a farm, so I signed books out to my stuffed animals. It is pathetic. But I did, I even had the stamps and cards and everything.” She said she has “always been a learner. I have always wanted to be a teacher.” Juanita characterized herself as the student who is very eager to learn, “I was one of those kids in school. I was in the front. I put my hand up for everything, and I was your typical A student who was a teacher pleaser, teacher wannabe.”

Tracey described the type of thinking she loved to do as a teacher when working with other teachers on improving her classroom: “We had to talk big, then specifically what are we going to do next week in our lessons, then big again, then specific. It was really messy, and the people I was working with were really intense. They really had my head going.” Tracey noted that thinking about big questions and making sense of things for herself are personally motivating for her, and she drew from those experiences when developing her own philosophy of what makes effective professional learning.

Almost all the professional developers in the study described themselves as learners who like to help teachers, sometimes with something specific. Sherry, Gwyneth, and Maxwell also describe specific motivation beyond their identities as learners and helpers. Sherry described caring about the Broad Areas of Learning in the curriculum and critical thinking skills on a principled level, because she saw them as important for building a good society. She articulated a commitment to student learning of those skills to improve the quality of citizenship and

societal action. Gwyneth described a love of anything “interdisciplinary” that had been with her from university classes, to her work as a teacher and leader, because she was always seeing connections. Maxwell articulated a calling beyond learning, which he linked to his personal cultural identity. He described noticing how certain students were marginalized by schooling while he was engaged in an early professional development role: “It was just a matter of sitting back and seeing that our First Nations and Métis kids were being marginalized in a way that I had not seen before.” Maxwell articulated, “I guess it just became personal. It hit home that this is the opportunity that I have, through professional development, to start to change things.” He later described it in the most personal terms, through a discussion of being a parent:

I am the father of a child - I only have one child. There is a 70% change he is not going to graduate on time, just because of his background, his ethnicity. I said to myself, how come? How come we are okay with that as a system? Why is this acceptable? I have challenged many people in conversations we have had when they say, “Why are we focusing on our First Nations and Métis kids?” Is there any other group of students where we’d be okay with that same achievement level? And nobody can give me an answer!

For each of the participants, there was an explicit link between their personal identity as a learner and their desire to lead professional learning that helps teachers. Over half linked it to change and growth in classroom practice, and two, Maxwell and Sherry, to specific transformations in society that could be powered through education. Both Maxwell and Sherry described specifically looking for their impact because they felt strongly that change was needed, but all the professional developers described caring about impact because they wanted learning to occur. Without a formal push for evaluating professional learning, many described engaging in formative evaluation because they wanted to learn if what they were doing was effective.

Understanding personal identity provides information about why self-assessment might be a part of facilitation process even when systems do not demand information about the impact of professional learning.

### **Identity as a teacher**

When asked how they assess their own impact as a professional developer, many study participants made comparisons to how they assessed their impacts as teachers. Dean noted that gathering that data as a professional developer is much harder, because you get less daily informal observational information: “It is so different from teaching where you are getting most of that feedback right away, mostly informally.” When considering what to collect for his science leader professional development, he thought back to his teaching practice and declared formative data more useful than summative data:

I would never collect as much summative assessment as I did in the past. I would do it if I were still teaching, because people would be telling me I had to, more so than me needing it. It would be far more what the kids need. I feel like I don’t have it in me right now to figure out what other assessment data could we have been collecting from the science leader crowd in terms of yes, we need this for PD and we need that.

For all teachers, being confident your students are learning is a struggle. However, for the professional developers, the problems are exacerbated because they rarely see the application of the new learning in context. Tracey articulated that the daily observation of her classroom made it easier to gauge her impact in the classroom, “Whereas in a classroom, you are better able to gauge that. You are with those students, those learners, and you are actually seeing their products.” She noted that often teachers go away after they learn something in professional learning and the facilitator does not watch them practice or get to see the end result, making it

difficult to know if the new practice is working. Tracey described knowing her impact as a professional developer as “one of my struggles,” and noted it was not as big an issue when she was an experienced teacher: “When I was in the classroom, I always pretty much knew how well I was doing. I helped that kid learn, or I did not help that kid.”

Sherry agreed the observational data is critical in terms of what to do next, but she described how coaching in a classroom lets her observe students as a teacher and the teacher as a coach, giving her even more information than she had as a teacher. Sherry explicitly linked co-teaching and observing with providing useful observational data about her impact:

I did a little bit of modeling and we did a co-teaching session where we took a concept and we assumed prior knowledge in our student which they lacked.

Interviewer: Right. I have made that mistake many times.

Sherry: We were really surprised by that. It had to do with bias in non-fiction. We were just very surprised at their limited awareness of bias and vocabulary. What we found was that to move on with our lesson plan, we had to take that step back right then and elicit whatever fragments we could from the students. We found we could actually spin the web using the tool, by using it anywhere. I find that there is that point, too, where their creativity is engaged and they start extending the use of the tool into their own practice in ways that then become instructive to me. Those are signs of a really high quality PD experience, for me.

Sherry linked the evidence of student use of the critical thinking to her feeling of success as a professional developer. Observation also convinced her that teachers were using the tool well and Sherry described how she also learned things she could take back to her own teaching:



What was really interesting to see in her work is that she started transforming the tool beyond what I had ever done myself or could have anticipated and using it in ways that then became instructive to me. She was very subtle with their use, whereas I tend to be more explicit. Using the same tool I am speaking about, when I was working with another teacher, we discovered together that one of the tools wasn't simply just a way for students to record their knowledge. It could actually be used as a knowledge generating tool, and they could plug into it anywhere, wherever they had something to work with.

Sherry was the sole coach in the study due to the number of coaches returned to the classroom, as a result it is unclear how much the opportunity to engage in observation and application in your own classroom might influence how professional developers to gauge their own impact. Amy, Pat, and Gwyneth, who observed during learning walks or informal classroom visits, all referenced triangulating observations to help them see how learning was applied in the classroom. However, only Gwyneth characterized the learning walks as providing her and other teachers in her school with information about their own impact in pursuing professional learning goals.

Referencing thinking about personal teaching when planning or assessing professional learning was common to all study participants, except Pat, who referred to it much less commonly. Pat also stated he had been out of the classroom for an extended time and thought primarily as an administrator and superintendent, which may account for the difference. References to experiences as a teacher often came up when the professional developers were describing why they made a specific decision. They used teaching to describe the information they had, the assumptions they were making, and a general frame of reference.

## **Relationship with professional learning**

Beyond personal identity and comparisons to teaching, past experiences in professional learning also influenced the professional developers' beliefs about the value and impact of professional learning, leading them to make specific decisions in their assessment processes. Julie made explicit connections between her experiences and relationship with professional learning and her practices as a professional developer:

I think back to when I was a classroom teacher and I took part in different professional learning opportunities. By talking about some of the content presented in a particular session, it would help me clarify my thinking process, or make the foundation more solid, in other words solidify my beliefs. Then I was able to explain to other people, "This is what I am doing and this is why I'm doing it in such and such a way." I try to remember that it was useful to me as a teacher.

Michelle was less universally positive and mentioned being motivated by both positive and negative professional learning experiences. She noted that she was the type of teacher, like Julie, who regularly volunteered to participate in professional learning. She noted "it was just always something that I did and was interested in" even as she was "simultaneously unhappy" with the professional learning she experienced.

Sherry described her best professional learning as life-changing and noted it permanently changed her classroom:

The most transformative professional learning I ever had took place over a four day period 16 years ago. I went to California, and I went to four days of extensive workshops in three tools with Dr. Richard Paul. I nearly dropped dead of exhaustion at the end of each of

those days. Those tools are not hard; my students use them every day. It is my curriculum – that is all we do.

Tracey described being volunteered to work with a teacher mentor working on a research project. Her principal figured Tracey would try anything and Tracey described working on the project with the mentor as shaping her thinking through questions and commitment to growth: “I just think she is a master teacher and she is constantly looking at her practice. Constantly, at how can I be more effective? I don’t think we can ever get to a place in education where we say ‘Eh, that is it. We got it.’” Tracey described the process of referring back to her best professional learning as a teacher to help her commit to an ongoing, continuous growth as an educator. She also noted the professional development she experienced changed and improved her teaching in a way that influenced the type of professional learning she tried to offer: “PD changed for me then, too. I wasn’t just hearing ‘Oh, this is a cool way to teach geometry,’ which I enjoyed, it was ‘What do you think about the way kids can learn?’” Tracey described wrestling with the big questions of education in embedded professional learning environments, and credits that learning with leading her to her work as a professional developer: “I think one of the reasons why I got this job is because we were doing classroom embedded professional development in the school I was in, and it was phenomenal. But it takes a lot of commitment, and work, and relationship – it takes a lot to support that happening.”

Amy described similar job embedded learning as very influential on her as a teacher and professional developer:

Another thing that has been pretty valuable for me is that I spent a year teaching in the U.K. My seventh year of teaching, sixth maybe. From an assessment perspective, the way they do things over there. The way it works there is that students who are in year eleven

graduate in May. School doesn't end until July. Everything is offered in departments, and the whole English department, we basically ended up having the hour where you would have normally taught the year eleven students for that whole department to come together for three months at the end of the school year. It was awesome. We would get all the data – it was all about data there, too. We would get all the data from the year elevens, and it would be planning for next year based on that.

Like Michelle and Gwyneth, she also expressed frustration with some parts of professional learning: “There was huge accountability there. I had lots of frustrations about that system. I spent all my time basically planning my excuses for why a student wouldn't reach the target. If I could have spent the time helping the kid, the kid probably could have made it.” Amy notes that her experiences taught her how important it was to consider results and engage in dialogue, but it also affected her beliefs about what professional learning should do to support teachers.

The professional developers who saw their own experiences as professional learners as significant used those experiences to plan professional learning they lead. They tried to develop processes that reproduced the key elements they value or avoid elements they thought were problematic.

### **Identity as a professional developer or leader**

For some of the professional developers in the study, a sense of identity as an effective professional developer or leader relates to their decision to self-assess and helps to frame the focus of those assessments. As a precursor personal condition, some participants identified strongly with their roles and others did not, but it did not seem strongly related with specific parts of an assessment process.

Juanita noted she had no specific interest in being a professional developer, she got into the role because it was an opportunity to learn. “I don’t think I ever said I want to do this, to give PD. I did it because I wanted to learn and grow, and if that is part of it, so be it. I will do a good job because that is just me. I will work until it is good.” She noted she assessed her impact because she wants to be an effective facilitator.

Michelle noted that she is a professional developer mostly because she enjoys it: “And I just also really like going to PD so it turns out I also really like being part of the back end of it.” However, unlike Juanita, she did not describe a motivation to be a strong facilitator as propelling self-assessment. More commonly, she described a desire to help teachers with instructional issues and diagnose points of difficulty.

Dean described liking influencing, setting a vision and influencing others when he referred to a discussion he had with a colleague about leading:

He used to be an X. Math guy, totally anal. He would have fit at the Ministry in the days when we had the teacher leader binders and it was the pages with all the handouts. I just don’t quite miss those days. When we started this Science 10 and we had the teacher leader team, I came into the room and said, “So we need to develop PD to support this. What does that look like?” I just pitched a few ideas. I knew where I wanted to be at the end of the day. X and X have been very perceptive. A few others have picked this up. They describe me as someone who already has the vision for where we are going to be at the end of the day. But X says, “Every time you have a meeting like that, everyone is convinced it is their idea, and they are 100% on board.” Part of it is because I have a big, solid vision for where Science curriculum overall is going. Nobody is going to pick a lot of holes in it so I am very calm and relaxed about it. I know when you bring a group of

teachers together, or the science leaders, if you don't give them the first 20 to 25 to vent, your day is not going anywhere.

Dean described the power of planning PD with teacher leaders and reading the room well. He determined his effectiveness as a leader by comments about his influence rather than seeking a formal evaluation process.

Gwyneth described a comparatively wide number of influences that shaped her identity as a leader and connected them to specific learning. She mentioned three external presenters, two of whom she referenced the division bringing in. In each case, she described specific lessons that shaped her thinking and identity and the learning giving specific guidance about how to conceptualize professional learning: "It was informative because it helped me focus on the explicit steps that I needed to make for the quality of instruction to improve. Not just the quality of instruction, but the climate of learning. Those are different." Gwyneth said her learning about leadership and professional learning taught her "how to construct it [professional learning] and why." She described the learning as "highly influential," because there "were a number of people I read or got to see in person and they stuck with me." For each person Gwyneth mentioned, she explained a specific lesson, for example, with Karen Hume she learned how to "to sequence things and those tiny little nuances and details I would labor on to be sure they were really right." As a result, Gwyneth's process and the questions she asked herself changed:

Was this the right sequence? If I undertook this thing first, would it result in a better product, or better outcome at the end? Karen (Hume) would do that effortlessly....She was very rigorous.

Allan, who I mentioned, was very rigorous as well. Those people, in terms of the learning front, were very important. There were some committees that I belonged to that I think

contributed to my sense of how those discrete plans for professional development fit in the larger scheme of leadership. I was fortunate to be on committees that looked at the *Leadership for Learning Framework*. I think that was very foundational. Being able to look at how all of those moving pieces fit together. And I can think back to leadership meetings and to seminars where we really pressed in on that, and that really was a moment for me. Critical friends. I think that if I think back to, you know, learners along the way who are with you, they have a very similar need and those kind of people pressed me forward.

Gwyneth confirmed she received substantial support around strategic planning, around the “thinking of the why” and she saw examples that helped her consider the sequencing of professional learning. However, her leadership identity was not shaped with formal influences about the assessment or evaluation of professional learning, although she valued good questions she was asked:

Interviewer: I’m not hearing you say that somebody taught you how to evaluate your impact, but that you learned a lot of things related to it that helped you to figure that out. Do I have that right?

Gwyneth: Yeah, and I guess what I sought to do on my own, although I did have colleagues along the way and other folks asking good questions--in particular a really influential superintendent when I was first appointed made a bit of a difference. In terms of feedback and assessing were always important to me in the classroom so I hoped that I always applied those same practices—

Interviewer: So, that was a personal passion and a lens—

Gwyneth: A personal passion and a lens to always be asking, you know, how do we know that we're having the impact that we're having?

Interviewer: So that was a key question for you?

Gwyneth: Key question. How do we know? And so, that was always a key question for conversations, be they about my own leadership practice—how do I know? Just as I would have a visual in my office on the learning, in early years of my practice—less now, honestly—but in my early years I would have *the Leadership for Learning Framework* and other frameworks and I would explicitly think about actions that I undertook and how they aligned. And I did it differently. One year I had the whole circle and I put little things on it. One year I cut up the leadership for learning framework and I tried to track doing each part.

As a part of the theoretical sample, Gwyneth was asked directly about the role of her identity as a professional developer and leader and provided a more elaborate version of what other participants confirmed; professional development process were workshopped, mentored, and supported in several organizations (including school divisions and the Saskatchewan Professional Development Unit) to support excellence in professional learning. When participants referenced these supports, they also referenced thinking about professional learning in a more nuanced way that was more aligned with current theory on professional learning. Very few participants referenced any similar thinking about program evaluation or self-assessment, and they referred to their identities as leaders or professional developers when asked about their thinking regarding their evaluation processes. Self-assessment and evaluation were mostly discovered, not taught, and were not significantly referenced contributors to professional identity.



Personal and systemic causal conditions are important in understanding the process of self-assessment, because they are the modifiers described by the professional developers. The conditionality they used in their decision-making descriptions indicates it is contextual and contingent on many factors in each situation in a way that is antithetical to descriptions of formal evaluation processes, both because it is emergent and it omits structural elements critical in program evaluation. Facilitation and self-assessment were intimately linked to the descriptions, indicating there may be much to learn from thinking of assessment of professional learning as formative and internal. However, the key details in understanding how professional developers engage in self-assessment of their own impact lies in connecting the pre-cursor conditions with the core elements of the self-assessment process.

### **Core phenomenon**

A core phenomenon, or core category, describes the parts of the process being examined in a grounded theory study (Charmaz, 2014; Creswell, 1998; Strauss & Corbin, 1990; 1994). Where possible, the stages are described with in vivo language so the titles of core categories and the resultant theory reflect the thinking process faithfully. In this case, three of the stages of the process are the words of the participants (wondering, finding out, and action), but the fourth, sensemaking and attribution, use words from research largely because it was the stage in the process that the participants were much less explicit about.

The process of self-assessment of professional learning had four key stages as the professional developers described them. Since the decision to evaluate anything was described as largely optional, a decision to assess had to occur. Professional developers described this stage as *wondering*. Once a specific question or wonder was undertaken, data were gathered to *find out* the answer. The data were considered to determine what it might mean, in a

*sensemaking* process, and finally, the professional developer decided what to do, if anything, at the *action* stage. Each of these stages could be extended or very brief, and not all stages occurred, as a wonder could be abandoned at any time. Given the number of barriers to self-assessment and its optional nature, the process occurred when professional developers deemed it as necessary and the pre-cursor conditions favored an evaluative process.

## **Wonders**

Tracey described the questions that power professional learning and evaluation of professional learning as *wonders*. Like the majority of other participants, she sometimes overtly characterized professional learning process as inquiry process, driven by questions from the participants and facilitators. When describing a data sharing and analysis process she used called a concensogram, which was focused on the division's main instructional focus, Tracey noticed teacher leaders rated the categories differently than administrators. She described wondering why the different groups rated differently and asking the teacher leaders to speculate about the difference when they rated the Concensograms after the administrators had done so. Tracey hypothesized about causes and potential related factors and considered the potential significance of the difference.

Like a formal research question, a wonder, as Tracey conceived it, is a direction for inquiry, data collection, and theory generation. However, unlike a research question, the process of wondering might be a passing thought. Participants discussed the process of wondering as a regular occurrence when planning, leading, or debriefing professional learning, but it did not always lead to specific action or data collection. Like Tracey, many participants characterized the wondering process as pedagogical. They mentioned inquiry stances or dispositions as

elements of good learning and teaching, and questioning or wondering as a subset of the wondering behavior or as an instructional process.

Maxwell and Michelle both articulated wondering as an activity within professional learning processes. Michelle described a whole professional learning cycle as teacher-driven inquiry, and Maxwell described provoking wondering in teachers to sustain learning beyond a session. He described having his audience predict the difference in outcomes for Indigenous and non-Indigenous students to motivate them to think about “the why” and their own beliefs:

I like to have them predict things, to get them thinking about the why. I like them to experience more in an inquiry-based [way] – I will give them something, have them ask questions about it and then perhaps support them with some of the answers to some of the questions they have, but I will never answer all of their questions.

I want them to leave and find other ways to have their questions answered. I don’t want to be going, yeah, yeah, yeah, yeah. I want that inquiry to maintain itself.

Participants often connected facilitation and wondering. Jennifer and Julie described a facilitation process designed to get teachers to wonder about their reading data, even in the webinar process which they characterized as difficult to be responsive with. Pat also described having teachers ask themselves inquiry questions about reading data:

If we are looking at reading scores, working through *SaskReads* [Saskatchewan Reads document] and the various instructional strategies and experimenting to see what kind of impact we can have, we would want them to take a look at the data they had, not only from the students at the end of the year last year, but from the beginning of this year and the year we are in as well. Is there summer slide? What is going on? We are talking, as teacher

groups, about knowing each individual student by name and by need. We can use the data we have to inform our instruction.

For many participants in the study, wondering was a part of a process of professional learning that was connected with facilitation process where participants consider their own data and speculate about causes then make plans for action. Maxwell summarized it when he connected wondering with facilitation: “I always leave these things with that big wonder, how was that for the participants? You don’t get a chance to figure that out unless you have a mechanism to get that feedback.” For Maxwell, wondering sparked self-assessment.

Sherry, who both planned professional learning and engaged in coaching, was typical in her articulation of an inquiry stance for planning professional learning: “I always have unconsciously, but I have now consciously in the inquiry process, started differentiating based on those ways of knowing, using different ones at different stages in the process.” When she discussed an individual coaching process, she described a compelling need for the teacher, then converted it to a why or wonder:

I think I try and take a bit of a divining rod approach at the beginning, when I am talking to somebody. I think it is essential that I discover their point of crisis, or passion, or need.... I try to understand it deeply in their own terms. Then I need to be able to suggest something that they recognize, quite immediately, connects. They need that confidence that I get it, and that I am willing to play it in a way that will make sense to them.

Sherry described herself as motivated by “deep inquiry questions” and described coaching as a process that moves teachers and students towards being stronger critical thinkers. She connects this directly with a value for student learning:

They see the teacher, in relation to knowledge, as a quester not an owner. They see a different relationship between teacher and student, between student and knowledge, between student and student, and between student and community. I am very interested in strategies that help to ensure more people in the room are operating at full power.

Gwyneth described wonders slightly differently than the other participants, in ways closer to my personal theorizing about what sparks my self-assessment as a professional developer. Her description is not of a general point of curiosity about something in professional learning, but more specifically about something that causes her cognitive dissonance:

Gwyneth: I think you could see a (pause) learning fissure. I call them learning fissures - little tiny cracks or philosophical bumps that bubble to surface. Some you have some great pedagogy going on in domain 3 [of the Danielson Framework], like culturally responsive [instruction]. If you looked at my building, I could tell you that the level of culturally responsive practices were laudable, and we were making great improvements. We were implementing, you know, interactive strategies and different ways of thinking. But our classroom environment didn't always support that work and our classroom environment didn't always reflect that work. So simple things –

Interviewer: So a learning fissure is a dissonance?

Gwyneth: A disconnect. What you see as leader is that it is not as impactful as it could be, right? And I always work from a place of positives. There were great things going on, but what if!? What if we kicked it up to this level?! So simple things like physical structure, the layout of the classrooms, and the way we were behaving with our children. They are all one. I noticed that if conferencing with our kids was really important; I was wondering, how were we getting to that physically?

For me, as for Gwyneth, the desire to engage in any form of assessment or evaluation is sparked by desire for improvement and disconnects between that desire and what I learn is happening. Like all participants, questions were at the heart of attempts to understand their own impact. Even participants with relatively few opportunities to work intensively with teachers, like Dean, described wondering what happened with the learning or why certain things happened the way they did. Wondering was a key element of the praxis of evaluation for the professional developers.

### **Extending wonders**

Hypothesizing or stating interpretations of data was a starting point for formal and informal versions of evaluating professional learning that participants described. However, hypothesizing was not an automatic trigger for data collection about the impact of professional learning. Only a fraction of all wonders triggered a process of deeply considering. Sometimes a wonder might be stated, then followed by a belief about likely causal connections. For example, Tracey speculated about the impact of the coaching structures the learning facilitators use. She wondered what they were doing when they meet with teachers and asked, “What are structures that are in place? Tell me about where your coaching plays a role in this.” Tracey looked at responses, and if she had easy explanations, she did not extend the evaluation process. Only wonders that extended themselves went beyond initial data collection. The data collection participants described as helpful was consistently connected to wondering, then extended beyond it. Often, that wondering occurred many times when considering a chunk of data.

For every data piece, Tracey described a personal process of speculating about potential reasons for an occurrence. With the Concensogram, she asked herself questions including: “Were the learning facilitators working with teachers who were not as far along? Was the

difference caused by personality? Did administrators think the results were pretty good, considering?” She also described talking with the other facilitators and considered the effectiveness of the regular conversations between the two groups, given it had been a focus for the year. Finally, Tracey described structures where wonders help to drive the work: “All five learning superintendents and two administrators sit down on the learning facilitator planning group. We all bring our wonderings and reflection, and plan based on that. We have a year plan and then use what we hear and wonder about to plan responsively.” Tracey makes specific connections between wondering and ongoing formative assessment, where professional learning instructional plans are adapted based on interpretation of data, and hypothesis drives choice.

Similarly, Juanita described using survey data to plan for following sessions and articulated her wonders about the data relative to her main questions about how to deliver the most effective professional learning:

The survey data was important for the core team to plan. It was not about lunch or how much you liked it. We asked questions we intended to use or care about – that is one of my soap boxes. People need to see something come from data they share or they won’t be honest in the future. We asked about what people attended and what they wanted it to look like. I want to know how to build next and which sessions are becoming a community of practice.

Often, the professional developers described considering data briefly then attributing it to a specific reason and moved forward to action. Sometimes they were less confident about what the data likely meant and described engaging in a process to find out more. Again, this process might be extensive or it might be a quick conversation to assert a plausible explanation or knowledge claim before moving on. Typically, participants framed possible explanations as

statements in response to a wonder. For example, Juanita noted that “the survey taught me we need to teach teachers how to give feedback. I deleted a few because they focused on the ability of the facilitator and I did not want them to hear complaints about a fellow staff member.” Her observation and action were nearly simultaneous.

Sometimes, generating a reason, some data, and questions occurred rapidly in a different sequence. For example, Pat said that part way through his year at his new school, he determined a specific need from the information, but he was still speculating about the context the need occurred in: “We need a common sense of identity and a sense of shared beliefs. What we had heard coming in was that the school was a school with a lot of behavior issues. Was it because students were not engaged?” He started with a conclusion about what the school needed, moved to a conversation piece of data he collected, then to a wonder he had about the reasons. Similarly, in response to his wonder about a need for shared beliefs, he described a piece of data he saw as evidence of lack of a common understanding of shared beliefs, and then immediately described a planned response: “They said they had shared beliefs but couldn’t find them. We posted them in a main area where everyone could find it.”

For the study participants, wondering was not the sole initiation of formal data collection. Participants named types of data like provincially mandated reading assessments or organizationally required agenda assessments. But they typically described these pieces as giving some of the information they needed for decision-making, not all of the necessary information, and they were often skeptical of the value of the data. While formal data collection was related to wonders for some participants, they did not directly connect the large-scale student data to professional learning



During the study, no participants described completing a formal evaluation from start to finish, even when they included some of the elements of evaluating professional learning, like Terry or Michelle's references to logic models or Guskey's (2000) levels of evaluation. The process of wondering provided the main impetus to collect information about personal impact, although two administrators also cited reporting to stakeholders and two people who contracted out professional learning services described gathering data about the impact of professional learning as a part of their typical process as a professional learning unit. Wondering sparked self-assessment, but the participants did not report it led to formal evaluation.

### **Something to wonder with**

One interesting element about the participant description of wondering was that it had to be sparked by something. Three common elements seemed to commonly spark wonders:

1. Personal experiences that caused a long-term quest to understand something, such as a professional development experience that extended into a career-long wonder about the best way to support teacher change;
2. Collaborative wonders that came from colleagues describing perspectives or interpretations that differed from the professional developers and provoked questions; and
3. Facilitation processes that yielded formative data that sparked questions about the quality or nature of participants' learning.

Each element of wonder sparking is significant, because it helps to describe why a professional developer might engage in some form of self-assessment. Some professional developers in the study reported all three examples, and some reported fewer. Using facilitation

practices that yielded data seemed to be a particular key. Participants without those processes tended to end wonders with incidental conversational data, rather than engage in a full process of self-assessment.

### **Process of data collection: “finding out”**

The process of finding out had two main triggers as described by the participants. The first was the use of planned assessment tools during or after professional learning, and the other was gathering predominantly conversational or observational data in response to a wonder. The gathering process was typically informal and predominantly formative, while the planned assessments typically served multiple purposes.

#### **Planned assessment**

Participants shared the types of data they used to find out how teachers and leaders were affected by professional learning. Survey, exit slip, self-reporting of use, and data generated from facilitation processes were common sources of the data that participants described. Much of the data was focused on participant satisfaction, which was considered good evidence because teachers, as professionals, know when they are learning. Some data focused on participant understanding of new concepts or participant self-reports of use or barriers, largely because those were main goals articulated by the professional developers. These forms of data were self-reported, but participants typically paired them with other forms of data to triangulate.

Formative assessment data from facilitation was commonly referenced as a source of information about participant understanding and barriers to learning. Data sharing processes, like Juanita’s sharing of the survey results with distributive leaders, Dean’s sharing of participants’ responses with other participants, or using the Concensogram in Tracey’s facilitation were examples of times where professional developers gathered additional

information about participant thinking and skills to compare to other data collected. These types of data collection were highly correlated with self-assessment processes and yielded more wonders, specific attribution, and action.

Planned observational data was also a common way to triangulate data used to “find out”. Some professional developers like Michelle, Dean, Tracey, Amy, Maxwell, Julie and Jennifer described watching participants for behavioral clues, but others also gathered formal observations, like Gwyneth’s learning walk process, or Sherry’s co-teaching. Michelle described observation as the most convincing evidence she had that she had supported a teacher in instructional growth, noting that she had to “see it” to be confident.

Planned data collection was the most likely to be connected to formal data analysis process, with groups or with the other professional developers. Amy described the value of the group debrief in sharing what each facilitator had observed, and all three administrators in the study described the value of teachers seeing data sharing walls, posted goals, or newer practices such as shared social media. They also each described a progression of professional learning founded on data from classroom observation correlated with data processes that helped them to see what their teachers were understanding about the content of the professional learning processes.

### **Gathering information about a wonder**

Wonders also sparked a separate form of data collection that was more instinctive. Although the process might be planned, wonders came up, then data was rapidly considered and interpreted. Sometimes participants’ experiences led them to make quick conclusions about the likely response to a wonder, and the participant moved on, but at other times, there was no

immediate answer. If there was no immediate answer, participants described trying to collect data to determine what was happening and why. Amy described the process she used when facilitating a process with a group of superintendents and administrators. She was concerned it “might turn negative” since it was about barriers. In the planning of the session, the three-person facilitation team had wondered if this section was likely to be an issue and how they would respond if they thought there were problems. As a result, Amy monitored the learners closely by observing them and listening to their conversations:

I lead the one part on the barriers. There were three superintendents and the rest were administrators, five or six of whom were new. We were concerned it would move to the negative. They knew each other well and did not need to trade horror stories. I said, “I know you have all been leading things when the tone sends to the negative.” Body language and heads nodding told me they understood purpose to generate solutions. So many conversations went to Ministry and funding, but participants usually redirected their own groups. I was intentional in setting up the conversation with purpose, because everyone has had that experience of things going off the rails.

Amy believed that conversations staying focused was possible, and that she was responsible to ensure it happened. Based on her experiences, she believed that many participants would likely get back on task without her intervention, but some groups might need her support. Amy described observing body language and listening to conversations for cues, in a practical process to find that things might be “going off the rails.” She also looked at visuals on the paper where her professional learners were writing, a process deliberately set to both help the learners and provided formative data for the facilitators. Amy noted that one group wrote nothing and said she was “intentional in getting in that group. That was the group getting on the Ministry

thing. As soon as I came over, it was almost like proximity, one member said ‘we can’t do anything about this.’ Then they were in a good conversation, which I could hear. They redirected.”

The participants frequently described feeling responsible for the quality of learning within a session and believed that it would not improve over time without their actions. They were not consistent in describing problems as always controllable. However, they used conversations and observation to give them the information a teacher in a classroom uses to make classroom management decisions or to decide if something needs to be taught. The process of attribution, where they quickly determined if they might need to act was often invisible to the participants, who described going from a wonder to action. When asked more specifically, they described evidence they collected and their specific sense of control or responsibility for action.

Once they were asked to clarify and expand on their thinking, participants added many factors that influenced a decision to gather information, make sense of it, or act on it. Amy described how she often gathers information so she can respond:

As a teacher or leader, I do that all the time. When I am facilitating in this building or the administrators group, I really struggle with that because the opportunity to flex the plan is not there. I have enjoyed listening to X think aloud about that with the facilitator community. In our building we just don’t have the time to make substantial changes to what the activity is. With administrators, we have 100 people in the room and engagement is difficult, and you can’t please everybody. We are facing a big challenge from a movement from policy and procedure, and we have moved to learning. I am not sure

people were kept in the loop, and they are not in the same mindset of volunteers who really want to learn.

Amy noted sometimes her role plays a factor in her power: “It is easier to gather data and be a responsive PD leader working ‘side by side’ than having a role where you are perceived as a mole or controlling. The decision to include administrators as planners was a decision to make us leaders from the middle.” Amy jumped off from her thinking about reduced responsibility and control when leading from the middle to a whole new wonder: “It makes me think about succession planning. How different will that look five to seven years from now?”

Juanita described a similarly complex set of data she is considering when trying to find out what the teacher leaders she was working with needed. After videotaping herself, then observing a section of the recording, she noted a series of quick things she does, starting with creating an opening for a specific participant by saying, “You are thinking.” Juanita said she made the statement “when I could see he had something important to say, but did not know how to share it. He is the most experienced leader and I depend on him a lot. He is respectful, but I want him to share.” When pressed, she said she considers the reactions of various participants specifically because of their personalities, relative influence in the group, and circumstances in their schools. Juanita described wanting to know his reaction because she perceived it as likely to be typical of a group of teachers and apt to influence others. When asked again how she knew, she stated she knew that from experience, then commented, “I spend a lot of time monitoring facial expressions and body language. Body language is a good ‘tell’ about if people understand or they feel marginalized. I look for feeling and try to find the person I can watch to see if I am making sense.” Juanita used body language as a cue to compare to what people are

saying out loud about their understanding or barriers, and said, “I use body language to know where people are at and where I am at.”

Professional developers in the study described similar incidental information gathering as finding out. If Juanita wondered what one person was thinking and did a quick check in the space of a few minutes, this same pattern was repeated by each of the participants throughout the space of a sequence of professional learning. Although all participants described formal data collection about professional learning they lead, the wonders and quick checks were described often, and used as equally valid data about what teachers or leaders were thinking, feeling, or understanding. Julie described those incidental conversations as a form of member checking, and Maxwell and Jennifer described processes where a call from a teacher or leader, in the absence of other data, seemed to be an indication of success.

### **Attributing and sense-making through the facilitation lens**

Trying to determine what is happening and what you should do in response, if anything, can be a lengthy or near instantaneous process as described by the professional developers. While the main purpose of a formal evaluation process is to summarize findings relative to a specific given purpose, the intent of the professional developers in the study was often to decide what do immediately or in the next part of the process of facilitating professional learning. Attribution theory indicates a person might decide what is happening, then ask questions like: Can this be controlled? Do I have control over it? Will it continue over time if I do nothing? (Seifert, 2004; Weiner, 2010). In many cases, professional developers described deciding what was happening, then choosing a specific action. They rarely overtly described an attribution thinking process, which the exception of determining things they could not control.

Like other participants, Maxwell looked through data and interpreted it, determining what it likely meant and what the significance may be. He was looking at data from a recent presentation about culturally competent practice and noted that feedback related to Indigenous issues goes “to heart first, then head.” Maxwell noted he needs to disconnect the feedback from emotional responses, because “it hurts when I see the feedback that the mindset is not further than when I started.” Maxwell described his main goal as seeing belief systems and mindsets change, so he looked for statements that might indicate that change. During the analysis process, he looked at each question on the feedback form separately, considering what each respondent wrote. As he read each comment, for example, “hearing the voices of all was very enlightening and powerful,” he made specific observations like “there was no real personal voice.” He related the statements to individual events in the professional development like the talking circle, like that no one passed. He noticed specific references to elements like the “exercise” and paid particular attention to combinations of verbs, like “understand and feel” which he said was evidence that a person was specifically affected. Maxwell’s process of making sense of data has five steps as he described them:

1. Looking for words and thinking about their connotation
2. Making implications for emotional and intellectual impact
3. Placing response on “a continuum based on the goals of the professional learning”
4. Dismissing some responses as too vague to be turned into evidence of anything in particular
5. “Look to see what themes or trends I can pull out to find the activity is valid, or it needs to be changed.”



Maxwell articulated that valid equals “an experience that impacts them so they are willing to change things in their classroom. Something that goes the extra step.” He accepted that as evidence of personal impact because “people need to be more than just enjoying PD. I want a shift or movement in mindset, a challenge. A small shift in mindset is not enough. A change in practice is important. A large shift in mindset does that.”

Dean used a similar process of looking at the evidence he collected of leaders’ descriptions of teaching about the Broad Areas of Learning and Cross Curricular Competencies from the renewed curriculum. Dean wondered if school divisions were providing enough support for teachers to fully implement curriculum in a way that supported students in reaching the Broad Areas of Learning and Cross Curricular Competencies. Like Maxwell, he looked for combinations of words, but unlike Maxwell, he used a process of constant comparison to his experiences working in school divisions. While Maxwell’s example was of a conference in a different province, so he could not triangulate data with other things he knew about his audience members, Dean’s was close to home in his own province. The participants indicated that the more information they had about the people giving responses, the more likely they were to triangulate the data with other data points. Like Dean, many participants said teacher feedback at the end of a session is an insufficient amount of information to use to determine impact, although like Jennifer, they described that form of data collection as the “most common.”

Dean’s concern that school divisions may not be providing sufficient support so teachers were fully implementing curriculum in a way that supported the Broad Areas of Learning and Cross Curricular Competencies seemed to him to be consistent with the data he reviewed. Although the leaders said they felt it had been a focus, their descriptions of the activities were missing some of key ideas and phrases Dean said would convince him there was sufficient focus

and understanding. When he saw “very general statements without specifics” he asked, “Are the science leaders integrating these?” When he read descriptions of PD, he compared them to what he had observed, noting “I didn’t see it happening when I was out doing PD in divisions.” When he saw a general phrase about covering the idea he noted that, “It tells me very little is happening.” He concluded that the answers are mostly evidence that “they knew it existed and where it was in the curriculum document.” Dean’s process included looking at the statements, relating it to things he has seen in divisions, and then drawing conclusions. Sometimes he speculated about causes, for example when he stated that the leaders do not have expertise in science, and even when they do, it is typically in one senior science. As he kept reading through the data, he noted he was “more confident in outcomes and indicators because the responses were similar and addressed concerns he often hears.”

As he neared the end of his data, Dean’s data analysis process moved from looking for individual statements and analyzing those to combining groups of statements to rendering judgement. At the end of the process, Dean scanned the list, and he concluded that “there is a lot of work going on around assessment, especially regarding as, of, and for.” He noted “lots of references to inquiry”, but he “questions the value given the answers on the front matter questions.” Given the data, Dean stated he is not sure science leaders are using the resources. He speculated this was because “some are not aware, some look regularly at documents, and some do not.” Dean concluded that “content has eaten up the discussion and thinking time.”

Even in the absence of data, professional developers drew conclusions about the success of professional learning. Jennifer noted that when no feedback was collected and none was offered, she assumed that was evidence of success: “We didn’t get any feedback, so that is good feedback. If they were unhappy, they let you know. If they didn’t understand, they would ask.”

Similarly, Julie noted that a feedback process in December had been positive, and there was not feedback of any kind for the new professional learning materials, so all must be well because “no one contacted us.” However, she noted that since the professional learning was online, she had no observations to give her baseline data: “Unless we are face to face with a group of people, we have no idea of the impact.”

Like Amy, Maxwell, and Dean, Tracey summarized a pattern of thinking about data that reflected looking for clues about the goals of the professional learning. She demonstrated a process that extended initial interpretation to thinking deeply about what the data means in context:

1. Connecting the data “to everything else I know.”
2. Considering why it is this way.
3. Wondering about “what would alter or change it.”
4. Wondering about frequency and circumstances “how common something is – what structures are in place, what schools.”
5. Relating the data to her to personal beliefs “about compelling why – reflection and growth.”

Moments after considering the data she collected, Tracey moved to considering her personal agency, to “thinking about what I can do.” She stated she “feels ownership for it as learning facilitator work” or in “school visits” and when working with administrators. Tracey also noted she is always thinking about how to remove personal and systemic barriers for her participants: “Reflective practice is the main goal of all conversation – one is removing barriers, and other is overcoming them.” Many of Tracey’s thinking processes referred specifically to

thinking about what she controlled and did not, if something might change if she did nothing, and what she could do in response. Both Juanita and Michelle described similar ways of interrogating problems to see what to do next to increase the impact of the facilitation processes.

### **Likely interpretation**

Some study participants did not describe steps in their process of thinking about data. Instead, they made intuitive links between data and what Amy called “likely” interpretations. Pat, Sherry, Gwyneth, and Amy all used likely interpretation, sometimes considering two options before settling on one. Even when asked probing questions, they did not break thinking down more explicitly. However, they were no less likely to gather or triangulate, and their interpretations were similarly filtered through various pre-conditions they articulated. When asked, to you, “What is the best evidence that the professional learning is working?” Pat’s reply was a typical blend of adding up different data sets and content to reach a conclusion, including the observation that data and professional learning can be separated:

We don’t always connect the data to the professional learning. Our staff needs to understand data better for that to work. Observation in the classroom, or formal observation in supervision cycle tells you what is happening. We have look-fors identified and you can see who is using it and who isn’t yet. We encourage and we have professional learning. Our professional learning does not always allow the time – that is on us as administrators. In our next PD session we are taking this data to teachers and using it.

Participants predominantly used data about what their professional development participants were understanding, believing, or doing as a measure of their own effectiveness. They paired data about understanding with data about teachers’ impressions, and with data about barriers they could see, putting particular emphasis on addressing barriers. Juanita, Tracey, Pat,

Gwyneth, Michelle, Sherry, and Amy all described multiple attempts to find information about barriers and articulated versions of the sentiment that if barriers were not addressed, the understood professional learning would not transfer into action. The other study participants articulated similar ideas about barriers, but were more focused on several specific barriers they thought were likely to occur, like teacher motivation.

Like the processes of wondering and finding out, most of the sensemaking and attribution process was focused on the act of facilitating. While the professional developers were interested in knowing teachers tried the new learning and it affected students, they articulated barriers to gathering and using that data and to connecting the data they had directly to professional learning. Sherry, Gwyneth, and Pat were the sole exceptions, as they all referenced specific student results and observations of teacher practice. Juanita also referenced that key data in some examples, but it was not the focus of the type of data used for the talk aloud process.

### **Action**

The final stage of the typical assessment process detailed by the professional developers is the action stage. The stage typically followed the process of attribution quickly, and has just a few key elements. Once the problem articulated by the data had been attributed to a particular cause, the professional developers decided if action was appropriate, and if it was, what to do. Like teachers, the professional developers decided on and enacted relatively complex actions quickly.

Pat identified that his staff needed to better understand their reading results and be able to take instructional actions based on individual student results. He articulated that the problem was complex, and could include not understanding the data, not valuing the data, physical spaces and time to gather data, and not believing that students could learn with right time and supports.

Pat discussed everything from school schedules to the architecture of the school when considering how to set things up so the school-controlled barriers to low performance in reading could be addressed. The description of his plan was couched in causal conditions and posited outcomes, but because he described believing in the high quality of teacher professionalism and his own ability as a leader to address the problem, he detailed steps he was undertaking to address the issue with professional learning:

Not all our teachers were doing their own assessments; they might have the EAL teacher do them. We need teachers to see the value in a formative way to increase student outcomes. I think teachers will be disappointed to see how far we are behind level. Sharing the data will create dissatisfaction, but also create a desire to act, because our teachers want our students to be successful. Then we can co-construct and also give people who are beyond the change a chance to self-direct or differentiate. If they are showing their data, sharing their reading, and excited to attend PD, it tells you that person is moving forward without your help.

Breaking down a paragraph of Pat's thinking reveals details in the complexity of the assessment and response. Pat said he "knew" teachers were not doing their own assessments based on observations and conversations with teachers. He also knew, based on what teachers said, that they perceived checking reading levels as something completed for, or compelled by, others. Pat described connecting what "he knew" to theory he had encountered in a previous division in the 13 Parameters, namely that improved reading started with a belief that all students can learn, and we need to know them by name and need to help them. Pat articulated that his experiences told him the research was trustworthy, and he saw beliefs as the first step.

Next, Pat concluded that his staff would be disappointed with the results. He said his experience that teachers wanted students to succeed and cared about their teaching helped him reach the conclusion. He described specific members of staff who were likely to be the most invested and might support others based on his past observations. Next, Pat planned specific processes of professional learning, including co-constructing actions and differentiating for teachers who already had the skills everyone was learning. In an earlier interview, Pat explained why processes like co-constructing build common understanding and respect for adults' professionalism while placing him as a co-learner. He also stated he had past experiences where the most capable learners were required to listen to what they had already learned and he felt it was a common problem with professional learning in many divisions he had worked with. Pat articulated repetition caused disengagement for potential teacher leaders, so he needed to avoid it. Finally, Pat articulated what he would accept as early behavioral indications that would tell him the professional learning had been effective: "If they are showing their data, sharing their reading and excited to attend PD, it tells you that person is moving forward without your help." Pat's description of the evidence he would accept matched his earlier description of successful autonomous professional learning, something he described as a main focus of all the professional learning he leads. He directly connected evidence he gathered with his actions, often in iterative cycles.

The complexity of Pat's interpretation of causes and planning of responses was typical of professional developers in the study. He described many factors affecting his thinking and changes to his plan, which he implemented as necessary, for example: "Around breaks, pressure on teachers goes up because of behaviors with students. Sometimes we switch to topics that are less stressful." Pat's chosen actions were informed by data he had access to, including data he

had deliberately gathered and incidental data, predominantly from conversations with teachers. Pat also informed his choices through his beliefs about professional learning and his experiences as a leader and teacher.

Like Pat, Juanita's decision-making process was complex and informed by many things beyond self-assessment. Her description of her chosen actions took under one minute and referred to a short video clip of a meeting, but had many motivations. First, she described pausing, or using wait time because: "I need to know they understand, so I will wait after summarizing to see if they understand. I am waiting to see, does this make sense? I am trying to get them beyond teacher thinking to structuring an agenda. I offered an additional description and affirmation for X, who needs it." Juanita wanted the group to understand and decide the agenda for themselves. She also noted she was wrestling with her own personality: "It is hard to give up control and it is sometimes easier to do things myself. I need to be a part of the group and influence it, but I want an effective format." She described the form the group is discussing and the potential barrier it represents if the participants don't understand it the same way. She knew from past experience this was a likely issue, so she responded: "The form is designed to provide clear information, and I am repeating back so there is consistency of message. We don't always communicate well, so I work very hard at saying this is what I have shared with everyone else. It is the only way to guarantee consistent, common information." Next Juanita considers the facial expressions and comments of the meeting participants, then briefly describes a process she wants to use. Like Amy, Juanita selected the process to avoid issues: "I describe a process to avoid venting and ensure conversation is about things that we control. I knew it was important because I did not get to the X group on time and the feedback was negative. I describe it so other AP [Advanced Placement] coordinators will know what to do. I gave the example of the system



level AP question to clarify.” Juanita described her decision-making process as a deliberate response to a distributive leadership model when leaders need to be empowered but also need to be clear about the common message and how to respond to issues:

I am using an explicit model – lots of “I do” initially, now “we do” as we have more skills. I do a lot of trouble shooting and think aloud with my leaders so they will be equipped to deal with issues. If I don’t bring it into the conversations, they don’t know how to deal with things. The work I do is very complex, and dependent.

Juanita had a constant iterative process of bringing in a lot of information from other leaders, addressing questions, summarizing and clarifying that she uses throughout the meeting. She described herself as making “micro adjustments based on the people and context.” Juanita noted “established relationship makes things more open” than they might otherwise be, and she articulated a follow up check “with key leaders to see if things worked after the meeting.”

For many professional developers, the most complex actions occurred when connected with data about student results. When the professional developers considered necessary actions, they went beyond considering what needed to happen in the professional learning and connected their thinking more holistically to the entire system. In addition to changing facilitation, they considered advocating, changing structures, developing allies, and supervision. Most commonly they related these ideas to their self-assessment when they referred to student data. Maxwell described “constantly looking at student data” in exclusion of professional learning in specific, then initiating a wondering to action cycle with a cognitive dissonance (like Gwyneth’s learning fissures): “When I see numbers, I consider how it happened and why it happened. I think about reasons like we retained more students, and then the graduation rate dropped.” Maxwell connected the drop in graduation rates, structures, advocacy, allies, and professional learning,

rather than thinking professional learning alone, and he described his role in acting using many other leadership functions. Michelle noted student data is “tracked by teachers in x division assessments and provincial assessments in 2018” and that the school divisions themselves would need to respond with actions because of the variety of factors involved.

Gwyneth’s responses perhaps best illustrated the process of determining the impact of an initial action and then engaging in a myriad of interdependent actions. She described “different layers of work” including advocating changes to division structures including: learning fairs, developing allies in critical friends in other schools, and processes of looking at data with her data team to advocate for needed resources. Her school broke away from the standard professional learning plan for the division based on Gwyneth’s attribution of staff feedback:

You always have to take your learners from where they are. My adult learners in my particular building were at a place where they were craving more. Their particular desire for higher level thinking was such that it honestly wasn’t being satiated.

So the next likely professional development opportunity for them as a staff was a forum that they had expressed, that given their needs, would not be satisfying. I said, “Okay, if that is the case, then we are really going to pitch something that aligns with our A3 [strategic plan] and our practices around reflection and give us feedback for ourselves that meets these needs.”

Gwyneth’s responses included taking her staff to two other buildings for systematic learning walks, having her data team present their journey, posting the school data publicly, and creating summaries of their learning journey for sharing. The action required permission of senior administrators and professional developers, other administrators willing to be critical friends, and groups of teachers prepared by gradual release for the observation of their peers.

She had to consider a wide variety of actions to make changes in teacher practice and student results more likely.

The most significant changes were in who was “finding out”, interpreting, and choosing actions. While Gwyneth described extensive conversations with her data team and other administrator, teachers were the primary agents:

We started with ourselves, learned alongside other administrators in learning walks and practiced the questions with data teams. Then once we had rehearsed it for ourselves, we were brave enough to ask those questions with colleagues from two other schools....

A bus dropped you off, you went into the school and walked the walks, and then you went back to the bus and to the next schools. In preparation for that work we looked at their A3s [strategic plans], because we wanted to understand, what were we looking for? At the same time that we wanted inspiration, they wanted feedback. They wanted to know, “Could you tell what our A3 was? Could you tell what our CIT was? Could you help us find evidence of it?” At the same time as helping them, we were helping us, because in preparation for it, we looked at their A3s and were very purposeful.... Then we brought all three schools back, and we had grade alike sharing opportunities where you could have your Padlet [digital tool for a shared document] with you and you could share and every group or cohort from all three schools had a site where they were creating a shared learning space.

Then they had a talking circle at the very end where they could put names to faces and faces to names: “Oh that was your classroom. I see on this Padlet that someone had a question about this.” You could post questions on the Padlet as well. We had a walk around survey so you could do that. It was very good in moving some of our instructional

practices forward that were centered on conferring more with our kids and making sure that our cultural responsiveness was really as alive as we thought it was.

Interviewer: And why was it good at that? How did you know it was good at that?

Gwyneth: It was good because it helped others to tell us what our rooms spoke to. Others could articulate, “It seems to me like your room might be grounded in these beliefs. It seems to me that you might use these instructional practices and this pedagogy in your room. Am I correct?” It was a way of vetting, “Yes, that is what I do in that room. Or I can’t believe you missed this. Why did you miss this?” It was because it wasn’t explicit enough. “How did you miss that my whole CIT [collaborative learning team] was focused on conferring? There was no space for conferring, so I didn’t know it was important to you.”

Gwyneth’s re-telling of the process indicates substantial rethinking of standard process and a purposeful shift in who was collecting data and why. Over the course of the interview and in the extant documents she shared, she illustrated hours of professional learning planning and other actions she planned to gather and respond to data from teachers’ classroom environments and student results. Her data collection process and facilitation process were interleaved, illustrating many sequences of wondering, finding out or data gathering, interpreting, and acting.

### **Conclusion**

The internal self-assessment of the professional development process is entirely focused on how the professional developer can use it in later iterations of leading professional learning. If the professional developer did not have a wonder, or the professional learning process was likely to be what Michelle called a “one-off,” less self-assessment was conducted because the only opportunity to respond was in the initial session. Rendering judgement was highly contextual

and situated in the professional developer's experiences. It had only passing resemblance with formal external evaluation, although the aim of the process, to provide high quality, effective professional learning, was the same. Understanding the conditions that influence the self-assessment process and theorizing about the process professional developers evolved provides useful information about why professional learning evaluation is relatively rare and what responses might be helpful in addressing the issue.

## **CHAPTER 5**

### **Introduction**

The goal of improving the quality of education is at the heart of professional development efforts paid for by school divisions. In its least directive purpose, professional learning helps teachers explore their interests and passions related to teaching, and in its most prescriptive, professional training, teachers and educational leaders are instructed in specific mandated patterns they are compelled to follow. Regardless of its purpose, professional learning has the focus of transformation and growth, with the ultimate goal of improving student outcomes, just as learning in the health care system is ultimately focused on better patient outcomes. The goals of professional learning, both the immediate goals for teachers and the long term goal of improving outcomes for students, will only be met if the professional learning offered is actually successful in causing change. Knowing that professional learning is resulting in teachers using new practices is essential in making the supporting case for why school divisions would choose to provide funding and support for professional learning.

Professional learning in Canada and the United States has been widely critiqued as ineffective in transforming teacher practice (Joyce & Showers, 2002; Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009), and newer theory about “best-practice” in professional learning has compelled professional learning to transition to new job-embedded or extended forms (Broad & Evans, 2006; Darling-Hammond et al., 2009; Timperley, Wilson, Barrar, & Fang, 2007). The professional developers in this study described intensive work, largely of their own initiative, that they engaged in to understand if the facilitation methods they were using were helping teachers. In particular, the study participants expressed a desire to assess if better

understand occurred given the concepts they were teaching, and their questions were partially informed by newer standards for professional learning. The process of trying to understand if facilitation works is a part of a larger, K-12 educational focus on changing professional learning to make it more likely to be successful.

Connecting changed professional learning to changed teacher or leader practice and changed student results remains elusive. Like the researchers who commonly speak to educators about professional learning (Guskey 2000; Reeves 2010), the professional developers in the study articulated a series of barriers that rendered attempts to determine the impact of professional learning rare. However, understanding the self-assessment processes of professional developers provides essential information about how evaluation might be built into the K-12 education system. Understanding the self-assessment process also highlights the elements the professional developers saw that must be addressed before that systematic evaluation could happen. As the key stakeholders in planning and delivering professional learning, the professional developers are best positioned to describe what is currently happening, and the most likely to utilize any changes that are recommended. A strong conceptualization of the thinking of professional developers offers the opportunity to break the cycle of critiquing the nature or quality of professional learning and commenting on the ongoing lack of evaluation of professional development.

This chapter elaborates the outline of the self-assessment process, the core phenomenon of described in chapter four. Each stage of the self-assessment process is summarized and related to the implications for other stages. Implications, recommendations stemming from the implications, and reflections are discussed next.

## **Theory of professional developers' self-assessment processes**

The professional developers in this study routinely engaged in a process of self-assessment that had commonalities with, but was not the same as, program evaluation. Like most evaluation of professional learning, the professional developers focused on knowing if the learning was positively received. They also used formative assessments to check if the professional learning they led was changing teacher understanding or beliefs, and they checked for potential barriers often. Little attention was paid to knowing if teachers tried the new practices, or if adopting the new practices improved student results. The professional developers' attempts to understand their own impacts were almost entirely self-initiated, and began with the process of wondering. The phases included: wondering, finding out (data collection), sensemaking and attribution (interpretation), and taking action.

Examining the strategies and reasoning of the professional developers in this study provides useful new information that illustrates what motivates assessment processes, and describes what professional developers consider that tells them whether more information is needed or if things are sufficient as they stand. The examination of the self-assessment process, when contextualized in conditions discussed in chapter four, also highlights the push and pull factors impacting the likelihood of any form of self-assessment or evaluation. Finally, a close examination of the elements of the process reveals what information the process is likely to demonstrate, and where the process omits or avoids judgement for decision-making purposes.

### **Wondering**

Wondering was a self-initiated process of inquiry that the professional developers all engaged in at times. The process of wondering began with questioning what was happening, or would happen, and considering why (see Figure 5.1). Some wonders were brief and easily



dismissed, and other wonders persisted across years and were investigated opportunity to respond was in the initial session. Rendering judgement was highly contextual.

The goal of improving the quality of education is at the heart of professional development efforts paid for by school divisions. In its least directive purpose, professional learning helps teachers explore their interests and passions related to teaching, and in its most prescriptive, professional training, teachers and educational leaders are instructed in specific mandated patterns they are compelled to follow. Regardless of its purpose, professional learning has the focus of transformation and growth, with the ultimate goal of improving student outcomes, just as learning in the health care system is ultimately focused on better patient outcomes. The goals of professional learning, both the immediate goals for teachers and the long term goal of improving outcomes for students, will only be met if the professional learning offered is actually successful in causing change. Knowing that professional learning is resulting in teachers using new practices is essential in making the supporting case for why school divisions would choose to provide funding and support for professional learning.

Professional learning in Canada and the United States has been widely critiqued as ineffective in transforming teacher practice (Joyce & Showers, 2002; Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009), and newer theory about “best-practice” in professional learning has compelled professional learning to transition to new job-embedded or extended forms (Broad & Evans, 2006; Darling-Hammond et al., 2009; Timperley, Wilson, Barrar, & Fang, 2007). The professional developers in this study described intensive work, largely of their own initiative, that they engaged in to understand if the facilitation methods they were using were helping teachers. In particular, the study participants expressed a desire to assess if better understand occurred given the concepts they were teaching, and their questions were partially

informed by newer standards for professional learning. The process of trying to understand if facilitation works is a part of a larger, K-12 educational focus on changing professional learning to make it more likely to be successful.

Connecting changed professional learning to changed teacher or leader practice and changed student results remains elusive. Like the researchers who commonly speak to educators about professional learning (Guskey 2000; Reeves 2010), the professional developers in the study articulated a series of barriers that rendered attempts to determine the impact of professional learning rare. However, understanding the self-assessment processes of professional developers provides essential information about how evaluation might be built into the K-12 education system. Understanding the self-assessment process also highlights the elements the professional developers saw that must be addressed before that systematic evaluation could happen. As the key stakeholders in planning and delivering professional learning, the professional developers are best positioned to describe what is currently happening, and the most likely to utilize any changes that are recommended. A strong conceptualization of the thinking of professional developers offers the opportunity to break the cycle of critiquing the nature or quality of professional learning and commenting on the ongoing lack of evaluation of professional development.

This chapter elaborates the outline of the self-assessment process, the core phenomenon of described in chapter four. Each stage of the self-assessment process is summarized and related to the implications for other stages. Implications, recommendations stemming from the implications, and reflections are discussed next.

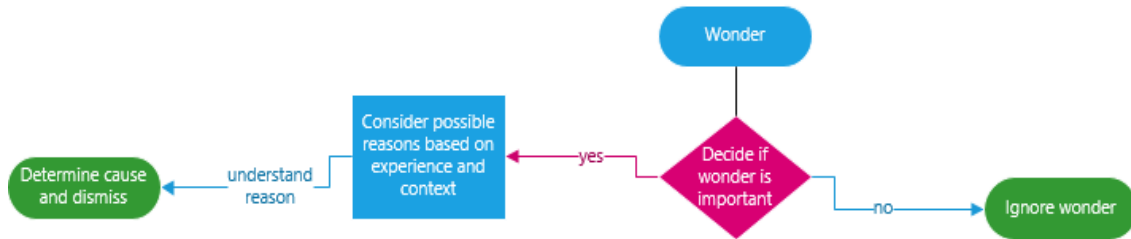
### **Theory of professional developers' self-assessment processes**

The professional developers in this study routinely engaged in a process of self-assessment that had commonalities with, but was not the same as, program evaluation. Like most evaluation of professional learning, the professional developers focused on knowing if the learning was positively received. They also used formative assessments to check if the professional learning they led was changing teacher understanding or beliefs, and they checked for potential barriers often. Little attention was paid to knowing if teachers tried the new practices, or if adopting the new practices improved student results. Unlike formal evaluations, no professional developers described any particular mandate or specific deliverables like evaluation reports related to their questions, although some questions were at the heart of major priorities for their school divisions. Some described sharing their impact with co-workers or superiors, or writing year-end summaries about their schools. Wondering was a personal process related to a desire to be effective, rather than a systemic initiative designed to evaluate division or stakeholder successfulness in changing teacher practice or student results.

Once a wonder occurred, professional developers described considering the wonder's importance (see Figure 5.1). Some wonders were passing thoughts easily dismissed. If the professional developer engaged with the wonder because it was personally or professionally significant, the next step was considering if the question could be answered based on the professional developer's past experiences or context. Since professional developers relied heavily on their own experiences, many determined potential causes quickly and dismissed potential alternatives. When cause was intuited or remembered, the wonder was typically dismissed. However, sometimes the wonder became a nagging problem or issue, and the

professional developer decided it was worth investigating. Self-assessment occurred only when there was no easy answer or explanation.

Figure 5.1 – Wondering as a trigger for self-assessment



The process of early dismissal of a wonder is significant because of its implications for change in professional development. Commonly agreed on explanations, organizational culture, instincts, and past practice were all sufficient to end wonders, making the process of self-assessment unlikely to challenge school division actions, professional learning processes, or priorities. As professional learning processes are undergoing major changes in North America, some wonders were about cognitive dissonance with professional learning practices, but they did not result in systematic investigations of impact. The professional developers' wonders frequently focused on practical details related to changing their facilitation processes or addressing barriers, but they did not focus on transformational change. In addition, the pace of change and scarce resources results in professional developers challenging their major wonders as too complicated or beyond their control, resulting in minor modifications of professional learning rather than major envisioning of new forms of professional learning or more efficacious processes. If evaluation of professional learning was a natural part of the instructional cycle as it was in the classroom, it is possible the process of wondering could lead to more systemic

responses. However, in the absence of evaluative thinking, wonders were akin to formative assessment; they were useful questions that lead to changes of processes in progress. The process of self-assessment helped to refine professional learning process but was not focused on assessing substantial changes to teacher practice.

### **Finding out**

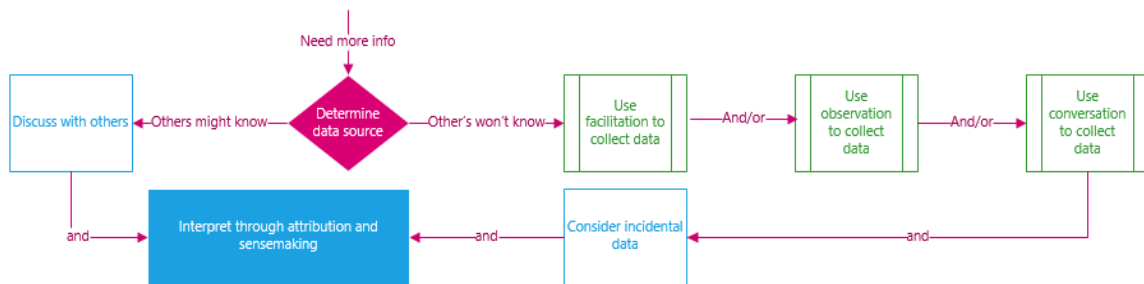
Once a professional developer determined more information was needed, the process of data collection began (see Figure 5.2). Sometimes a formal plan for collecting data was established, such as Gwyneth's photos and learning walks, but more commonly, processes were either informal or built into facilitation. If participants thought others might be able to answer the wonder, they usually began data collection with discussion (see Figure 5.2). Professional developers described asking co-facilitators or calling trusted members of the potential participant group for their feedback. Some participants checked with mentors, or considered what mentors or superiors suggested, but this approach was less common. The participants also consulted with others in forms of pre-assessment to find out necessary information that sprung from wonders.

Starting with consultation or conversation was significant for several reasons. First, it was consistent with the professional developers' concern with positional authority, and placed them as interested co-workers. Second, it meant professional developers made decisions based on the common thinking of the group they were already a part of. The early use of conversation and consultation was helpful in building cohesion and meeting self-described needs, but it was more likely to spark modification than revolution.

When others did not know the specific answer to the wonder, or the wonder needed to be answered during or after the professional learning, the professional developers chose from a variety of processes for collecting data. Using facilitation processes and having conversations

during or after the professional learning, even incidental conversations, were much more commonly reported than observations (paradoxically, observations were frequently described as the most convincing data). Sometimes the selection of various types of data processes was pre-planned, but even when it was, it often changed in response to data. A professional developer might wonder about a likely barrier in beliefs, design a process to determine if the barrier was an issue, find the barrier, then implement a process to respond (see Figure 5.2). The cycle of data collection, attribution, and action was a flexible response that continued until the professional developer felt the wonder was sufficiently addressed.

Figure 5.2 – Finding out by collecting data



The process articulated by the professional developers was highly responsive, and they explained it as essential to tailoring the professional learning to the specific participants and ensuring that participants left with the key skills and an understanding of what the professional learning was designed to teach. Unlike a formal evaluation process, *finding out* was immediate and the ideas were typically used. However, the processes were based primarily on professional developers' perceptions of the data and limited to things the professional developers had experienced or seen. As a result, the data gathering was uneven and little attention was paid to issues related to accuracy compared to issues like usefulness.

Professional developers were typically isolated in trying to determine their own impacts, although they described recruiting what Gwyneth described as “critical friends” to help them think and observe. Because they made limited use of frameworks for thinking about data collection and interpretation, the professional developer focused on learning about facilitation in a way unanticipated in early theory in the study. Professional developers were specifically motivated by being good facilitators. They each defined “good facilitator” or “good leader” personally, then choose to actively pursue improvement based on the definition. Over half referenced the theory in newer models of professional learning and described issues related to distributive models, which were exacerbating complexities in data gathering processes.

The implications of their understanding of newer process and desire to succeed were significant. The professional developers were intrinsically motivated to change and improve, and valued being perceived as helpful to teachers. They gathered information formatively and made changes rapidly. If their definitions of “good facilitator” or “good leader” included practical assessment process that did not challenge teacher autonomy, they showed every sign it would be used, modified to fit specific circumstances, and refined over time, just as their facilitation processes were. They repeatedly described adopting new facilitation processes because they experienced them, and could see when and why they would be useful.

The professional developers gathered substantially more information about how learning was understood and what barriers might be a problem than research indicated they might. Four of the professional developers indicated they saw Michelle’s “knowing/doing gap” and could not accept on faith that learning would be used unless it was well understood, aligned with teacher beliefs, and likely to succeed based on barriers the teacher would face. The professional developers saw teacher engagement and agency with professional learning, the quality of

learning achieved, and the deliberate reduction of barriers as critical pre-cursors to the possibility of use in the classroom. They needed to be convinced that wonders in these areas were addressed before considering wonders related to use, and selected tools that yielded formative data that allowed them to address those concerns. Formal evaluations with an extended process, completed after professional learning is complete did not address the most pressing need to *find out*, making it likely they would be rejected even if formal processes were better understood. The pace of formative processes make them much better suited to finding out answers to the common wonders the professional developers articulated because the information came in time to be responded to in situ.

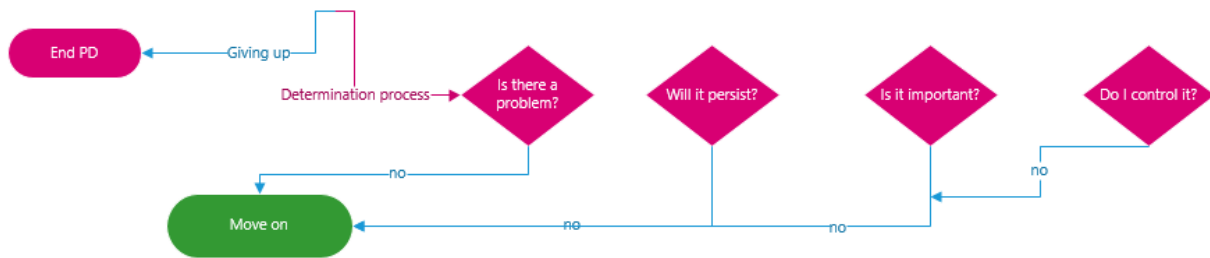
### **Interpretation: attribution and sensemaking**

The process of data gathering about facilitation and barriers was elaborate, but the process of attribution and sensemaking was brief (Figure 5.3). Even when asked in several ways, many professional developers in the study did not describe a process for turning data into evidence. They attended to many of the elements of program evaluation standards instinctively, but omitted or glossed over accuracy standards. The professional developers were wont to state a reason without describing alternatives, although they described details about the milieu that created a rich, layered picture of context. Once a “likely” reason was found, the professional developer focused on a determination of the most appropriate courses of action. Some of the professional developers did not describe how they decided, jumping from data collection to summary judgement and then spending more time endeavoring to respond. Those who articulated thinking about action explained an attribution process (see Figure 5.3). The process began with considering data from the wonder to decide if there was a problem. If there seemed to be an issue or a problem, then the professional developers considered three factors: would it persist,



was it important, and could they influence it. A “yes” to all questions, even when it was qualified, typically resulted in action (see Figure 5.3). A “no” to any question usually resulted in the professional developer expressing frustration and moving on or complaining, then giving up. The focus was on deciding if acting was appropriate more than analyzing and interpreting data (see Figure 5.3). Participants did not describe processes for inspecting, modeling, coding, or clarifying data, although some used processes of summarizing or generalizing. Even when prompted, they did not discuss how different types of data might require different ways of trying to make sense of them. Interpretations were rapidly attributed based on experiences, and the participants moved on. Figure 5.3 illustrates the professional developer’s rapid process of determining if action could be taken, which was typically substituted for the program evaluation process of turning data to evidence to render judgement.

Figure 5.3 – Interpreting data



The thinking processes of the professional developers typically lacked a formal plan for interpretation, missing processes of interrogating data, which likely resulted in less systematic, accurate, or fulsome determinations and explorations. While two professional developers initiated processes of considering data with teachers, they did not describe research supported

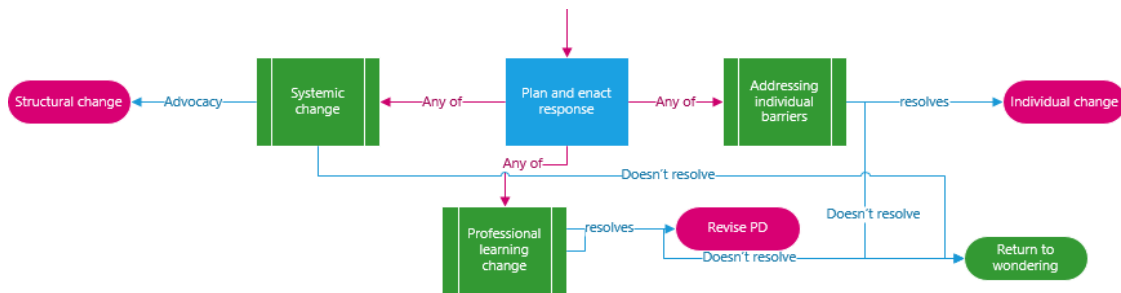
models for making sense of data in the way they did for facilitation processes, indicating a lack of learning related program evaluation and data analysis or a lack of interest in the process may have been an issue. Interpretations frequently occurred without data about teachers' use of the new practices or potential impacts on students, so judgements were missing two key elements in the logical links that might allow a conclusion that professional learning was effective in affecting the classroom.

### **Action**

Taking action (see Figure 5.4) was the mainstay of the professional developers' self-assessment, because they gathered information with the primary goal of deciding what to do immediately and in the long-term. In contrast, many forms of program evaluation, except for utilization-focused program evaluation, are focused in providing information that others might act on. If the determination process of the professional developers was cursory, the action phase was the opposite; it was a complex, contextually-based calculation. The professional developers often interwove responses to multiple perceived issues related to the professional learning, including what systems needed to do, what individuals in the professional learning might need, and potential changes to be made to the professional learning processes. At this stage, professional developers most closely aligned with newer principles of professional learning, as their responses and next learning steps were as targeted at the lived realities of their adult learners as they could possibly make them. The professional developers targeted many of their actions as specific responses to perceived issues (see Figure 5.4). The processes they used to enact changes are worthy of study in their own, as they are composed of factors related to the theory about leadership, cultures of local and provincial systems, and change. Once professional developers decided they had the agency to respond to their wonders, they typically responded in

a wide variety of ways, not one. Figure 5.4 illustrates some of the main responses professional developers used during the taking action phase.

Figure 5.4 – Taking action



The professional developers had great opportunity to respond to what they believed was occurring, and they also had access to key actors as system insiders. As a result of their access and motivation, they reported the utility of their self-assessments as high. However, without support from their employers and the education system to learn about and use systemic evaluation process, it is possible the accuracy of their assessments was affected. In addition, in the interviews and artifacts, very few references to systems were designed to set goals and provided a structure for monitoring them, such as the provincial hoshins (frameworks for K-12 Saskatchewan educational planning during this study). Except for references to the overall goals, which occurred about half the time, there was little evidence that the provincial strategic planning process was related to the professional developers' self-assessments. The structures of the provincial processes were not referenced by the professional developers, and the processes used provincially did not come to mind for most of the professional developers when they were asked about who wanted to know about their impacts. Only two mentioned them, and they did not refer to them when given additional opportunities to do so. The self-assessment processes

remained largely independent of system goals and were predominantly targeted at personal efficacy in processes facilitating and leading. As a result, the self-assessment processes were more likely to yield improved capacity in helping teachers understand than to impact division or province wide improvement goals for student learning.

### **Implications**

Understanding how professional developers self-assess remains critical because their assessments are the primary form of evaluation of professional learning that occurs provincially. Like many places in Canada and the United States, evaluation of professional learning is rare, even though efforts to engage in continuous improvement education are rampant. For many educators, the logic of how education might improve is imprecise. Assuming students are not magically transformed, there needs to be some new or research-supported instructional methods or structural changes that might cause teaching and learning to improve. The mechanism for how the new methods will become commonplace and used is often the logical leap; strategic plans go from what they would like to see changed in education systems to lagging indicators in the student data they intend to use as evidence of student progress. Mechanism for the change, and leading indicators of the impacts of professional learning may be omitted all together or referenced in a general way that indicates the acknowledgement that professional learning will be necessary without any indication of a particular plan. The common leaps over planning for mechanisms of change are not sufficiently addressed by a body of research that articulates clear evidence about best mechanism, leading measures, and likely progressions. It is unclear what, exactly, will need to happen as a part of the professional learning so the new ideas are understood and seen as important, and what supports will need to happen afterward that will result in teachers changing their classrooms rooms successfully. Although there are many

theories about how this might occur, there is a genuine lack of generalizable evidence over time (Desimone, 2009). Even if there is some degree of certainty about the value of some new method for the classroom, a successful process for learning about it and using it is still uncertain. The absence of processes for evaluating or assessing professional learning (Guskey, 2000) means practitioners are offering professional learning to create change without knowing much about what impact it has or is likely to have.

While the need for strong evaluation is pressing, it remains elusive (Guskey, 2000; Reeves, 2010). Calls for more evaluation of professional learning have occurred in a vacuum of understanding about how the leaders of professional learning are conceiving of, or engaging in, current processes. Understanding the thinking process of professional developers offers critical insight into how professional learning is constructed and forms of insider evaluation are conducted. The information about current process allows for finding common ground between what researchers suggest and what circumstances are likely to create, and gives insight about what might make assessment or evaluation process easier and more successful.

This section on findings contains extended comparisons to three theorists who suggest ways we could think about evaluating professional learning. Desimone (2009; 2011) explains the logic of why we invest in professional learning, and Guskey (2000; 2002) and Reeves (2010) explain what to consider when evaluating professional learning and how to go about it. Each comparison helps to articulate what the focus of the self-assessments in the study was, and how it was the same or different than the conceptualizations in research and educational theory. Using the comparison to the logic of professional learning, and then potential models for assessment and evaluation conducted in the field, this section describes key implications for potential changes in the practice of determining the impact of professional learning.

## **Significance of the leaps in professional learning logic**

Desimone's (2009) description of the logic of professional learning provides an interesting counterpoint to the professional developers' description of their self-assessment practices. She articulated the theory of change assumed by all forms of professional learning, which contends that professional learning alters teacher knowledge, understanding, and beliefs and the theory of instruction, namely that teachers will then change their practices in a way that "influences student achievement" positively (p. 184). The theory of change includes two logical links, both of which were the focus of self-assessment of the professional developers. First, that "teachers experience effective professional development" and second, that "the professional development increases teachers' knowledge and skills and/or changes their attitudes and beliefs" (p. 184). The professional developers in this study all articulated a personal theory of effective professional learning composed of personal experiences. Some also included a substantial research base, describing how prior methods of professional learning were challenged by specific authors. Most described "effective professional learning" in a way consistent with the agreement among researchers described in the literature review. The professional developers also described a feeling of responsibility for the quality of the professional learning, and stated they controlled the quality, within the structural limitations of issues like time and budget constraints. The decision to self-assess came from a personal desire to be an effective and useful facilitator for teachers, and then indirectly, to further the education of students.

Because the professional developers wanted to be strong facilitators, they articulated the two logical links in the theory of change and deliberately collected information. They all described asking teachers if they thought the professional development was helpful or effective at meeting their needs, and sought to compare other data, like feedback from peers, to teachers'

responses. Using their formative assessment processes, and comparing those deliberate formative data points with incidental and purposeful conversation and observation, the professional developers checked for changes in beliefs, attitudes, barriers, skills, and/or understanding, depending on their goals and contexts. All the professional developers expressed the belief that professional learning, if constructed powerfully, changes teacher understanding or beliefs, and they sought evidence they were creating those changes.

Where the professional developers and Desimone's description of the logic of professional learning diverge is in the theory of instruction. Only a few professional developers in the study explained checking the logical link that "teachers use their new knowledge and skills, attitudes, and beliefs to improve the content of their instruction or their approach to pedagogy, or both" (p. 184) and no professional developers overtly monitored that the change fostered student learning, which was the final logical link, and the essential purpose. Like Dean argued, there was an element of drinking the Kool-Aid. The professional developers articulated it was hard to gain access to observe all the teachers they had worked with, unless those teachers were in their personal schools because they were instructional coaches or administrators. Even when they could, they did not systematically collect evidence specifically to see if the new practices changes students' results, although both Sherry and Gwyneth described more general observations of students and collection of substantial data about teacher practice.

Table 5.1 – Desimone’s logic of professional learning as lived in professional developers’ self-assessment

<b>Desimone’s Theory</b>	<b>Desimone’s Logic steps</b>	<b>What professional developers did</b>	<b>Professional developers’ stated purpose</b>
<b>Theory of change:</b> Professional learning alters teacher knowledge, changes and beliefs	Teachers experience effective professional development	<ul style="list-style-type: none"> <li>● Asking teachers if they thought the professional development was helpful or effective at meeting their needs</li> <li>● Comparing to feedback and observation</li> <li>● Feeling responsible to the quality of the professional learning</li> </ul>	<ul style="list-style-type: none"> <li>● Help teachers (all)</li> <li>● Improve perceived or actual quality of professional learning (all)</li> <li>● Personally grow as a leader or facilitator (all)</li> </ul>
	The professional development increases teachers' knowledge and skills and/or changes their attitudes and beliefs	<ul style="list-style-type: none"> <li>● Stating a personal theory of what effective professional learning was that was composed of personal experiences.</li> <li>● Some participants included a substantial research base, describing how prior methods of professional learning were challenged by specific authors.</li> <li>● Checking for increase using facilitation strategies</li> <li>● Observing for increase during facilitation</li> <li>● Asking about increase of change during or after facilitation</li> <li>● Checking for barriers to change before, during and after facilitation</li> <li>● Asking for teacher self-reports of change</li> </ul>	<ul style="list-style-type: none"> <li>● Help teachers (all)</li> <li>● Improve student learning (some)</li> <li>● Change the system (some)</li> <li>● Reduce barriers to change (all)</li> </ul>
<b>Theory of instruction:</b> Teachers will then change their	Teachers use their new knowledge and skills, attitudes, and beliefs to	<ul style="list-style-type: none"> <li>● Not often checked due to opportunity, potential infringement on teacher professionalism, and timing.</li> <li>● Checked in coaching situations</li> </ul>	<ul style="list-style-type: none"> <li>● Not appropriate to check up on or judge teachers</li> <li>● Perceived as valuable but</li> </ul>



practices in a way that influences student achievement	improve the content of their instruction/ pedagogy	<p>or by learning walks when professional developer was embedded in a school</p> <ul style="list-style-type: none"> <li>• Described as most easily assessed by observation in classrooms</li> </ul>	<p>assumed if professional learning was effective</p> <ul style="list-style-type: none"> <li>• Potentially inappropriate or problematic</li> <li>• Useful to know</li> </ul>
	The change in teacher practice fosters student learning	<ul style="list-style-type: none"> <li>• Professional learning rarely directly connected to student progress through specific assessment</li> <li>• Difficult to prove causality, but could be associated with related data like reading data or attendance data</li> </ul>	<ul style="list-style-type: none"> <li>• Useful to share with stakeholders like parents, can create “buy-in” with other teachers</li> <li>• Hard to gather</li> </ul>

The gaps in the theory of instruction are significant for several reasons, but failing to connect the learning to its theoretical goal is foremost among them. As Desimone noted, all parts must be evaluated for a claim of impactful professional learning to be substantiated. That meant that although the professional developers attempted to evaluate their impact, they assumed that if the professional development is effectively constructed, it resulted in changed student learning. The assumption is problematic, because it means professional developers, and to a greater extent, school divisions, routinely assume professional learning is related to improvements in student learning, provided teachers described it as helpful or enjoyable.

Desimone’s critique that educators assume professional learning will help students, which she substantiated primarily by examples from research studies, appears to hold true for practitioners of professional development in the field, as they rarely described assessing the elements required for the theory of instruction. Those elements that professional developers most directly controlled, namely the quality of the professional learning and its success in

creating a change of understanding or beliefs, were the focus of their self-assessment. The elements of the theory of instruction, which fell outside the direct control of the professional developers, were far less likely to be systematically assessed, and when they were, the need to assess ran afoul of the professional developers' perceptions of professionalism. Amy's explanation she provided the professional learning, but the teacher makes the choices about classroom instruction, underscores the professional developer's articulation that the locus of control is appropriately placed with the teachers, not the professional developers, and should not be a focus of any professional developers' self-assessment.

### **Common ground for self-assessment and evaluation**

To understand the self-assessment process of professional developers as a potential tool for determining the impact of professional learning, self-assessment needs to be understood in the context of definitions of evaluation. Guskey defines evaluation as "the systematic investigation of merit or worth" (2000, p. 41) and characterizes it as being purposeful and thoughtful. The process described by the professional developers was typically not a systematic process. It was self-initiated based on a particular interest, and was not extended to the teachers' or leaders' use of the new learning. It also did not make direct connections to the impact on students. Instead, it was tightly focused on the facilitation process, and the professional developer's wonders about its impact on teacher understanding or beliefs. The professional learning self-assessment processes were usually professional developers' generalized assessments of professional learning, rather than a systematic investigation of merit or worth. However, the key commonalities that the professional developer's self-assessment process have with program evaluation provide fertile ground for developing evaluative processes.

Both evaluation and professional developers' self-initiated assessment focus on a shared purpose - understanding the perceived merit of a professional learning process. The insider nature of the self-assessment provides specific benefit related to the assessments of merit. While a program evaluator might be externally contracted and provide a variety of reports to a client who may act on the results, the professional developer is focused on using the information gathered formatively to shape professional learning and has the main agency in acting on any findings. Both the systematic assessment and the response to findings are needed for the improvement of a professional learning program, meaning that regular use of program evaluation or research studies would not be sufficient to see change and growth. Second, both forms of considering the merit or worth start with something that needs to be determined, weigh the inputs and actions, and assess results to determine next steps. The foundational elements, taught as a formal curriculum to program evaluators and derived by professional developers, have common process elements.

Proof of causality in professional learning, specifically that professional learning causes changes in student results, is very difficult to determine (Guskey, 2000; Hill, 2013; Hirsh, 2013). Professional developers in the study did not try to prove causality and articulated there were too many factors to consider. Most participants routinely checked on teachers' understanding to see if it improved during the professional learning experience, usually formatively within the experience itself but sometimes with instruments like post-event surveys. However, only some participants checked to see if teachers improved practice in the way suggested in the experience, and few connected professional learning to specific improvements to student learning. Since student learning or engagement were rarely connected to professional learning, trying to systematically discount other possible factors in student improvement was impossible. However,

program evaluation processes might provide missing elements that would make the standard of likelihood that Guskey (2000) suggests attainable. A combination of the professional developers' formative assessments and action elements with more specific and deliberate pre-planning for data gathering and interpreting would help to combine the essential elements into a more successful method for regular monitoring of professional learning.

Reeves (2010) offers a potential solution that is a blend between formal evaluation and the self-assessments described by professional developers in the study. It includes four elements:

1. Plan for what you will accept as evidence before you start, and deliberately collect it before, during, and after professional learning.
2. Focus on assessment for learning, rather than evaluation, and consider teachers, building administrators, system administrators, policy makers, and parents as sources of data and ideas.
3. Create preflight checklists for teachers, principals, and system administrators before action occurs so the goal and processes are clear to everyone.
4. Use effective formative feedback often. (p. 57)

His second element was a main focus of the self-assessment of the professional developers, but the other three were partially evident or missing in the descriptions of self-assessment. For extended professional learning, goals and process were described and shared, and formative feedback occurred during professional learning. Deliberate plans for what evidence would be collected before the start of professional learning were not usually stated, but they sometimes evolved.

Combining Guskey, Reeves, and Desimone in comparison to the self-assessment reveals a common pattern. Professional developers assessed the things most within their control, which required less support from others or from the systems they were a part of. This study did not attempt to deeply examine the nature of the systems they worked in, instead focusing on examining the process of self-assessment rather than the context. Ethnographic or phenomenological studies of related questions could provide useful additional richness related to the systems professional developers work within. It is possible the professional developers feared discovering their professional learning was ineffective, so did not investigate all the logical links to avoid uncovering a problem. They expressed knowing their impact was important to them. However, it seems unlikely they wanted no negative information given the time and effort the professional developers spent in self-initiated investigating related to Guskey's (2000) first three levels and Desimone's (2009) theory of change. It is also possible systemic inertia made changing instruction unlikely, or lack of interest from others convinced the professional developers it was not worth their time. Systemic inertia and lack of interest seem potentially related given the barriers the professional developers reported. Possibly, professional developers' concern for respecting teacher professionalism and other elements of educational culture made evaluating uncomfortable or inappropriate. Regardless, they had an interest in knowing, since they investigated their impact informally and without being required to. However, the investigations did not withstand something beyond themselves. Because professional developers were interested in self-assessing impact, changes to other factors could likely allow for more fulsome evaluate process with at least some support from professional developers.

### **Gathering data about teachers' use of new knowledge and skills**

Some resolution of the impasse between professionalism and the theory of instruction is essential for any form of self-assessment or evaluation to determine if professional learning is working as it is intended. The minority opinion, in the form of professional developers like Sherry or Gwyneth, may provide the best guidance. Both women felt observing teacher knowledge and use of new skills was essential and actively involved teachers in both goal setting and observation stages, running counter to the more common thinking that assessing professional learning stopped checking understanding, appreciation, and barrier reduction. Knight (2014) noted that if observing teacher practice is formative assessment for growth, rather than summative measurement of success, it can increase autonomy and accountability, and be the substrate for constructing more effective professional learning.

Professionalism and checking changes in instruction and results do not need to be mutually exclusive. In Knight's (2014) coaching process, and Dufour et al.'s (2010) professional learning communities, teachers pick targets based on student need, then gather their own data about their impact much like the processes articulated by this study's professional developers. Movement to these forms of professional learning aligns with new theory of effective professional learning, and provides utility for assessing teachers' use of new skills. Unfortunately, both process are plagued by the same problems that hurt professional developers trying to gather evidence, what Sherry described as teachers not knowing what good evidence is and not planning to collect it. A much stronger understanding of evidence of change in teacher practice, and corresponding evidence of improvement in student results, is necessary for both teachers and professional developers to resolve issues with evidence collection. Professional resources designed to address the need (Dufour, Dufour, Eaker, & Many, 2010; Hume, 2009; Peery, 2011) could be a strong

support, provided that divisions, their teachers, and other stakeholders agreed that change is important. Professional learning processes, provided they contain learning about gathering, analyzing, and interpreting data, can help overcome concerns that assessing use is antithetical with teacher agency.

### **Data Literacy**

The largest issue with teachers and professional developers collaborating to collect evidence is that neither the majority of professional developers, nor most teachers they work with, feel comfortable with a full data collection process. Jimmerson & Wayman (2015) summarized the key areas where educators are requesting help:

Educators articulated professional learning needs related to data use in six main areas: (a) asking appropriate questions of data (to guide analysis and use); (b) accessing and operating district data systems; (c) data literacy/interpretation; (d) fitting data use with day-to-day practice; (e) sharing information via collaboration; and (f) knowledge codification. (p.1)

Regardless of the methods employed to increase the level of evaluation in professional learning, increased data literacy is pre-requisite. However, Jimmerson & Wayman also described a lack of “clear plans for addressing data use capacity through professional learning supports,” (p. 2) indicating school divisions and other organizations do not yet believe data use skills are critical. The researchers, who recommended knowledge mobilization skills, stated that knowledge is only useful if codified, essentially that having new knowledge in a field is not helpful until the knowledge is widely understood so it can be applied (Levin, 2011; Sudsawad, 2007; Weiss, 1979). The argument that the knowledge gained by professional developers may not always be useful holds merit; some described explicit exchanges with mentors, co-leaders, or

in workshops, but acquiring information about best facilitation strategies was not typically described as systematic, except within the provincial facilitator's cadre. As a result, useful learning about the impacts of various professional learning processes did not move readily around various K-12 education structures, although the people doing the learning had substantial expertise in designing processes for sharing and discussing new insights about teaching and learning. Formal professional learning activities designed to help practitioners collaborate to increase data literacy and share the implications are essential elements in mobilizing the new learning provincially during these initial stages.

### **Reducing barriers**

The professional developers in the study confirmed many of the barriers to evaluation that research articulated. They described issues with lead time before professional learning and time for planning, and frequently described problems with the current funding levels that Garet et al., (2001) articulated. The professional developers also agreed there was little time to capture, process, and respond to data (Honig, Copland, Rainey, Lorton, & Newton, 2010, p. 2), let alone make effective professional learning decisions based on data or monitor results. They were able to gather information about teachers' valuing professional learning, understanding the key ideas, or experiencing reduced barriers formatively in sessions, but had difficulty accessing data about changes to teacher practice and student results. The administrator and senior administrators in the professional developer participant group described how establishing a professional learning culture within the school had to be a deliberate process and was often a barrier as described in research (Cole, 2004; Fullan, 2007; Reeves, 2010). Their thinking on the issues of school culture might be related to the fact they often had the best opportunity to observe teachers.



Guskey's levels of professional learning evaluation (2000) included evaluation of barrier reduction that were atypical in other models, but were a main focus for the professional developers. As they described planning for self-assessment of goals, engaging in facilitation, or structuring multi-day or year professional learning, the professional developers often illustrated monitoring and attempting to address barriers in much greater detail than researchers articulated. They saw the barriers as critical early in the facilitation process, to prevent issues that might cause teachers to reject the new learning conceptually or practically. As a result, they frequently bracketed descriptions of their own processes with contextual information about the participants, the timing, the employers, the working conditions and other elements of the lived experiences of their own participants and circumstances. The descriptions of the professional developers indicate the significance of gathering data about potential barriers and planning to reduce them is much more critical in practice than researchers typically theorize.

While the overall planning for evaluation of professional learning was not systematic as described by the practitioners, the attempts to reduce barriers were. The professional developers were deliberate and methodical in anticipating likely issues and planning processes to address them, then assessing to see if the barriers they controlled were being reduced. As the professional developers described the process in the talk-alouds about their own choices, they often named more than four factors related to potential barriers that affected each choice. As Juanita described it, the detailed, minute assessments related to barriers were essential in ensuring the goals of the professional learning processes were met, and it was possible that teachers might use the learning in their classrooms. The thinking of professional developers extended beyond practitioners thinking about daily realities to theorists thinking about the larger context. It reflected a common experience that the quality of professional learning is conditional,

contingent on the people it is designed for and circumstances it occurs in, not just adherence to theory of “best-practice.” By extension, quality assessments of professional learning need to account for significance on contextual factors in moderating the impact of any particular professional development.

Desimone’s (2011) critique that professional development results are often ignored and hard to gather appeared to be partially true for the professional developers in the study, who describe many systemic barriers that prevented it. However, her contention there was no focus on “the processes that make it work” (p. 68) was not accurate for the participants. They spent substantial time trying to see if professional development was working, and had elaborate structures for doing so. However, they mostly attempted to see if the process was leading to increased understanding or a change in beliefs, fulfilling only half of Desimone’s four burdens for evaluation.

Suggesting that school divisions and educational partners engage in program evaluation as a regular part of professional learning is impractical. Professional developers and other leaders are not in a position to engage in that type of evaluation. According to Hirsch (2013), practitioners may not understand the need for evaluation of professional learning. Additionally, “administrators and evaluators lack the skills, experience, and ability to effectively evaluate the true impact of professional development” (Mark, 2014, p. 1648). While the majority of professional developers in the study were fairly experienced, they articulated concerns about the skills to determine true impact. Hill et al. described the professional learning as locally developed, with a short time span of use, and “little or no formal evaluation”, which matched the descriptions of the professional developers (2013, p. 476). Even when professional learning was used more than once, like in Amy’s example, or for an extended time, like in Gwyneth’s, a rapid,

simple process of evaluation would be necessary to meet the needs of organizations like the school division. Full program evaluation would not be generalizable, and it would take too much time given the pace of professional learning. While divisions and other stakeholders need to understand much more about how to evaluate professional learning, they are very unlikely to adopt the formal evaluation process program evaluators use, even if is usability oriented. A simplified version will need to be developed and customized for school divisions to address lack of experience and short time span. Learning about how to use an adapted structure to determine the quality of professional learning is within the scope of K-12 educators, but continued calls for full evaluation mean that change is untenable and rejected.

### **Comparing self-assessment and program evaluation standards**

Stufflebeam (1999) developed a checklist to accompany the *Program Evaluation Standards* for final, summative meta-evaluations. It helps practitioners think through what they need to attend to in thinking about evaluation, and attends to the main standards including utility, feasibility, propriety, and accuracy. Patton's (2013) version with a utilization focus, attempts a similar function. While neither are functionally similar enough to education to be useful in the internal assessment processes professional developers might use, they can be helpful in determining what is commonly omitted altogether in professional developers' process. The professional developers in the study did not mention program evaluation standards, but they repeatedly referred to logic and concerns related to utility, and relevant elements of feasibility and propriety. However, except for context analysis, which they all attended to directly and frequently, and documentation, which they described as only personally relevant, they did not typically refer to accuracy standards. In particular, concerns related to defensible information sources, valid information, reliable information, systematic information, justified conclusions,

and analysis for qualitative and quantitative data were conspicuously absent when compared to the indicators on Stufflebeam's checklist. The lack of systematic planning for specific evaluation of professional learning as described by Guskey (2000) or Reeves (2010) presents a similar concern. The combination of lack of thinking about accuracy in data collection and analysis and the limited systemic planning renders the self-assessment inadequate for determining the impact of the professional learning, even if more data were collected about teacher use and student results.

### **Recommendations**

The processes of self-assessment used by professional developers based on their own interest in effective facilitation and leadership reject researchers' arguments that professional learning assessments focus almost entirely on participant satisfaction, but they confirm that a genuine process for fully assessing professional learning is not in place. Professional developers agreed that most professional learning offered is not assessed for teacher use or student impacts, and there is no comprehensive plan for determining if professional learning is working, either provincially or with specific stakeholders. Understanding the thinking of professional developers, however, allows for nuance in recommendations for improving the frequency and efficacy of the professional learning assessments. This section describes a prerequisite burden for research guidelines about what works in professional learning, then describes recommendations for mechanisms that could promote change, including: a) clarifying logic, b) clear goals, c) a systematic approach to data collection and interpretation, d) engaging in observation of using learning, and e) addressing issues that arise from insider evaluation. The five recommendations provide steps possible in the context of the provincial education system and comparatively easier to implement than full-blown program evaluation for stakeholder

groups like school divisions. In that sense, the recommendations are a compromise; they will not be the equivalent of a strong research study, but offer a genuine opportunity for useful information as a part of regular offerings of professional learning.

### **Research guidelines to help professional developers**

Prior to the recommendations for practitioners, this section discusses the pressing need for better direction from researchers. Research has suggest major changes to theory about what makes effective professional learning, but has yet to establish supported evidence about the generalizable design information professional developers need (Borko, 2004; Hill et al., 2013), so school divisions and other professional learning providers will need better direction from academia. Research is failing to provide clear guidance because it continues to focus on evaluating specific programs, rather than large scale comparisons of forms and process that can be generalized for planning purposeful professional learning (Desimone, 2009; Wayne, Yoon, Zhu, Cronen, & Garet, 2008). If comparisons of the new forms of professional learning occurred more fully in research, then practitioners could plan their professional learning from those principles, confident that they were planning from guidelines about what was most likely to improve teacher understanding and use of professional learning. Professional developers in the study were clearly interested in what was likely to be effective, and would be intrinsically motivated to do additional learning. Two caveats apply; the research would need to be easily available online without access to scholarly journals, and written in plain language so it could broker knowledge. (Levin, 2011; Sudsawad, 2007; Weiss, 1979). The deliberate translation of research into easy to use formats is a precursor to avoiding the theory to practice gap (Sudsawad, 2007). With research-based guidelines in hand, professional developers would plan based on already established influences, rather than general ideas, and could focus less of their attention of

self-assessments on facilitation process and more on transfer of teacher understanding into teacher practice.

### **Making the implicit logic of professional learning explicit**

The professional developers typically described the types of general foci, for example “learning about teaching reading.” The goals were sometimes connected to strategic plans, but the plans did not use the implicit logic of professional learning. A division might have a goal of improving math scores, and look at math rubric results once a year. With lagging indicators about student learning and no specific plan connecting the targets for student learning with a progression of professional learning goals and plans, strategic action is difficult or potentially ineffective.

School divisions and other stakeholder groups need to routinely connect the student learning target with specific evidence. Questions that ask about the intended future state and start with data could help divisions in considering key details:

1. What is our most important student learning goal (future state)?
2. What would convince us we have achieved it?
3. Why are we not already achieving it? (pre-assessment).

If practitioners believe a student need is so strongly connected to instruction that divisions would use limited professional learning time to address it, then senior leaders and professional developers (not just professional developers) need to clearly understand what happens and what change they would like to see. Once again, guiding questions in the planning stage can help divisions address key details in the logic that might be overlooked:

1. What do our teachers often do now? (pre-assessment).

2. What additional educator behaviors and beliefs would be likely to make a significant difference (instructional change)?
3. What factors are preventing educators from believing and doing those things now (barriers)?

Planning for evidence collection, as suggested by Reeves (2010) and Guskey (2000), requires considering what the best evidence is before planning the professional learning. It is also critical that divisions use a mix of leading and lagging measures. Leading measures are things that are likely to change first, convincing the planning team that changes are underway, even if they have yet to impact student results. Lagging measures are types of evidence that are visible once the change has occurred. Research suggests that school divisions might typically only consider if teachers liked the math PD based on the exits slips after an event (Hill et al., 2013, Hirsch, 2013; Reeves 2010). Using Desimone's (2009) logic of professional learning and Guskey's levels of evaluation, good evidence collection needs to avoid assumptions that liking will result in improved student results. Instead, each element of the logic of professional learning needs to be specifically assessed, which requires a plan that draws from each of the levels as Guskey described them.

Table 5.2 – Planning to use leading and lagging measures

<b>Guskey Level</b>	<b>Desimone’s logic</b>	<b>Potential evidence</b>
Participant reaction	<b>Theory of change:</b> teachers experience effective professional development	<ul style="list-style-type: none"> <li>● <b>Leading:</b> evidence from participant self-reporting (conversations, observations, survey)</li> <li>● <b>Lagging:</b> evidence from participant self-reporting collected after time has passed, connected to other levels</li> <li>● <b>Leading:</b> Observation of professional learning and comparison to theory of effectiveness</li> <li>● <b>Leading:</b> formative assessment during professional learning</li> </ul>
Participant understanding	<b>Theory of change:</b> the professional development increases teachers' knowledge and skills and/or changes their attitudes and beliefs	<ul style="list-style-type: none"> <li>● <b>Leading:</b> evidence from participant self-reporting (conversations, observations, survey)</li> <li>● <b>Leading:</b> formative assessment during professional learning</li> <li>● <b>Lagging:</b> measures like survey designed to assess participant response to types of learning (conceptual, procedural, metacognitive and factual)</li> </ul>
Barrier reduction	Omitted	<ul style="list-style-type: none"> <li>● <b>Leading:</b> evidence from participant self-reporting before, during, and after professional learning (conversations, observations, survey)</li> <li>● <b>Leading:</b> formative assessment during professional learning</li> <li>● <b>Leading:</b> leader self-assessments of likely barriers given current structure</li> <li>● <b>Leading:</b> observation of participants using and not using new learning</li> <li>● <b>Lagging:</b> survey or focus group about barrier reduction during and</li> </ul>



		after implementation windows
Participant use	<b>Theory of instruction:</b> teachers use their new knowledge and skills, attitudes, and beliefs to improve the content of their instruction/ pedagogy	<ul style="list-style-type: none"> <li>● <b>Leading:</b> evidence from participant self-reporting before, during, and after professional learning (conversations, observations, survey)</li> <li>● <b>Leading:</b> observations of performance tasks during professional learning</li> <li>● <b>Lagging:</b> observation of use in the classroom or school</li> <li>● <b>Leading:</b> student feedback before and during use or non-use</li> </ul>
Improves student learning or engagement	<b>Theory of instruction:</b> the change in teacher practice fosters student learning	<ul style="list-style-type: none"> <li>● <b>Leading:</b> student feedback before, during, and after use</li> <li>● <b>Lagging:</b> teacher self-reporting of impacts based on observations and conversations</li> <li>● <b>Lagging:</b> measures of change, including things like classroom products, achievement measures, engagement measures and non-academic measures (like attendance)</li> </ul>

Having a mix of leading and lagging measures for each stage is essential, because it ensures planners know what is working and what is not, and allows for time to respond. If teachers do not understand the new instructional approach or believe it no better than the one they are already using, there is no point in proceeding, because the professional learning will not change student results. Professional developers need to plan to check early and often at each leading stage, rather than assuming change in lagging measures will occur if participants report they enjoyed the learning.

Reeves (2010) suggests the plan needs to be in place from the start, have clear checklists and other tools to allow everyone to conceptualize and observe for the new behaviors, and

include lots of formative feedback. His suggestions articulate some of the common gaps that occur after the professional learning event, and school divisions could use questions based that spring from Reeves (2010) to clarify their thinking:

1. What, exactly, does it look like when this new instruction is in place?
2. Who will observe for the new instruction and how will we know what the observers find?
3. How will teachers get and give feedback about the new instruction?
4. What additional supports will be in place if someone does not succeed with the new instruction initially? How will we know if our support is sufficient?
5. How will we gather feedback about the impacts of the instruction from stakeholders like teachers, leaders, student, and parents? What will we do with the feedback once we have it?

Understanding each of the logical links is one of the most significant changes that could be made to better support assessment and evaluation of professional learning. Jumping from “teachers like the PD event” to “student learning improved” seems questionable when stated overtly, but was described as common by the professional developers in the study. Better understanding the mechanisms by which instruction changes is the critical first step before any of the other helpful improvements, like making clearer goals, having a systematic plan, or observing more often.

### **Developing strong goals**

Goals for educational improvements in Saskatchewan are set without thinking specifically and deeply about the associated professional learning, much as they are in Canada (Broad & Evan, 2006). Saskatchewan stakeholders might be hoping, for example, that Indigenous students achieve the same number of credits as non-Indigenous high school students without describing what specific changes would need to occur at the classroom level and how those changes would

likely be accomplished (<http://publications.gov.sk.ca/documents/11/85636-Education%20Strategic%20Sector%20Plan%20Matrix%202014-2020.pdf> ). Since evaluation of professional learning from within divisions is focused almost entirely on participant reaction to professional learning (Guskey, 2000; Ingvarson et al., 2005; Broad & Evans, 2006; Desimone, 2009; Desimone 2011; Hirsh, 2013; Mark, 2014), there is no easy mechanism for describing how professional learning might even connect to student results, let alone be the main mechanism for changing classroom practice. When any goal for improvement in education is set, senior leaders need to think about how they anticipate the change might occur and why (Hirsh, 2013). If any part of the change is at the classroom level (and many things are), specific goals need to be articulated. The goals must reflect observational data that confirms the need for change is related to teacher practice, and needs to explicitly describe the current and future state. Examples of sector strategic planning (<http://publications.gov.sk.ca/documents/11/85636-Education%20Strategic%20Sector%20Plan%20Matrix%202014-2020.pdf>) articulated current and future states that contain little to no information about teacher practice but assume its transformation as a condition of improvement. Evidence of student growth will need to be more than just attendance and graduation rates; school divisions must hold themselves accountable for deeper learning and engagement for all students (Fullan, Rincón-Gallardo, & Hargreaves, 2015), measured by tools like *Our School* and connected specifically to professional learning.

The status quo of senior leaders determining goals independent of professional learning is unlikely to result in changes to classroom practice. The professional developers in the study described independently setting goals for professional learning related to a general topic like improved reading picked by senior leaders, then choosing by themselves if they were interested in monitoring the results. The study revealed that most of the professional developers did not

believe the senior decision makers were actively monitoring the impact of professional learning or understood how to do so. Framed positively, they described the lack of involvement as trust, but many also described it as a lack of awareness. If schools are genuinely interested in forward progress, they need to be actively invested in professional learning. The investment needs to reflect the many hours required for instructional change, and rapid switch from topic to topic that research decries.

The many professional developers will also need to create much more specific goals to evaluate based on Desimone's (2009) theory of change or use Reeve's (2010) assessment process. Clarity of goal is a precursor to conducting an effective assessment (Ingvarson et al., 2005) and the majority of professional developers in the study described many goals or a general idea in place of a measurable and measured goal. Part of the professional learning about assessment must be focused on setting goals for professional learning that are achievable and connect to teacher classroom practice and student results.

### **Need for a systematic approach**

It is difficult to distinguish the impact of professional learning "from other factors that promote school and individual improvement" (Hirsh, 2013, p. 12) so a systemic approach will provide the most helpful information about the impact of professional learning. Guskey states (2000) that a perfect causal link is not necessary, and likelihood is sufficient for many divisions. However, a reasonable likelihood is rarely established with current context for two key reasons. The first is that professional developers in the study were rarely systemic in their data collection and did not plan for detailed data collection regarding each goal of the professional learning. Second, they mostly collected data focus on teacher appreciation, understanding, and barriers. Information about teacher use in the classroom and student impacts

were unavailable, as they were for Dean, Terry, and Michelle because they did not work in school divisions, or difficult to connect to the professional learning, as they were for the majority of other professional developers in the study. Even the professional developers who described trying to connect to students' data and teacher observation, like Tracey, Juanita, Sherry, Maxwell, Pat, or Gwyneth expressed concern about trying to establish what had caused any changes they observed, and noted that access was still an issue. The professional developers did not describe infrastructure or resources designed to use the data they possessed (except for some participants who mentioned data teams) and they articulated an absence of clear guidance around how to use data-based decisions in their professional learning (Hirsh, 2013; Taylor et al., 2015). Even if the professional developers knew which programs might be worth offering to impact learning, most did not describe the capacity to gather and use data about teacher practice, which is essential to seeing a relationship between professional learning and student results (Hirsh, 2013; Sherman, Tibbetts, & Condelli, 1997).

A systematic approach to planning for, collecting, and analyzing data is an essential addition to the self-assessment process described by the professional developers, but the current practice saw professional developers often skipped over analysis processes in favor of action. In the last decade and a half, many more studies about how school divisions use data are available, thanks to the current interest in data-based decision-making (Jimmerson & Wayman, 2015; Mandinach et al., 2008; Spillane, 2012; Wayman & Jimerson, 2014; Young, 2006). However, only a few of these studies are directly connected to professional learning, once again highlighting the gap in the logic of how teaching might improve. School divisions would be well served to use many of the professional resources designed to help schools engaging in data-based

decision-making by extrapolating related concepts to professional learning planning, assessment, and delivery.

### **Responding to issues of insider evaluation**

Since the professional developers in the study engaged in self-assessment independent of any specific direction from more senior leaders and formal evaluation was not referenced, there were few attempts to address potential concerns with the fact that insiders were the people determining the impacts. The traditions of professional development and training did not expect change; decades of accepted practice assumed professional developers would lead something episodic with no actual expectation that the training would even impact a teacher's classroom (Reeves, 2010; Stein, Smith, & Silver, 1999). Evaluating the impact of professional learning was highly improbable. However, new attempts to engage in more effective processes are still at the experimental stages, making in an ideal time to interest practitioners in considering efficacy of the new methods. Ingvarson, Meiers, and Beavis (2005) note a lack of clarity about what facilitation methods are being used and how they are being assessed is a significant barrier for professional developers. In this study, many professional developers were very specific about strategies, but vague about goals and change theory of professional learning, indicating professional developers could not design fulsome evaluations without learning more about the connection between forms of professional learning, purpose, and ways to measure. Given that the desire to know impact and the process of learning to assess it are new, providing professional learning on the subject to anyone working as a professional developer is a smart choice, and may improve the quality of the insider assessments.

Professional developers also raised personal and organizational assumptions often when they were referencing organizational structural realities. Their personal assumptions developed

over years of praxis (Owens & Rogers, 1999) and were well established, but those beliefs were also sometimes challenged by the process of wondering. It is unlikely the professional developers would decide major changes were needed as insiders with common assumptions, but they were interested in questioning. Without an external framework for an assessment process, the prospect of change would be limited to professional developers and systems with both professional agency and a focus on genuine change, something school divisions are not typically noted for. Having a common structure designed to help with data gathering, analysis, and interpretation could help address natural system-confirming biases, and provide an outlet for using the professional developers' inquiring minds.

### **Increasing observation**

Few of the professional developers described getting the opportunity to observe, and those who did were the most likely to assess their own professional development for teacher use and student impacts. All the professional developers articulated some version of the idea that observing teachers engaged in the new behaviors was the best evidence they had that the professional learning succeeded in helping teachers adopt the new practice. Lawless and Pellegrino (2007) reinforced the value of a unified professional development evaluation plan coupled with direct observation, and other researchers underscored the unique value of observations (Broad and Evans, 2006; Knight, 2014; Stein, Smith, and Silver, 1999; Reeves, 2010). Reeves described observation as essential because it examines one classroom and student at a time, rather than an aggregate, providing a more direct line to the decisions of teacher and leaders (2010). Desimone agreed, noting observations “enable coaches and principals to determine whether teachers are using reform practices perfunctorily or effectively” (2011, p. 71). Observation offers the best easy evidence of degree of success in a form that professional

developers are used to, given their experiences as teachers. However, many participants expressed concerns about teachers feeling judged or creating inappropriate supervision-structures. The literature on coaching can provide guidance about potential ways to structure observations so they are informative without crossing the line into supervision, and participants like Gwyneth and Sherry provided examples from their school-based experiences that can be helpful in illustrating potentially useful approaches.

Observation may be the most desirable way to assess if professional learning is being used in the classroom, but doing more observation is not an easy change. Desimone's statement that "observation is burdensome and expensive" (2009, p. 190) articulated a potential concern, especially given recent provincial reductions in resources, however, models where observation is conducted in schools within the school day or as a part of professional learning can reduce the expense. Although they did not focus on expense, professional developers in the study articulated that access and culture were significant, and research confirmed these barriers can be pernicious (Cole, 2004; Fullan, 2007). While they are significant, access and culture are within the control of school divisions, and could change if divisions believed that observation was essential to knowing if changes in student results were likely given the type of professional learning offered.

## **Conclusion**

The resulting theory and recommendations described here offer an opportunity to develop a routine method of assessing professional learning. As a part of the common work in school divisions, assessment and evaluation process could influence the quality of professional learning, participant satisfaction, and ultimately, how we teach and the quality of learning.



## CHAPTER SIX

### Theory summary

This chapter is a brief summary of the conditions and theory described in the fifth chapter. It is a succinct overview of the theory resulting from the research questions followed by reflection and recommendations for further research.

The main research question of this study was, “How do professional developers describe and demonstrate the process of self-assessment?” In response, participants described largely self-initiated processes focused predominantly on improving personal leadership of professional learning. They used facilitation processes, talking to others, and personal observations as their main forms of data collection, and focused on taking action based on their interpretations.

The sub questions of the study were:

1. How do professional developers decide which strategies to use to determine the impact of their professional learning?
2. What influences the perpetuation of evaluation methods (or lack thereof)?
3. How is the praxis of professional learning influenced by the professional developer’s experience?

The study found that the professional developers rejected the word “evaluation”, and a minority referenced formal evaluation processes. They relied on experiences and conversations with other leaders to shape their methods and their conclusions. At any stage of the self-assessment process, professional developers stopped self-assessing if they believed they had the information they were looking for or could not impact issues they were discovering. If they chose to persist

to the taking action stage, they often tackled multiple, complex actions to respond to what they found.

### **Conditions that influenced self-assessment or evaluation processes**

The study participants in general described two types of conditions, personal and systemic, that influenced the likelihood and composition of self-assessments. The personal conditions shaped what participants decided to assess and why, and increased the likelihood of assessment, since most assessment was inspired by decisions of the professional developer or leader. Of the systemic conditions, everything except theories of professional learning decreased the frequency or depth of self-assessment. A desire to lead high quality professional learning increased it.

#### ***Highlights:***

- The professional developers in the study routinely engaged in a process of self-assessment that had commonalities with, but was not the same as, program evaluation.
- Like most evaluation of professional learning, the participant initiated self-assessments focused on if teachers or leaders provided positive feedback about the professional learning. Formative assessments were also used to check if the professional learning they led was changing teacher understanding or beliefs, and to check for potential barriers. These practices occurred predominantly in face-to-face professional learning, and were decreased if professional learning was distributed or online.
- The professional developers' attempts to understand their impacts were commonly self-initiated, although one organization studied used a common template for planning and assessment that was reflected in the work of a number of participants.

## Process for self-assessment

### *Highlights:*

- Participants agreed most professional learning in school divisions occurs without assessment.
- No participant described a full, formal process for evaluation required by an employer, and the term “evaluation” was frequently replaced with the word “assessment” by almost all participants, indicating the concept of evaluation was problematic in some way.
- When assessment did occur, it was typically initiated by the participant, usually with the purpose of improving facilitation skills, rather than ensuring professional learning impacted student results.

## Theory

There were four key stages that participants used to self-assess. They did not always go through all the stages, but when they did, the process occurred in the sequence of wondering, finding out, attributing, and taking action:

1. ***Wondering:*** A wonder was an inquiry, passing thought, learning fissure, or general question. The wonder could spark further investigation or could be dismissed. A simple explanation or feeling that the question was not important ended any process of self-assessment. Participants reported many wonders that led nowhere and a few that resulted in self-assessment.
2. ***Finding out:*** Finding out was about gathering data. Sometimes a formal plan for data collection occurred, but most methods of gathering information were rapid and

prototyped, designed to give easy information that could be acted on. Conversations, even incidental ones, were the most common form of data collection and could occur with other leaders, participants, and non-participants. The next most common method was a facilitation strategy designed to have teachers demonstrate or share their learning, analyze data, or reveal their beliefs. Formal data collection, like surveys, feedback forms, and sharing circles were also used, but were less common.

3. ***Attribution and sensemaking:*** Attribution and sensemaking combines as the process of interpreting data, and they were nearly instantaneous as described by study participants. Participants quickly considered information and its potential meaning. There was little description of coding processes, themes, alternative hypothesis, or other elements typically found in research. Sometimes triangulation was used to provide nuance, and the experiences and knowledge of the professional developers was always used to provide additional factors and considerations. Participants paid particular attention to potential evidence of barriers and the implications of the barriers on teacher actions.
4. ***Action:*** Problems were sometimes dismissed as unimportant, passing, or not within the control of the participants, but the participants almost always found something they discovered that they could act on if they reached the action stage. Actions were much more complex than described in research and went beyond changing facilitation. They included changing structures and systems, advocating, addressing individual barriers, and restarting the wondering process based on new questions. More than one action was typically initiated based on interpretation of data, making it hard to see what action “made the difference” but more likely that problems would be addressed.

## **Further research**

### **Implications of processes like coaching on self-assessment and evaluation**

One of the limits of the sample in this study was the substantial reduction in data from coaches when many were returned to the classroom. Since professional learning theory indicates that coaching plays a substantial role in helping teachers use what they are learning (Knight, 2007; 2014) it is very possible data around observation and assessment of use would have been different had there been a more substantial pool of instructional coaches to draw from. The implications theory is unlikely to have been substantially affected, but related implications and recommendations likely would be.

More research related to coaches and assessment or evaluation processes would also be helpful because of the relationship between the professional developers' concerns about teacher autonomy and professional discretion, and their concerns that teachers may be judged or feel judged as a part of assessment and evaluation processes. One of the challenges of instructional coaching (Knight 2014) is that coaches need to preserve trust and agency in order for coaching to succeed. Understanding more about processes coaches use to gather, analyze, and make judgements about data without rendering judgements on teachers has the potential for adding nuance to methods for future assessments of professional learning.

### **Study of “accuracy” of self-assessments**

This study was not designed to render judgement on the quality of professional learning or the quality of self-assessment that the professional developers engaged in. Its purpose was to

articulate the current process as completely as possible in order to help understand how self-assessment is currently conducted by professional developers and what factors impact the process. However, subsequent studies could further refine or more fully articulate the process of wondering, finding out, attributing, and acting if there was more information about how well those activities were completed. Like the potential of rendering judgement on teachers, the potential of rendering judgement on professional developers presents problems.

The process of rendering judgement also requires standards to assess against, and neither research standards nor program evaluation standards are wholly appropriate to the types of assessment process (Reeves, 2010) that would be practical and informative in the K-12 education sector. Since there are no largely accepted standards, a synthesis of possible options expressed by Guskey, Reeves, Desimone, and others could help provide a template and basis for judgement. If standards equivalent to the *Learning Forward Standards* for professional learning (2011) were commonly circulated, it is likely professional developers and their relative employers would be better positioned to conduct self-assessment.

## **Reflections**

### **Timing**

The timing of the study was problematic as it occurred in two years of substantial reductions in the general operating grants for school divisions in the province. When the research began, many potential participants were available in a wide variety of stakeholder groups. In areas such as the Ministry of Education, gradual reductions in roles including curricular consultants occurred over time, although the reduction in professional learning offering was relatively recent. In school divisions, the start of the study saw professional developers in roles of administrator, superintendent, coordinator, consultant, and coach. The

first year data collection was accelerated as school divisions reduced their central office staffs, and ethics approval was completed early to start interviewing as the proposal was with committee.

It was the second year, once ethics were reviewed, that was problematic. Some school divisions laid off all coaches or consultants, and nearly one quarter of active participants were returned to classroom, changed jobs, or retired. As a result, elements of theory did not reach full saturation in some instances because the theoretical sample was not as extensive as it might have been to challenge categories. For example, the study might have included more coaches, to see if their opportunity to observe changed theory around causal factors or attribution and data collection process. However, with dramatic province-wide reductions in the number of coaches and much increased workload on remaining central office staff, recruitment of additional coach participants was not possible. Additional data collection specific to the coach role should be a focus of additional study of professional developers' self-assessment processes.

### **Change to personal theory**

I began this process thinking evaluation of professional learning was rare, and believing it needed to be more common. While completing the study, I have concluded that self-assessment of professional learning is common, and that many processes used for that self-assessment are elaborate. The changes in professional learning in my time in education have done more than move teacher learning from "sit and get" workshops to increasingly embedded and extended process. Combined with learning about formative assessment, they have created circumstances where many professional developers use facilitation to gather information to respond to and for self-improvement. Because professional developers are also thinking about barriers, they are

planning professional learning that is much more responsive than I, or the research, might have imagined.

## **Conclusion**

Professional learning continues to be a main mechanism school divisions and other stakeholders use to forward educational agendas and transform student results. Employees with the responsibility for teaching their peers, including administrators, teacher leaders, and central office staff try to offer professional learning that is effective and appreciated. These professional developers plan with, and respond to, self-generated information about teacher preference, barriers to use, and understanding regularly. They are missing systemic assessment processes that routinely support effective decision-making, particularly regular gathering of information about teacher use of professional learning and the relationship between professional learning and changes to student results. Juanita noted the challenges given the number of factors that impact student learning and described wishing that if she could “change one thing,” all professional learning would “be assessed to determine its impact immediately and then later after some time has passed.” While this laudable goal is not current practice, the self-assessment processes professional developers described indicate the next steps. Understanding the way professional developers think about, and assess, the professional learning they lead provides a framework for doing more than calling for increases in effective professional learning and evaluation of professional learning; the process of self-assessment provides a framework for growth, and helpful details that promote assessment processes.



## References

- Adler, P. A., & Adler, P. (1994). Observational techniques. *Handbook of Qualitative Research*, 1(3), 377-392.
- Annells, M. (1996). Grounded theory method: Philosophical perspectives, paradigm of inquiry, and postmodernism. *Qualitative Health Research*, 6, 379-393.
- Annells, M. (1997). Grounded theory method, part I: Within the five moments of qualitative research. *Nursing Inquiry*, 4(2), 120-129.
- Bertrand, M., & Marsh, J. A. (2015). Teachers' sensemaking of data and implications for equity. *American Educational Research Journal*, 0002831215599251, 861-893.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Blank, R. K. (2013). What research tells us: Common characteristics of professional learning that leads to student achievement. *Journal of Staff Development*, 34(1), 50-53.
- Blitz, C. L., & Schulman, R. (2016). *Measurement instruments for assessing the performance of professional learning communities*. Regional Education Library. Retrieved from [https://ies.ed.gov/ncee/edlabs/regions/midatlantic/pdf/REL\\_2016144.pdf](https://ies.ed.gov/ncee/edlabs/regions/midatlantic/pdf/REL_2016144.pdf)
- Birman, B. F., Desimone, L., Porter, A. C., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership*, 57(8), 28-33.
- Bonner, A., & Tolhurst, G. (2002). Insider outsider perspectives of participant observation. *Nurse Researcher*, 9(4), 7-19.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3-15.

- Broad, K., & Evans, M. (2006). *A review of literature on professional development content and delivery modes for experienced teachers*. Toronto, CAN: University of Toronto, Ontario Institute for Studies in Education.
- Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, 8(1), 137-152.
- Charmaz, K. (1995a). Between positivism and postmodernism: Implications for methods. *Studies in Symbolic Interaction*, 17, 43-72.
- Charmaz, K. (1995b). Grounded theory. In J. Smith, R. Harré, & L. Langenhove (Eds.), *Rethinking methods in psychology* (pp. 27-65). London, UK: Sage.
- Charmaz, K. (2014). *Constructing grounded theory* (2<sup>nd</sup> ed.). London, UK: Sage.
- Charmaz, K., & Mitchell, R. G. (1996). The myth of silent authorship: Self, substance, and style in ethnographic writing. *Symbolic Interaction*, 19(4), 285-302.
- Chase, S. (2005). Narrative inquiry: Multiple lenses, approaches, and voices. In Denzin, N. & Lincoln, S. (Eds.), *The Sage handbook of qualitative research* (3rd ed.), 651-679. Thousand Oaks, CA: Sage.
- Cole, P. (2004). *Professional development: A great way to avoid change* (Vol. 140). Melbourne, AUS: Incorporated Association of Registered Teachers of Victoria.
- Creswell, John W. (1998). *Qualitative inquiry and research design*. Thousand Oaks, CA: Sage.
- Czarniawska, B. (2004). *Narratives in social science research*. London, UK: Sage.
- Cuban, L. (1993). *How teachers taught: Constancy and change in American classrooms, 1890-1990*. New York, NY: Teachers College Press.
- Darling-Hammond, L., & Richardson, N. (2009). Research review/teacher learning: What matters. *Educational Leadership*, 66(5), 46-53.

- Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession*. Washington, DC: National Staff Development Council.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, 24(2), 81-112.
- Desimone, L. M. (2011). A primer on effective professional development. *Phi Delta Kappan*, 92(6), 68-71.
- Draucker, C. B., Martsolf, D. S., Ross, R., & Rusk, T. B. (2007). Theoretical sampling and category development in grounded theory. *Qualitative Health Research*, 17(8), 1137-1148.
- Dufour, R., Dufour, R., Eaker, R., & Many, T. (2010). *Learning by doing: A handbook for professional learning communities at work*. Bloomington, IN: Solution Tree Press.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256-273.
- Earl, L., & Timperley, H. (2015). *Evaluative thinking for successful educational innovation*. Paris, France: Organization for Economic Co-operation and Development Publishing.
- Elmore, R. F., & Burney, D. (1997). *Investing in teacher learning: Staff development and instructional improvement*. Community School District 2, New York City. New York,

NY: National Commission on Teaching and America's Future and Consortium for Policy Research in Education.

Fournier, D. M. (2005). Logic of evaluation—Working logic. *Encyclopedia of Evaluation*, 238-242.

Fullan, M. (2007). Change the terms for teacher learning. *Journal of Staff Development*, 28(3), 35-36.

Fullan, M., Rincón-Gallardo, S., & Hargreaves, A. (2015). Professional capital as accountability. *Education Policy Analysis Archives*, 23(15), 1-22.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945.

Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs. forcing*. Mill Valley, CA: Sociology Press.

Glaser, B. G. (2001). *The grounded theory perspective: Conceptualization contrasted with description*. Mill Valley, CA: Sociology Press.

Glaser, B. G., & Holton, J. (2004, May). Remodeling grounded theory. *Forum Qualitative Sozialforschung/ Forum: Qualitative Social Research*, 5(2). Retrieved from

<http://www.qualitative-research.net/index.php/fqs/article/view/607/1315>

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.

Guskey, T. (2009). Closing the knowledge gap on effective professional development. *Educational Horizons*, 87(4), 224-233.

Hume, K. (2009). *Evidence to action*. Toronto: Pearson.

- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.
- Hawley, W. D., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice*. San Francisco: Jossey-Bass, 127–150.
- Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research consensus, crossroads, and challenges. *Educational Researcher*, 42(9), 476-487.
- Hirsh, S. (2013). The impact factor. *Journal of Staff Development*, 34(5), 10-12, 14, 16.
- Hirsh, S. and Killion, J. (2009, March). When educators learn, students learn: Eight principles of professional learning. *The Phi Delta Kappan*. 90(7), 464-469.
- Honig, M. I., Copland, M. A., Rainey, L., Lorton, J. A., & Newton, M. (2010). *Central office transformation for district-wide teaching and learning improvement*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.
- Hood, J. (2007). Orthodoxy versus power: The defining traits of grounded theory. In A. Bryant & K. Charmaz (eds.). *Handbook of Grounded Theory*. London, UK: Sage, 2-15
- Hume, K. (2009). *Evidence to action*. Toronto: Pearson.
- Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives*. 13(10), 1-26.
- Jimerson, J. B., & Wayman, J. C. (2015). Professional learning for using data: Examining teacher needs and supports. *Teachers College Record*, 117(4), 1-34.
- Joyce, B. & Showers, B. (2002). *Student achievement through staff development* (3<sup>rd</sup> ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945-980.
- Kirkpatrick, D. L. (1959). Techniques for evaluation training programs. *Journal of the American Society of Training Directors*, 13, 21-26.
- Knight, J. (2007). *Instructional coaching: A partnership approach to improving instruction*. Thousand Oaks, CA: Corwin Press.
- Knight, J. (2014). *Focus on teaching: Using video for high-impact instruction*. Thousand Oaks, CA: Corwin Press.
- Lawless, K. A., & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of Educational Research*, 77(4), 575-614.
- Learning Forward. (2011). *Standards for professional learning*. Oxford, OH: Author.
- Leiblich, A. (1998). *Narrative research: Reading, analysis and interpretation*. Thousand Oaks, CA: Sage Publications.
- Levin, B. (2011). Mobilising research knowledge in education. *London Review of Education*, 9(1), 15-26.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry* (Vol. 75). London, UK: Sage Publications.
- Mandinach, E. B., Honey, M., Light, D., & Brunner, C. (2008). A conceptual framework for data-driven decision-making. *Data-driven school improvement: Linking data and learning*. New York, NY: Teachers College Press, 13-31.
- Mann, H. (1865). *Horace Mann on the crisis in education*. Yellow Springs, Ohio: Antioch Press.

- Mark, C. (2014). Evaluating and funding the professional development program. *Adult and Continuing Education: Concepts, Methodologies, Tools, and Applications*. USA: IDI Global, 1685-1705.
- Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision-making in education*. Santa Monica, CA: Rand.
- Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. Forum: *Qualitative Social Research*, 11(3). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/1428/3027>
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco, CA: Jossey-Bass.
- Mills, J., Bonner, A., & Francis, K. (2006). Adopting a constructivist approach to grounded theory: Implications for research design. *International Journal of Nursing Practice*, 12(1), 8-13.
- Morse, J. M. (2011). Molding qualitative health research. *Qualitative Health Research*, 21(8), 1019-1021.
- Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91(3), 328–346.
- Owen, J. M., & Rogers, P. J. (1999). *Program evaluation: Forms and approaches*. Thousand Oaks, CA: Sage.

- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Patton, M. Q. (2013). *Utilization-focused evaluation checklist*. Retrieved from [http://www.wmich.edu/sites/default/files/attachments/u350/2014/UFE\\_checklist\\_2013.pdf](http://www.wmich.edu/sites/default/files/attachments/u350/2014/UFE_checklist_2013.pdf)
- Peery, A. (2011). *The data teams experience: A guide for effective meetings*. Englewood, CO: Lead + Learn Press.
- Polkinghorne, D. E. (1995). Narrative configuration in qualitative analysis. *International journal of Qualitative Studies in Education*, 8(1), 5-23.
- Polkinghorne, D. E. (2007). Validity issues in narrative research. *Qualitative Inquiry*, 13(4), 471-486.
- Reeves, D. B. (2010). *Transforming professional development into student results*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Schwille, J., Dembélé, M., & Schubert, J. (2007). *Global perspectives on teacher learning: Improving policy and practice*. Paris, France: International Institute for Educational Planning, UNESCO.
- Seifert, T. (2004). Understanding student motivation. *Educational Research*, 46(2), 137-149.
- Sherman, R., Tibbetts, J., & Condelli, L. (1997). *Evaluating professional development: A framework for adult education*. Washington, DC: Pelavin Research Institute.
- Smith, E. R., & Tyler, R. W. (1942). *Appraising and recording student progress*. New York: Harper & Brothers.



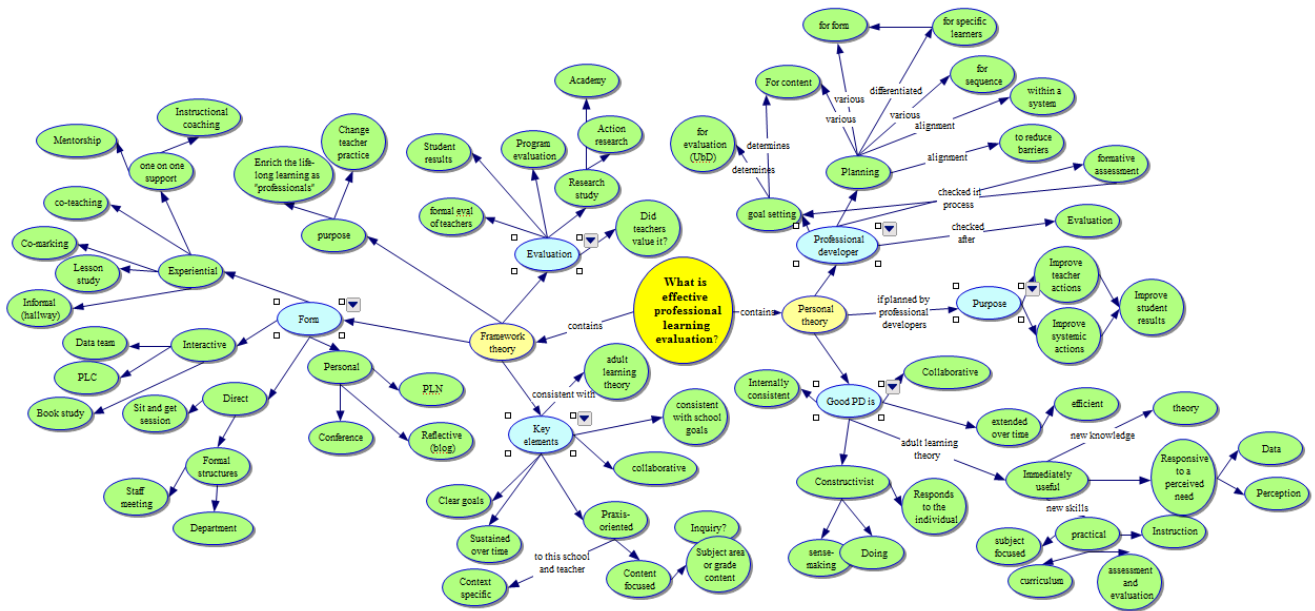
- Stufflebeam, D. L. (1999). *Program evaluations metaevaluation checklist*. Retrieved from [https://wmich.edu/sites/default/files/attachments/u350/2014/program\\_metaeval\\_short.pdf](https://wmich.edu/sites/default/files/attachments/u350/2014/program_metaeval_short.pdf)
- Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education*, 118(2), 113–141.
- Steeves, L., Berryman, M., Carr-Stewart, S., Kovach, M., Laliberté, C., Sharon Meyer, S., Merasty, B. Sloboda, A. & Stelmach, B. (2014). *Seeking their voices: Improving Indigenous student learning outcomes*. Regina, SK: Saskatchewan Instructional Development and Research Unit.
- Stein, M. K., Smith, M. S., & Silver, E. A. (1999). The development of professional developers: Learning to assist teachers in new settings in new ways. *Harvard Educational Review* 69(3), 237-269.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. Denzin & Y. Lincoln (Eds.) *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage, 273-285.
- Sudsawad, P. (2007). *Knowledge translation: Introduction to models, strategies, and measures*. Austin, TX: Southwest Educational Development Laboratory, National Center for the Dissemination of Disability Research.
- Taylor, J. A., Getty, S. R., Kowalski, S. M., Wilson, C. D., Carlson, J., & Van Scotter, P. (2015). An efficacy trial of research-based curriculum materials with curriculum-based professional development. *American Educational Research Journal*, 52(5), 984-1017.

- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2008). *Teacher professional learning and development*. Perth, Western Australia: International Academy of Education.
- Trochim, W. (2006). *Introduction to evaluation*. Retrieved from [www.socialresearchmethods.net/kb/intreval.php](http://www.socialresearchmethods.net/kb/intreval.php).
- Tunison, S. (2016). *Cultivating knowledge*. Rotterdam: Sense Publishers.
- van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. London, CAN: Althouse.
- Villegas-Reimers, E. (2003). *Teacher professional development: An international review of the literature*. Paris: International Institute for Educational Planning.
- Wayman, J. C., & Jimerson, J. B. (2014). Teacher needs for data-related professional learning. *Studies in Educational Evaluation*, 42, 25-34.
- Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garec, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, 37(8), 469-479.
- Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist*, 45(1), 28–36.
- Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 426-431.
- Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research*, 11(4), 522-537.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design (2<sup>nd</sup> Ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development.
- William, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.

Young, V. M. (2006) Teachers' use of data: Loose coupling, agenda setting, and team norms.  
*American Journal of Education*, 112, 521–548.

## APPENDIX A:

### MAIN THEMES OF RESEARCH IN PROFESSIONAL LEARNING EVALUATION



# APPENDIX B:

## MIND MAP OF POTENTIAL METHODOLOGIES



**APPENDIX C:**  
**DESCRIPTION OF STUDY**

My name is Wendy James, and I am doctoral candidate at the University of Saskatchewan. I am conducting a study, called *Know Your Impact*, to try to understand how educators who lead professional learning self-assess the impact of their own work.

The study is about the professional developer's thought process and does not attempt to assess the impact or quality of professional learning in any division or organization.

I am hoping to learn more about how educators self-evaluate the professional development they facilitate through two interviews. That first interview will take about an hour and is designed to gather basic information about how participants think about the professional learning they lead.

The second part of the study is me attending a part of a professional learning activity the participant is leading. I will record a small part that the participant selects so she or he can watch later. Then we'll meet again, and the participant can talk me through what his or her thought process was during the section that was recorded. The video recording will just occur when the participant would already be leading professional learning, but the discussion will require another hour of the participant's time.

I would like your permission to interview one or more people in your division, and to record a part of a professional learning process. I would appreciate it if you could give my name, contact information, and this description to some potential participants.

**APPENDIX D:**  
**INTERVIEW AND TALK ALOUD QUESTIONS**

**Interview 1**

*Main Questions – Evaluation of Professional Learning*

- What do you hope to influence as a professional developer?
- Describe how you know if the professional learning you lead is making a difference in teacher practice or in student learning.
- If you could change something about how you evaluate your impact right now, what would it be?
- Who is interested in knowing the impact of the professional learning you lead? How do you share the impact?
- Professional Development
  - How do teachers like to learn?
  - What are some common structures of the professional development you lead?
  - What prevents you from doing what you would like to do in your professional development practice?
  - What do you do during a facilitation to help you figure out what your learners need next?
  - Of all the information you might have about the impact of the professional learning you facilitate, what type is most common? Most useful?
  - What type of information do you wish you had, but you don't? What prevents your access to that information?

*Ask if it doesn't come up – context in interview 1:*

- Tell me about your career before your current job.
- If you have a typical day, what does it look like? If not, why are your days varied?
- Why have you chosen to work as a professional developer?
- Please describe a typical professional development event or experience that you lead or supported.

## **Interview 2: Talk Aloud Protocol**

Prior to interview two, I will have the opportunity to observe the participant leading a professional development experience or planning a process. I will keep field notes and copies of any artifacts, and will create a video recording focused on the participant. The participant will have the opportunity to watch the tape prior to interview two to provide opportunity for reflection and make it more likely that the focus can shift beyond reaction to seeing oneself on video (Knight, 2014). During the second interview, I will ask open-ended probes designed to encourage the participant to think aloud about the professional learning and his/or her thinking as the facilitator or leader. As such, the questions will need to be broad and contextual, and I will need to vary them depending on the observation and resulting video recording. When probing to encourage a talk aloud, the interviewer asks open questions that are situation dependent.

However, probes often include questions like the following:

- Why did you say that?
- What were you thinking when you...?
- I noticed that you paused/re-directed there. Can you tell me why?
- I'd like to watch that part again. Can you explain what is happening?



**APPENDIX E:**  
**PARTICIPANT CONSENT FORM**

Project Title: Know Your Impact  
Researcher: Wendy James, Graduate Student

College of Education  
University of Saskatchewan  
306-244-1474  
[lwj118@mail.usask.ca](mailto:lwj118@mail.usask.ca)

Supervisor:  
Dr. Jay Wilson, Professor  
College of Education  
University of Saskatchewan  
306-966-7617  
[jay.wilson@usask.ca](mailto:jay.wilson@usask.ca)

**Purpose(s) and Objective(s) of the Research:**

The goal of the research is to understand how professional developers self-evaluate the professional development they lead.

**Procedures:**

- We will meet three times formally.
- The first meeting will be an opening interview and I will create a written transcript of our discussion for you to consider before our second interview. I will also take field notes and take pictures.
- The second time, I will watch you plan or deliver a professional learning session and take more notes. We will record video of a short section of it and then I will give it to you to watch.
- Once you have done that, we'll meet again and I will ask you to talk me through why you made the decisions you did in the recorded session. I would appreciate copies of any planning or evaluation documents you would like to share.
- Each time we meet to talk, it will take about an hour. The length of time I observe a professional development planning or evaluation activity will depend on your preferences, but it will be a fraction of the time it would usually take to do the activity. You will also need to watch the video of your facilitation or planning once by yourself.

- We will meet in a place that is convenient for you. You will be the participant in this study – others may be present in the professional development session or planning, but they will not be the focus of the data collection.
- Please feel free to ask any questions regarding the procedures and goals of the study or your role.

### **Potential Risks:**

- There are no known or anticipated risks to you by participating in this research and you will have the right to see and amend the collected data. You will make the decision about the level of confidentiality that is appropriate.
- You should only answer questions you are comfortable with.

### **Potential Benefits:**

You may learn more about your own process in thinking about professional learning through the opportunity to reflect.

### **Confidentiality:**

- Because the participants for this research project have been selected from a small group of people, many of whom are known to each other, it is possible that you may be identifiable to other people on the basis of what you have said. I will attempt to remove anything that would clearly identify you, and you may ask me to remove other statements you think might identify you.
- You will make the decision if you wish to be identified by your name or a pseudonym you select.
- The research project results and associated material will be safeguarded and securely stored by Jay Wilson at the University for a minimum of five years (if the study is published) post publication. When the data is no longer required, it will be appropriately destroyed.

### **Right to Withdraw:**

- Your participation is voluntary and you can answer only those questions that you are comfortable with. You may withdraw from the research project for any reason, at any time during the data collection or analysis without explanation or penalty of any sort.
- Should you wish to withdraw, you may contact me to let me know you would like to withdraw. I will destroy the data immediately if that occurs.

- After the point where I share the findings with my committee and others, it is possible that some form of research dissemination will have already occurred and it will not be possible to withdraw your data.

**Follow up:**

I will share a copy of the summary of the findings with you.

**Questions or Concerns:**

Contact the researcher(s) using the information at the top of page 1.

This research project has been approved on ethical grounds by the University of Saskatchewan Research Ethics Board. Any questions regarding your rights as a participant may be addressed to that committee through the Research Ethics Office [ethics.office@usask.ca](mailto:ethics.office@usask.ca) (306) 966-2975. Out of town participants may call toll free (888) 966-2975.

**Consent:**

There are several options for you to consider if you decide to take part in this research. You can choose all, some, or none of them. Please put a check mark on the corresponding lines that grants me your permission:

I grant permission to be audio taped: Yes: \_\_\_ No: \_\_\_

I grant permission to be videotaped: Yes: \_\_\_ No: \_\_\_

I wish to remain anonymous: Yes: \_\_\_ No: \_\_\_

I wish to remain anonymous, but you may refer to me by a pseudonym: Yes: \_\_\_ No: \_\_\_

The pseudonym I choose for myself is:

**Continued or On-going Consent:**

- I will ask you verbally at each additional meeting if you still consent.
- Your signature below indicates that you have read and understand the description provided; I have had an opportunity to ask questions and my/our questions have been

answered. I consent to participate in the research project. A copy of this Consent Form has been given to me for my records.

\_\_\_\_\_  
Name of Participant                      Signature                      Date

\_\_\_\_\_  
Researcher's Signature                      Date

A copy of this consent will be left with you, and a copy will be taken by the researcher.

