Fall 2022

Evaluation from Both Sides Now: Towards an Epistemology of Evaluation Practice

Heather D. Codd
Claremont Graduate University

Evaluation from Both Sides Now: Towards an Epistemology of Evaluation Practice

By

Heather D. Codd

Claremont Graduate University

2022

Approval of Dissertation Committee

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Heather D. Codd as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in Psychology.

Leslie A. Fierro, Ph.D., Chair

Claremont Graduate University

Senior Fellow

Stewart Donaldson, Ph.D.

Claremont Graduate University

Distinguished University Professor

Kendall Cotton Bronk, Ph.D.

Claremont Graduate University

Professor

Dr. Christina A. Christie, Ph.D.

University of California, Los Angeles

Wasserman Dean and Professor

Abstract

Evaluation from Both Sides Now: Towards an Epistemology of Evaluation Practice

By

Heather D. Codd

Claremont Graduate University: 2022

Throughout its history, the evaluation field has developed numerous theories. These theories, or evaluation theory as they are collectively known, are integral to the knowledge of the discipline and represent the field's collective understanding of how evaluation can and should be practiced. Yet, research suggests that the influence of evaluation theory on evaluation practice is minimal. This finding has left the field questioning what knowledge, if not evaluation theory, guides practitioners?

Some theorists propose that evaluation practice is influenced by practical knowledge, a diverse knowledge base inherent to the doing of evaluation. Practical knowledge is a blend of explicit procedural knowledge and tacit ways of knowing that inform practitioners' perceptions and guide their practice. It is developed through practitioner's interactions with and theorizing of practical problems as well as insights drawn from experience.

Despite its perceived importance to evaluation practice, limited research has been conducted on practical knowledge in evaluation. Consequently, the field lacks an understanding of the underlying knowledge, or epistemology, of evaluation practice. The current study addressed this research gap by investigating practical knowledge in evaluation with a specific focus on the ways of knowing that underlie and guide practitioners.

A multiple methods research design, consisting of two studies, was implemented. The first study used a convergent mixed methods design to analyze qualitative and quantitative data collected through document review (*N*=3) and the repertory grid technique (*N*=22). The findings presented a holistic epistemology of evaluation practice grounded in four inter-related ways of knowing: knowing self, knowing others, knowing the discipline, and knowing the common good and equity. Important insights into the underlying dimensions of the ways of knowing, including creative processes that actuate knowing (i.e., knowing in action) and methods evaluators use to express what has become known through their practice (i.e., expressions of knowing) were also interpreted from the findings.

The second study examined how the epistemology of evaluation practice identified in the first study is integrated into graduate programs in evaluation. Qualitative data were collected through semi-structured interviews with evaluation faculty (*N*=11) as well as alumni and advanced students (*N*=16) from six graduate programs in evaluation based in the United States. The findings suggest that the epistemology is reflected in the evaluation educational programs, albeit to varying degrees. Knowing the discipline was the most evident among the programs while knowing others and expressions of knowing were moderately evident. Knowing the common good and equity, knowing self, and knowing in action were less apparent, suggesting that these may be areas for further development.

This research advances the scholarship on evaluation practice and provides valuable insights into the relationship between evaluation theory and practice. Additionally, the findings present a holistic representation of the epistemology of evaluation practice that challenges existing assumptions about how the field conceptualizes practice and knowledge construction.

Last, the findings provide valuable information that can further professionalize evaluation by refining the field's understanding of the unique nature of evaluation practice and informing evaluator education, reflective practice, and professional development.

Dedication

I dedicate my dissertation to my husband, Paul, who is my greatest ally and a continuous source of encouragement and support. I am very lucky to have you in my life. I also dedicate this work and give special thanks to my parents, Melvin and Marilyn, who taught me to work hard for what I want and provided me with immense support during my studies and research. Last, I dedicate this work to the young minds in my life, some of whose academic journeys have coincided with mine and another who is just getting started. I look forward to seeing what great things you will accomplish.

Table of Contents

Evaluation from Both Sides Now: Towards an Epistemology of Evaluation Practice

Chapter 1: Introduction and Literature Review

The role of theory in evaluation is multi-faceted (Donaldson & Lipsey, 2006). Evaluators leverage social science theory to design evaluation studies and garner evidence (Chen & Rossi, 1980; Donaldson & Lipsey, 2006; Mertens & Wilson, 2019). Program theory helps evaluators and evaluation partners conceptualize evaluands, including their various components and causal mechanisms (Chen, 2015; Chen & Rossi, 1983; Donaldson & Lipsey, 2006). Last, evaluation theory educates the field on the goals of evaluation, roles of evaluators, and procedures involved in conducting evaluation studies (Alkin, 2013; Mertens & Wilson, 2019).

While the importance of theory is widely recognized, evaluation theory is distinct in that it represents the unique knowledge of the discipline (Christie & Lemire, 2019; Rog, 2015; Shadish, 1988, p. 1). The knowledge of a discipline, according to Chinn and Kramer (2018), is "knowledge that has been collectively judged by standards shared by members of the disciplinary community and that is taken to be a valid and accurate understanding of elements and features that comprise the discipline" (p. 4). Thus, evaluation theory is essential to distinguishing evaluation from other disciplines and evaluators from other professionals (Chinn & Kramer, 2018; Mertens & Wilson, 2019; Shadish, 1988; Shadish et al., 1991).

Evaluation theory consists of various prescriptive approaches that articulate the actions required to conduct evaluations in a manner consistent with professional standards and principles (Alkin, 2013; Christie, 2003a; Mertens & Wilson, 2019; Shadish et al., 1991). As such, it stands to reason that these theories readily translate into the everyday practice of evaluation professionals. However, empirical evidence suggests that the influence of evaluation theories on

practice is minimal (Christie, 2003b; Christie & Masyn, 2010; Shadish & Epstein, 1987). This disconnect has left researchers (e.g., Christie, 2003b, p. 92) questioning what knowledge, if not evaluation theory, guides evaluation practice.

One potential source of knowledge proposed in the literature is practical knowledge (Christie, 2003c; 2012; Christie & Masyn, 2010; Christie & Rose, 2003; Kundin, 2010; Schwandt, 2005; 2008; Shadish et al., 1991; Tourmen, 2009). Schwandt (2008) describes practical knowledge as "the tact, dispositions, and considered character of decision making called for in various situations faced in 'doing' the practice" (p. 33). It is developed through practitioner's interactions with practical problems and the insights they draw from these experiences (Guzman, 2009; Schön, 1983). Some scholars argue that practical knowledge may be more influential on evaluation practice than theoretical knowledge in some instances (Christie, 2003a; Kennedy, 1982).

Guzman (2009) explains that practical knowledge consists of explicit and tacit dimensions. Practical knowledge is explicit when it involves the application of procedural knowledge (i.e., knowledge of how to do a specific action) to perform an action in a consistent way (Guzman, 2009, p. 92). Tacit practical knowledge, on the other hand, is based on knowing (Nicolini, 2003) or the ways that practitioners perceive and draw meaning from the situations they encounter in their practice (Guzman, 2009, p. 92). These ways of knowing help practitioners understand how to interpret and approach the various demands of their practice, particularly those that are complex and contextual.

Practical knowledge represents a shift from an empirical view of knowledge development to one that views knowledge development as a function of professional practice (Chinn &

Kramer, 2018; Schön, 1983; Zander, 2007). From this perspective, practice is not just a series of steps that are implemented and reproduced but a dynamic process of perceiving, theorizing, and experimenting while building understanding. Moreover, knowledge is not an object applied to practice but rather a form of situated practice that involves "different epistemological and ontological assumptions" woven into human action (do Nascimento Souto, 2013, p. 63).

Despite its perceived importance to evaluation practice, limited research has been conducted on practical knowledge in evaluation. Thus, the field lacks an epistemology of evaluation practice. This deficit has numerous implications for evaluation practice, education and professional development, and the evaluation scholarship.

First, the field's current understanding of the knowledge, and processes of knowledge construction, that guides evaluation practice is limited and biased. Evaluation is often characterized by western epistemological approaches (e.g., empiricism, rational-cognitive knowing) without consideration for other ways of knowing that are required for and benefit practice (House, 1980; do Nascimento Souto, 2013). This characterization can skew how the field conceptualizes practice, mislead evaluators seeking to develop their evaluation capacity, and alienate practitioners who engage in diverse ways of knowing.

Second, evaluation educators lack a conceptual tool to integrate practice into evaluator education (Raelin, 2007). Research suggests that evaluation students can develop an overreliance on theory, while practice-related competencies such as interpersonal skills and reflective practice have been identified as training gaps (Chouinard et al., 2017; Galport & Azzam, 2017). The integration of theory and praxis are essential to preparing future evaluators and ensuring communities have access to qualified evaluators.

3

Last, the wisdom of practitioners has not been fully tapped. Consequently, opportunities to advance evaluation theory and practice are not being leveraged by the field (Cousins & Earl, 1999; Christie, 2003a; Datta, 2003; Shadish et al., 1991; Smith, 1993; Worthen, 1990). Failure to capitalize on these opportunities limits the field's growth and the extent to which it can influence changes (e.g., attitudinal, programmatic, policy) that support social betterment (Greene, 2013; Henry & Mark, 2003; Mark et al., 1999; Shadish, 1994).

The current research helped to address this gap by examining practical knowledge and the ways of knowing in evaluation. By doing so, the research advances the scholarship on evaluation theory and practice and provides new insights into evaluator education and professional development. The remainder of this chapter (1) grounds the research in the extant literature on evaluation practice and theory, (2) introduces the concept of practical knowledge and explains its role in professional practice, (3) presents the ways of knowing in nursing as a model for conceptualizing the ways of knowing in evaluation, and (4) summarizes the current study, including its purpose, research questions, and significance to the field.

Review of Relevant Literature

This section begins by describing the origins and intent of evaluation theory as well as the various assumptions underpinning them. This is followed by a review of the current thinking on theory and practice, including the conceptual and empirical research, as well as information on practical knowledge and its perceived benefits and challenges. Last, is a presentation of the ways of knowing in nursing as a framework for conceptualizing professional practice; this is accompanied by observations on the framework's potential relevance to evaluation.

Evaluation Theory

Shadish et al. (1991) define theory as "a body of knowledge that organizes, categorizes, describes, predicts, explains, and otherwise aids in understanding and controlling a topic" (p. 30). Formal theories are those which are explicit and widely communicated through scholarly work and education. They are often, albeit not always, developed and validated through scientific methods and, therefore, viewed as objective and factual (Kimball, 2016).

In its relatively brief history, the field of evaluation has generated numerous theories "that specify what a good or proper evaluation is and how evaluation should be done" (Alkin, 2013, p. 4; Shadish et al., 1991). These *evaluation theories* are unique from other formal theories in some respects. First, they are prescriptive, serving as models or guiding frameworks as opposed to generalizations that describe and predict (Alkin, 2013, p. 4). Shadish et al. (1991) argue that the ideal evaluation theory would be descriptive and explain why certain practices lead to specific results across evaluation settings (pp. 30-31). Consequently, there have been growing calls for further research on evaluation theories to facilitate the development of descriptive theories (Alkin, 2004; Christie, 2003a; Henry & Mark, 2003).

Second, many evaluation theories are informed by the experiences and approaches of knowledgeable practitioners. Current theories are indicative of the "interplay among problems uncovered by practitioners, the solutions they tried, and traditions of the academic discipline of each evaluator" (Christie, 2012; Leviton, 2015; Shadish et al., 1991, p. 31). As such, practice, or practical knowledge more specifically, has informed the field's theory base.

Scholars have categorized evaluation theories using the elements these theories tend to have in common. Christie and Alkin (2013) developed the Evaluation Theory Tree in which they

arrange theories along three branches based on the theory's main emphasis (i.e., methods, use, values). Mertens and Wilson (2019) extended this work by adding a social justice branch and mapping each branch onto major paradigms in evaluation as described in Table 1. Importantly, Mertens and Wilson (2019) illuminate how one's conceptualization of evaluation extends beyond practical experience and academic training – evaluation theories are influenced by "evaluator's beliefs about themselves and their roles" as well as the philosophical assumptions that underlie their worldview (p. 35).

Table 1

Major Paradigms in Evaluation

Paradigm	Branch	Description
Postpositivist	Methods	Focuses primarily on quantitative designs and data, may use mixed methods but quantitative methods dominate.
Pragmatic	Use	Focuses primarily on data that are found to be useful to stakeholders, advocates for the use of mixed methods.
Constructivist	Values	Focuses primarily on identifying multiple values and perspectives through qualitative methods; may use mixed methods, but qualitative methods dominate.
Transformative	Social Justice	Focuses primarily on viewpoints of marginalized groups and interrogating systemic power structures through mixed methods to further social justice and human rights.

Note. From *Program Evaluation Theory and Practice* (2nd ed., p.42), by D.M. Mertens and A.T. Wilson, 2019, The Guilford Press. Copyright 2019 by The Guilford Press.

House (1980) presents a categorization of evaluation approaches based on alignment with "variation in the assumptions of liberalism" (p. 46). House argues that major evaluation models are all based on a form of subjectivist ethics. Models based on the utilitarian form give priority to maximizing happiness and assume that "any activity that maximizes happiness is the "right" thing to do (p. 47). Conversely, models based on the intuitive/pluralistic form do not rely on

explicitly defined ethical principles (p. 47), but rather base ethical judgments on "intuition and experience" (p. 49). Table 2 describes the evaluation models associated with each form, including influential theorists, studies, or organizations presented in parenthesis.

Table 2

House's (1980) Categorization of Major Evaluation Models by Form of Subjectivist Ethics

Utilitarian	Intuitionist/pluralist
• Systems analysis (Rivlin) • Behavioral objectives (Tyler)	• Art criticism (Eisner) • Professional review (National Study of School Evaluation, Council on Program Evaluation)
• Decision making (Stufflebeam, Guttentag, Patton) • Goal free (Scriven)	• Quasi-legal (Wolf, Stenzel) • Case study (Stake, McDonald, Guba)

Despite similarity in ethics, the models differ in epistemology (House, 1980). The models that fall under the utilitarian form are grounded in an objectivist epistemology that assumes that objectivity can be achieved through scientific methods and instruments (p. 50). Here, objectivity is equated "with the procedures for determining intersubjectivity" or freedom "from bias or distortion" (House, 1980. p. 55). House (1980) defines prediction as the ultimate aim of the objectivist orientation (House, 1980, p. 57).

The intuitionist/pluralistic approaches embrace a "subjectivist epistemology" that is more interested in relating the evaluation to the experiences of the audience than establishing truth (House, 1980, p. 56). Here, improving understanding is the aim and there is less concern for intersubjectivity and reproducibility (House, 1980). Models rooted in the subjectivist

epistemology tend to use more "general" than "specific" research methodologies which have explicit details and procedures (House, 1980, p. 56).

However, for the purposes of this study, the most relevant distinction between the two epistemologies relates to their treatment of theory and practice. According to House (1980), in the objectivist epistemology "there is a rigid separation of observer and facts- highly abstract theory separated from application" (p. 57). Conversely, in the subjectivist epistemology, utility is based on the observer's interests and "theory and practice are blended together" (p. 57).

Relationship Between Evaluation Theory and Practice

Though a wide range of evaluation theories exist today, the nature and extent to which these theories relate to real-world evaluation practice is a topic of ongoing discussion. The extant literature provides various interpretations of the relationship between evaluation theory and practice – with some theorists questioning the necessity of evaluation theory to evaluation practice (Chelimsky, 1998; Scriven, 1991). Scriven (1991), for example, argues that theory is "something of a luxury" for an evaluator (p. 360). Alternatively, House (1980) describes evaluation theories as "idealizations" as opposed to models one would adhere to in real-world practice (p. 21).

Schwandt (2009) discusses a theory-practice distinction common to "western theoretical and scientific traditions of knowledge" (p. 30). From this perspective, theory is understood as "the realm of contemplation and reflection" and the intelligence upon which action is informed, while practice is associated with performativity, execution, and application (Schwandt, 2008, p. 29). Here, Schwandt (2009) is drawing attention to the tendency to view the relationship between theory and practice as unidirectional and hierarchical.

8

Others describe the relationship between theory and practice as bidirectional. Chelimsky (2012) explains that early theorists understood the real-world limitations of their theories but believed that "as evaluation practitioners gained experience in applying methods and principles, they would feed their new learning back to inform theory on a continuing basis" (p. 93). Similarly, Christie (2003a) describes a mutually reinforcing relationship whereby "practice informs theory, and in turn, theory influences practice" (p. 2). These perspectives reject the hierarchical nature of theory and practice, albeit still presenting them as distinct.

Regardless of interpretation, there are concerns regarding the extent to which evaluation theory informs practice, often referred to as a theory-practice gap (Chelimsky, 2012; Chouinard et al., 2017; Christie, 2003b). Rog (2015) argues that many practitioners view theory "as the purview of academia and do not always see it as relevant to their daily work nor do they always see how their work would inform our theories" (p. 224). Indeed, this sentiment is consistent with the literature, which suggests that the influence of evaluation theories on practice is minimal.

Research on Evaluation Practice and Theory. Several scholars have emphasized the need for research on evaluation practice (Alkin, 1991; Shadish et al., 1991; Smith, 1993), in particular the connection between evaluation theory and practice. Researchers have met this call, to at least a limited degree, since the mid-1980s (Chandler, 2001, as cited in Datta, 2003; Chouinard et al., 2017; Christie, 2003b; Christie & Masyn, 2010; Shadish & Epstein, 1987). Each of these studies are described next, culminating in a need to further understand an existing theory-to-practice gap in the field.

Shadish and Epstein (1987) examined the relationship between practice and evaluation theory empirically. They used survey research to compare the actions of practitioners with those

prescribed by evaluation theorists. Rather than focus on a specific aspect of practice, which the authors argued previous studies had done, they broadened their inquiry to include multiple dimensions of practice. These dimensions included "methods, when to evaluate, the role evaluators should play, sources and kinds of questions to ask, how to construct dependent variables, and how to facilitate use" (p. 556).

The researchers collected data from a random sample of evaluators (N=318) drawn from two professional evaluation associations at the time of the study (i.e., Evaluation Network and Evaluation Research Society, which later combined to form the American Evaluation Association). Participants in the study completed a comprehensive survey that collected data regarding participants' background and a recently conducted evaluation. The latter was examined using 74 questions pertaining to various aspects of practice such as decisions, perceived role, and actions taken. These items were intended to represent advice from various evaluation theorists, including Donald Campbell, Lee Cronbach, Michael Scriven, Robert Stake, Peter Rossi, Joseph Wholey, and Carol Weiss (Shadish & Epstein, 1987, p. 559). Following this line of inquiry, participants were asked how representative their recent evaluation was of their typical practice as well as for information about potential theoretical influences (e.g., scholarly works, evaluation concepts) on their practice (Shadish & Epstein, 1987, pp. 558-559).

The analysis of data shared by respondents yielded four practice patterns: (1) academic or evaluation conducted for scientific and theoretical purposes, (2) stakeholder service, (3) decision-driven (i.e., evaluator decided to engage in the evaluation only after assessing the evaluation's potential benefits and costs), and (4) outcome or evaluation intended to judge a program's merit and effectiveness (Shadish & Epstein, 1987, pp. 578-579). Shadish and Epstein

10

(1987) also investigated which evaluation theory predicted these patterns. The academic pattern was statistically significantly associated with Cronbach and Campbell's work ($R = 0.40$; F = 3.70; df = 8,159; $p = 0.0006$), and the stakeholder service pattern was statistically significantly associated with Stake ($R = .037$; F = 3.16; df = 8,159; $p = 0.0024$).

Shadish and Epstein (1987) argue that the identification of four patterns is evidence of "at least some relationship between theory and practice" (p. 584). However, the authors also found that most practitioners (71%) were not familiar with the scholarly work of evaluation theorists (i.e., key readings and theoretical concepts). These findings suggest that theorists' actual influence on participants' practice was low.

Christie (2003b) investigated the relationship between evaluation theory and evaluation practice by comparing the reported practices of evaluation theorists (N=8) and evaluation practitioners (N=138). Participating theorists represented a diverse range of theoretical perspectives: Robert Boruch, Huey-tsyh Chen, Michael Patton, J. Bradley Cousins, Elliot Eisner, Ernest House, David Fetterman, and Daniel Stufflebeam (p. 9). Practitioner participants were not all "evaluators per se" (p. 10). Rather, some participants were administrators who were involved in evaluation work as part of their roles with a program. Christie argued this blend of evaluators and non-evaluators was characteristic of real-world practice and those typically involved with conducting evaluations (p. 10).

Christie (2003b) collected data regarding the respondent's self-reported evaluation practice from both the participating theorists and the "evaluator" sample using a survey instrument that was developed in collaboration with the sample of evaluation theorists. Specifically, theorists submitted survey items that described the manifestation of their evaluation

theory in practice with respect to methods, use, and valuing. These items formed the basis of a "theory-to-practice assessment instrument," which consisted of 38-items (p.11).

Multidimensional scaling was used to examine the similarity and dissimilarity between the theorists' and practitioners' reported practices. Generally, the results suggest that evaluation theory was not integral to how the practitioners carried out an evaluation. Christie (2003b) found that only 36% of practitioners were within meaningful proximity to a theorist. Additionally, it appeared that practitioners tended to incorporate portions of different theories in their practice instead of following a particular theory.

Using the data collected by Christie (2003b), Christie and Masyn (2010) conducted a latent profile analysis to describe the self-reported practices of evaluators (N=138). Four classes of practice emerged from the analysis. The first, and largest class, was the "indistinct pattern of practice" which accounted for 38% of the sample (p. 244). Scores from these participants were not closely associated with any of the theoretical approaches assessed by the theory-to-practice assessment instrument (Christie, 2003b, p. 11).

The second class Christie and Masyn (2010) identified was "methods-focused," which included 28% of participants who indicated "strong and distinct preferences for using particular methods" (p. 241). The third (user-focused) and fourth (robust pattern of practice) patterns each accounted for approximately 16–17% of participants (p. 244). The user-focused pattern was primarily concerned with the utility of data and viewed the evaluator's role as facilitating use. Participants falling under the robust pattern of practice were also concerned with utility and facilitating use, however, their reported practices were the most closely aligned with the practices of the theorists who developed the survey items (Christie & Masyn, 2010, p. 242). As

such, this group was determined to be most "engaged in evaluation practices that are rooted in theory" (Christie & Masyn, 2010, p. 242).

Overall, Christie and Masyn (2010) found that most participants' practice did not closely align with theory. The only exception being those participants who fell within the smaller "robust pattern of practice" group (Christie & Masyn, 2010, p. 244). The authors proposed that practitioners rely more readily on working knowledge, defined as "an organized body of knowledge of evaluation procedures, which is used spontaneously and routinely in the context of their work" (p. 249).

Chandler (2001, as cited in Datta, 2003) interviewed a sample of evaluation practitioners (*N*=17) to identify the factors that influence practice decisions in general. The sample included individuals of varying levels of evaluation experience, employment domain, nationality, and preferred evaluation approach (Greene, 2001, as cited in Datta, 2003). The results indicated that the client's needs and context were most important to decision-making, while evaluation theories did not significantly influence evaluation practice (Chandler, 2001, as cited in Datta, 2003).

Chouinard et al. (2017) used qualitative methods to examine the field experiences of students (*N*=5) enrolled in an evaluation theory and practicum course. The authors were interested in how students understood and used evaluation theory to navigate the demands of practice (p. 280). Data were collected through reflective journaling, one-on-one interviews, and a focus group. The findings indicated that students' understanding of theory and practice "ultimately could not provide the guidance nor the technical direction they required for an engagement with the exigencies of evaluation practice" (Chouinard et al., 2017, p. 283). Despite

this, the findings suggest a positive relationship between students' perceived need for theory and the complexity of evaluation contexts (Chouinard et al., 2017).

Potential Sources of the Theory-Practice Gap. The studies discussed in the previous section support the notion of a theory-practice gap in evaluation. Additionally, they suggest that practice is influenced by various factors. The conceptual literature helps to elaborate on these findings by offering information on the focus of evaluators, the theoretical knowledge of the field, and the nature of evaluation work.

The Primary Focus of Evaluation Practitioners. Evaluation is often conceived to be a pragmatic discipline, primarily focused on meeting a societal need (Alkin, 2003; Christie, 2003a, 2012; King, 2003; Shadish et al., 1991; 1994). As such, factors specific to the context often dictate the goals and procedures of evaluation studies (Alkin, 2003; Chandler, 2001, as cited in Datta, 2003; Greene et al., 2001; House, 1980, King, 2003; Tourmen, 2009). As King (2003) explains:

> Evaluators tend to concentrate on conducting studies for clients rather than validating the theoretical assumptions underlying various evaluation models. Like pilots in midair whose immediate priority must be to fly their airplanes successfully, evaluators "in midair" must concentrate on doing what is necessary to complete evaluation studies effectively. They constantly face pressure to "do" evaluation expediently, inexpensively, and excellently— hence, the development of a comprehensive set of standards (Joint Committee, 1994) to guide professional practice rather than a focus on traditional theory building as in the academic disciplines. To our clients, validating evaluation theory is not cost-effective. (p. 58)

As such, evaluators often work within the confines of pragmatic demands and contextual constraints, which may limit their ability to apply evaluation theories to their practice.

Theoretical Knowledge of Practitioners. There is a high degree of professional pluralism in the evaluation field (House, 1993). Each theory embodies a unique conceptualization of how and why evaluations are conducted, whose needs are to be met, and what outcomes should be expected. These theories form a complex web of values, methods, and principles.

The AEA Evaluator Competencies indicate that a competent evaluator "selects evaluation approaches and theories appropriately" (King & Stevahn, 2020, p. 52). Inherent to this expectation is the assumption that practitioners are highly knowledgeable of many, if not all, evaluation theories. However, the literature suggests that many evaluators lack formal evaluation education, which is where one typically gains exposure to evaluation theories (Christie, 2003a; Christie et al., 2014; Shadish & Epstein, 1987).

Evaluators lacking knowledge of evaluation theories are poorly positioned to select the most appropriate theory for an evaluation. Furthermore, there is no established context for evaluation theories or guidance as to how one might assess the suitability of one theory over another (Alkin, 2003). As such, applying evaluation theories to practice may be challenging for any evaluator, regardless of training or background.

The Nature of Evaluation Work. Schön (1983) argues that perceptions of professional practice are based on technical rationality, an epistemology of practice embedded in how professionals are educated and the knowledge they are expected to utilize in practice. Prescribers of this epistemology are of the view that science can render standardized approaches (i.e., theoretical solutions and techniques) for resolving the problems and demands encountered by

practitioners (Schön, 1983; Thompson & Thompson, 2018). Schön (1983) argues that technical rationality does not acknowledge the "swampy lowland" of professional practice, where problems are messy, confusing, and ill-suited for standardized approaches (p. 39).

The literature suggests that evaluation work is often conducted in the "swampy lowland" (Schön,1983, p. 39). Schwandt (2003; 2005) discusses the "rough ground" of evaluation practice "where values, personalities, evidence, information, feelings, sensitivities, emotions, affect, ambiguities, contradictions, inconsistencies, and so forth are simultaneously in play" (p. 99). Tourmen (2009) also emphasizes the complexity of evaluation work, which "consists of major and minor choices, back-and-forth movements, difficulties in decision making, compromises, contexts that are not completely under control nor easy to foresee, and so on" (p. 28).

Given the complex and unpredictable nature of evaluation work, it stands to reason that theories and standardized approaches often fall short in real-world practice. Chelimsky (2012) claims that "practitioners now recognize that they cannot count on theory to help them with contextual issues involving people, subject-matter complexities, history, or politics," which accounts for most, if not all, evaluation contexts (p. 94). In such instances, practitioners must utilize the knowledge inherent to their practice to make sense of their experience and theorize a means of moving forward (Chelimsky, 2012; Schön, 1983; Schwandt, 2005; 2009).

Practical Knowledge

Practical knowledge represents an epistemology of practice that sees knowing and knowledge development as a function of professional practice (Schön, 1983; Schwandt, 2005). Schön (1983) explains, "when we go about the spontaneous, intuitive performance of the actions of everyday life, we show ourselves to be knowledgeable in a special way" that is "tacit, implicit

in our patterns of action and in our feel from the stuff with which we are dealing" (p. 49). This knowledge is embedded in how we perceive ourselves, others, the contexts, and "the connections among action, consequence, and situation" (Argyris & Schön, 1974, p. 7; Christie & Rose, 2003; Hillier, 1998; Kimball, 2016; Schön, 1983; Schwandt, 2005, 2008; Usher & Bryant, 2014).

Guzman (2009) posits that practical knowledge consists of explicit and tacit dimensions. Practical knowledge is explicit when it involves the application of explicit procedural knowledge (i.e., knowledge of how to do a specific action) to perform an action in a consistent way (Guzman, 2009, p. 92). Conversely, tacit practical knowledge involves knowing (Nicolini, 2003) or the ways that practitioners perceive themselves and their practice (Guzman, 2009, p. 92). It enables practitioners to carefully navigate practice by informing them of what to pay attention to, what questions to ask, how to relate to people, and what judgements to make (Schwandt, 2008).

Schwartz and Sharpe (2010) identify six characteristics of practically knowledgeable practitioners. These characteristics include (1) understanding purpose coupled with a desire to address the needs of clients, (2) an ability to improvise, and balance and interpret multiple principles and guidelines, (3) social perceptiveness and understanding nuances, (4) empathy, (5) emotional intelligence, and (6) experience in one's practice (pp. 25-26). These characteristics suggest that practical knowledge is the cumulation of various forms of awareness, attributes, capabilities, and motivations.

The role of practical knowledge in evaluation is acknowledged in the literature. Shadish et al. (1991) claim that:

    all evaluation practitioners are nascent evaluation theorists. They think about what they

    are doing, make considered judgments about which methods to use in each situation,

weigh advantages and disadvantages of choices they face, and learn from successes and

failures in their past evaluations. (p. 35)

Schwandt (2009) describes practical knowledge as both "a cognitive ability and a way of literally

being present in a situation" (p. 33). It is self-constitutive and closely related to "our entire

'being'—our gestures, emotions, orientation, stance, and perspective, as well as our ways of

understanding and questioning" (Schwandt, 2005, p. 101). Christie (2003c) and others (e.g.,

Chelimsky, 2012; Tourmen, 2009; Weiss, 2013) have also discussed the personal conceptions

that guide evaluators' practice as they navigate evaluation contexts.

The Benefits of Practical Knowledge

There are numerous benefits associated with practical knowledge. First, and arguably

most importantly, it enables us to cope with "complexity, uncertainty, instability, uniqueness,

and value-conflict," which practitioners often encounter (Schön, 1983, p. 39; Schwandt, 2003;

2005). This includes reframing situations such that a theoretical path of action becomes apparent.

In doing so, practitioners become better positioned to leverage evidence-based practice and close

the gap between theory and practice (Carper, 1978; Chinn & Kramer, 2018; Hillier, 1998).

Second, practical knowledge is highly contextualized and can be situated within a

specific time and space. This provides practitioners with the flexibility to determine the variables

at play, what those variables mean in the here and now, and what is 'right' and 'wrong' in a

specific instance (Schön, 1983). The value of flexibility and responsiveness to evaluation work

has been emphasized by numerous evaluation theorists (e.g., Patton, 2008; Stake, 2003; Weiss,

1998).

Third, when explicated, practical knowledge can support professional effectiveness and learning. Through critical reflection and reflexivity, practitioners can develop greater awareness of their practice, including the philosophical assumptions and values that shape it (Brookfield, 2017; Jewiss & Clark-Keefe, 2007; Mezirow, 1990; van Draanen, 2017). This awareness can highlight misalignment between professional practice and goals and clarify learning needs.

The Challenges of Practical Knowledge

Despite these benefits, the literature identifies problems associated with practical knowledge. Namely, it is often implicit and may be difficult for evaluators to articulate or even access (Christie, 2003a; Guzman, 2009; Hillier, 1998; Kimball, 2016; Schön, 1983; Schwandt, 2008; Usher & Bryant 2014). Van Manen (1999, as cited in Schwandt, 2008) argues that practical knowledge is:

> Not primarily in the intellect of the head but rather in the existential situation in which [a] person finds himself or herself… something that belongs phenomenologically more closely to the whole embodied being of the persona as well as the social and physical world in which this person lives. (p.12)

Similarly, Schwandt (2008) describes practical knowledge as "one's ability to be present in and handle a situation" (Schwandt, 2008, p. 31). In the absence of intentional reflection, practitioners may not be aware of why they are acting, feeling, or thinking a certain way. This limits the extent to which practitioners can learn about and from their practice, communicate knowledge to others, and creates challenges for empirical inquiry on practical knowledge.

Additionally, practical knowledge may be based on unvetted biases and assumptions (Argyris & Schön, 1974; Chelimsky, 2012; Christie, 2012; Kennedy, 1982; Taylor et al., 2001).

19

Unlike formal theory, practical knowledge may not be scrutinized, which can lead to systematic error. As Jewiss and Clark-Keefe (2007) explain:

> left unchecked, evaluators may inadvertently steer the project into the path of their own personal issues and bias, delimiting attention to a range of possible data sources, making less than could be made of important phenomena, and overemphasizing or glorifying aspects of programs or certain participants' perspectives. (p. 338)

Last, practical knowledge may not be viewed as a legitimate form of knowledge in disciplines, such as evaluation, with an empiricism-orientation (House, 1980; Schön, 1983; Schwandt, 2005). This optic may lead some practitioners to avoid situations that do not lend themselves to a scientific solution or manipulate practice situations to fit formal theory (e.g., ignore or dismiss data, force the mold, explain away discrepancies; Schön, 1983, p. 44). Both actions limit the extent to which practitioners meet their professional responsibilities and contribute to the well-being of society.

Ways of Knowing in Nursing

Some fields have taken steps to explicate the knowledge inherent to their respective practices. Carper's (1978) seminal research on the ways of knowing in nursing has had a profound impact on the discipline's epistemological and ontological underpinnings (Johns, 1995; Zander, 2007). As part of this research, Carper (1978) identified four discrete yet inter-related ways of knowing: empirics, ethics, personal, and aesthetics (p. 23).

Chinn and Kramer (2018) expanded on Carper's work by proposing a fifth way of knowing, emancipatory knowing, and identifying patterns and dimensions underlying each way of knowing (i.e., critical questions, creative processes, formal expressions, authentication

processes, and integrated expressions in practice). Collectively, these ways of knowing represent the core knowledge needed for mastery in nursing and, therefore, are considered essential components of nursing education and competent practice (Carper, 1978; Chinn & Kramer, 2015; Johns, 1995; Zander, 2007).

Acknowledging that nursing is a distinct discipline and field from evaluation, there are some similarities between the two. First, nursing and evaluation share a similar overarching purpose, which is to contribute to a better society (Chinn & Kramer, 2018; Greene, 2013; Mark et al., 1999; Shadish, 1994). Second, there are similar nuances inherent to both forms of professional practice. For instance, practitioners in both fields must (1) maintain a strong empirical orientation while being adaptive to the context, (2) make decisions about what interests will and will not be attended to, (3) focus on specific parts of a program (or patient) while staying attuned to the whole, and (4) use assessment and reasoning to make evaluative judgment.

The next section provides a description of the ways of knowing in nursing as well as their potential relevancy to evaluation. Prior to this, some clarification regarding the use of the terms 'knowledge' and 'knowing' is warranted. Chinn and Kramer (2018) define knowing as "ways of perceiving and understanding Self and the world" that are fluid and internal to people (p. 3). Like Schwandt (2005), Chinn and Kramer (2018) link knowing to ontology, arguing that knowing is "particular and unique to our existence and to each individual's personal reality" (p. 4). Once knowing is communicated, it becomes knowledge, which can be specific to an individual or shared by a community (Chinn & Kramer, 2018, p. 3).

Empirics

The empirics way of knowing is referred to as the "science of nursing" and is grounded in knowledge produced through scientific study and research (Carper, 1978, p. 23; Chinn & Kramer, 2018). This way of knowing is demonstrated through scientific competence and the use of empirical information to conceptualize, structure, and validate. Empirical knowing is generally expressed through facts, formal and thematic descriptions, theories, and models (Chinn & Kramer, 2018, p. 15). Among the ways of knowing in nursing, empirics is thought to be the most accessible and easiest to communicate.

Evaluation is often described as "the systematic collection of information" and evidence to inform judgments, decisions, and improvements (Patton, 2008, p. 29). The approaches used to systematically collect information are empirical, involving the application of scientific methodology and evaluation theories (Fourier, 2005; Wilce et al., 2021). While evaluation theories are generally not scientifically tested, they are "empiricist in orientation" (House, 1980; p. 62). Additionally, the importance of research knowledge is heavily emphasized in both the AEA evaluator competencies (King & Stevahn, 2020) and guiding principles for evaluators (AEA, 2018a).

Ethics

Carper (1978) refers to ethical knowing as "the moral component" of nursing (p. 29). This way of knowing "involves making moment-to-moment judgements about what ought to be done, what is good and right, and what is responsible" (Chinn & Kramer, 2018, p. 7). According to Chinn and Kramer (2018), ethical knowing in nursing involves clarifying and exploring as practitioners seek to identify what's most important in context. When a conflict between ethical

principles arises, nurses are expected to think through and justify tradeoffs (Chinn & Kramer, 2018). The demands of ethical knowing require awareness of professional codes of ethics and other relevant social values as well as the ability to engage in ethical reasoning (Carper, 1978; Chinn & Kramer, 2018).

Professional ethics are imperative to evaluation practice. The key components of ethical evaluation practice are described in AEA's guiding principles for evaluators (AEA, 2018a). Some evaluators may also be subject to codes of ethics due to their place of employment or professional background. Articles regarding ethical dilemmas are also published in evaluation journals (e.g., *American Journal of Evaluation*) to educate and raise awareness of ethical situations in the field.

Similar to nursing, evaluation ethics will come into conflict in some contexts (AEA, 2018a). Morris (2002) has also questioned the relevancy and transferability of ethical codes to real-world practice. As such, knowledge of broader societal values and ethical reasoning are often needed, especially when evaluators are faced with making trade-offs (Morris, 2002).

Personal

Personal knowing refers to knowing oneself in the context of one's practice (Carper, 1978; Johns, 1995). This knowing involves awareness and authenticity of one's emotions and biases as well as an ability to cope with anxiety and maintain a sense of self (Johns, 1995, p. 229). Knowing oneself allows for meaningful experiences with others and, in particular, the therapeutic use of self (i.e., use of one's perceptions and disposition to advance a therapeutic process; Chinn & Kramer, 2018; Punwar & Peloquin, 2000).

The literature connects personal knowing to reflective practice, which is an approach to practice that involves learning "through and from experience" by drawing on and validating knowledge, attending to the affective and value dimensions of practice, and challenging assumptions (Boud et al., 1985; Finlay, 2008, p. 1; Schön, 1983). The importance of reflective practice to evaluation is referenced in numerous evaluation theories (e.g., Utilization-Focused Evaluation, Patton, 2008; Empowerment Evaluation, Fetterman, 1994; Evaluative Inquiry for Learning in Organizations, Preskill & Torres, 1999) and statements regarding evaluation quality and ethics (e.g., AEA Evaluator Competencies, Guiding Principles for Evaluators, Statement on Cultural Competence in Evaluation).

Aesthetics

Carper (1978) describes aesthetics as the "art of the practice" (p. 25), which is driven by an empathetic acquaintance and a willingness to act on what's meaningful in the moment (Chinn & Kramer, 2018). As we become more open and sensitive to the interests and values of others, we gain new insights into how we can modify our practice to better align with those values and interests (Carper, 1978, p. 17). In this regard, empathy and presence foster innovation and professional effectiveness.

Aesthetics knowing is similar to strategies proposed by Donaldson et al. (2002) to manage excessive evaluation anxiety. Active listening and developing an understanding of how evaluation partners are feeling are among the strategies proposed by Donaldson et al. (2002). The authors also suggest that understanding and attending to the concerns of individuals and communities can help evaluators predict and potentially avoid sensitivities in future evaluations.

Responsive Evaluation may also be related to aesthetic knowing. Among other goals, this evaluation approach aims to be responsive to stakeholder's information needs, values, and ways of understanding (Stake & Abma, 2005). Responsive Evaluation also encourages evaluators to prioritizes what is currently happening in the program rather than focus exclusively on what should happen in the future (i.e., program intents).

Emancipatory

As previously mentioned, emancipatory knowing was recently added to the ways of knowing in nursing framework to reflect the "praxis of nursing" (Chinn & Kramer, 2018, p. 5). Chinn and Kramer (2018) describe emancipatory knowing as the ability "to recognize social and political problems of injustice and inequity, to realize that things could be different, and to piece together complex elements of experience and context to change a situation as it is to a situation that improves people's lives" (pp. 71-72).

Emancipatory knowing is facilitated by critical reflection on broader societal issues (Chinn & Kramer, 2018). Critical reflection involves critiquing one's knowledge by questioning the authority and credibility of the assumptions that influence our perspectives and experiences (Brookfield, 2017; Hora & Smolarek, 2018; Mezirow, 2009). This critique builds awareness and helps practitioners to understand how issues related to inequity and power relate to practice situations (Chinn & Kramer, 2018). As practitioners envision a better world, they become clear of their sense of purpose and modify their practice accordingly (Chinn & Kramer, 2018). The outcome of this process is praxis, which is described as "value-grounded, thoughtful reflection and action that occurs in synchrony and that integrates ontology and epistemology" (Chinn & Kramer, 2018, p. 296).

As mentioned, social betterment is often viewed as the ultimate purpose of evaluation (Greene, 2013; Henry & Mark, 2003; Mark et al., 1999; Shadish, 1994). However, our pathways to social betterment are complex and reliant on our ability to influence our social context (Henry & Mark, 2003; Mason & Azzam, 2018). The dimensions of emancipatory knowing could be helpful in unpacking this pathway and providing the evaluation field and practitioners with greater insight into the mechanisms of social change. In addition, these lenses can help situate evaluation within its social context and illuminate how the field is and is not informing changes that lead to social betterment (Hooper, 2010).

A summary of the ways of knowing in nursing, including the underlying dimensions, is provided in Appendix A. As mentioned, this framework provides a useful lens in which to conceptualize practical knowledge in evaluation, including ways of knowing. The next section describes the current study, which focused on examining the ways of knowing in evaluation.

The Current Research

Research Purpose and Research Questions

The overarching goal for this research was to garner new insights about practical knowledge in evaluation, in particular, how evaluators perceive and interpret meaning from their experiences and how this shapes their practice (i.e., the ways of knowing in evaluation). By doing so, the researcher aimed to shed new light on (1) the knowledge and sense-making processes that underlie evaluation practice, (2) the relationship between theory and practice, and (3) evaluation as a distinct discipline. In this pursuit, the following research questions were addressed:

1. What are the ways of knowing in evaluation?

2. What dimensions underlie the ways of knowing in evaluation?

3. How do the ways of knowing manifest in practice?

As this is applied research, it was deemed beneficial to examine the ways of knowing in evaluation in an applied area of the discipline. As mentioned, the ways of knowing in nursing are integrated into nursing educational programs to help prepare nurses for practice. As such, this research examined the presence of the ways of knowing in evaluation in graduate programs in evaluation by addressing the following question:

4. How, if at all, are the ways of knowing in evaluation integrated into graduate programs in evaluation?

Summary of Methodology

The research was based on a multiple methods design, which consisted of two studies. The first three research questions were addressed by the first study, which was a fixed, convergent mixed methods design (Creswell & Plano Clark, 2018). Data was collected through a document review and the repertory grid (rep-grid) technique, analyzed separately, and then converged in a synthesis phase (Creswell & Plano Clark, 2018).

The final research question was addressed through a qualitative study. Data was collected using semi-structured interviews with faculty and alumni/advanced students (*N*=27) from six evaluation graduate programs in the United States (U.S.). Directed content analysis was employed to identify instances of ways of knowing and its underlying dimensions in evaluation programs.

A Note on Reflexivity

As a practitioner and student of evaluation, I was not a neutral participant in the research process. I had preconceived notions, assumptions, and 'baggage' (Finlay, 2002, p. 698) related to evaluation practice and evaluator knowing based on my professional experiences and worldview. I also had goals for the research, both personal (i.e., fulfill a degree requirement, support career advancement) and communal (i.e., support improved practice and training, contribute to the knowledge of the discipline), which intertwined the research with other aspects of my life and identities.

Since I could not remove myself from the study, my best option was to situate myself into the research and be transparent about my positions. A reflexive action plan (Appendix B; van Draanen, 2017) was implemented simultaneously with this research and in collaboration with my dissertation chair and a critical friend. The plan created space in the research process to clarify my values and positions as they related to evaluation practice and evaluator knowing and fostered progressive subjectivity (Mertens & Wilson, 2019). A key component of this plan is my positionality statement, which is provided in Appendix C.

Chapter 2: Methodology

To address the research questions presented in Chapter 1, the current study used a multiple methods design consisting of two studies that were implemented sequentially. Study 1 used a concurrent mixed methods strategy to collect quantitative and qualitative data using document review (Strand 1) and rep-grid (Strand 2) methods. These strands investigated how evaluators perceive and interpret meaning from their experiences and how this shapes their practice (i.e., the ways of knowing in evaluation). Findings from this study addressed the first three research questions.

Study 2 used qualitative methods to collect data using semi-structured interviews. These data addressed the final research question by providing insights into how the ways of knowing in evaluation were integrated into six evaluation educational programs. Figure 1 provides a graphical map of the studies.

Figure 1

Map of Multiple Methods Design
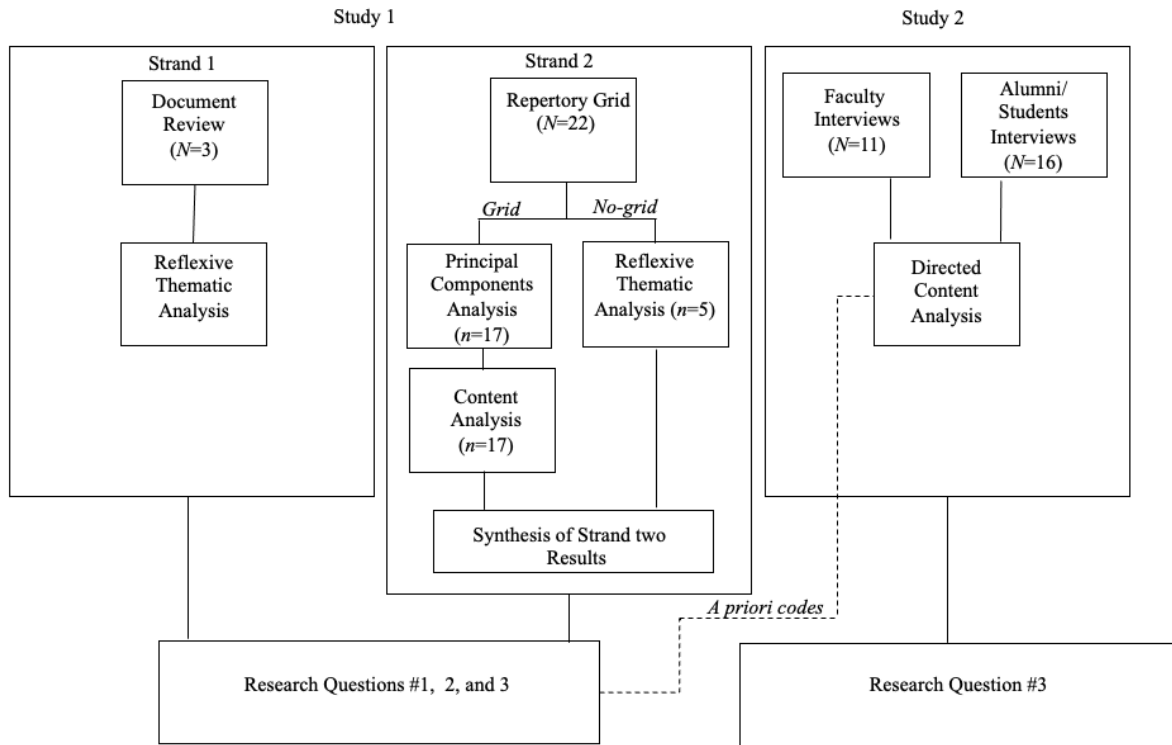
This chapter begins by describing the interpretative framework for this research. Following this, is a comprehensive overview of both studies.

Interpretative Framework

Given the interest in examining how practitioners know and develop knowledge as part of their practice, a constructivist framework was appropriate. A constructivist paradigm assumes that individuals have the capacity to create knowledge (Mertens & Wilson, 2019). Specifically, human processes allow individuals to interpret meaning and multiple realities from their

experiences (Creswell, 2013; Mertens & Wilson, 2019; Patton, 2002). These interpretations are subjective and normative to the individual but are real to them, nonetheless (Creswell, 2013).

Consequently, it is the goal of researchers "to understand the complex world of lived experience from the point of view of those who live it" to reach some understanding (Mertens & Wilson, 2019, p. 543). This goal was a key consideration when determining the methods for this study, particularly how best to collect and analyze the data.

Study 1

Strand 1: Document Review

The following section describes the design and procedures for the document review component of Study 1.

Document Sample. The study's data collection began with a document review to collect information regarding what evaluators need to know to practice evaluation and how they make sense of their work. Patton (2002) indicates that documents "constitute a particularly rich source of information about many organizations and programs" (p. 293). In the U.S. evaluation context, there are three leading sources of information about professional evaluation practice: the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2018a), 2018 AEA Evaluator Competencies (AEA, 2018b), and AEA Statement on Cultural Competence (AEA, 2011). These documents have been developed and updated through rigorous collaborative processes and, to the researcher's knowledge, are the closest approximation of the profession's collective values and perspectives on what constitutes evaluation practice and practitioners.

According to Bowen (2009), "the quality of the documents and the evidence they contain" is more important than the number of documents reviewed (p. 33). The reviewed

documents provided a broad interpretation of the nature of quality and ethical evaluation practice (King & Stevahn, 2020). Additionally, these documents were developed in consultation with a diverse representation from the field and, therefore, embody a multitude of perspectives. Last, and importantly, they reflect what evaluation practice in the U.S. is centered on and what knowledge, qualities, and behaviors are encouraged from evaluators (Schwartz & Sharpe, 2010).

AEA Guiding Principles for Evaluators. The Guiding Principles "reflect the core values" of AEA and are "a guide to the professional ethical conduct of evaluators" (AEA, 2018a, p. 2). There are five principles in total, which include systematic inquiry, competence, integrity, respect for people, and common good and equity (AEA, 2018a, p. 1). Each of these principles is accompanied by sub-statements to emphasize and elaborate on their meaning. For example, the respect for people principle helps evaluators conduct their work and communicate in a manner that upholds the dignity and self-worth of stakeholders. The common good and equity principle emphasizes the importance of considering the broader societal interests at stake in any evaluation.

AEA Evaluator Competencies. In 2018, the AEA established a set of evaluator competencies as part of ongoing efforts to professionalize the field (Stevahn et al., 2020). These competencies provide evaluators with a framework "for contemplating knowledge, skills, and abilities deemed necessary for effective practice" (Stevahn et al., 2020, p. 76). As described by King and Stevahn (2020), 49 competencies are organized into five domains: professional practice, methodology, context, planning and management, and interpersonal. The competencies are to serve three primary uses (1) enhance evaluation practice, (2) structure evaluation education and training, and (3) advance social justice (King & Stevahn, 2020, p. 60).

AEA Statement on Cultural Competence. The AEA Statement on Cultural Competence reflects the significance of cultural competence to evaluation and informs evaluators of AEA's expectations with respect to culturally responsive and sensitive evaluation practice (AEA, 2011). It outlines the essential practices for cultural competence in evaluation, all of which rely heavily on evaluators' awareness, perceptions, and sensemaking. These practices include acknowledging "the complexity of cultural identity," recognizing "the dynamics of power" and "bias in language," and employing "culturally approach methods" (AEA, 2011, pp. 7-8). Collectively, the statement provides a basis for evaluators to develop a stance on culture, which will better position them to understand, acknowledge, and integrate culture into their practice.

Procedures. The reviewed documents are in the public domain and were accessed through the AEA website (https://www.eval.org/). Each document was analyzed using reflexive thematic analysis (RTA; Braun & Clarke, 2006; 2008) to "interrogate dominant patterns of meaning surrounding" regarding what evaluators know and how they make sense of their practice (Braun et al., 2018, p. 8). Descriptive and reflexive field notes were taken throughout the analysis to capture observations about the research context, documents, and procedures and elaborate on emerging theories and the researcher's positions and role.

As per the Braun and Clarke (2006) perspective, the analysis began with a pre-analysis reflection to clarify the principles and perspectives relevant to the analysis. Braun and Clarke (2008; 2019) acknowledge the researcher's role in identifying and interpreting themes and emphasize the importance of being explicit about the choices made before and throughout the analysis process. The pre-analysis reflections were informed by questions provided by the authors, which include (1) what counts as a theme? (2) what type of analysis will I do? (3) how

will themes be identified? (4) at what level will themes be identified? (5) what epistemological paradigm will shape my analysis? (6) what questions will guide my analysis?

As part of this process, the researcher clarified that themes would capture an important element of evaluator knowing such as a basis for knowing (i.e., something knowing is grounded in), a process of knowing (i.e., how knowing occurs), or a manifestation (or expression) of knowing. Additionally, the use of semantic (i.e., explicit) and latent (i.e., implicit) codes was predetermined. A summary of the pre-analysis is provided in Appendix D.

Data Analysis. The RTA was carried out in six recursive phases consistent with the Braun and Clarke (2006; 2019) perspective (see Figure 2). These phases are presented in sequential order for communicative purposes. However, the implementation of these phases, particularly phases two through five, was highly iterative and fluid.

Figure 2

Six Phases of RTA (Braun & Clarke, 2006; 2019)

Phase 1	
Become familiar with the data	
Phase 2	
Inductive coding through line-by-line analysis	
Phase 3	
Cluster codes to develop candidate themes	
Phase 4 and 5	
Revise and define final themes	
Phase 6	
Report final themes	
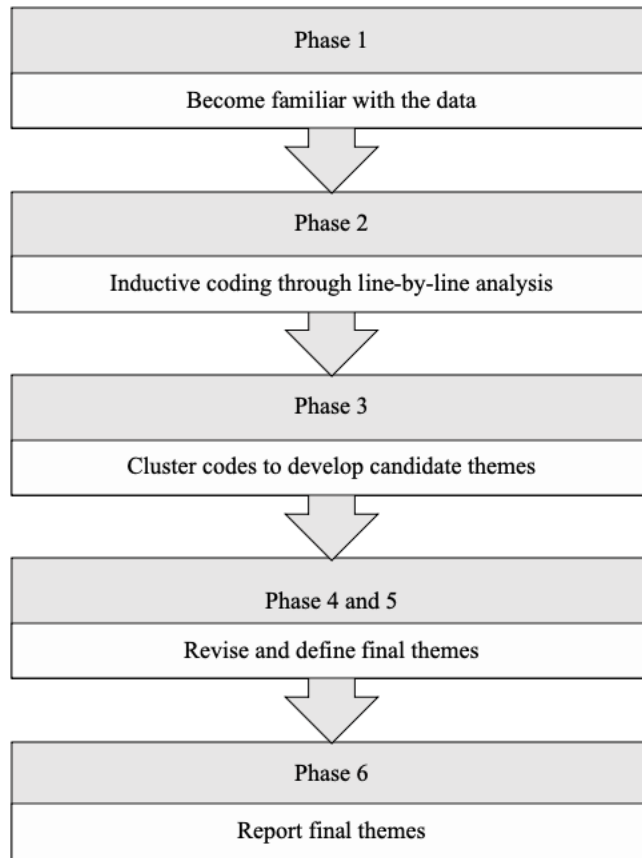
Phase 1. In Phase 1, the researcher became familiar with the data by reviewing each document for points or patterns of interest, connections, inconsistencies, or obscurities that could inform the subsequent phases (Braun et al., 2019). This process involved two full reviews of the documents, supplemented by a few partial reviews (i.e., returning to specific sections). The insights drawn from these reviews were captured in the researcher's field notes.

The researcher was familiar with the documents prior to the analysis since these documents were discussed as part of the researcher's doctoral studies and incorporated into other

scholarly work. Additionally, these documents are referenced and adhered to as part of the researcher's professional practice. That said, the researcher has never systematically reviewed the documents as part of a research study or examined them from a lens specific to knowing and knowledge generation. Unsurprisingly, the review fostered new insights about the nature of the documents and underlying assumptions related to what evaluators must know or infer from their practice.

The researcher drew two key insights from this phase. First, the ambiguity or fuzziness (Schwartz & Sharpe, 2010) of many of the concepts and terms discussed in the documents became apparent. Schwartz & Sharpe (2010) argue that concepts are fuzzy when they lack a concrete and consistent meaning. Rather, their meaning is relative and dependent on what seems more or less right in a given context. For instance, what it means to "demonstrate integrity" AEAb, 2018b, p. 2) or "behave with honesty and transparency" (AEA, 2018a, p. 3) could vary significantly depending on the context. As written, the documents assume that evaluators can interpret and apply these fuzzy concepts in practice.

Second, the weight of the requirements pushed onto evaluators led to some uncertainty regarding their achievability. Both independently and collectively, the documents are layered with principles, actions, and knowledge that could easily conflict in a practice setting, such as collecting data using "credible, feasible, and culturally appropriate procedures" (AEA, 2018b, p. 2) or "recognize and balance the interests of the client, other stakeholders, and the common good while also protecting the integrity of the evaluation" (AEA, 2018a, p. 4) These layered specifications provided a renewed appreciation for the complexity of evaluation work and the need to balance multiple expectations and frameworks.

Phase 1 also highlighted some similarities and differences between the documents. One obvious distinction is the style in which the documents are written. The evaluator competencies and guiding principles are presented as lists, summarizing specific knowledge, skills, and ethical considerations of evaluation practice. These lists are accompanied by minimal background or contextual information and provide limited examples or suggestions regarding their practical implications. The Statement on Cultural Competence, on the other hand, is written in a narrative style and with an educative tone. Unlike the other documents, which are more action-oriented, cultural competence is presented as a life-long undertaking.

Phase 2. A more rigorous engagement with the documents occurred in Phase 2, which involved an inductive, line-by-line analysis of each document to identify semantic and latent codes (Braun &Clarke, 2006; Braun et al., 2018, p. 846). The researcher worked through the documents three times, returning to specific sections on numerous occasions.

For this phase, and all subsequent phases, the analysis was largely conducted in MAXQDA Analytics Pro 2020 (version 20.2.1). The aim of the initial pass was to capture data that related to or provided information about knowledge or knowing in evaluation in preparation for a subsequent, more rigorous examination. To accomplish this, excerpts of interest were coded with the generic label 'knowing.' The decision to code excerpts as opposed to single words was to ensure the analysis captured the meaning underlying the data.

The second passes were intended to interrogate the excerpts to unearth the meaning captured in the data set. Guiding questions, developed in the pre-analysis, were utilized to help the researcher maintain an intentional and consistent focus and to facilitate reflexivity (Figure 3).

Figure 3

Guiding Questions for Study 1, Strand 1 RTA

> 1. How does knowing or knowledge relate to this segment?

> 2. What does this mean with respect to what and how evaluators know? How do I know that this is the case?

> 3. Who would agree and disagree with my interpretation? How, if at all, does this impact what I am seeing in the data?
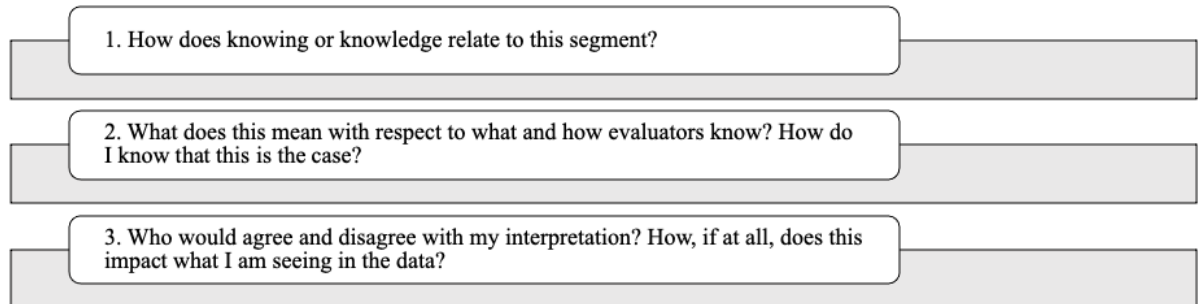
Semantic (i.e., explicit) and latent (i.e., implicit) codes were applied to the data (Braun & Clarke, 2006). To the extent possible, exact words in the document were used as labels for the semantic codes. In other instances, and for the latent codes, labels were based on the researcher's assessment of the best descriptor. There were several instances when it was not entirely clear whether a code should be semantic or latent. In these instances, the researcher thought of the code types on a continuum and labelled codes semantic or latent based on which type the segment was most like. Multiple codes were applied to the segments that involved more than one source or process of knowing.

The third pass of the data set aimed to test the initial analysis. The relevancy and adequacy of the excerpts as well as the accuracy of codes and labels was assessed. This process involved one full pass and several partial passes as there was often a need to return to previous excerpts to ensure consistency and accuracy. No excerpts were removed from the data set; however, refinements were made to code labels and some ($n=19$) codes deemed inaccurate or irrelevant to the topic were removed.

Phase 3. In Phase 3, candidate themes were developed iteratively using the codes as the basic building blocks. This process, which is commonly used in RTA, involves collating similar codes into "coherent clusters of meaning that tell a story about a particular aspect of the dataset" (Braun et al., 2019, p. 855). Initially, the codes were clustered by document to help assess the prevalence of the resulting themes across the data set. Each code was reviewed to determine their meaning or conceptual significance. Codes with similar meaning or concept were clustered and assessed to identify a single organizing concept (Braun & Clarke, 2006, 2019).

Preliminary memos were used to transform clusters into candidate themes. These memos identified and described the relevancy of the organizing concept to evaluator knowing and, to the extent possible, elaborated on potential relationships between the embodied codes. These memos were captured in the researcher's field notes and accompanied by initial code maps.

Following this, the prevalence of each candidate theme was assessed. As determined in the pre-analysis, a candidate theme was considered prevalent if it appeared in each document at least once, which was the case for each theme. Frequencies for each theme were calculated to understand how the data was distributed across themes.

Phases 4 and 5. In Phases 4 and 5, the candidate themes were reviewed and assessed to determine their soundness (Braun & Clarke, 2006). This assessment was conducted independently and collectively with the researcher's dissertation chair and critical friend. For the independent assessment, each candidate theme was examined against the raw data to ensure that it was grounded in the data. This process involved reviewing the codes and data assigned to each candidate theme to check the accuracy of its interpretation and it its relevance and quality (i.e., is

this informative) to the theme. Three codes were removed from the code system because they did not fit within the candidate themes they had been assigned to or other candidate themes.

Second, relationships and patterns within themes and among codes were analyzed using the code relations and code map functions in MAXQDA. These tools depict the intersection of codes in the data and helped to facilitate reflection on why codes overlap. Specifically, the researcher assessed whether co-occurrences indicated a relationship between codes and how this did or did not inform the candidate themes.

Phase 6. In this final phase, the themes were re-examined in light of the research questions, previous analysis, and relevant literature to ensure that they individually and collectively fit with the data and addressed the research questions (Braun et al., 2019). The data assigned to each theme was re-examined to finalize the themes, sub-themes, and patterns. No changes resulted from this re-examination.

Tables summarizing themes and subthemes were finalized. Additionally, a conceptual framework and narrative of each theme was created to represent the underlying patterns and relationships (Braun et al., 2019). These results are presented in Chapter 3.

Strand 2: Rep-grid Technique

Similar to Strand 1, the purpose of this strand was to better understand how evaluators perceive and interpret meaning from their experiences and how this shapes their practice (i.e., the ways of knowing in evaluation). Quantitative and qualitative data were collected using the rep-grid technique. This method is well-grounded in Personal Construct Theory (PCT; Kelly, 2003) and used extensively in research (Bell, 2003). Background information on PCT and the rep-grid technique is provided in Appendix E.

Participants. As mentioned in Chapter 1, the ways of knowing in nursing are perceived to be the core knowledge needed for high-quality practice and, therefore, are likely to be most apparent in experienced and competent nurses. Thus, it stands to reason that the ways of knowing in evaluation will be most apparent in evaluators who are also experienced and competent. Consequently, a purposeful sampling approach was used to obtain a sample of experienced evaluators acknowledged by the field for exemplary practice (Patton, 2002).

Two techniques were used to develop the sampling frame. First, an intensity sampling strategy (Creswell, 2013; Patton, 2002) was used to identify recipients of AEA's Outstanding Evaluation Award (*N*=23) and the Alva and Gunnar Myrdal Evaluation Practice Award (*N*=24). Recipients of the Outstanding Evaluation award have been recognized for conducting an evaluation that is "exemplary of its kind and a potential model for other evaluators doing similar kinds of work," while recipients of the Evaluation Practice award are recognized for "outstanding evaluation practice" and making "substantial cumulative contributions to the field of evaluation through the practice of evaluation and whose work is consistent with the AEA Guiding Principles for Evaluators" (AEA, n.d.-a, Award Description). Hence, it was reasonable to assume that award recipients embodied the desired levels of experience and proficiency.

Award recipients are reported publicly and were retrieved from AEA's website at https://www.eval.org/About/Awards/Past-Award-Winners. Recipients' emails were retrieved via

a web-based search. In total, 40 award recipients were invited to participate in the study, of which 12 agreed.[1] The participants were equally split across the two awards.

In addition to the intensity sampling strategy, a snowball sampling approach was used. The rationale for including this approach was two-fold. First, the approach helped expand recruitment efforts to ensure data saturation was achieved. Second, and importantly, the snowball approach helped to mitigate the bias inherent to the sampling frame due to the field's history of recognizing some perspectives over others. Of primary concern was the underrepresentation of evaluators who are Black, Indigenous, or People of Color (BIPOC) whose scholarship and contributions to the field have been historically minimalized or unrecognized (Shanker, 2019).

The snowball sample was developed by asking each participant from the intensity sample to identify exemplar practitioners. To encourage diversity in their identifications, some participants were asked to considered practitioners whose practice was similar and different from their own; however, this instruction was not always needed. This approach yielded an additional 10 participants to the study, bringing the total sample size to 22 participants.

Both samples were subject to two inclusion criteria: (1) U.S. residency and (2) ten or more years of experience practicing evaluation. The study did not collect demographic

---

[1] The researcher was unable to obtain valid email addresses for some awardees (n=7) and, therefore, was unable to invite them to participate in the study.

information from participants. However, some background information was acquired through participants' descriptions of their practice and other dialogue.

All members of the sample were university educated and had obtained a master's degree or higher. Academic backgrounds varied but were largely situated in the social sciences. Only three (13.6%) identified as BIPOC all of which were identified through the snowball sampling technique.

Participants were employed in various sectors (see Table 3). Most participants were primarily employed by private firms (n=7) or academic institutions (n=7) or independent consultants (n=6). Many members of the sample were or had been practitioner-scholars for a large proportion of their careers.

Table 3

Distribution of Participants Across Employment Sectors (N=22)

Primary Employment Sector	Number of Participants
Private firm	7
Academic institutions	7
Independent consulting	6
Government	2

Participants utilized a diverse range of evaluation approaches. Utilization-focused, Empowerment Evaluation, Social Justice, and collaborative/participatory approaches were shared by numerous participants. However, some participants also aligned their practice with Culturally Responsive Evaluation, Evaluation Capacity Building, and Indigenous Evaluation. Participants also practiced evaluation across a variety of content areas. Criminal justice,

education, social services, and health were the most discussed areas. Other areas common to a small number of participants were community engagement and public policy.

        Procedures. The value of the rep-grid method is dependent on the effectiveness of the elements in drawing out constructs that are informative to the study and the quality of the researcher's facilitation. As such, the first step following the Institutional Review Board's (IRB) approval of the research was to test the procedures and researcher's proficiency. Toward this aim, the researcher conducted a small pilot study that emulated the rep-grid planned measures and procedures. A sample (*N*=4) of evaluators participated in the pilot. Similar to study participants, pilot participants were well experienced (i.e., at least 10 years of evaluation experience) and highly regarded by members of the field.

        Generally, the pilot study results were positive in that the technique gathered information about how participants engage in and construe their practice. Some modifications to the rep-grid protocol were made. Namely, clarification of the focus of the rep-grid (i.e., topic) and the prompts best suited to focus participants' thinking. Additionally, the pilot study highlighted the diversity in evaluation approaches and practice across participants. This finding emphasized the importance of allowing for a participant-led approach to completing the rep-grid (i.e., defining the elements) as opposed to one where the researcher standardizes the measures across participants.

        Participants in the sampling frame were sent an email inviting them to participate in the study (see Appendix F). The invitations included the (1) purpose of the study, (2) participation requirements (i.e., specific activities, estimated time commitment), (3) perceived benefits to the field, and (4) a link to an expression of interest survey, hosted on Qualtrics survey software. The

expression of interest survey (see Appendix G) included questions to assess the individual's eligibility (i.e., U.S. residency and at least 10 years of evaluation experience) and a letter of informed consent. At the end of the survey, participants were redirected to an interview booking webpage hosted on Calendly.com.

The rep-gid technique was administered over Zoom, a video-conferencing platform. Each session began with a general explanation of the study and rep-grid technique and an introduction to the topic of interest. The topic, "my evaluation practice," was written on the top of each participant's grid and accompanied with the qualifying phrase, in terms of what you do and pay attention to when conducting an evaluation (Jankowicz, 2004). This topic was used because it was deemed generic enough to avoid priming participants while specific enough to focus their thinking.

Once the participant indicated that they were ready to begin, I implemented the rep-grid method as per the protocol in Appendix H. Participant data was recorded in blank grids created in Microsoft Excel (Version 16.58). Zoom's share screen option was utilized during each session to allow participants to view their grid as it was being completed. On average, each rep-grid was completed in 90 minutes. Descriptive and reflexive field notes were taken after each interview to capture observations about the participants and methodology as well as elaborate on emerging theories and the researcher's positions and role.

All interviews were recorded with permission for transcription purposes. Following the interview, each participant was debriefed and emailed a copy of their grid to review for accuracy and completeness. One participant requested a minor modification to the wording of a construct, which was honored by the researcher.

Measures. Each rep-grid consisted of four measures (1) elements, (2) constructs, (3) ratings, and (4) a summary construct. Figure 4 shows a completed grid from the study to illustrate where each measure was recorded on the grids. Further information on each measure is provided in the subsequent paragraphs.

In this study, elements were specific actions taken by participants during an evaluation. Participants defined the elements of their practice by listing eight to 12 things they typically do when conducting an evaluation. Examples of elements in Figure 4 include 'getting input from stakeholders' and 'collecting data.'

Constructs are bi-polar contrasts that individuals use to distinguish different elements of their practice (Fransella, 2015; Jankowicz, 2004; Kelly, 1991). In this regard, constructs are not concepts but opposing possibilities that represent how an individual conceptualizes different actions or tasks (Kelly, 2003, p. 10). Participants' constructs were elicited using triadic elicitation, a technique common to the rep-grid method. Specifically, the researcher presented participants with a randomly selected triad of elements and asked them to identify which two elements are similar and why (Jankowicz, 2004)[2]. The explanation of why the elements were similar represents one pole of the construct (i.e., Pole A), which was recorded on the rep-grid (Kuipers & Grice, 2009). Participants were then asked why the two similar elements differed

---

[2] Element triads were randomly generated using RANDOM.ORG, a web-based random numbers generator.

from the third element (Jankowicz, 2004). The articulated difference is noted in the second pole (i.e., Pole B). Combined, these two poles represent one bi-polar construct (Kuipers & Grice, 2009).

To use an example drawn from Figure 4, when presented with the elements 'facilitate sense making,' 'making sure evaluation expectations are aligned with improvement,' and 'reflecting on my own world view,' the participant indicated that the former were similar because the 'focus is on organization's work and purpose; action or strategies,' which was documented under Pole A. When asked why 'reflecting on my own world view' was different, the participant noted for that element, 'the focus is on my purpose and work,' which was documented under Pole B. Together, these poles form the bi-polar construct 'focus is on organization's work and purpose; action or strategies…focus is on my purpose and work.'

Ratings measured the relationships between elements and constructs. The study used a five-point rating scale, allowing for adequate discrimination without overburdening participants (Jankowicz, 2004). As described in Table 4, elements deemed by participants to be very similar or similar to Pole A were scored with a 1 or 2 respectively, while elements deemed very similar or similar to Pole B were scored with a 5 or 4, respectively. Elements considered equally similar to both poles were assigned a score of 3.

To return to our example, once the construct 'focus is on organization's work and purpose; action or strategies…focus is on my purpose and work' was elicited, the participant was asked to rate each element using the rating scale. As depicted in Figure 4, the participant rated the element 'collecting input from stakeholders' as 1, indicating that it is very similar to 'focus is on organization's work and purpose; action or strategies' (Pole A). Alternatively, they rated the

element 'seeking opportunities to automate/innovate data collection' as 5, indicating that it is very similar to 'focus is on my purpose and work' (Pole B).

Table 4

Rating Scale Used to Score Elements Against Constructs

Scores	Description
1	Very similar to Pole A
2	Similar to Pole A
3	Equally similar to both poles
4	Similar to Pole B
5	Very similar to Pole B

The final measure was a summary construct that was presented to participants to capture an overall score of how fundamental or not each element was to their evaluation practice. Participants used the same general rating scale provided in Table 4. However, in this instance, a score of one or two indicated that an element was very similar or similar to the statement 'fundamental to my evaluation practice,' where a score of four or five indicated that an element was similar or very similar to the statement 'not so fundamental to my evaluation practice.' A score of three indicated that an element was equally similar to both ends of the summary construct. Each element was rated on the summary construct and recorded on the grid.

As depicted in Figure 4, the participant scored most of the elements with a 1 or 2, suggesting that these elements are very similar or similar to the statement 'fundamental to my evaluation practice.' The only exceptions were the elements 'reflecting on my own world view' and 'seeking opportunities to automate/innovate data collection,' which were scored as equally similar to both ends of the summary construct (3) or similar to the pole 'not so fundamental to

my evaluation practice' (4), respectively. In other words, the elements with lower scores are more fundamental to the participant's evaluation practice, while the elements with higher scores are less fundamental.

Figure 4

Sample Rep-grid Completed by a Study 1 Participant

TOPIC: My evaluation practice IN TERMS OF what you do and pay attention to when conducting an evaluation.

Measure 1 (elements)

Construct Pole A	1 Getting input from stakeholders	2 Collecting data	3 Conducting analysis and interpreting the data	4 Facilitate sense making	5 Understanding the perspectives and worldviews of clients	6 Reflecting on my own world view	7 Making sure evaluation expectations are aligned with improvement	8 Seeking opportunities to automate/innovate data collection	Construct Pole B
focus is on organization's work and purpose; action or strategies (4, 7)	1	2	2	1	2	5	1	5	focus is on my purpose and work (6)
information about the research questions (2,8)	5	1	3	4	5	2	4	1	information about socio-cultural influences that can impact the data (5)
about outputs and outcomes (1,7)	2	5	5	1	3	5	1	5	about the process of evaluation (3)
about helping people to learn technical evaluation knowledge (1,4)	1	5	2	1	2	5	2	5	about using my technical evaluation knowledge (2)
recognizing the demands characteristics that come into the evaluation and how it impacts people (5,8)	1	3	5	1	1	5	1	5	recognizing the demand characteristics and how they impact you (3)
intention is understanding not necessarily improvement; engaging how clients understand the world and how to work within that (4, 5)	3	3	2	1	2	4	5	3	intention is on improvement; how to make the data useful (7)
fundamental to my evaluation practice	1	1	2	2	1	3	1	4	*not so fundamental to my evaluation practice*

Measure 2 (constructs) — rows of Construct Pole A / Construct Pole B

Measure 4 (summary construct) — final row (fundamental to my evaluation practice)

Measure 3 (ratings)

Data Analysis. Since rep-grids collect both quantitative and qualitative data, multiple methods were employed for analysis purposes. The majority of the data were analyzed at the individual-level, using principal component analysis (PCA), and at the aggregate-level, using content analysis (Honey, 1979). However, some variation in the analysis was needed to accommodate the outcomes of data collection. Namely, five participants did not complete a rep-grid and, therefore, could not be included in the PCA or content analysis. Rather, the qualitative data obtained from these participants was analyzed using reflexive thematic analysis (RTA).

Principal Components Analysis. All completed rep-grids (n=17) were subject to PCA to identify the underlying components and relationships inherent to each individual rep-grid (Kirkwood, 1988; Kuipers & Grice, 2009). The identified components represent the structure of the elements and constructs obtained by each rep-grid (Kuipers & Grice, 2009). Each PCA was conducted in Idiogrid (Version 2.4), which is a software designed for administering and analyzing rep-grid data.

The goal of the PCA was to reduce the data to the fewest number of components that could account for an adequate amount of the variability in each rep-grid (Kirkwood, 1988). The minimum number of components was determined by examining the cumulative proportion of variance explained, eigenvalues, scree plots, and loadings. Generally, components needed to have an eigenvalue great than 1 (Kaiser, 1960) and three loadings greater than 0.4 (Guadagnoli & Velicer, 1988) to be retained. An exception to the latter was made when a component on a rep-grid consisted of two highly correlated ($r = 0.75$) constructs. Since the component was interpretable and the constructs were not highly correlated ($r<0.37$) with other constructs on the

51

grid, the component was kept as part of the analysis (Worthington & Whittaker, 2006; Yong & Pearce, 2013). The results of each PCA were summarized in a table.

The components yielded from each rep-grid were interpreted and labeled accordingly. Frequencies were calculated to determine the number of unique components. Descriptions for each component were developed and included in a summary table presented in Chapter 3.

Reflexive Thematic Analysis. Data collected from participants who did not complete a rep-grid were analyzed using RTA (Braun & Clark, 2008; 2019). Similar to Strand 2, the analysis began with a pre-analysis reflection to clarify how themes would be developed and assessed as well as the perspectives that would guide the analysis (Braun & Clark, 2008). For example, themes were defined as prevalent patterns of meanings that capture important elements of how participants know or construe their practice.

Given that RTA was being applied to a sub-set of the data, the stipulations regarding how and when a theme was determined to be a prevalent pattern was purposefully flexible. Of particular concern was dismissing a theme that was not common across participants of the sub-set but may be relevant to the broader data set. As such, the analysis captured themes that were both prevalent across the data (i.e., two or more participants) or within a specific participant's data. This allowed the latter to be considered as part of the broader findings.

The pre-analysis also involved developing guiding questions for the analysis, which are included in Figure 5. These questions were integrated into all analysis phases as a means of guiding and justifying decisions and focusing the development of themes. A summary of my pre-analysis is provided in Appendix I.

Figure 5

Strand 2 RTA Guiding Questions

1. What does this tell me about this participant's practice? What does this mean with regards to what and how they know?

2. How is this similar or different from other codes, themes, and participants? What assumptions am I making to draw these comparisons?

3. How do I know my interpretation is credible? What other evidence do I have? What, other than my interpretation, could this mean?

The RTA followed the six phases outlined by Braun and Clarke (2006; 2019). Given that these phases have already been described, a high-level summary of these phases is sufficient. The researcher became familiar with the data by reviewing participant's transcripts and field notes (Phase 1). The field notes provided insights into why the rep-grid technique may not have suited these participants and the potential implications for the RTA. For example, one participant tended to describe their practice using a narrative style, whereby information on elements would be woven into stories about their experiences as an evaluator. For another participant, the elements of their practice were so inter-related it was challenging to explicate a single construct from a combination of elements.

An inductive, line-by-line analysis of the interview transcripts was conducted to identify semantic and latent codes (Phase 2; Braun &Clarke, 2006; Braun et al., 2019). These codes were clustered to form candidate themes related to what and how these participants know (Phase 3). Memos of each theme were developed that described the key organizing concept and its relationship to the codes.

53

In phases 4 and 5, the candidate themes were reviewed and assessed for their relevancy and soundness by comparing them to the raw data (Braun & Clarke, 2006). The memos were refined to document the final themes and provide supporting evidence from the data (Phase 6). Summary tables of the results are presented in Chapter 3.

Content Analysis. Content analysis (Honey, 1979) was conducted to aggregate the constructs across all rep-grids (Jankowicz, 2004). Honey's (1979) approach assumes that each individual construct provided by a participant is related to their overall perspective on the topic (Jankowicz, 2004). This overall perspective was captured during the administration of each rep-grid by asking participants to rate each element on the summary construct 'fundamental to my evaluation practice…not so fundamental to my evaluation practice.' These ratings allowed the researcher to calculate scores that aided comparisons across grids, namely percentage similarity scores and High-Intermediate-Low Index.

The content analysis was carried out in three steps. The first step was to calculate percent similarity scores for the constructs on each individual rep-grid. A percentage similarity score is a measure of how closely the ratings on a particular construct match the ratings on the overall summary construct (Jankowicz, 2004). In other words, these scores indicate the extent to which a construct relates to the individual's overall opinion of a topic (i.e., summary construct). Higher 'similarity scores' indicated greater agreement between the individual construct ratings and the overall summary construct (Jankowicz, 2004), suggesting that the individual construct is more important to the respective participant.

Percentage similarity score were calculated for each construct on every rep-grid by:

1. Calculating pairwise differences or the difference in ratings between each individual construct and the summary construct for each rep-grid (Jankowicz, 2004, p. 104).

2. Calculating the sums of pairwise differences (SPD) or the sum of the difference in ratings between each individual construct and the summary construct.

3. Transforming the SPDs into percentage similarity scores by dividing it by the product of the largest possible difference (LPD) based on the rating scale and the number of elements (E), and then multiplying by 200 (since constructs have two poles, the range of possible percentages is spread over a 200-point scale as opposed to a 100-point scale; Jankowicz, 2004, p. 115). The formula described is {SPD/[LPD x E]} x 200.

Figure 6 provides a shortened version of a rep-grid displaying the ratings for one construct and the summary construct, the pairwise differences between these ratings, and the rating scale. The SPDs is equal to 14. Since a five-point rating scale with scores from 1 and 5 was used, the LPD between any one construct and the summary construct is four (i.e., maximum rating – 1). Figure 6 also shows that there are 10 elements in this rep-grid.

Figure 6

Shortened Version of a Rep-Grid

TOPIC: My evaluation practice IN TERMS OF what you do and pay attention to when conducting an evaluation.

Construct Pole A	1 Determine what can be produced within budget	2 Determine what can be produced within the timeline	3 Identify and understand who is involved	4 Assess skill and perspective requirements	5 Chart the questions and methods	6 Determine a theoretical framework	7 Determine the right facilitative approach	8 Determine the mechanics of data collection	9 Determine the right communication dimensions	10 Analyze data	Construct Pole B
less influence on other elements (9,10)	4	4	5	5	5	5	3	3	2	3	this is pivotal; has a primary influence on other elements of the evaluation (3)
Fundamental to my evaluation practice	4	4	1	2	3	2	3	3	4	3	Not so fundamental to my evaluation practice
Pairwise Differences	0	0	4	3	2	3	0	0	2	0	

Based on these data, the percentage similarity score for the construct is 70% or:

$$\text{Percentage Similarly Score} = \left[\frac{\text{SPD}}{[\text{LPD x E}]}\right] \text{x } 200$$

$$= \left[\frac{14}{[4 \text{ x } 10]}\right] \text{x } 200$$

$$= \quad 70\%$$

The second step was to categorize each similarity score into a High-Intermediate-Low (H-I-L) index to accommodate variation in the metrics used by each participant to determine their ratings. Essentially, the H-I-L index is intended to divide the similarity scores for each rep-grid into the highest, intermediate, and lowest thirds (Jankowicz, 2004, p. 174). Constructs labelled H or I are important to how a participant construes their practice, while those labelled L are less important (Rojon et al., 2018). There does not have to be an equal number of scores in each third.

To do this, the researcher first determined the range of the data by subtracting the minimum percent similarity score from the maximum percent similarity score. Next, the range was divided by three to calculate each third of the H-I-L classification. Scores were classified as H-I-L based on which limit they fell within. For example, the range of scores in Table 5 is 50 (maximum score of 62.5 – minimum score of 12.5). 50 divided by 3 is 16.7, which was added to the minimum score to determine a low limit of 12.5 – 29. As indicated in the Table 5, two scores fell on or between the lower range and were categorized as L. Next, the intermediate limit was calculated by adding 16.7 to the top of the lower range (29) resulting in a limit of 29.3 to 46.

57

Only one score in the table fell within the intermediate limit and was categorized accordingly.

The remaining scores, which exceeded the intermediate limit were categorized as H.

Table 5

Example of Participant's Similarity Scores and High, Intermediate, or Low Scores

Constructs	% Sim. Score	H-I-L Categorization
recognizing the demand characteristics that come into the evaluation and how it impacts people…recognizing the demand characteristics and how they impact you	62.5	H
focus is on organization's work and purpose; action or strategies…focus is on my purpose and work	50.0	H
about outputs and outcomes…about the process of evaluation	12.5	L
information about the research questions…information about socio-cultural influences that can impact the data	37.5	I
about helping people to learn technical evaluation knowledge…about using my technical evaluation knowledge	50.0	H
intention is understanding not necessarily improvement; engaging how clients understand the world and how to work within that…intention is on improvement; how to make the data useful	25.0	L

Step three was to sort the constructs of each individual rep-grid into categories using the core-categorization procedure (Jankowicz, 2004). Constructs were allocated into emergent categories and sub-categories based on commonality. The constructs in each category were organized by H-I-L. Categories consisting of mostly (i.e., 60% or greater) H and I constructs were retained, while categories containing mostly L constructs (i.e., 60% or greater) were removed from further analyses (Rojan et al., 2018). The categories retained were subsequently examined to determine their meaning. The categories were summarized in a table, including the

number of constructs encompassed by the category and samples of associated H constructs. This

table is presented in Chapter 3.

Strand 1 and 2 Synthesis

In this final phase of analysis, the findings from the document review and the rep-grid

data were merged for the purposes of triangulation and elaboration. The results from each

analysis approach were summarized in a joint display table (Creswell & Plano Clark, 2018).

Comparative analysis was conducted to identify points of convergence (triangulation) and

divergence between the findings. Points of convergence were combined to create a richer and/or

more elaborate description and response to the research questions. Further analysis was

conducted to explore and explain points of divergence. The synthesis informed the development

of a framework representing the ways of knowing in evaluation, which is presented in Chapter 3.

Study 2

Design

Study 2 addressed the final research question: Howw, if at all, are the ways of knowing in

evaluation integrated into graduate programs in evaluation? The researcher conducted a

descriptive study whereby qualitative data were collected using semi-structured, key informant

interviews. Key informants for this study were faculty, recent alumni (i.e., graduated within the

last five years), and advanced students of graduate programs in evaluation. These informants

were sampled from a sub-set of graduate programs in evaluation based in the U.S.

Prior to data collection, pilot tests were conducted to examine the feasibility and

effectiveness of the planned data collection procedures and interview protocols. A sample of

evaluation professors ($n=2$) and advanced students ($n=2$) participated in the pilots. The pilot

results aided in the refinement of questions and prompts to better capture participants'

experiences.

      Selection of Graduate Programs. Graduate programs were identified through a two-step

process. First, the researcher reviewed LaVelle's (2018) directory of evaluator education

programs and selected programs that required four or more evaluation-specific courses as part of

their curriculum.[3] Graduate programs that met this threshold were added to the study's sampling

frame.

      While the directory was a useful starting point for identifying programs, it is not

comprehensive. As such, a second step was added to the process to help ensure a broad

representation of evaluation programs. Specifically, the researcher reviewed the recipients and

alma maters of AEA's Marcia Guttentag New Evaluator Award (2011-2021). Recipients of this

award have completed a master's or doctoral degree within five years and been deemed

promising new evaluators (AEA, n.d.-b). It was assumed that evaluation programs that trained

award recipients deliver quality evaluation education and, as such, were added to the study's

sampling frame.

---

[3] Claremont Graduate University was not included on the list due to its proximity to the researcher.

In total, nine educational programs were identified through the two-step process. The researcher could not accurately predict the sample size needed to address the research question. As such, an initial sample was drawn from five randomly selected programs.

The researcher evaluated saturation from two perspectives (Saunders et al. 2018). Thematic saturation was understood as the point when the ways of knowing in evaluation educational programs had been adequately represented (Starks & Trinidad, 2007). Data saturation was determined when novelle, meaningful insights had waned, and the data was repetitive (Grady, 1998). Roughly two-thirds through data collection, the researcher became aware of data redundancies; however, some unique insights persisted that were informative to the ways of knowing. Therefore, an additional program was randomly selected for the study bringing the total number of programs to six.

As data collection continued, novelle and relevant insights waned. The researcher also felt that the data collected from the sample had adequately captured the ways of knowing in evaluator education.[4] The remaining scheduled interviews were completed, and no new participants were recruited. For confidentiality purposes, the programs included in the study will not be disclosed in any reporting or communications resulting from the study.

---

[4] Based on the researcher's field notes, data saturation was achieved around interviews 8 (faculty) and 13 (alumni/advanced students).

Specific details regarding program curriculum were not analyzed as part of the study. However, some general observations can be made about the programs. Curricula for each program required courses in evaluation, research methods, and statistics. Specific courses in evaluation theory and practice were a requirement across all programs. Additionally, each program provided opportunities for practical experience in evaluation through courses and project work with some programs requiring internships or practicum placements.

Key informant sampling. All core evaluation faculty (n=18) and recent alumni (n=25) from each program were invited to participate in the study. Core faculty were defined as faculty who taught required courses in evaluation and had advisory responsibilities for students pursuing research and/or careers in evaluation. These faculty were identified with the assistance of each program's program director or other key program contacts. Of the faculty invited to participate in the study, 11 agreed.

Alumni who graduated within the last five years from the program were identified through a dissertation database search in ProQuest. Of those invited, 11 ultimately participated in the study. The researcher was interested in obtaining a similar number of alumni across programs; however, the recruitment resulted in uneven representation for various reasons. One participant was the only recent graduate (i.e., graduated within the last five years) confirmed by the researcher. Additionally, minimal or no response to the study invite was received from two other programs. As such, five advanced students, identified through program websites and faculty participants, were recruited to offer additional insights on the underrepresented programs. In total, 27 faculty, alumni, and advanced students participated in the study. Table 6 provides a breakdown by program and group.

Table 6

Participants by Program

Program	Faculty	Alumni	Advanced Students	Program Totals
1	3	1	1	5
2	2	1	1	4
3	3	-	3	6
4	2	3	-	5
5	1	3	-	4
6	0	3	-	3
Group Totals	11	11	5	27

Procedures

Qualitative data were collected using semi-structured interviews conducted over Zoom. The purpose of the interviews was to collect information regarding the knowledge acquired by students studying evaluation at the graduate-level. Potential informants were sent an invitation email (Appendix J & K). Among other topics, the email introduced the researcher, described the study, and requested their participation in an interview. The invite also included a link to an informed consent letter (Appendix L & M). At the end of the letter, interviewees were directed to an interview scheduling page that allowed them to schedule an interview at their convenience.

The interviews began with a brief overview of the study and the focus of the interview. Pertinent information regarding participant confidentiality and data security were also discussed. Each interview, including debrief, lasted approximately 60 minutes. Summaries of discussions, emerging themes, observations, and reflections were captured in the researcher's field notes. The interviews were recorded with permission and transcribed for analysis purposes.

Measures

The aim of the interviews was to measure if and how the ways of knowing in evaluation were integrated into the educational experiences offered by the selected programs. Interview guides consisting of open-ended questions and prompts were developed and utilized for the faculty and alumni/advanced student interviews.

The faculty interview guide consisted of four sections focused on participants' experiences and approaches to teaching and advising evaluation graduate students. Participants were asked about the courses they have taught as well as other aspects of their role that involved educating students. Participants were also asked about the strengths and limitations of their program with regards to preparing students for practice and their personal goals and aspirations for preparing future evaluators. A copy of the faculty interview guide is provided in Appendix N.

The alumni and advanced students interview guide began with a warm-up section to gather background and contextual information as well as develop rapport. Importantly, this section gathered data regarding the knowledge and experience that students brought to the program so it could be taken into consideration as part of the analysis. Section two of the interview guide aimed to gain an understanding of participant's experiences while in the program. Of particular interest were those experiences that influenced and shaped participants' evaluation knowledge and practice. Similar to faculty participants, alumni and advanced students were asked to share their perceptions on the strengths and limitations of the program. A copy of the interview guide is provided in Appendix O.

64

Data Analysis

      The data were analyzed using a directed (deductive) content analysis based on the Elo

and Kyngäs (2008) perspective. This analysis is typically used to test a theory in a new context

or garner new information about an existing theory (Elo & Kyngäs, 2008; Hsieh & Shannon,

2005). In the current study, the analysis was intended to examine the presence of the ways of

knowing in evaluator education. The analysis was conducted over a series of steps, which are

summarized in Figure 7.

Figure 7

Directed Content Analysis Process

Step 1
Establish a priori codes based on Study 1 results

Step 2
Structured data coding based on a priori codes

Step 3
Review and describe the coded data

Step 4
Create summary memos
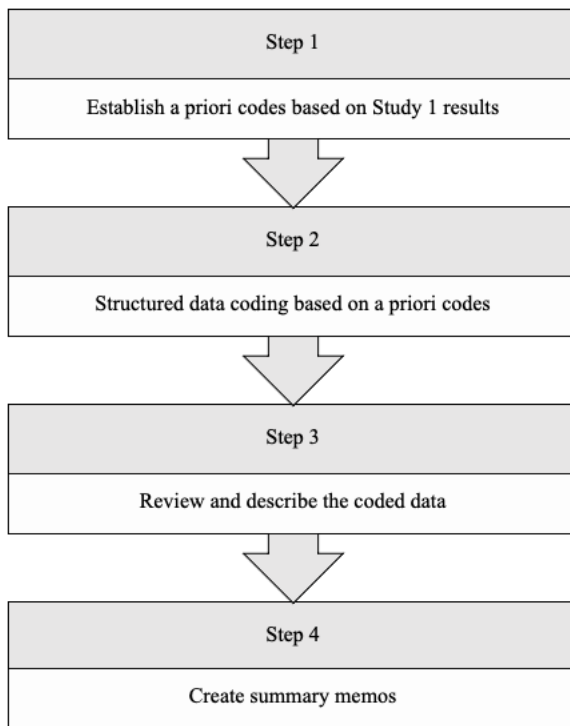
The first step in the analysis was to establish a priori categories based on the findings from Study 1. Specifically, the ways that evaluators know and understand their practice as well as the underlying dimensions involved in knowing informed six overarching a priori codes. These a priori codes were organized in a matrix and accompanied by descriptions and potential examples to aid the analysis. The latter were drawn from the researcher's field notes.

The a priori codes were tested on a sample of data to examine their suitability and feasibility for the analysis. The pilot resulted in some minor modifications to one code (i.e., knowing in action). Specifically, it appeared that potential examples accompanying this code were too specific for the data and were removed. The final code matrix is provided in Table 7.

Table 7

Study 2 A Priori Code Matrix

Codes	Description	Potential Examples
Knowing self	Knowing grounded in one's awareness of and connection to the genuine self (Chinn & Kramer, 2018). This knowing is developed through rigorous and honest self-examination and critical reflection on one's history, socialization, culture, assumptions, and perspectives (Mezirow, 1990).	Discussion/activities that lead to greater self-awareness. Might include social identity maps, journaling, individual and dyadic reflection.
Knowing others	Knowing that stems from one's ability to understand the perspectives and experiences of others; fostered through empathy and gaining close proximity to the lived experiences of others, while recognizing that one can never fully understand others (AEA, 2011; Schwartz & Sharpe, 2010).	Discussion/activities that enhance students' ability to relate to others. Might include perspective-taking, training in facilitation, engagement, and conversation, and emphasis on understanding others' histories and world views.
Knowing the discipline	Knowing acquired by the integration of formal theories, scientific methods, and professional values and principles into practice (Chinn & Kramer, 2018).	Training related to evaluation theory, research methods, evaluation procedures, evaluation ethics.
Knowing the common good and equity	Knowing that evolves from one's appreciation of a fair and just world and sensitivity to what is possible. This includes "the ability to recognize barriers that create unfair and unjust social conditions and to analyze complex elements of the sociopolitical context to change a situation that improves people's lives" (Chinn & Kramer, 2018, p. 290).	Discussion/activities about the history of evaluation and its role in society and relationship to marginalization, structural racism, and inequity.
Knowing in action	Creative processes and strategies that help evaluators to perceive and understand their practice.	

Expressions of knowing	A way of expressing what has become known that is integrated into how evaluators engage in practice.	Training on reporting and communications and model development

Next, a structured, line-by-line analysis of the interview transcripts was conducted in MAXQDA Analytics Pro (Version 20.2.1). Consistent with this approach, data that corresponded with an a priori codes were coded (Elo & Kyngäs, 2008). Two complete passes and several partial passes of the data were required to complete this analysis. Data that were not assigned a code were reanalyzed to identify data that may have been missed in the initial coding process (Assarroudi et al., 2018). While this step did not result in any additional data, it enhanced the trustworthiness of findings by reducing the risk that pertinent data was missed during coding.

Following this, the data assigned to each a priori code was reviewed to interpret meaning, seek points of convergence and divergence, and explore nuances. Emergent sub-codes were assigned to the data to describe the contents of the a priori codes, namely instances that illustrated the ways of knowing or the underlying dimensions. Specific program components that corresponded with these instances were noted when possible.

Instances where alumni and advanced student perspectives differed from faculty perspectives were examined and documented. The a priori codes and sub-codes were compiled in a table and accompanied by exemplars from the data. An abbreviated version of this table is provided in Appendix P.

Last, summary memos for each program were created, which described if and how the ways of knowing were and were not captured by the data. Supporting evidence was used to ground the assessment of the data. These memos were combined into a single report, which is presented in Chapter 3.

Chapter 3: Results

Study 1

Study 1 addressed the first three research questions: (1) what are the fundamental ways of knowing in evaluation? (2) What dimensions underlie the fundamental ways of knowing in evaluation? And, (3) how do these dimensions manifest in practice? The results are first presented by strand to highlight the findings garnered by each data source (i.e., documents and practitioners). Following this, the results of both strands are synthesized to address the research questions.

Strand 1

In Strand 1, Reflexive Thematic Analysis (RTA) was carried out on data obtained from three foundational documents of evaluation. As discussed, these documents reflect what the AEA community views as required and important to exemplary evaluation practice. As shown in Table 8, a total of 98 segments were coded and comprised the data set. The AEA (2018) Evaluator Competencies yielded the most coded segments ($n=47$), followed by the Statement on Cultural Competence ($n=28$), and the Guiding Principles ($n=23$). The large number of coded segments retrieved from the Evaluator Competencies was not surprising given the focus and style of the document. Essentially, that document is a list of skills, actions, attributes, and qualities for competent evaluators, many of which are accompanied by an underlying knowledge source(s).

Table 8

Frequency of Coded Segments by Document

Document	Number of Coded Segments
AEA Evaluator Competencies	47
AEA Statement on Cultural Competence	28
AEA Guiding Principles	23
Total	98

A total of 22 unique codes were assigned to the data. These codes represent various knowledge areas (e.g., methods, ethics), processes (e.g., adhering, clarifying) and expressions (i.e., creating logic models) related to knowing. The codes were clustered into three overarching themes: (1) the ways of knowing in evaluation, (2) knowing in action, and (3) expressions of knowing. A map of themes and codes is provided in Figure 8.

Figure 8

RTA Themes and Codes
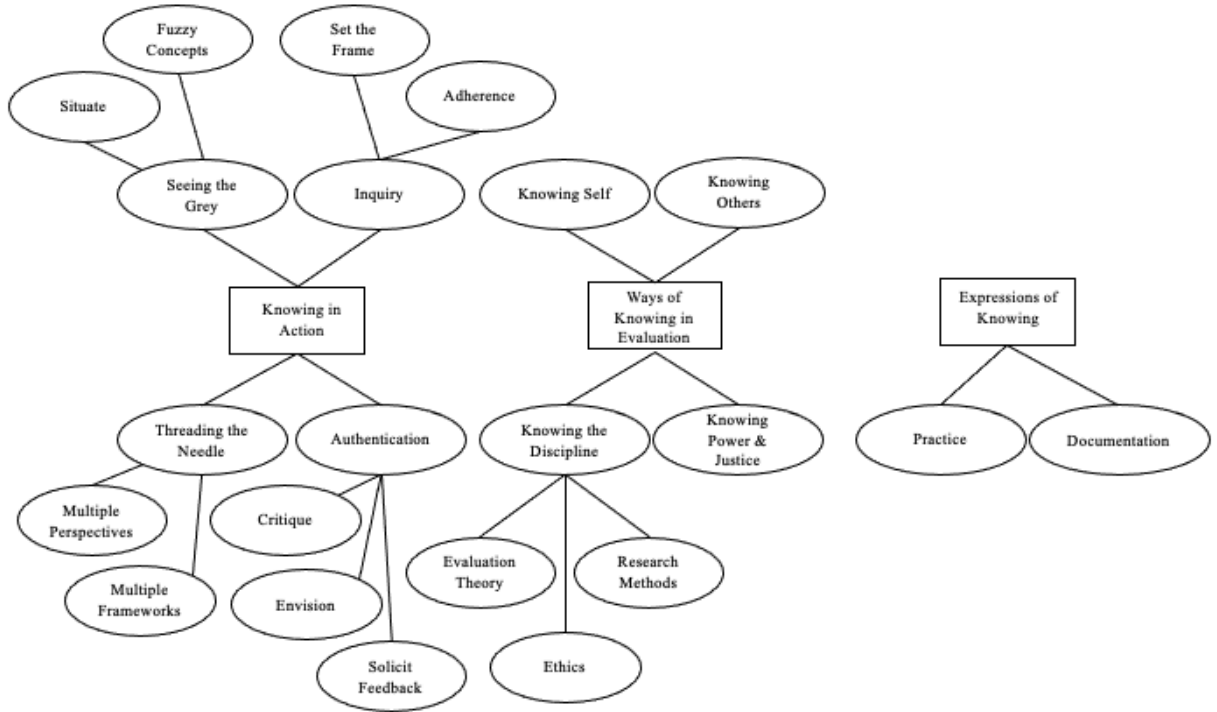
Ways of Knowing in Evaluation. The first theme is the ways of knowing in evaluation, which represents what one must know and understand to practice evaluation. It encompasses four sub-themes (1) knowing self, (2) knowing others, (3) knowing the discipline, and (4) knowing justice and power.

Knowing Self. Knowing self is grounded in one's awareness of and connection to the genuine self (Chinn & Kramer, 2018). These data connect knowing self to various aspects of evaluation practice. Potentially, the most obvious connection is to professional improvement. Self-knowing can help evaluators to identify "personal areas of professional competence and needs for growth," "improve practice," and ensure the evaluation team, "possesses the education,

abilities, skills, and experiences required to complete the evaluation competently," and "work in the cultural context of the evaluation."

However, the data goes beyond knowing self for professional improvement to encouraging evaluators to engage in self-exploration to enhance their self-awareness. Self-awareness involves understanding one's "background" and "privilege and positioning." For example, the Guiding Principles state that evaluators should "maintain a high degree of self-awareness and self-examination to better understand how their own backgrounds and other life experiences serve as assets or limitations in the conduct of an evaluation." Evaluators are also encouraged to discuss "the values, assumptions, theories, methods, results, and analyses that significantly affect" their interpretation of the evaluation findings and be explicit about their "values, perspectives, and interests concerning the conduct and outcome of the evaluation."

Knowing Others. Knowing others stems from one's ability to understand the perspectives and experiences of others (Schwartz & Sharpe, 2010). Overwhelmingly, the message presented in the data is the need "to gain an understanding of" and clarify "the diverse perspectives, stakeholder interests, and cultural assumptions" involved in an evaluation. In this sense, knowing others involves both breadth and depth in one's understanding of the people involved with and impacted by the evaluation.

The data offers many practical reasons for knowing others, including setting the stage for the client's engagement in "designing, implementing, interpreting, and reporting evaluations" and positioning them to "understand, interpret, and critique the work." Additionally, evaluators must consider "contributors to human behavior" and "biases, stereotypes, and lack of shared

world view" to draw valid conclusions. Importantly, knowing others helps evaluators respond "respectfully to the uniqueness of the evaluation context."

Knowing the Discipline. Knowing the discipline involves the integration of formal theories, scientific methods, technical knowledge, and professional values and principles into practice (Chinn & Kramer, 2018). This form of knowing could be viewed as most akin to theoretical knowledge since it demands an understanding of "evaluation approaches and theories" and "program logic and program theory. The data also indicates that evaluators need to understand "a wide range of evaluation theories and methods to design and carry out an evaluation that is optimally matched to the context."

The data also emphasized an evaluator's ability to carry out evaluation studies, which involves evaluation-specific procedural knowledge. For example, evaluators identify "evaluation purposes and needs," determine "evaluation questions," and use "systematic evidence to make evaluative judgments." In some respects, this knowledge helps to differentiate evaluation from other disciplines.

Yet, evaluation work is also supported by knowledge that is not specific to the field, such as research methods, project management, and communication. For instance, the data indicates that "evaluators conduct data-based inquiries that are thorough, methodical, and contextually relevant." Evaluation also involves negotiating and managing a "budget, resources, and timeline" and communicating in "meaningful ways that enhance the effectiveness of the evaluation." To some extent, these non-evaluation-specific skills illustrate the transdisciplinary nature of evaluation.

74

Last, knowing the discipline involves adhering to ethics. For instance, evaluators must "abide by current professional ethics, standards, and regulations (including informed consent, confidentiality, and prevention of harm) pertaining to evaluation participants" and "adhere to the highest technical standards appropriate to the methods being used while attending to the evaluation's scale and available resources." As such, evaluators must attend to the ethical principles of their field and research.

The data indicates that ethics also informs how evaluators behave. For example, an evaluator "acts ethically through evaluation practice that demonstrates integrity and respects people from different cultural backgrounds and indigenous groups" and treats "fairly, the range of perspectives and interests that individuals and groups bring to the evaluation."

Knowing Justice and Power. Knowing justice and power falls under the broad assumption that evaluation is value-based and "entails a view of society" (Hamilton, 1977, p. 25; Greene, 2013). This sub-theme is grounded in values for equity and justice and a commitment to advancing these values through practice. Of the sub-themes, knowing justice and power appears to be the most closely associated with evaluation's ultimate purpose of social betterment (Greene, 2013; Henry & Mark, 2003; Mark et al., 1999; Shadish, 1994).

Knowing justice and power involves being socially conscious. According to the data, socially conscious practice involves attending "to the ways power and privilege affect evaluation practice" and "systems issues within the context." Additionally, evaluators should be "aware of marginalization" which requires an understanding of the intersection between people's histories, culture, identities, and experiences, and must embrace "an ethical commitment to fairness and

equity." Last, the data note that evaluators are required to be aware of "the goals of a democratic, equitable, and just society."

Awareness alone is not sufficient in advancing changes to attitudes, programs, or policies that will foster a better society. As such, the data also emphasize a proactive stance towards justice and inequity. This stance includes the ability to "identify and make efforts to address the evaluation's potential threats to the common good," "consider intended and unintended social consequences in the overall assessment of their work," "use their power to promote equality and self-determination" and "maximize the benefits and reduce unnecessary risks or harms for groups and individuals associated."

The ways of knowing in evaluation theme accounted for just over half (n=98 or 51%) of coded segments. Among its sub-themes, knowing the discipline (n=33) and knowing justice and power (n=26) were the largest sub-themes, followed by knowing others (n=16) and knowing self (n=15). Table 9 summarizes each sub-theme.

Table 9

Summary of Ways of Knowing in Evaluation Sub-themes

Subtheme (*frequency*)	Description	Example Excerpts
Knowing the discipline (*33*)	Knowing acquired by the integration of formal theories, scientific methods, and professional values and principles into practice (Chinn & Kramer, 2018).	• "Applies the foundational documents adopted by the American Evaluation Association that ground evaluation practice." • "The culturally competent evaluator draws upon a wide range of evaluation theories and methods to design and carry out an evaluation that is optimally matched to the context."
Knowing justice and power (*26*)	Knowing that evolves from one's appreciation of a fair and just world and sensitivity to what is possible. This includes "the ability to recognize barriers that create unfair and unjust social conditions and to analyze complex elements of the sociopolitical context to change a situation that improves people's lives" (Chinn & Kramer, 2018, p. 290).	• "Identify and make efforts to address the evaluation's potential risks of exacerbating historic disadvantage or inequity." • "Consider intended and unintended social consequences in the overall assessment of their work."
Knowing others (*16*)	Knowing that stems from one's ability to understand the perspectives and experiences of others; fostered through empathy and gaining close proximity to the lived experiences of others, while recognizing that one can never fully understand others (AEA, 2011; Schwartz & Sharpe, 2010).	• "Strive to gain an understanding of, and treat fairly, the range of perspectives and interests that individuals and groups bring to the evaluation, including those that are not usually included or are oppositional." • "Clarifies diverse perspectives, stakeholder interests, and cultural assumptions."

Knowing self (*15*)	Knowing grounded in one's awareness of and connection to the genuine self (Chinn & Kramer, 2018). This knowing is developed by being present and engaging in rigorous and honest self-examination and reflection on one's history, socialization, culture, assumptions, and perspectives (Mezirow, 1990).	• "Identifies personal areas of professional competence and needs for growth." • "Cultural competence in evaluation requires that evaluators maintain a high degree of self-awareness and self-examination to better understand how their own backgrounds and other life experiences serve as assets or limitations in the conduct of an evaluation."

Additional analysis on code relations yielded some informative insights about the ways of knowing in evaluation sub-themes. Namely, the sub-themes are interrelated, suggesting that evaluators often rely on two or more ways of knowing when engaged in practice. Figure 9 depicts a map of knowing self, knowing others, knowing the discipline, and knowing justice and power. The connecting lines between sub-themes indicate co-occurrence (i.e., instances where sub-themes were applied to the same data segment), with thicker lines and closer proximity of codes indicating higher instances of co-occurrence.

Figure 9

Co-occurrence of Ways of Knowing in Evaluation Sub-themes
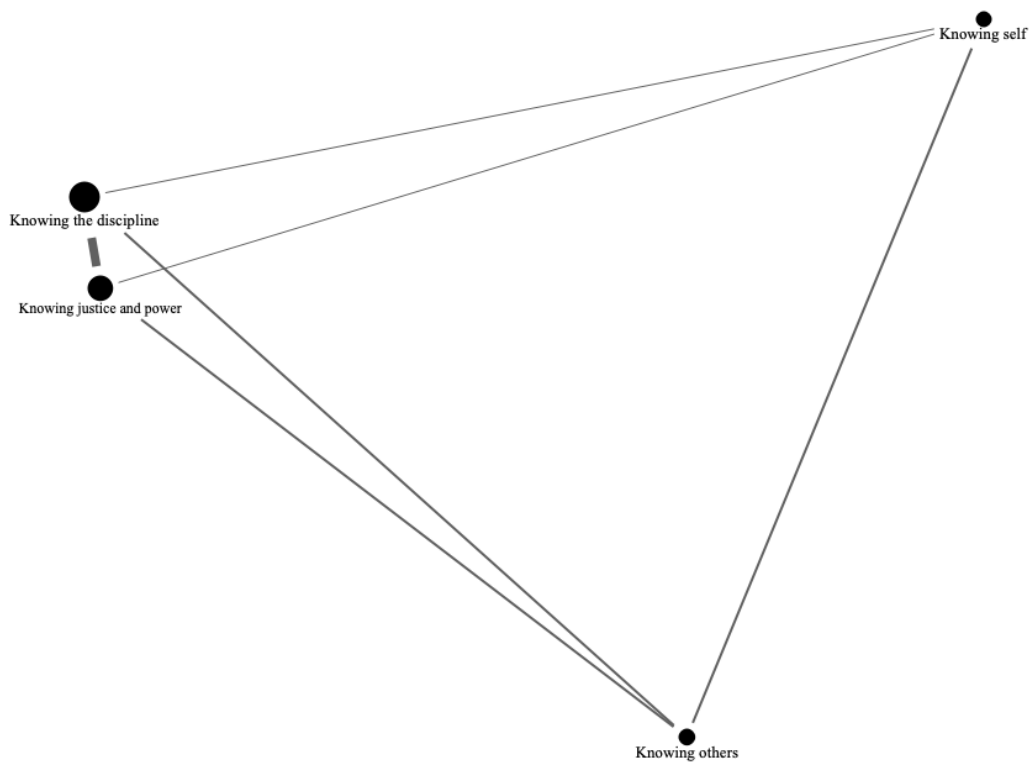
As depicted in Figure 9, these sub-themes co-occur in the data to varying degrees. The thick line between knowing the discipline and knowing justice and power indicates a high degree of overlap relative to other ways of knowing. This could suggest a complementary aspect or unique importance about these sub-themes.

Combined, knowing self, knowing others, knowing the discipline, and knowing justice and power represent the ways of knowing in evaluation inherent to the data set. These ways of knowing provide the foundation for the remaining themes, knowing in action and expressions of knowing.

Knowing in Action. The second theme, knowing in action, describes the creative processes that foster knowing (Chinn & Kramer, 2018). Schön (1983) argues that while "we sometimes think before acting, it is also true that in much of the spontaneous behavior of skillful practice we reveal a kind of knowing, which does not stem from a prior intellectual operation" (p. 51). This knowing is driven by dynamic processes that are inherent in our practice and informed by our ways of knowing (Chinn & Kramer, 2018; Schön, 1983, p. 50). Knowing in action consists of four sub-themes, which are described below.

Seeing the Grey. The first sub-theme is seeing the grey, which relates to how evaluators interpret the nuances of their work to make judgments about their practice. As identified in the pre-analysis, the data includes numerous 'fuzzy' concepts that lack a universal definition and need to be interpreted in context (House, 2015, p. 91; Schwartz & Sharpe, 2010). Table 10 provides some examples of fuzzy concepts included in the data.

Table 10

Examples of Data Segments that Involve Fuzzy Concepts

Data Segment	Fuzzy Concepts
"Identify and make efforts to address the evaluation's potential threats to the common good especially when specific stakeholder interests conflict with the goals of a democratic, equitable, and just society."	Common good; equitable; just
"Strive to maximize the benefits and reduce unnecessary risks or harms for groups and individuals associated with the evaluation."	Benefits; unnecessary risks or harms; groups/individuals associated with the evaluation
"Promotes evaluation use and influence in context"	Use; influence; context
"Collects data using credible, feasible, and culturally appropriate procedures."	Data; credible; feasible; culturally appropriate
"Culturally competent evaluation emerges from an ethical commitment to fairness and equity for stakeholders."	Culturally competent evaluation; fairness; equity
"Promote full participation when evaluation activities are conducted in participants' primary or preferred languages. This includes consideration of culturally specific communication styles and mannerism."	Participation; languages; culturally specific communication styles and mannerisms

These fuzzy concepts are evidence of the greyness of our work, or what Schwandt (2003) refers to as the "rough ground" (p. 355). To properly apply these concepts, evaluators must be able to see the grey or determine what fuzzy concepts mean in a specific context. As part of this process, evaluators "maintain or fail to maintain core values of their practice in the face of contingencies, constraints and contradictions" (Schwandt, 2003, p. 355). As such, seeing the grey involves a conscious choice about what values are more or less important.

The data provides information on one process that evaluators use to see the grey, which is to situate their practice in the evaluation and broader contexts. For example, evaluators are to "design and carry out an evaluation that is optimally matched to the context," including the cultural context, describe "the program, including its basic purpose, components, and its

81

functioning in broader contexts," interpret "findings/results in context," and attend "to systems issues within the context."

Inquiry. The second sub-theme is inquiry, which relates to the evaluator's role in conducting evaluation studies to garner evidence. This role involves knowledge of research methods, evaluation theory, and procedural knowledge, which enable evaluators to collect and analyze data. For example, an evaluator conducts "data-based inquiries that are thorough, methodical, and contextually relevant," "determines and justifies appropriate methods to answer evaluation questions," and must be aware of the "many ways data can be analyzed and interpreted." However, systematic inquiry can only occur once evaluators have determined where to focus their inquiry, which can be challenging when working with complex evaluands that involve competing interests and diverse perspectives.

As such, the first step is to set a frame around what will become the focus of inquiry (House, 2015; Schön, 1983; Schwartz & Sharpe, 2010). According to the data, an evaluator "listens to understand and engage different perspectives" and "clarifies diverse perspectives, stakeholder interests, and cultural assumptions," which are important to understanding what is important to consider as part of the evaluation (AEA, 2018b, p. 3). Additionally, evaluators determine the "evaluation purposes and needs" and "evaluation questions," which set the frame for the evaluation.

Another dimension of inquiry discussed in the data is adherence. In the process of implementing an evaluation, evaluators are required to adhere to various standards related to methodology, culture, and ethics. Some of the standards from the data include the "technical standards appropriate to the methods being used," methods that "reflect the cultures in which

they were developed," the "foundational documents adopted by the American Evaluation

Association," and "professional ethics, standards, and regulations."

Threading the Needle. The third sub-theme, threading the needle, represents evaluator

efforts to balance expectations and find common ground among competing interests. Sometimes,

these efforts involve people and perspectives. For instance, evaluators must balance "the range of

perspectives and interests those individuals and groups bring to the evaluation, including those

that are not usually included or are oppositional," "multiple and sometimes clashing norms," and

"the interests of the client, other stakeholders, and the common good while also protecting the

integrity of the evaluation."

Other times, threading the needle relates more to the technical aspects of evaluation. For

instance, evaluators must determine how to collect and analyze data "using credible, feasible,

and culturally appropriate procedures" and "adhere to the highest technical standards appropriate

to the methods being used while attending to the evaluation's scale and available resources." In

this sense, evaluators must find a specific path or option that allows them to adhere to multiple

frameworks.

Authentication. The last sub-theme is authentication, which provides details regarding

how evaluators validate what has become known through processes such as critiquing,

envisioning, and soliciting feedback. Critiquing involves critically examining both the human

and methodological dimensions of evaluation studies. For instance, evaluators are to "assess and

make explicit the stakeholders', clients', and evaluators' values, perspectives, and interests

concerning the conduct and outcome of the evaluation" and "engage in ongoing critical

reflection on assumptions about what constitutes meaningful, reliable, and valid data and how these data are derived."

Envisioning, or the ability to imagine a future state, requires evaluators to "consider intended and unintended social consequences in the overall assessment of their work," identify "how evaluation practice can promote social justice and the public good," and "identify and make efforts to address the evaluation's potential risks of exacerbating historic disadvantage or inequity." The analysis indicates that envisioning and critiquing often overlapped (i.e., co-occurred) in the data. This may suggest that evaluators consider both the current and future possibilities when critiquing.

The final authentication process is soliciting feedback from partners and those impacted by the evaluation. For instance, evaluators "communicate methods and approaches accurately, and in sufficient detail, to allow others to understand, interpret, and critique the work," "explore with primary stakeholders the limitations and strengths of the core evaluation questions and the approaches that might be used for answering those questions," and "select or create data collection instruments that have been (or will be) vetted for use with the population of interest." Gathering feedback from others helps to authenticate the evaluation procedures and the knowledge that is generated through the evaluation.

Collectively, knowing in action, and its sub-themes represent the dynamic nature of knowing and the processes used by evaluators to understand their work. Similar to the ways of knowing, a large portion of the data was assigned to this theme (n=82 or 46%). Among the knowing in action sub-themes, authentication had the highest number of codes (n=38), followed

by inquiry (*n=18*), seeing the grey (*n=15*), and threading the needle (*n=11*). A summary of the

knowing in action sub-themes, including their frequencies is provided in Table 11.

Table 11

Summary of Knowing in Action Sub-themes

Subtheme (*frequency*)	Description	Excerpts
Authentication *(38)*	"Assessing the soundness" of what has become known (Chinn & Kramer, 2018, p. 287)	• "Uses systematic evidence to make evaluative judgments." • "Explore with primary stakeholders the limitations and strengths of the core evaluation questions and the approaches that might be used for answering those questions."
Inquiry *(18)*	The use of contextually and culturally appropriate research methods to garner evidence; stems from the belief that evaluation processes can lead to new understandings.	• "Employ data collection and analysis methods that address cultural differences in how knowledge is constructed and communicated." • "Data collection methods and tools reflect the cultures in which they were developed."
Seeing the grey *(15)*	The ability to see and contextualize the nuances of one's work. Interpreting meaning from amorphous concepts and contexts (Schwartz & Sharpe, 2010).	• "Identifies how evaluation practice can promote social justice and the public good." • "Self assess one's own privilege and positioning within that context."
Threading the needle *(11)*	Achieving symmetry and balance between multiple frameworks, principles, perspectives, and so on.	• "Recognize and balance the interests of the client, other stakeholders, and the common good while also protecting the integrity of the evaluation." • "Collects data using credible, feasible, and culturally appropriate procedures."

Expressions of Knowing in Evaluation. The theme expressions of knowing in evaluation relates to how evaluators express what has become known as a result of their work (Chinn & Kramer, 2018). According to these data, one way that evaluators express knowing is their practice. For example, when an evaluator "acts ethically through evaluation practice that demonstrates integrity and respects people from different cultural backgrounds and indigenous groups" they express ethical knowledge (i.e., knowing the discipline) and cultural knowledge (knowing others). Making "evaluative judgments" and "recommendations" are also ways that knowing is expressed through practice that was interpreted from the data.

Another way that evaluators express knowing is through documentation and theory development. The data discusses "program logic and program theory," and program descriptions, which are examples of ways evaluators share and communicate knowledge with others. These expressions of knowing are very much grounded in the discipline.

The number of segments coded as "expressions of knowing" was small relative to the other themes (*n*=6 or 1%). However, the data suggests several reasons why it is important to express what is known, such as knowledge sharing, facilitating learning and participation among evaluation partners, and supporting good practice. As such, this theme is directly linked to important values of the profession such as evaluation use and influence.

Summary of Strand 1 Findings

Collectively, the themes and sub-themes identified through the RTA form an initial framework describing the underlying knowledge and processes of knowing involved with evaluation practice (i.e., epistemology of evaluation practice) which is depicted in Figure 10. The ways of knowing in evaluation is the epistemological (i.e., underlying knowledge) basis of

evaluation practice. Each sub-theme represents an awareness or consciousness that informs evaluators' perceptions and judgments.

Evaluators' ways of knowing are supported by knowing in action, which are creative processes that activate and sustain knowing. Seeing the grey, inquiry, threading the needle, and authentication are processes that were interpreted from the data. However, these processes are likely just a sample of those used by evaluators to leverage their ways of knowing and create new knowledge.

The knowledge generated by evaluators are expressed through their practice. Expressions of knowing included in the data are evaluator's practice, the way we present ourselves (e.g., acts ethically) and relate to others, as well as documentation such as logic models, program theory, and program descriptions.

Figure 10

Initial Conceptual Framework: Ways of Knowing in Evaluation and the Dimensions of Knowing
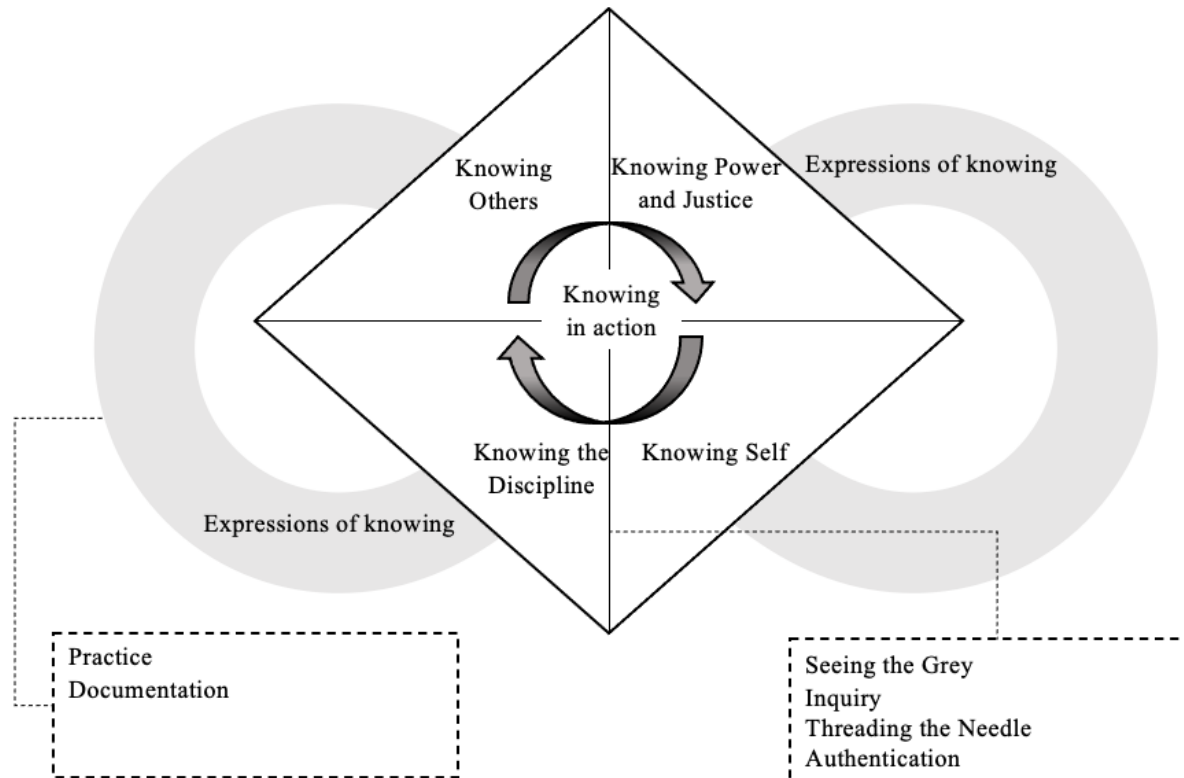
Strand 2

As mentioned in Chapter 2, the rep-grid data collected in Strand 2 was subject to principal components analysis (PCA) and content analysis (Honey, 1974), while the qualitative data was subject to a reflexive thematic analysis (RTA; Braun & Clarke, 2008; 2019). While the conceptual framework established in Strand 1 was not deliberately used to guide these methods, the researcher remained conscious of reoccurring themes and codes and used similar descriptive language (e.g., code labels) when it was appropriate to do so. The results of each analysis technique are described below.

PCA. Rep-grid participants elicited a total of 180 elements and 111 bi-polar constructs. The number of elements elicited by participants ranged from eight to 12. Most participants elicited six constructs, not including the summary construct provided to them. All constructs, including the summary construct, were included in the PCA (Jankowicz, 2004). As mentioned, PCA was used to identify the underlying components of each individual rep-grid. Each component from the PCA represents a set of constructs that form a pattern of variability in the data. Some data cleaning was required before the components could be interpreted.

Seven constructs were removed from the analysis due to low communality (<0.4.). One construct with a low communality (0.37) was retained because of its relevance to its respective component. An additional 12 constructs were removed for cross-loadings. Generally, the threshold for cross-loading was 0.4. However, five exceptions were made when constructs loaded highly (0.69 or greater) on one component and low on another component (<0.45). In all instances except one, components consisted of three constructs. An exception was made when two constructs that were highly correlated (r=-0.75) with each other, but not highly correlated

with other constructs loaded onto one component (Worthington & Whittaker, 2006; Yong & Pearce, 2013).

The analysis yielded a total of 25 components. Most rep-grids (53%) contained one unique component, 47% of rep-grids contained two unique components. A summary of the components, including eigenvalues, communalities, and loadings is provided in Appendix Q.

The components were rotated to aid interpretation. Since the constructs that formed each component were assumed to be related, an oblique (promax) rotation was used. However, in some instances an orthogonal (varimax) rotation was also applied to clarify the underlying meaning of the components. Table 12 provides the components by rep-grid.

Table 12

Summary of Components by Rep-Grid (N=17)

Rep-grid	Component 1	Component 2
1.	Perspective-taking	
2.	Role clarity	
3.	Seeking alignment	Facilitating use
4.	Fluidity	Managing ambiguity
5.	Engaging others	Creating an audit trail
6.	Role clarity	Authenticating
7.	Fluidity	Understanding personal value
8.	Facilitating use	Engaging others
9.	Seeking alignment	
10.	Making an evaluative judgment	
11.	Fluidity	Role clarity
12.	Enhancing validity	
13.	Adhering	
14.	Engaging others	
15.	Near- and far-sightedness	
16.	Recognizing boundaries	
17.	Recognizing boundaries	Sense-making

Sixteen unique components were identified. Table 13 provides a summary of each component, including their description based on the data, frequencies, and sample constructs. As indicated, the components role clarity, seeking alignment, fluidity, engaging others, facilitating use, seeking alignment, and recognizing boundaries were common to two or more rep-grids.

Table 13

Frequency and Description of Components

Components	Description	Frequencies	Sample Constructs (Pole A…Pole B)
Fluidity	The ability to engage in various styles of thinking and working and understanding, which is appropriate at a given time.	3	• creativity, thinking about the puzzle, aspiration, big picture thinking...experience and responsibility
Engaging others	Fostering participation among evaluation partners and communities.	3	• call for estimations of where they are and want to go...mechanisms/tools to help facilitate the group understanding where they are and/or where they want to go (have to invite collaboration)
Role clarity	Distinguishing the evaluators' and evaluation team's responsibilities from others involved in the evaluation (i.e., evaluation partners and communities).	3	• falls under my responsibility…clients' responsibility to lead
Seeking alignment	Examining the synergy between different aspects of the evaluation.	2	• alignment between expectations and resources...alignment between expectations and all matters related to the evaluation (ways of communicating, who are the decision makers)
Facilitating use	Attending to how the evaluation will be useful to partners and communities. Positioning evaluation partners and communities to use and act on the evaluation results.	2	• strategies to get at needs and intended use...technical specifications, what I am producing

Recognizing boundaries	Understanding and working within the limits of the task at hand.	2	• more of a calculation; playing with what's possible in order to get at what they know; adhering to their written words or something that has been prescribed...knowing is intuitive, create a simple and elegant process that captures the spirit of the work
Adherence	Taking actions and making judgments based on established standards and rules.	1	• requirement; compliance with laws and regulations; did I do what I was supposed to…voluntary; am I consistent with standards and practice in the field
Authenticating	Assessing validity and accuracy.	1	• have I got it right; validation...can I do this work; reflective
Creating an audit trail	Documentation of the evaluation, including what has become known or new knowledge that has resulted from the evaluation.	1	• audit trail and respecting community knowledge...process; modeling respect as a mechanism for creating an audit trail
Enhancing validity	Efforts to ensure the evaluation process will lead to accurate and representative conclusions.	1	• emphasis is on inclusion and voice…emphasis is technical and neutral process
Drawing the line	Making a determination based on an assessment of merit.	1	• somebody's expectations are influencing the evaluation...stopping those expectations; exiting the relationship
Managing ambiguity	Ability to move forward with the evaluation despite uncertainties.	1	• feel certain about, confident that it will be achieved...uncertain about what you can achieve (practical side of dissemination and will this rise to a level of interest for academic/policy community)

Near- and far-sightedness	Ability to focus on aspects of the evaluation that are relevant in the present and future.	1	• what is the data saying and how to make meaning from it...who you want to inform with the data you are collecting and how their information needs might be different
Perspective-taking	Examining an aspect related to the evaluation from a specific point of view. This includes others' viewpoints and one's own as well as different angles.	1	• focus is on organization's work and purpose; action or strategies…focus is on my purpose and work
Sense-making	Drawing meaning from observations.	1	• focused on the data (quality, interpretation, enough data)…More focused on how to best communicate (preparing a report, presentation)
Understanding personal value	Appreciating the value one brings to the evaluation. Recognizing the quality of one's practice.	1	• convincing clients on your credibility and value...delivering value; doing credible work

RTA. This sub-set of the data consisted of the interview transcripts of participants who did not complete a full rep-grid (*n=5*). The analysis resulted in two themes, each of which were common to the Strand 1 findings. The fist theme was the ways of knowing in evaluation, which consisted of three sub-themes (1) knowing others, (2) knowing the discipline, and (3) knowing self. The second theme was expressions of knowing, which included two sub-themes (1) practice and (2) documentation.

As presented in Table 14, knowing the discipline (*n=54*) was the largest sub-theme, followed by knowing others (*n=37*) and knowing self (*n=26*). The sub-themes under expressions of knowing, practice (*n=11*) and documentation (*n=8*), had relatively smaller frequencies.

Table 14

Frequencies of Strand 2 RTA Themes and Sub-Themes

Theme	Sub-theme	Frequencies
Ways of knowing in evaluation	Knowing the discipline	54
	Knowing others	37
	Knowing self	26
Expressions of knowing	Practice	11
	Documentation	8
	Total	136

Ways of Knowing in Evaluation. Given that these data only represent a sub-set of the participant data, the researcher exercised caution in what could be considered a finding. That said, these data offered some informative insights about evaluators' ways of knowing.

Participants discussed the importance of gaining familiarity with the evaluand and evaluation context, or as one participant described "a picture of how we all got there." The information considered to be important largely related to "the phenomenon and why it should be

96

evaluated," "the purpose behind it, how it was organized," and "the uses of the information if it's valid." One participant described, "I like to know the origins of the intervention, assuming that I'm looking at an intervention or a project or something of that sort. How did it come about? What is the theory behind it? What are the mechanisms?"

While information of interest was contextual in nature, how participants were perceiving it was closely aligned with the knowledge of the discipline. Understanding the evaluation context is integrated into several evaluation frameworks such as the Culturally Responsive Evaluation Framework (Frierson et al., 2002; Frierson et al., 2010) and the Centers for Disease Control and Prevention's Framework for Program Evaluation in Public Health (MMWR, 1999) and discussed in the literature.

Another aspect of knowing the discipline discussed in the data is the importance of developing an evaluation design and measures that are relevant to the context. Participants discussed the risks associated with casually applying standardized measures since "terms are going to mean different things for different groups of clients." As one participant explained, even a seemingly innocuous term like grit can be problematic:

> it's applied in North America, largely to poverty populations, ethnic minorities, and just to get their kids to school in the morning for a poor mother, and to get their child fed, or get the clothing laundered, everything takes a lot of grit…so grit can be quite different than the idea of practicing a forehand on a tennis court to get it right. Well, we need to know that instead of just taking, you know, the data reports on grit, however, it's measured, without knowing what it means, how it works, what the mechanism is, for increasing grit.

Certainly, poor measurement threatens the validity of the evidence garnered by an evaluation. However, as the quote above demonstrates, it also has implications on evaluation's role in reducing inequity and unnecessary harm (AEA, 2018a) and challenging patterns of marginalization (AEA, 2011), which relate to the ethical and social justice dimensions of practice.

According to the participants, examining and understanding the context occurs throughout the evaluation. Another participant explained that the final report should include both "the scientific results" and "situational information…these are the kinds of things you were entering into. And here's what we saw" so that clients gain a full picture of what has been learned through the evaluation.

What it means to know others was expanded and enriched through these interviews. Participants noted the importance of understanding people's histories, perspectives, and interests. As one participant explained, "people have stories, and you need to listen to those stories." Most often, listening and discussion were the techniques used for knowing others. However, one participant described a more formalized approached in which they were involved:

Prior to starting our data collection, we had three months of webinars, where it was mandated for all scientific or analyst team members…to come to our webinars so they could have literacy, so they could have indigenous literacy. So they could understand about indigenous data protections. So they could understand about data sovereignty with governance, so they could understand about indigenous theory and methods, they can understand about federal policy relating to American Indians, and education. And so they

could understand what 578 tribal governments look like, and that their governance

structures are all different.

As suggested in this quote, knowing others means becoming aware of the realities of those

involved with or impacted by the evaluation.

Knowing others is also needed to set the stage for engagement. One participant described

how a "pretty report is off-putting" for engagement:

one technique that I developed was to always have a scrappy looking cover on the draft

report. I don't have a fine-looking document…what I usually would do, I would take a

draft of the report and then I'd pencil mark the cover a bit, I'd cross out a couple words

and put a different word clearly in handwriting and that's what I give them, because what

am I telling them? I'm telling them, "We're not done yet."

Another participant argued "don't let the formal stages be your product alone, those are simply

mechanisms that open up opportunities for well-informed individuals to have a good discussion

about what we ought to do." Based on these perspectives, evaluation reports serve a purpose

beyond presenting the results of evaluation studies. And by knowing others, participants can

approach presenting results in a way that encourages dialogue and engagement.

All participants discussed relationship building in one form or another. One participant

indicated that "relationship building is built into a budget" and emphasized that relationships are

"reciprocal, culturally responsive, and honor sovereignty." Relationships are also built on respect

as one participant noted "you respect them. You don't try to outwit them. You don't try to work

around them." Treating clients respectfully is important because it "shows respect for what

they're thinking and shows respect for the problems they're having."

Interestingly, evaluators cannot assume that they know how to treat clients respectfully. Rather, understanding how to be respectful is contextual and must be determined on a case-by-case basis. One participant explained that they "sit back and listen and listen first, you know, and figure out how to be respectful."

Respondents noted that understanding the evaluation context helps them to determine if they should even engage in the evaluation. For one participant, this assessment involves understanding their purpose and practice:

> I'm concerned with social policy. So that would weighed very heavily for me on justification for evaluation. Has it been around long enough that the implementation can be documented and understood? Are the people who are involved in it open to an evaluation that may go in different directions?

As such, personal values emerged as a factor influencing practice.

Knowing self was central to one participant's practice. As they explained "how I typically approach navigate evaluation does not separate, personal from professional, it doesn't separate practical from theoretical…it's all related." Considering self was a key aspect of their decision to engage in a project, which involves being attuned with and listening to the "mental, physical, spiritual and emotional" aspects of themselves. Critical to this is their sense of trust towards clients and that they "have space," to engage in the work in a meaningful way.

Another important aspect of self was purpose. The same participant explained that "if I look back to when I was a little kid, to now, I was always about social justice." The participant described their approach to deliberately integrating this purpose into their practice by considering

"what are the cultural and community aspect" and determining "equity and resource sharing" and how projects "contribute to nation building, nation to nation work and honoring that."

Another participant commented "I've often wondered about schools even when I when I was a student, you know, purpose and things" and have "written fairly widely about schools in the States." They went on to explain that they are presently "really concerned about democracy" and that some schools are not providing a civic education that creates "a love for democracy," and teaches students how to participate in "discussions, voting, volunteerism." In this reflection, you can begin to appreciate the connection the participant was making between education, a sector they focus their work on, and supporting a democratic society.

Practice and documentation were discussed as ways that evaluators express and share what has become known through evaluation studies. How knowing is expressed though practice and documentation has been covered in previous sections. Examples include understanding how to respect clients and partners and using evaluation reports to facilitate discussion.

One way of expressing knowing that has not been discussed is publications. According to one participant "unpublished work is work that's never been done and so just coming up with a final report is not a satisfactory outcome for me. It has to be shared. It has to be in the literature." Other participants discussed how they have contributed to the literature through academic articles and books. The intent of publications is to develop the field's knowledge of evaluation practice by sharing what has been learned by participants and different approaches they have integrated into their work.

Table 15 provides example excerpts for each sub-theme identified in the data.

Table 15

Example Excerpts by Themes and Sub-themes

Theme	Subtheme	Description	Example Excerpts
Ways of knowing in evaluation	Knowing self	Knowing grounded in one's awareness of and connection to the genuine self (Chinn & Kramer, 2018). This knowing is developed through rigorous and honest self-examination and critical reflection on one's history, socialization, culture, assumptions, and perspectives (Mezirow, 1990).	• "How I typically approach navigate evaluation does not separate, personal from professional, it doesn't separate practical from theoretical…it's all related."
	Knowing others	Knowing that stems from one's ability to understand the perspectives and experiences of others; fostered through empathy and gaining close proximity to the lived experiences of others, while recognizing that one can never fully understand others (AEA, 2011; Schwartz & Sharpe, 2010).	• "You're trying to influence the people who are in charge of [disciplinary area], look at a topic with all the people who weigh in on it and you're going to do an evaluation the purpose of which is to change the way they're all doing business because you now understand something they don't understand and you need to tell them about it." • "Don't let the formal stages be your product alone, those are simply mechanisms that open up opportunities for well-informed

			individuals to have a good discussion about what we ought to do."
	Knowing the discipline	Knowing acquired by the integration of formal theories, scientific methods, and professional values and principles into practice (Chinn & Kramer, 2018).	• "We want to have some idea of what we're going to do, are we going to go out and do questionnaires with our, with clients? Are we going to, you know, how are we going to collect data, all that but exactly how we do the measures? Exactly how we do the analysis, you know, comes along is something we see as part of the development of the study approach, as we learn more and more about the situation."
Expressions of Knowing	Practice	A way of expressing what has become known that is integrated into how evaluators engage in practice. May include ways that evaluators modify or tailor the practice to address aspects of the evaluation they have become aware of.	• "And if I'm working in multiple tribes, to sort of sit back and listen and listen first, you know, and figure out how to be respectful."
	Documentation	A way of expressing what has become known through formal documents and other communications.	• "Just coming up with a final report is not a satisfactory outcome for me. It has to be shared. It has to be in the literature."

Content Analysis. A content analysis (Honey, 1974) was conducted on the grid data to further examine the underlying knowledge and ways of knowing inherent to the rep-grids on an aggregate-level. Percentage similarity scores were calculated for each grid to determine which constructs were more or less important to participant's overall perception of their evaluation practice. Constructs were indexed based on the H-I-L index. As described in Chapter 2, percentage similarity scores in the upper third of the range were indexed as H, medium range as I, and lower range as L. Constructs indexed as high or intermediate were determined to be of greater importance than constructs indexed as low and are therefore included in the final analysis.

The constructs were grouped into 12 categories using the core categorization procedure described in Chapter 2 (Table 16; Jankowicz, 2004). Constructs that could not be assigned to a category were removed (Jankowicz, 2004). Each category was assessed to determine the number of high or intermediate constructs it contained. As per the threshold described in Chapter 2, categories with less than 60% of constructs indexed as H or I, were removed. This process resulted in the removal of three categories (i.e., reflective thinking, distinguishing between relational and technical aspects of evaluation, and sense-making).

Table 16

Description and Frequency of Construct Categories from Content Analysis

Category	Description	Total Number of Constructs	Percentage H & I	Sample Constructs (Pole A…Pole B)
Perspective-taking	Examining an aspect related to the evaluation from a specific point of view. This includes others' viewpoints and one's own as well as different angles.	13	69%	• recognizing the demands characteristics that come into the evaluation and how it impacts people…recognizing the demand characteristics and how they impact you.
Engaging others	Fostering participation among many or all evaluation partners and communities.	10	70%	• seeking to involve all stakeholders…working with an insider group.
Recognizing boundaries	Finding solutions within the confines of the evaluation.	9	78%	• Provide a lens of importance; help us to understand what is important...Learning within an existing frame of what's important; already defined what's important.
Near and far sighted	Ability to focus on aspects of the evaluation that are relevant in the present and future.	9	78%	• focus is on the particular evaluation…focusing on the context and broader impact.
Role clarity	Distinguishing the evaluators' and evaluation team's responsibilities from others	8	88%	• think with me and gain understanding; facilitate

	involved in the evaluation (i.e., evaluation partners and communities).			understanding...my challenge to figure out.
Seeking alignment	Examining the synergy between different aspects of the evaluation	7	86%	• alignment between expectations and resources...alignment between expectations and all matters related to the evaluation (ways of communicating, who are the decision makers.
Purpose	Understanding the intent underlying one's practice.	5	60%	• making sure the evaluation will be used to drive change or improvement...determining how the evaluation will contribute to reducing inequities.
Facilitating use	Attending to how the evaluation will be useful to partners and communities. Positioning evaluation partners and communities to use and act on the evaluation results.	5	80%	• grounding the evaluation, the four stakes in the ground (what are we trying to do and who is that for...staying in touch, having a clear plan to make sure you are doing what you've agreed to do.
Making an evaluative judgment	Making a determination based on an assessment of merit.	4	100%	• participants, respondents are the priority...client is the priority.
Distinguishing between relational and technical	A way of perceiving an activity based on whether it involves	11	0	• about building relationships that will benefit the evaluation...Specifically

aspects of evaluation	technical evaluation skills or interpersonal/relational skills.			about analysis; drawing inferences from the data
Reflective thinking	Active consideration of one's beliefs or knowledge (Dewey, 1910)	5	20%	• value based; do I want to engage in the project...no value judgment; can I do it
Sense-making	Drawing meaning from observations.	4	25%	• thinking, processing, analyzing to make meaning...Doing the day to day evaluation activities

Study 1 Synthesis

       The results from Strands 1 and 2 were synthesized to gain a more complete and valid understanding of the ways that evaluators know what is needed to engage in evaluation practice and the associated dimensions of knowing. Specifically, points of convergence and divergence between the two strands were examined to see where the data overlapped, conflicted, and complemented one another. These points are summarized in Table 17.

Table 17

Study 1 Joint Display Table

Concept	Divergence		Convergence/Elaboration
	Strand 1	Strand 2	
Ways of knowing in evaluation	Four inter-related ways of knowing in evaluation emerged: knowing the discipline, knowing justice and power, knowing others, knowing self. Knowing the discipline and knowing justice and power received the greatest emphasis.	Various forms of underlying knowledge emerged including knowing the discipline and knowing others. Discipline knowledge and knowing others received the greatest emphasis. For the most part, underlying knowledge was integrated in practice.	The epistemology of evaluation practice is based on inter-related ways of knowing. In most instances, evaluators utilize multiple ways of knowing simultaneously while engaged in practice.
Knowing self	Supports professional improvement and self-awareness. Focused on examination of and critical reflection on one's values, history, culture, and positionality.	In the moment reflection intended to understand how one relates to the evaluation (e.g., is this project a good fit for me?) and critiquing one's practice. Discussed the role of purpose and personal values (i.e., social justice, democracy).	Evaluation practice is reliant on awareness and understanding of self.

Knowing others	Provided details on what should be known about others (e.g., perspectives, culture, interests, history). Knowing others helps to set the stage for engagement, supports validity, and helps evaluators respond to unique contexts.	Greater emphasis on the importance of knowing others compared to Strand 1. Discussed strategies about how to understand others' perspectives and interests as well as how to engage them in the evaluation activities.	Evaluation practice is reliant on awareness and understanding of others.
Knowing the discipline	Focused on understanding evaluation theory, research methods, and ethics.	Less focus on specific aspects of discipline-related knowledge. Rather, evaluators were interested in seeing alignment between different aspects of an evaluation. Generally, tasks that require discipline-related knowledge fall under the evaluator's role.	Knowledge of the discipline of evaluation is needed for practice. Both strands emphasized the importance of adhering to evaluation standards and principles.
Knowing justice and power	Emphasis on equity and power. Some emphasis on democratic values. Involves awareness and taking a proactive stance.	Greater diversity in social values: inclusion, giving voice, inequity, democracy, and advancing social change.	Evaluation requires an understanding of society (i.e., social consciousness) and is based on values related to the common good (e.g., equity, empowerment, participation).

Knowing in action	Four high-level processes emerged: seeing the grey, threading the needle, inquiry, and authentication.	Findings discussed techniques such as perspective-taking, near- and far-sightedness, recognizing boundaries, engaging others, seeking alignment, role clarity, and facilitating use.	Knowing in action presents in the form of high-level strategies and techniques.
Expressions of knowing	Discussion on documentation related to logic models and program theory; ethical practice; evaluative judgments and recommendations	Discussion on documentation referred to evaluation reports and publications.	Practice and documentation emerged as ways that evaluators express what they know or have become aware of through evaluative inquiries.

Points of Convergence. First, and importantly, four inter-related ways of knowing were inferred from the data: (1) knowing self, (2) knowing others, (3) knowing the discipline, and (4) knowing the common good and equity. The latter represents the overlap of principles of social justice, equity, and democracy across strands; addition of inclusion and social change in Strand 2; and alignment with AEA's fifth Guiding Principle (AEA, 2018a, p. 4). Collectively, these ways of knowing represent an epistemology of evaluation practice that inform and guide evaluators' practice.

The findings also suggested two underlying dimensions of knowing. Knowing in action are creative processes and strategies that help evaluators to perceive and understand their practice. In other words, these processes help them to develop solutions to problems, determine courses of action, and validate their conclusions and actions.

The second dimension relates to how evaluators express what has become known, typically in the process of conducting an evaluation. Two ways of expressing knowing were observed. Knowing is expressed in practice, such as making evaluative judgments and recommendations as well as tailoring practice to align with the needs and interests of evaluation partners and communities. An example from the data is understanding what constitutes respectful engagement with others and then tailoring one's practice accordingly. Knowing is also expressed through documentation such as logic models, program theory, evaluation reports, and publications.

Points of Divergence. Despite these points of convergence, some discrepancies between the two strands were found. The greatest point of divergence related to emphasis among the ways of knowing. While knowledge of the discipline was prevalent across both strands, it was the predominant way of knowing in Strand 1.

In Strand 2, the four ways of knowing were more evenly disbursed in terms of their relevance to the data. The researcher observed a greater presence of knowing others and self and less of a focus on knowing the discipline in these data. Despite some discrepancy between specific issues emphasized, the emphasis on knowing the common good and equity was consistent across strands.

Points of Elaboration. In most instances, discrepancies among strands allowed for elaboration of the findings. Specifically, the findings from both strands came together to provide a comprehensive account of the epistemology of evaluation practice. For example, Strand 1 focused on what should be known about others (e.g., history, culture) and Strand 2 provided details on the strategies that evaluators use to learn about and get to know others. Strand 1 identified discipline-related knowledge needed to conduct an evaluation and Stand 2 emphasized the importance of aligning these aspects during the evaluation process.

The findings also provided a more complete picture of knowing in action. Strand 1 provided examples of high-level processes, while Strand 2 described specific strategies that evaluators use. While it cannot be determined by the study, it is suspected that these processes and strategies are related. For example, perspective-taking could be a strategy used for seeing the grey, threading the needle, and authentication.

113

Importantly, the findings from both strands explained various ways that evaluators express knowing and how these expressions support knowledge development among clients. Strand 2 also discussed publications as a means of expressing knowing. The inclusion of publications as a way of expressing knowing introduces a learning loop, whereby the knowledge developed through practice contributes to the knowledge of the field, which then informs practice.

Study 1 Conclusions

As mentioned, the findings suggests that evaluation practice is grounded in four ways of knowing: (1) knowing the discipline, (2) knowing the common good and equity, (3) knowing others, and (4) knowing self. A description of each way of knowing is provided in Table 18.

Table 18

Description of the Ways of Knowing in Evaluation

Ways of Knowing	Description
Knowing the discipline	Knowing acquired by the integration of formal theories, scientific methods, and professional values and principles into practice (Chinn & Kramer, 2018).
Knowing the common good and equity	Knowing that evolves from one's appreciation of a fair and just world and sensitivity to what's beneficial for all members of society. This includes "the ability to recognize barriers that create unfair and unjust social conditions and to analyze complex elements of the sociopolitical context to change a situation that improves people's lives" (Chinn & Kramer, 2018, p. 290).
Knowing others	Knowing that stems from one's ability to understand the perspectives and experiences of others; fostered through empathy and gaining close proximity to the lived experiences of others, while recognizing that one can never fully understand others (AEA, 2011; Schwartz & Sharpe, 2010).

Knowing self	Knowing grounded in one's awareness of and connection to the genuine self (Chinn & Kramer, 2018). This knowing is developed by being present and engaging in rigorous and honest self-examination and reflection on one's history, socialization, culture, assumptions, and perspectives (Mezirow, 1990).

Two underlying dimensions of knowing were interpreted from the findings. The first, knowing in action, are creative process and techniques that actuate knowing, leading to new theories and understandings. Four high-level processes and techniques were interpreted from the findings (see Table 19).

Table 19

Summary of Knowing in Action - Process and Specific Tasks

Processes	Descriptions
Seeing the grey	The ability to see and contextualize the nuances of one's work. Interpreting meaning from amorphous concepts and contexts (Schwartz & Sharpe, 2010).
Inquiry	The use of contextually and culturally appropriate research methods to garner evidence; stems from the belief that evaluation processes can lead to new understandings.
Threading the needle	Achieving symmetry and balance between multiple frameworks, principles, perspectives, and so on.
Authentication	"Assessing the soundness" of what has become known (Chinn & Kramer, 2018, p. 287)
Techniques	
Perspective-taking	Examining an aspect related to the evaluation from a specific point of view. This includes others' viewpoints and one's own as well as different angles.
Near- and far-sightedness	Ability to focus on aspects of the evaluation that are relevant in the present and future.
Recognizing boundaries	Finding solutions within the confines of the evaluation.
Seeking alignment	Examining the synergy between different aspects of the evaluation
Role clarity	Distinguishing the evaluators' and evaluation team's responsibilities from others involved in the evaluation (i.e., evaluation partners and communities).

115

Engaging others	Fostering participation among many or all evaluation partners and communities.
Facilitating use	Attending to how the evaluation will be useful to partners and communities. Positioning evaluation partners and communities to use and act on the evaluation results.

The second dimension is expressions of knowing, which consists of the ways evaluators express what has become known through their experiences. For example, knowing is expressed through practice when evaluators tailor their approaches to align with the evaluation context and draw the line with respect to their involvement in an evaluation. Knowing is expressed through documentation created as part of, or as a result of, an evaluation. Specific examples of such documentation include program theory and models, evaluation reports, and publications.

These expressions of knowing provide concrete examples of how the ways of knowing in evaluation manifest in practice and how these manifestations can contribute to the knowledge of the field. The latter contributes to evaluators' ongoing learning and professional development, thus reshaping their ways of knowing.

Collectively, the ways of knowing in evaluation, knowing in action, and expressions of knowing form an epistemology of practice for evaluation depicted in Figure 11. The epistemology assumes that evaluation practice involves the integration of multiple ways of knowing, and strategies to theorize practice and express what has become known. The implications of this epistemology are explored in Chapter 4.

Figure 11

Epistemology of Evaluation Practice Framework

Knowing in Action

- Seeing the grey
  - Situate
  - Interpret fuzzy concepts
- Inquiry
  - Set the frame
  - Adherence
- Threading the needle
  - Multiple perspectives
  - Multiple frameworks
- Authentication
  - Critique
  - Envision
  - Solicit feedback
- Perspective-taking
- Near- and far-sightedness
- Recognizing boundaries
- Seeking alignment
- Role clarity
- Engaging others
- Facilitating use

Knowing the common good and equity

Knowing self

Knowing others

Knowing the discipline

Expressions of Knowing

- Practice
  - Practice tailored to the unique
    context and cultural context
  - Evaluative judgements and recommendations
- Documentation
  - Program theories and models
  - Evaluation reporting
  - Publications

Study 2

      Study 2 addressed the final research question: How, if at all, are the fundamental ways of knowing in evaluation integrated into graduate programs in evaluation? A directed content analysis (structured) was conducted on the transcripts from the faculty and alumni/advanced student interviews (Elo & Kyngäs, 2008). This analysis was based on a priori codes identified in Study 1 (see Table 7). As shown in Table 20, a total of 470 segments of data were coded.

Table 20

Frequency of Code Assignments with Sample Excerpts from Interviews

Code	Frequency	Sample Excerpts
Knowing the discipline	301	• "So we talked a lot about the evaluation tree. If you're familiar with that, and the different branches of the tree, I think there's like, use, methods, valuing. And we did discuss some of those different evaluation theories, ways of planning evaluations, kind of comparing and contrasting them in when it might be appropriately appropriate to select one over another kind of based on your, your project." (alumni/student) • "I took a lot of quantitative and qualitative research methods, courses, mixed methods, research courses, and then a ton of evaluation theory courses." (alumni/student)
Knowing others	52	• "As an evaluator, you're going to be critical of your client. At some point, you're going to have to say, you know, this, I know you had high hopes for this, this does not work very well. You know, that's what the data suggests. And so learning how to do that in this context can inform that as well." (faculty) • "I think being able to communicate effectively with stakeholders was a skill that I learned, specifically in that evaluation, one intro course, but then also in my research assistantship, with some of the projects I was on." (alumni/student)
Expressions of knowing	40	• "…so that ties back to what they did in the fall semester, like now trying to apply the knowledge of, you know, formulating a logic model in the real-life situation." (alumni/student)
Knowing the common good and equity	38	• "There was one week in the class where we talked about what are things we need to consider when we are designing and implementing an evaluation, if we're in a community that has been historically marginalized. And we are being asked to do this work with an equity focus lens. So, who's at the table? And … what are issues of power dynamics? And what how might we attend to issues of power within these social contexts?" (faculty)

Knowing self	26	•	"I have students journal, to think about the ways in which this work is applicable to their identity as evaluators, and how it might continue to influence them. Not just in the courses but beyond. And so I would say yes, in terms of their evaluation identity, we are explicitly having conversations and writing about the ways in which their work is impacting who they are and how they see themselves, and also the work that they want to do once they leave." (faculty)
Knowing in action	13	•	"The challenges that pop…up for me are, first of all the tolerance for ambiguity." (faculty)

Further analysis of the data assigned to each a priori code resulted in 17 emergent sub-codes (Table 21). The sub-codes highlighted key program components and educational experiences that corresponded with the ways of knowing in evaluation and their dimensions (Figure 11).

Table 21

Study 2 A Priori Codes and Emergent Sub-Codes

A Priori Codes	Emergent Sub-codes
Knowing self	Self-awareness
	Professional identity
	Positionality
Knowing others	Communication
	Interpersonal skills
	Relationship building
Knowing the discipline	Evaluation design/procedures
	Evaluation theory
	Research Methods (including statistics)
	Evaluation ethics
	History of evaluation
Knowing the common good and equity	Role of evaluation/evaluators
	Social consciousness
Knowing in action	Managing ambiguity
	Managing complexity
Expressions of knowing	Logic Models/Theories of Change
	Reporting/Presenting findings

Knowledge of the Discipline

Generally, the knowledge of the discipline is well integrated into graduate education programs in evaluation. Predominantly, participants discussed the inclusion of evaluation theory, research methodology and statistics, and evaluation design as key areas of this knowledge. Students gained this knowledge through coursework and practical experiences such as internships and assistantships.

Students were exposed to various evaluation theories and often referenced the Mertens and Wilson (2018) theory tree as a learning tool. The theoretical training of a small number of students was focused on Utilization-Focused Evaluation (Patton, 2008), which alumni found problematic when attempting to bridge theory and practice:

> there wasn't a lot of focus on specific evaluation theories, and kind of more of that fundamental knowledge. We covered a little bit of the history of evaluation, but not totally. And so I think there are some gaps in my knowledge about that area…working day to day as an evaluator, I do find myself thinking back to that, and thinking, oh, boy, I'm not really sure what specific theory to apply here, or what, what's best.

Evaluation ethics is one aspect of the knowledge of the discipline that respondents did not consistently agree was/was not included in the programs. Generally, faculty were consistent in their view that evaluation ethics was "built in" to the program and something they are "always covering." One faculty described a strategy they use to develop students' ethical knowledge:

> they get two letters from the novice evaluator who's in a dilemma that they have to provide advice to…using the guiding principles…And then they, they look at the advice that each other has provided. And they talk about that advice.

However, alumni and student reflections on their exposure to ethics varied. Some alumni recalled "a lot of discussion throughout the other courses about ethics." Others discussed specific course activities related to evaluation ethics. For example, an alumni participant recalled a class where they used fake scenarios to practice dealing with an ethical situation such as when a client asks to have a negative finding removed from an evaluation report:

122

we got into groups, and we kind of had to talk through the scenarios of how did we

respond to the stakeholder…and the professor talked through some things, he had a lot of

experience with being an evaluator, so he would often share things like, um, you know,

one thing I would say, in this situation, if we can frame this, as, you know, maybe areas

of growth or areas of enhancement, sometimes just rewording it that way can really help

ease maybe any anxiety over including that information.

Other alumni indicated that "we didn't really focus a lot with the guiding principles" or

referenced their exposure to research ethics:

I think ethics was related kind of do like human subjects, kind of pieces and but a little

less about, yeah, kind of maybe some of the complications or like sticky situations you

might come into contact with as an evaluator and working for like a client.

In addition to ethics, participants noted that budgeting, project management, and

managing large data sets were technical skills that they felt were not included in the programs

but are important for evaluators. For example, one alumni/advanced student participant stated, "I

don't think on the program our program did a very good job teaching us about project

management or teaching us about budgets." A faculty participant expressed a similar concern

stating, "how do people really understand how to budget evaluation when they come out of

here?"

Among the students who raised data management as a shortcoming of their program,

different observations were shared. Some alumni/advanced students felt that their programs

assumed students brought this expertise with them. Others felt that exposure to data management

came down to chance, as one alumni/advanced participant explained:

even data management pieces we learned on the projects we were on, but if you didn't

have a big data set to manage in your project, or multiple data sources over multiple

years, you know, for a given project you were assigned to, then you probably didn't walk

away with those skills.

Knowing Others

Knowing others was most often discussed in the context of relationship building and

communication with clients. As one faculty explained:

I frame it to students that it is a social and relational process. And that, and don't get me

wrong, like the methodological skill, the technical, the technical knowledge of evaluators

matters. But if you can't get that social and relational part, working in a good way in your

evaluation, and you don't respond to that with your design and approach, you may think

that you come away with rigorous data that reflects important issues with the program,

but your stakeholders may not see it that way.

For the most part, participants identified practical experience as the primary area where

students are exposed to "interpersonal dynamics," learn to "work with other people" and develop

"a lot of interpersonal skills of just even how to work with, you know, programs and directors

and participants." One alumni/advanced student explained:

I think being able to communicate effectively with stakeholders was a skill that I learned,

specifically in that evaluation…but then also in my research assistantship, with some of

the projects I was on…There are a lot of times where, you know, we might be talking to

stakeholders, and we kind of almost have to calm their worries about evaluation…And so

124

being able to kind of effectively communicate, what the project is and the intent and

purpose and all of that, I think is really important.

Relational skill development was also included in some courses, particularly those that were

practice-based and involved working in teams and interfacing with clients. In some instances,

students also had specific opportunities "practicing giving constructive feedback" and on "how

to give critical feedback kindly."

That said, some caution regarding the extent to which knowing others is meaningfully

integrated into students' experience is necessary. For one, some key strategies for knowing

others, such as perspective-taking and learning how to facilitate and engage others, that emerged

in Study 1 were not present in the data. Additionally, participants often referenced aspects of

knowing others, such as interpersonal communication, as program shortcomings. One faculty

commented:

I think we could do better at it. And what would better look like? Maybe a little bit more

intentionality? And like using, and like building even the vocabulary for good

communication, like what does it mean to be in a conflict resolution? What does it mean

to be in negotiation? What does it mean to hold space? I would love if we were doing a

little bit more specific work around holding space.

Another student indicated that interfacing and interacting with clients are valuable skills

that they developed prior to starting their doctoral studies. From their perspective, it's

"something that doesn't get touched on as much as, you know, actually working with clients and

so on, because it's different than just doing a research project with your advisor." A faculty

shared a similar sentiment stating that "we could be more intentional about building those skills."

125

Knowing the Common Good and Equity

Two interesting patterns were observed in the data assigned this code. First, it appears that some programs are more intentional than others about developing students' awareness of social issues (e.g., structural racism and inequity) and understanding of evaluation's role in advancing the common good (e.g., justice, power, and equity). For example, some programs introduce "students to what evaluation is historically" and "the role of evaluation in society as we think about the ways in which evaluation can help attend to those grand issues."

An alumni described their efforts to deliberately integrate social justice and equity-focused evaluation into their teaching and mentoring:

I was upfront with the students and with our clients, particularly given the programs that we are working with this semester, that the paradigm in which I operate from with this class, and I think generally in my own work is we are doing this from an equity focus lens, from a culturally responsive lens. Because we're also working in communities where many of these clients are not the clients themselves, but the participants of these programs who have experiences with being exploited. And so I'm leaving, knowing we don't want to leave, and we don't want to leave our clients and the participants of these programs worse off than when we entered. And it was important for me to be explicit about that.

An emphasis on the common good and equity was also shared by some alumni/advanced students who indicated that the socio-political context of evaluation was a "major theme in…our evaluation theory courses" and that their program "opened my mind and sort of heart to

126

vocabulary of social justice and greater awareness of a critical lens on evaluation." One alumni/advance student participant recalled:

> we did cover some of the ways that evaluation can be used as a tool to help in these areas and these kind of current topic areas, but then we also did discuss how evaluation can be a detriment in these areas, and how they cannot really help us progress. And so, I remember that conversation and it being an interesting one to kind of talk through both sides and how evaluation can be a help and a hindrance as well.

The second pattern related to differing views between faculty and students regarding students' exposure to matters pertaining to justice and equity. Almost all faculty felt that social justice and equity was integrated into their program's curriculum. However, some alumni/advanced students suggest that their exposure to these issues was a result of their interaction with "other students and some of the faculty but not the faculty in my program" and "seminars, when I would go to colloquia and sort of interact with the larger, scholarly community." Another alumni/advanced student indicated that "real discussions of, you know, power and having evaluation as a tool for challenging systems of power…was not clearly discussed or, you know, taught."

Knowing Self

Of the ways of knowing, knowing self was the least prevalent in the data. There were instances where alumni/advanced students gained self-awareness or developed a stronger sense of professional identity. Some programs incorporate reflective practice into their program, either in the form of workshops or as part of discussions on personal biases. In some instances, students were encouraged to explore their positionality and "what power and privilege they do have and

how they manage it" through journaling or as part of qualitative methods courses. One program also has students write papers that describe how "different theories and approaches that they've encountered in the class that they find serve their context and disciplinary area and their particular values related to evaluation."

The data indicated that many of alumni/advanced student participants brought "in their own lived professional experiences," and, therefore, high degrees of self-awareness. This was particularly the case among alumni/advanced students with previous career experience in areas such as counselling, education, and social work, which tend to prioritize critical reflection and self-examination. One alumni/advanced student participant explained, "reflection has always been a part of my work. And in many places, it's been required in my work…it's hard to really see anything specifically tied to the program, because I've experienced my own reflective practice as just being an ongoing thing." As such, these participants may have been less sensitive to program efforts that aimed to develop their self-awareness.

However, there is evidence that self-knowing was not prominent across the data set. As one participant stated, "it wasn't a component of the program. There was definitely no identity exploration, or even really examination of positionality." The researcher also observed that information about knowing self was rarely volunteered and, in many instances, related to isolated activities that were not carried forward throughout the duration of the program.

Expressions of Knowing

The data suggests that programs engaged in activities to help students learn how to express knowing, mainly through documentation such as logic models, evaluation reports, and evaluation plans. One faculty participant discussed teaching a course specific to evaluation

reporting and communications. As part of this course, students "develop a communications plan… design oral presentations, they do design data visualizations, both quantitatively and qualitatively… design a creative project…practice turning scientific language into accessible language, explaining technical things in ways that are accessible, they study accessibility guidelines and do some report writing." For the most part training on how to produce evaluation reports and other documentation such as logic models were included in course work and practical experience opportunities.

Knowing in Action

Instances of knowing in action tended to relate to inquiry and managing the complexity and ambiguity of evaluation contexts. Students' capacities for inquiry, particularly their development of research and procedural knowledge, were generally well supported by programs as discussed in a previous section. However, one faculty noted that the lack of "a clear objective sense of what is the right way to go about something and measure something" and that the "valuing and methodological approach is debatable" and "something that's negotiated" can be "very disorienting for students."

Student's reaction to the ambiguous nature of evaluation practice was carried over into other topics of discussion. For instance, one faculty observed that not all students have the same "tolerance for ambiguity," which can limit the value-added from their practical experiences. Specifically, some students faced with uncertainty will "shut down" while others "just jump in." Another faculty observed that students are often asked to "read the directions and bake the cake at the same time," which can foster uncertainty and a lack of confidence.

129

Chapter 4: Discussion

In his seminal book, *Evaluating with Validity*, House (1980) claims that "the social import of evaluation is enormous; its self-understanding relatively minute" (p.11). Since then, the field has grown in many ways, including with the emergence of new evaluation theories and frameworks as well as the advancement of initiatives to professionalize the field (e.g., competencies, educational programs). Yet, the extent to which self-understanding has improved is unclear.

Shadish (1998) tells us that "evaluation theory is who we are" (p.1). Shadish goes on to describe evaluation theory as "a set of diverse theoretical writings held together by the common glue of having evaluation practice as their target" (p. 2). Unfortunately, only a minority of evaluators have had the resources, motivation, and privilege to publish their theories (House, 1980; Shanker, 2019). Thus, what has been equated as the "knowledge base" and "professional identity" of evaluation is largely void of the collective wisdom of practitioners (Bowman, 2021; Shadish, 1998, p. 1).

Unsurprisingly, research on evaluation suggests a theory-practice gap in evaluation. The literature provides numerous reasons for this gap, many of which were discussed in Chapter 1. However, none of these reasons propose that the gap may relate to how the field conceptualizes evaluation theory and practice.

Various disciplines, including evaluation, acknowledge the existence of informal theories based on the practical knowledge of practitioners (Bowman, 2021; Christie, 2003c; Christie & Rose, 2003; Shadish et al., 1991). While not formalized through publication, these theories are believed to have a substantial influence on practice, enabling practitioners to deal with the

vitality and complexity of their work (House, 1980; Schon, 1983). But, unlike other disciplines, the evaluation field has not taken steps to examine, learn from, or legitimize these theories.

This research aimed to fill this void by examining practical knowledge in evaluation, including the ways of knowing in evaluation. The intent was not generalizability, but rather exploring if, and what, practitioner-based knowledge may have been left on the table. This section situates the research findings in the extant literature on evaluation practice, theory, and education and discusses their potential implications for the field. The perceived strengths and weaknesses of the research are also disclosed.

An Epistemology of Evaluation Practice

Presumably all theories, including informal, practice-based theories, are drawn from a knowledge base. Thus, explicating the underlying knowledge (i.e., epistemology) of evaluation practice is essential to understanding these theories. Based on the study's sample, the researcher interpreted that evaluation practice is grounded in four ways of knowing – knowing self, knowing others, knowing the discipline, and knowing the common good and equity. These ways of knowing were organized into a framework depicting an epistemology of evaluation practice (Figure 12).

Figure 12

Epistemology of Evaluation Practice Framework

Knowing in Action

- Seeing the grey
  - Situate
  - Interpret fuzzy concepts
- Inquiry
  - Set the frame
  - Adherence
- Threading the needle
  - Multiple perspectives
  - Multiple frameworks
- Authentication
  - Critique
  - Envision
  - Solicit feedback
- Perspective-taking
- Near- and far-sightedness
- Recognizing boundaries
- Seeking alignment
- Role clarity
- Engaging others
- Facilitating use

Knowing the common good and equity

Knowing self

Knowing others

Knowing the discipline

Expressions of Knowing

- Practice
  - Practice tailored to the unique context and cultural context
  - Evaluative judgements and recommendations
- Documentation
  - Program theories and models
  - Evaluation reporting
  - Publications

132

The research findings suggested several features of these ways of knowing, which may offer some useful insights into evaluation practice. First, evaluation practice may involve multiple ways of knowing (Bowman, 2021; House, 1980). Evaluators often emphasize systematic inquiry and scientific methodology as the genesis for knowledge development. Yet, much of what the field views as essential to useful and impactful evaluation (e.g., use, collaboration, cultural responsiveness) is dependent on ways of knowing that are not rooted in scientific discovery.

Second, while these data suggest that each way of knowing is unique, they did not appear to be independent. Rather, they co-occurred suggesting that evaluators simultaneously engage multiple ways of knowing to understand and navigate their practice. A co-occurrence between knowing the discipline and knowing the common good and equity was particularly prevalent in the study, which emphasizes the importance of aligning methods with equity and justice principles. Importantly, this finding suggests that no single way of knowing is sufficient for evaluation practice (Carper, 1978, as cited in Chinn & Kramer, 2018).

Third, evaluator knowing is both cognitive and social-emotional in nature. Similar to House (2015) who argued that "evaluators evaluate with their whole person, not only methods," the research findings suggest that thinking and feeling are both important contributors to practice (p.127). Moreover, Schwartz and Sharpe (2010) argue that emotions, rather than cognition, stimulates action. For example, empathy for other's discomfort is often what motivates individual to take actions to alleviate or address the situations they encounter.

Last, while the framework treats each way of knowing as equally valid and influential, knowing self is positioned at the core based on insights drawn from this study and the literature

suggesting that knowing self is central to all knowing (Polanyi, 1966; Wilken, 2002). The personal aspects of professional practice have been a controversial topic in evaluation. The topic often surfaces in discussions on evaluator bias and systematic error and is considered as a factor that needs to be controlled (Abma, 2009; Chouinard et al., 2017; Chelimsky, 2012; Davidson, 2005; Jewiss & Clark-Keefe, 2007; Scriven, 1998). However, some authors posit that self-knowing, or self-awareness, is an important component of one's practice. Symonette (2004), for example, discusses "inside-out" and "outside-in" work as critical to one's journey towards cultural competence (p. 99). The author explains:

> critical, yet woefully underdeveloped, segment of needed capacity-building work involves microfocused assessment and evaluation processes that undergird and support inside-out work. Such work calls for a mindfully conscious self—with expansively refined lens and filters—that enables accurately discerning, navigating, negotiating, and understanding the shifting sociocultural terrain using appropriate codes of engagement. (pp. 99-100)

Mezirow (1991) explains that our meaning perspectives stipulate the principles upon which individuals interpret and draw meaning (p. 2). This includes how we conceptualize evaluands, the issues we see and how we think they should be addressed, how we relate to others, and the connection between our practice and purpose (Weiss, 2013).

Implications. This section discussed four ways of knowing interpreted from the study's sample. These ways of knowing challenge some existing canons of evaluation and may provide new insights into evaluation practice. Further opportunities to test and build from these ways of knowing are proposed as a future direction.

134

Several characteristics of the ways of knowing are discussed, including the centrality of knowing self to evaluator knowing. Despite the long-held belief in evaluator neutrality and objectivity, it is unlikely that one can simply 'switch off' their meaning perspectives (Scriven, 1998). Nor would it be helpful to do so since these perspectives are needed to make sense of the world. What is advisable is for evaluators to develop a critically reflective stance to deepen their self-awareness and understanding of the various ways their meaning perspectives intersect with their practice (Mezirow, 1991). Suggestions for doing so are provided as a future direction.

Theorizing Practice

While the ways of knowing in evaluation identified in this study, are a basis for an epistemology of evaluation practice, knowing in action and expressions of knowing elaborate on the underlying dimensions of evaluator knowing. These dimensions could play a critical role in actuating knowing, sharing what has become known through evaluative processes, and building the field's knowledge base. Additionally, expressions of knowing may provide insights into how knowing manifests in practice in the form of practice and documentation.

Knowing in action provides an explanation of how practitioners perceive and theorize their practice. The study interpreted both high-level strategies (e.g., seeing the grey, authentication) and techniques (e.g., perspective-taking, near and far-sightedness) that aid in theory development. Many of these strategies and techniques align with some of the practical knowledge characteristics discussed by Schwartz and Sharpe (2010), such as the ability to balance and interpret multiple principles and guidelines as well as understand nuances (pp. 25-26).

Absent from the findings is a sense of logical order as to how knowing in action unfolds in practice. Rather, knowing in action appeared to resemble the active-reactive-interactive-adaptive spirit proposed by Patton (2013). From this perspective, understanding evolves and changes throughout the evaluation process, as does the evaluator.

Knowing in action may also provide insight into how evaluators develop and carry out various evaluator competencies. For instance, interpersonal competence is an evaluator competency (AEA, 2018b) and is also emphasized in evaluation theories such as Interactive Evaluation Practice (King & Stevahn, 2012), Empowerment Evaluation (Fetterman, 1994), and Utilization-Focused Evaluation (Patton, 2008). However, as this research suggests, interpersonal competence can be challenging to acquire (Dewey et al., 2008; Galport & Azzam, 2017).

Warner (2020) argues that developing interpersonal competencies is both a "challenge and opportunity for educators of evaluators" (p. 433). The author proposes the integration of service-learning and its core components of "curricula instruction, field experience, and reflection" into evaluator education as a strategy for developing interpersonal competence (p. 436). Some of the knowing in action techniques (e.g., perspective-taking and critiquing) interpreted from the study's sample could support service learning efforts, helping evaluators appreciate others' positions and how to communicate and resolve conflict effectively.

Thompson and Thompson (2018) describe theories as consisting of open (explicit) and closed (implicit) knowledge. The expressions of knowing presented in the findings provide some insights into open and closed knowledge in evaluation. Clearly, some expressions of knowing are more explicit than others.

Documentation, such as program theory and scholarly articles, are examples of explicit expression of knowing interpreted from the findings. Other examples from the literature include Fitzpatrick et al. (2009) and Hutchinson (2018), which examine evaluators' experiences in the field, including lessons learned. As open knowledge, these materials can contribute to shared understanding and develop evaluation capacity among evaluation and other interested partners as well as members of the evaluation community.

However, knowing expressed through practice tends to be closed. This is problematic for various reasons. Knowledge that is implicit is typically not subject to scrutiny, which would highlight problematic assumptions, biases, and areas for improvement. Additionally, closed knowledge can easily go unnoticed and, therefore, have a limited impact on learning and capacity building.

Implications. The research findings offer suggestions regarding how practitioners know and theorize their practice. The strategies and techniques interpreted from the findings provide additional details to the proposed ways of knowing and, potentially, connections between practice theories and evaluation theories. Acknowledging that this study was exploratory in nature, and limited to a specific sample of evaluators recognized as exemplar practitioners by the field, future research is needed to continue to advance our understanding of the ways of knowing. With this in mind, opportunities to further examine these connections are discussed as a future direction.

Consistent with the literature (Guzman, 2009; Thompson & Thompson, 2018), evaluator knowing is likely expressed both explicitly and implicitly. Thus, the onus is on evaluators and evaluation teams to create space for dialogue and venues to examine and explicate the knowing

inherent to their practice (Thompson & Thompson, 2018; Mezirow, 1990). Avenues to foster reflection among evaluators are discussed as a future direction.

Educational Programs in Education

Within the evaluation field, it is widely accepted that evaluators require specific training to develop the skills and knowledge needed to conduct evaluation work. There are many routes one can take to develop their evaluation knowledge and skills (e.g., self-study, professional development workshops; Christie et al., 2013; LaVelle & Donaldson, 2010; Stufflebeam, 2001). However, no route is likely as comprehensive and rigorous as a graduate-level educational program in evaluation such as those included in this study. In the words of some study participants, these programs are intended to create "experienced practitioners" who can "take these ideas, these concepts, these tools, and apply them" and to develop "innovators and people who improve the world through evaluation."

Given the rigor and aspirations of these programs, one would assume that if the ways of knowing and its dimensions are relevant to evaluation practice, they would be present in these programs. And the findings suggest that these elements are integrated into the sample, albeit to varying degrees. Namely, graduate programs tend to concentrate on aspects related to knowing the discipline, such as scientific methodology and evaluation theory. Students appear to receive relatively less exposure to the other ways of knowing (i.e., knowing others, knowing self, and knowing the common good and equity).

In many ways, these findings are not surprising and potentially intentional. For one, these programs educate students in an academic discipline and are housed in institutions of higher learning. Thus, there are standards and expectations regarding program content and the academic

138

abilities of students that programs must adhere to. Additionally, university education is grounded in technical rationality, which is a western epistemology that views scientific knowledge as the most legitimate form of knowledge (Schön, 1983). From this perspective, scientific theory and technique provide the tools needed to solve real-world problems (Schön, 1983). The influence of technical rationality on evaluation has been raised by numerous evaluation scholars (Bowman, 2021; House, 1980; Schwandt, 2005).

Last and importantly, scientific methodology and evaluation theory are fundamental to the field (Chelimsky, 2012; Rog, 2015; Shadish, 1988). They encompass the knowledge and capabilities needed for social inquiry and are a resource for how to conduct evaluation studies that are accurate, useful, and culturally responsive. It is worth noting that none of the study's participants questioned the importance of methods and theory to evaluation practice.

What did surface in the data were questions regarding the emphasis of programs on knowing the discipline, particularly evaluation theory, relative to other ways of knowing. Some alumni/advanced student participants suggested a saturation of theory, or as one participant put it, "reading 10 articles from each theorist, I don't know how much that really benefits you." Alumni/advanced student participants also referenced other knowledge and skills not integrated into their respective programs, such as conflict resolution, relational skills, and self-awareness, that they believed would be beneficial to students. Many of these competencies relate to the other ways of knowing presented in the framework, which were not as apparent in the data.

Implications. On the surface, the findings of Study 2 may not seem problematic. Developing sound methodologists with expansive theoretical knowledge should be a key motivation for these programs. However, the literature offers some cause for concern.

First, students may acquire good prepositional knowledge of evaluation (i.e., awareness of prepositions and steps involved in conducting an evaluation) but not procedural knowledge (i.e., understanding of how to do an evaluation). Some may argue that procedural knowledge is gained during fieldwork and other practical experiences. Here, students have opportunities to apply their prepositional knowledge and develop procedural knowledge from their experiences. However, the study findings indicate that the extent to which procedural knowledge is developed during fieldwork is likely more situational than general.

Second, students may develop a theory-based perception of practice (i.e., practice is the implementation of theory), which does not translate to real-world evaluation contexts. Chouinard et al. (2017) discuss the tensions experienced by evaluation students when attempting to apply theory while navigating the social and political nature of evaluation projects. Similarly, Rog (2015) raises concerns with "by the book practice," which causes "knee-jerk reaction to design and implement" without consideration of the constraints and needs of a context (p. 231).

Last, the findings align with literature suggesting that the evaluation curriculum continues to privilege scholarly knowledge and western ways of knowing. According to Bowman (2021):

> How evaluation is theorized, taught, and transmitted through practice and scholarly activities needs to directly address privilege, decenter it, and destabilize and eradicate the marginalization harm of the nonprivileged. The behavior of evaluators needs to leave behind the bankrupt idea that a Western academic and capitalistic perspective is neutral. It is harmful and does not treat all ways of knowing as equally valid and valuable evidence. (p. 322)

If the evaluation field aspires to address social justice and inequity, decolonizing evaluation education is a necessary step. This includes opening students to the numerous and valid ways of knowing and the legitimacy of knowledge created through practice.

Future Directions

The research offers numerous theoretical and practical benefits to the field. First, the findings present a conceptualization of evaluation practice as a dynamic process of perceiving and theorizing. This conceptualization provides an alternative way of framing future research on evaluation, including the relationship between evaluation theory and practice, which is viewed as an area of much-needed inquiry (Cousins & Earl, 1999; Chouinard et al., 2017; Christie, 2003a; Datta, 2003; Shadish et al., 1991; Smith, 1993; Worthen, 1990). Future research could examine connections between practice theories and evaluation theories. Such studies may lead to new insights into the connection between theory and practice, which could refine and advance evaluation theories.

Second, and importantly, the researcher hopes the present study will catalyze further inquiry into the epistemology of evaluation practice. The framework presented in this manuscript is a good start to understanding the epistemology of evaluation practice, but it needs to be tested on a broader sample of evaluators. Additionally, more work is needed to leverage the wisdom of evaluators who are BIPOC as they were not sufficiently represented in the study. This work may lead to developments and refinements of the study's framework, new frameworks, or a greater appreciation for existing frameworks that will enrich the field's overall understanding of evaluation practice.

Third, the research can inform evaluator education. The previous section discussed several implications associated with an overfocus on knowing the discipline and the absence or minimization of other ways of knowing in evaluation programs. Teachers of evaluation, who resonate with the findings, may find it useful in providing a structure and vocabulary to help integrate other ways of knowing into course and field work, dialogue, learning plans, and mentorship. It is the researcher's belief that a more holistic account of evaluation practice and evaluator knowing will facilitate the development of procedural knowledge, advanced loops of learning (i.e., double- and triple-loop; Argyris & Schön, 1974, Nicolini et al., 2003), and well-rounded practitioners.

Fourth, and similarly, the findings could help to inform evaluation capacity building efforts. Evaluators are often called on to develop evaluation capacity among partners and communities (Baizerman et al., 2002; Preskill & Boyle, 2008; LaVelle & Donaldson, 2021). In these instances, the framework could be used to orient audiences to evaluation practice in a way that is practice-based (i.e., a holistic and situational process of perceiving, theorizing, and developing knowledge).

Last, the study provides tools to foster reflective practice, which is considered by many to be essential to evaluation practice (AEA, 2011; AEA, 2018b). The framework offers a heuristic tool to structure critical reflection and self-examination without stifling originality and ownership by being overly instructive. Specifically, practitioners could use the framework to develop critical questions to help them examine their experiences from multiple angles and explicate closed (implicit) knowledge that influences how they theorize their practice (Johns, 1995). The framework can also aid evaluators in developing a deeper understanding of their

practice and how their practice relates to evaluation theories and emerging priorities within the field.

Study Strengths and Limitations

There are several strengths and limitations that are important to keep in mind when interpreting the study findings. These factors have been organized into three areas (1) validity, (2) methodology, and (3) reflexivity.

Validity

The study integrated several strategies to enhance the validity of the findings. First, the rep-grids were designed to allow participants to elicit their elements and constructs. This structure aligns with theory (i.e., Personal Construct Theory) and allowed participants to describe their practice in their own words, thus, controlling researcher bias (Bell, 2003; Jankowicz, 2004). Additionally, the researcher used several techniques to ensure that the rep-grids captured meaningful and accurate data, including triadic elicitation, laddering down, and qualifying phrases (Jankowicz, 2004).

Second, the rep-grid technique (Study 1) utilized member checking to help ensure the validity of the findings. Specifically, each rep-grid participant was provided a copy of their grid and asked to review it for accuracy and completeness. Numerous participants follow-up with the researcher to confirm the accuracy of their grid. One participant identified some language they felt misrepresented their perspective and worked with the researcher to develop new language for their grid.

Third, the use of multiple data collection methods helped to control for the biases and limitations inherent to each method (Creswell & Plan Clark, 2018). Additionally, the synthesis of

the Study 1 findings allowed the researcher to compare findings generated from different data collection methods (i.e., methods triangulation; Patton, 2002).

Despite these strengths, there are some limitations and cautions to report. Time constraints prevented the researcher from providing Study 2 participants with interview summaries for member checking purposes. This measure would have allowed participants to assess the accuracy of the researcher's interpretations and understanding of their experiences.

Additionally, resource constraints prevented the researcher from recruiting a second coder for Study 2's directed content analysis. A second coder would have positioned the researcher to check the accuracy of coding and interpretations as well as reduced the risk that pertinent information was missed. The researcher attempted to reduce the impact of this risk by using active listening techniques (i.e., paraphrasing and summary statements) to check on understanding and interpretations during the interviews. Additionally, the researcher aimed to be transparent about the data collected and how it was interpreted by incorporating raw data into the body of the manuscript and Appendix P. Despite these efforts, some caution towards the trustworthiness of the findings is warranted.

Methodology

Study 1 data collection was limited by the personal and environmental context, which prevented in-person data collection. As many theorists have argued, there is knowledge entrenched in practice that may not be easily communicated (Schwandt, 2008; Van Manen, 1999, as cited in Schwandt, 2008). Incorporating observations into the design would have allowed the researcher to explore the physical and contextual aspects of knowing, which were not captured by the research.

Additionally, Study 1's purposeful sampling strategies was only marginally effective in recruiting perspectives from evaluators who are BIPOC. To some degree, the researcher counteracted this outcome by ensuring that these voices were not lost in aggregation. However, these efforts were sometimes restricted by the need to ensure participant confidentiality. Consequently, the interpretations embodied in this study are subject to bias and should not be considered representative of the field as a whole.

Last, there were several shortcomings related to Study 2 data collection. The researcher was unsuccessful in recruiting adequate faculty representation from two of the programs, which may mean that key aspects of these programs were not captured. Additionally, some of the alumni participants in Study 2 were likely impacted by memory limitations, particularly those who graduated in the lower end of the years included in the study (i.e., between 2016 – 2018). To help counteract this, the researcher encouraged participants to be honest about what they could remember and did not pressure participants to provide responses when they felt unsure.

Importantly, it is suspected that the educational programs included in the study have undergone recent changes in response to the coronavirus pandemic and rising attention to social justice and structural racism. These changes may have resulted in some divergence in perspectives as some alumni would have completed their studies prior to these changes.

Reflexivity

One of the strengths of this study was the effort to maintain a reflexive stance. This stance was largely supported by a reflexive action plan (van Draanen, 2017) that helped the researcher remain aware of subjectivities and be explicit about their influence on the research and findings. The plan included the development of a social identity map, positionality

statement, reflective field notes, and dyadic reflections that helped to explicate these subjectivities and their influences on the study (Hertz, 1997).

In addition to supporting reflexivity, the action plan also supported credibility and confirmability through dyadic reflections. The researcher met with the dissertation chair and critical friend to discuss and scrutinize the research process and emergent findings. Often these discussions led to deeper reflections on the researcher's subjectivities and modifications to procedures to contribute to the soundness of the research process.

Conclusion

The relationship between evaluation theory and practice has been a topic of ongoing discussion in the field. Previous research has identified a theory-practice gap in evaluation, suggesting that evaluation practice is informed by an alternative form of knowledge inherent to practice. This research leveraged the extant research on evaluation theory and existing thinking of the epistemological foundations of nursing to investigate practical knowledge in evaluation and its relationship to evaluation theory.

The results of this research suggest that evaluation practice is rooted in knowledge of evaluation theory and other discipline-related knowledge as well as knowledge of the self, others, and the common good and equity. Collectively, these integrated ways of knowing described a holistic epistemology of practice that is unique to evaluators. It is hoped that this research will be a catalyst for further research and inform the education and professional development of evaluators.

References

Abma, T. A. (1998). Text in an evaluative context: Writing for dialogue. *Evaluation, 4*(4), 434-454

Alkin, M. C. (1991). Evaluation theory development II. In M. McLaughlin and D. Phillips (Eds.), *Evaluation and education at quarter century* (pp. 91-112). University of Chicago Press

Alkin, M. C. (2003). Evaluation theory and practice: Insights and new directions. In C. Christie (Ed.), The practice-theory relationship in evaluation. *New Directions for Evaluation, 97*, 81-89

Alkin, M. (2013). Comparing evaluation points of view. In M.C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed., pp. 3–9). SAGE Publications Ltd.

American Evaluation Association [AEA]. (n.d.-a). *AEA Outstanding Evaluation Award.* https://www.eval.org/About/Awards/AEA-Outstanding-Evaluation-Award

American Evaluation Association [AEA]. (n.d.-b). A*EA Marcia Guttentag Promising New Evaluator Award.* https://www.eval.org/About/Awards/AEA-Marcia-Guttentag-Promising-New-Evaluator-Award

American Evaluation Association [AEA]. (n.d.-c). *Evaluation 2022 proposal submission*. https://www.eval.org/Events/Evaluation-Conference/Conference-Proposal-Submission

American Evaluation Association [AEA]. (2011). *American Evaluation Association public Statement on Cultural Competence in Evaluation*. www.eval.org

American Evaluation Association [AEA]. (2018a). *Evaluator competencies*.

https://www.eval.org/Portals/0/Docs/AEA%20Evaluator%20Competencies.pdf

American Evaluation Association [AEA]. (2018b). Guiding Principles for evaluators.

https://www.eval.org/Portals/0/Docs/AEA_289398-

18_GuidingPrinciples_Brochure_2.pdf

Argyris, C., & Schön, D. A. (1974). *Theory in practice: Increasing professional effectiveness*.

Jossey-Bass Publishers

Assarroudi, A., Heshmati Nabavi, F., Armat, M. R., Ebadi, A., & Vaismoradi, M. (2018).

Directed qualitative content analysis: The description and elaboration of its underpinning

methods and data analysis process. *Journal of Research in Nursing, 23*(1), 42-55

Attia, M., & Edge, J. (2017). Be (com)ing a reflexive researcher: A developmental approach to

research methodology. *Open Review of Educational Research, 4*(1), 33-45

Bell, R. C. (2003). The Rep-grid Technique. In F. Fransella (Ed.), *International handbook of

personal construct psychology* (pp. 95-103). J. Wiley & Sons

Bergson, H. (2001). *Time and free will: An essay on the immediate data of consciousness*. Dover

Bolam, B., Gleeson, K., & Murphy, S. (2003). " Lay person" or" health expert"? Exploring

theoretical and practical aspects of reflexivity in qualitative health research. *Forum:

Qualitative Social Research, 4*(2), Article 26. https://www.qualitative-

research.net/index.php/fqs/article/view/699

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research

Journal, 9*(2), 27-40

Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*.

Routledge

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research

in Psychology, 3*(2), 77–101

Braun V., Clarke V., Hayfield N., & Terry G. (2019) Thematic analysis. In P. Liamputtong (Ed.),

Handbook of research methods in health social sciences (pp. 843-860). Springer

https://doi.org/10.1007/978-981-10-5251-4_103

Brookfield, S. D. (2017). *Becoming a critically reflected teacher* (2nd ed.). Jossey Bass

Carper, B. A. (1978). Fundamental patterns of knowing in nursing. *Advances in Nursing Science,

1*(1), 13–24

Centers for Disease Control and Prevention [CDC]. (1999). Framework for program evaluation

in public health. *Morbidity and Mortality Weekly Reports, 48*(11), No. RR-11

Chelimsky, E. (1998). The role of experience in formulating theories of evaluation practice.

American Journal of Evaluation, 19(1), 35–55.

https://doi.org/10.1177/109821409801900104

Chelimsky, E. (2012). Balancing evaluation theory and practice in the real world. *American

Journal of Evaluation, 34*(1), 91–98. https://doi.org/10.1177/1098214012461559

Chen, H. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated

evaluation perspective* (2nd ed.). SAGE Publications

Chen, H., & Rossi, P. H. (1980). The multi-goal, theory-driven approach to evaluation: A model

linking basic and applied social science. *Social Forces, 59*(1), 106-122.

https://doi.org/10.2307/2577835

Chen, H., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review, 7*(3), 283–302. https://doi.org/10.1177/0193841X8300700301

Chinn, P. L., & Kramer, M. K. (2018). *Knowledge development in nursing: Theory and process* (10th ed.). Elsevier, Inc.

Chouinard, J. A., Boyce, A. S., Hicks, J., Jones, J., Long, J., Pitts, R., & Stockdale, M. (2017). Navigating theory and practice through evaluation fieldwork: Experiences of novice evaluation practitioners. *American Journal of Evaluation, 38*(4), 493-506

Christie, C. A. (2003a). Editor's notes. In C. Christie (Ed.), The Practice-theory relationship in evaluation. *New Directions for Evaluation, 97*, 1–6

Christie, C. A. (2003b). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. In C. Christie (Ed.), The practice-theory relationship in evaluation. *New Directions for Evaluation, 97*, 7–36

Christie, C. A. (2003c). Understanding evaluation theory and its role in guiding practice: Formal, folk, and otherwise. In C. Christie (Ed.), The practice-theory relationship in evaluation. *New Directions for Evaluation, 97*, 91-93

Christie, C. A. (2012). Advancing empirical scholarship to further develop evaluation theory and practice. *Canadian Journal of Program Evaluation, 26*(1), 1–18

Christie, C. A., & Alkin, M. (2013). An evaluation theory tree. In M.C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed., pp. 11–57). SAGE Publications Ltd.

Christie, C. A., & Lemire, S. T. (2019). Why evaluation theory should be used to inform

    evaluation policy. *American Journal of Evaluation, 40*(4), 490–508.

    https://doi.org/10.1177/1098214018824045

Christie, C. A., & Masyn, K. (2010). Latent Profiles of evaluators' self-reported practices.

    Canadian Journal of Program Evaluation, 23(2), 225–254

Christie, C. A., Quiñones, P., & Fierro, L.A. (2014). Informing the discussion on evaluator

    training: A look at evaluators' course taking and professional practice. *American Journal*

    of Evaluation, 35(2), 274–290. https://doi.org/10.1177/1098214013503697

Christie, C. A., & Rose, M. (2003). Learning about evaluation through dialogue: Lessons from

    an informal discussion group. *American Journal of Evaluation, 24*(2), 235–243.

    https://doi.org/10.1177/109821400302400207

Cornelius, N. (2015). Personal construct theory, research, and practice in the field of business

    and management. In D.A. Winter and N. Reed (Eds.), *The Wiley handbook of personal*

    construct psychology (pp. 265-281). John Wiley & Sons, Ltd.

Costa, A. L., & Kallick, B. (1993). Through the lens of a critical friend. *Educational*

    leadership, 51, 49-49

Cousins, J. B., & Earl, L. M. When the boat gets missed: Response to M. F. Smith. *American*

    Journal of Evaluation, 20(2), 309–317

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five*

    approaches (3rd ed.). SAGE Publications

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research*

    (3rd ed.). SAGE

Datta, L. E. (2003). Important questions, intriguing method, incomplete answers. In C. Christie

  (Ed.), The practice-theory relationship in evaluation. *New Directions for Evaluation, 97*,

  37-46

Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*.

  Sage

Dewey, J. (1910). *How We Think*. DC Heath

Dewey, J., Montrosse, B., Schroter, D., Sullins, C., & Mattox, J. (2008). Evaluator competencies:

  What's taught versus what's sought. *American Journal of Evaluation, 29*(3), 268–287.

Donaldson, S. I., Gooler, L. E., & Scriven, M. (2002). Strategies for managing evaluation

  anxiety: Toward a psychology of program evaluation. *American Journal of Evaluation,

  23*(3), 261–273. https://doi.org/10.1177/109821400202300303

Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice:

  Developing practical knowledge. In I. Shaw, J. Greene, & M. Mark (Eds.), *The SAGE

  handbook of evaluation* (pp. 57–75). SAGE Publications Ltd.

  https://doi.org/10.4135/9781848608078.n2

do Nascimento Souto, P. C. (2013). Beyond knowledge, towards knowing: The practice-based

  approach to support knowledge creation, communication, and use for innovation. *RAI

  Revista de Administração e Inovação, 10*(1), 51-79

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced

  Nursing*, *62*(1), 107-115.

Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation practice, 15*(1), 1-15

Finlay, L. (2008). Reflecting on 'reflective practice'. Practice-based professional learning. *The Open University,* Paper 52. https://oro.open.ac.uk/68945/

Finlay, L. (2002). "Outing" the researcher: The provenance, process, and practice of reflexivity. *Qualitative Health Research, 12*(4), 531–545. https://doi.org/10.1177/104973202129120052

Fitzpatrick, J., Christie, C., & Mark, M. M. (2009). *Evaluation in action: Interviews with expert evaluators.* Sage

Fournier, D. (2005). Evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 140–140). SAGE Publications, Inc.

Fransella, F. (2015). What is a personal construct?. In D.A. Winter and N. Reed (Eds.), *The Wiley handbook of personal construct psychology* (pp. 1-8). John Wiley & Sons, Ltd.

Fransella, F., Bell, R., & Bannister, D. (2004). *A manual for rep-grid technique*. John Wiley & Sons

Frierson, H. T., Hood, S., & Hughes, G. B. (2002). Strategies that address culturally responsive evaluation. In J. Frechtling (Ed.), *The 2002 user-friendly handbook for project evaluation* (pp. 63–73). National Science Foundation

Frierson, H. T., Hood, S., Hughes, G. B., & Thomas, V. G. (2010). A guide to conducting culturally responsive evaluations. In J. Frechtling (Ed.), *The 2010 user-friendly handbook for project evaluation* (pp. 75–96). National Science Foundation

Galport, N., & Azzam, T. (2017). Evaluator training needs and competencies: A gap analysis. *American Journal of Evaluation, 38*(1), 80-100

Gherardi, S. (2009). Introduction: The critical power of the practice lens. *Management learning, 40*(2), 115-128

Ghere, G., King, J. A., Stevahn, L., & Minnema, J. (2006). A professional development unit for reflecting on program evaluator competencies. *American Journal of Evaluation, 27*(1), 108-123

Glaser B. G., & Strauss, A.L. (1967). *The discovery of grounded theory: strategies for qualitative research*. Adline

Grady, M. P. (1998). *Qualitative and action research: A practitioner handbook*. Phi Delta Kappa International.

Greene, J. C. (2001, November). *Introduction to the panel on understanding evaluators' practice decisions: The influence of theory, philosophy, values, experience, and individuality* [Conference session]. American Evaluation Association Annual Conference, St. Louis, Missouri

Greene, J. C. (2013). Making the world a better place through evaluation. In M. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed., pp. 208–217). SAGE

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*(2), 265-275

Guzman, G. (2009). What is practical knowledge? *Journal of Knowledge Management, 13*, 86-98

Henry, G. T., & Mark, M. M. (2003). Toward an Agenda for Research on Evaluation. In C.

Christie (Ed.), The practice-theory relationship in evaluation. *New Directions for
Evaluation, 97*, 69–80. https://doi.org/10.1002/ev.77

Hertz, R. (1997). Introduction: Reflexivity and voice. In R. Hertz (Ed.), *Reflexivity &
voice* (pp. vii–xviii). Thousand Oaks, CA: Sage

Hillier, Y. (1998). Informal practitioner theory: Eliciting the implicit. *Studies in the Education of
Adults, 30*(1), 35–52. https://doi.org/10.1080/02660830.1998.11730671

Honey, P. (1979). The rep-grid in action. *Industrial and Commercial Training, 11*, 452-459

Hooper, B. (2010). Falling forward: Lessons learned from critical reflection on an evaluation

process with a prisoner reentry program. In M. Freeman (Ed.), Critical social theory and

evaluation practice. *New Directions for Evaluation, 127*, 21–34.

https://doi.org/10.1002/ev.336

Hora, M. T., & Smolarek, B. B. (2018). Examining faculty reflective practice: A call for critical

awareness and institutional support. *The Journal of Higher Education, 89*(4), 553–581.

https://doi.org/10.1080/00221546.2018.1437663

House, E. R. (1980). *Evaluating with validity*. Sage Publications

House, E. R. (1993). *Professional evaluation: Social impact and political consequences*. Sage

Publications

House, E. R. (2014). *Evaluating: Values, biases, and practical wisdom*. IAP

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content

analysis. *Qualitative Health Research, 15*(9), 1277-1288

Hutchinson, K. (Ed.). (2018). *Evaluation failures: 22 tales of mistakes made and lessons learned*. SAGE Publications

Isaac, S., & Michael, W. B. (1997). *Handbook in research and evaluation: For education and the behavioral science* (3rd ed.). EdITS

Jacobson, D., & Mustafa, N. (2019). Social identity map: A reflexivity tool for practicing explicit positionality in critical qualitative research. *International Journal of Qualitative Methods, 18,* 1-12. https://doi.org/10.1177/1609406919870075

Jankowicz, D. (2004). *The easy guide to repertory grids*. Wiley

Jewiss, J., & Clark-Keefe, K. (2007). On a personal note: Practical pedagogical activities to foster the development of "reflective practitioners." *American Journal of Evaluation, 28*(3), 334–347. https://doi.org/10.1177/1098214007304130

Johns, C. (1995). Framing learning through reflection within Carper's fundamental ways of knowing in nursing. *Advanced Nursing Practice, 22*, 226–234

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141-151

Kelly, G. A. (1955). *The psychology of personal constructs: Vol.1. A theory of personality*. Norton & Company

Kelly, G. A. (1991). *The psychology of personal constructs: Vol. 1. A theory of personality.* Taylor & Frances/Routledge. (Original work published in 1955)

Kelly, G. A. (2003). A Brief introduction to personal construct theory. In F. Fransella (Ed.), *International handbook of personal construct psychology* (pp. 3-20). John Whiley & Sons, Ltd.

Kennedy, M. M. (1982). *Working knowledge and other essays*. Huron Institute

Kimball, E. (2016). Reconciling the knowledge of scholars and practitioners: An extended case analysis of the role of theory in student affairs. *Critical Questions in Education, 7*(3), 287–305

King, J. A. (2003). The challenge of studying evaluation theory. In C. Christie (Ed.), The practice-theory relationship in evaluation. *New Directions for Evaluation, 97*, 57–68. https://doi.org/10.1002/ev.76

King, J. A., & Stevahn, L. (2012). *Interactive evaluation practice: Mastering the interpersonal dynamics of program evaluation*. Sage Publications.

King, J. A., & Stevahn, L. (2020). Presenting the 2018 AEA evaluator competencies. In J. A. King (Ed.), The American Evaluation Association's program evaluator competencies. *New Directions for Evaluation, 168*, 49–61

Kirkwood, B. R. (1988). *Essentials of medical statistics*. Blackwell Science

Kuipers, K., & Grice, J. W. (2009). Clinical reasoning in neurology: Use of the rep-grid technique to investigate the reasoning of an experienced occupational therapist. *Australian Occupational Therapy Journal, 56*(4), 275–284. https://doi.org/10.1111/j.1440-1630.2008.00737.x

LaVelle, J. M. (2018). *2018 directory of evaluator education programs in the United States*. University of Minnesota Libraries Publishing. https://hdl.handle.net/11299/200790

LaVelle, J. M., & Donaldson, S. I. (2010). University-based evaluation training programs in the

United States 1980—2008: An empirical examination. *American Journal of*

Evaluation, 31(1), 9-23

LaVelle, J. M., & Donaldson, S. I. (2021). Opportunities and challenges ahead for university-

based evaluator education programs, faculty, and students. *American Journal of*

Evaluation, 42(3), 428-438

Leviton, L. C. (2015). Evaluation practice and theory: Up and down the ladder of abstraction.

American Journal of Evaluation, 36(2), 238–242.

https://doi.org/10.1177/1098214015573070

Mark, M. M., Henry, G. T., & Julnes, G. (1999). Toward an integrative framework for evaluation

practice. *American Journal of Evaluation, 20*(2), 177-198

Mason, S., & Azzam, T. (2019). In need of an attitude adjustment? The role of data visualization

in attitude change and evaluation influence. *American Journal of Evaluation, 40*(2), 249–

267. https://doi.org/10.1177/1098214018778808

McDowall, R.C. & Saunders, M. (2018). A novel use of Honey's aggregation approach to the

analysis of Repertory Grids. *Field Methods, 31*(1).

https://doi.org/10.1177/1525822X18806259

Mertens, D. M., & Wilson, A. T. (2019). *Program evaluation theory and practice: A*

comprehensive guide (2nd ed.). Guilford Press

Mezirow, J. (1990). *Fostering critical reflection in adulthood: A guide to transformative and*

emancipatory learning. Jossey-Bass Publishers

158

Mezirow, J. (2009). An overview of transformative learning. In K. Illeris (Ed.), *Contemporary theories of learning* (pp. 90-105). Routledge

Morris, M. (2002). Ethical challenges. *American Journal of Evaluation, 23*(2), 183–185. https://doi.org/10.1177/109821400202300206

Nicolini, D., Gherardi, S. and Yanow, D. (2003). Introduction: toward a practice-based view of knowing and learning in organizations. In D. Nicolini, S. Gherardi, & D. Yanow, D. (Eds.), *Knowing in organizations* (pp. 3-31). M.E. Sharpe

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage Publications

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage Publications

Patton, M. Q. (2013). *Utilization-focused evaluation (UF-E) checklist.* https://wmich.edu/sites/default/files/attachments/u350/2014/UFE_checklist_2013.pdf

Polanyi, M. (1966). *The tacit dimensions*. The University of Chicago Press

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*(4), 443-459

Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Sage

Punwar, A. J., & Peloquin, S. M. (2000). *Occupational therapy: Principles and practice* (3rd ed). Lippincott Williams & Wilkins

Raelin, J. A. (2007). Toward an epistemology of practice. *Academy of Management Learning and Education, 6*(4), 495-519

Rog, D. J. (2015). Infusing theory into practice, practice into theory: Small wins and big gains for evaluation. *American Journal of Evaluation, 36*(2), 223–238. https://doi.org/10.1177/1098214015573068

Rojon, C., McDowall, A., & Saunders, M. N. (2019). A novel use of honey's aggregation

approach to the analysis of repertory grids. *Field Methods, 31*(2), 150-166

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H. &

Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and

operationalization. *Quality & Quantity*, *52*(4), 1893-1907.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Schwandt, T. A. (2003). 'Back to the rough ground! 'Beyond theory to practice in

evaluation. *Evaluation, 9*(3), 353-364.

Schwandt, T. A. (2005). The centrality of practice to evaluation. *American Journal of*

Evaluation, 26(1), 95–105. https://doi.org/10.1177/1098214004273184

Schwandt, T. A. (2008). The relevance of practical knowledge traditions to evaluation practice.

In N.L. Smith & P.R Brandon (Eds.), *Fundamental issues in evaluation* (pp. 30–41).

Guilford Press

Schwandt, T. A. (2009). Toward a practical theory of evidence for evaluation. In S.I. Donaldson,

C.A. Christie, & M.M. Mark (Eds.), *What counts as credible evidence in applied*

research and evaluation practice (pp. 197-252). Sage Publications, Inc.

Schwandt, T. A. (2017). Professionalization, ethics, and fidelity to an evaluation ethos. *American*

Journal of Evaluation, 38(4), 546–553. https://doi.org/10.1177/1098214017728578

Schwartz, B., & Sharpe, K. (2010). *Practical wisdom*. Riverside Books

Scriven, M. (1991). *Evaluation thesaurus* (4th ed). Sage Publications.

Scriven, M. (1998, May 8-9). *The meaning of bias* [Conference session]. Proceedings of the

    Stake symposium on educational evaluation, Champaign, Illinois.

    https://files.eric.ed.gov/fulltext/ED432593.pdf

Senge, P.M. (1990). The fifth discipline: *The art and practice of the learning*

    organization. Doubleday/Currency.

Shadish, W. (1998). Evaluation theory is who we are. *American Journal of Evaluation, 19*(1), 1–

    19. https://doi.org/10.1016/S1098-2140(99)80177-5

Shadish, W. R. (1994). Need-based evaluation theory: What do you need to know to do good

    evaluation? *Evaluation Practice, 15*(3), 347-358

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation:*

    Theories of practice. Sage Publications, Inc.

Shadish, W. R., & Epstein, R. (1987). Patterns of program evaluation practice among members

    of the evaluation research society and evaluation network. *Evaluation Review, 11*(5),

    555–590.

Shanker, V. (2019). *Definitional tension: The construction of race in and through evaluation*

    (Publication No. 27735332) [Doctoral dissertation, University of Minnesota]. ProQuest

    Dissertations and Theses Global

Smith, N. L. (1993). Improving evaluation theory through the empirical study of evaluation

    practice. *Evaluation Practice, 14*(3), 237–242.

    https://doi.org/10.1177/109821409301400302

Stake, R. (2003). Responsive evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.),

    International handbook of educational evaluation (pp. 63–68). Springer Netherlands.

    https://doi.org/10.1007/978-94-010-0309-4_5

Stake, R., & Abma, T. (2005). Responsive evaluation. In S. Mathison (Ed.), *Encyclopedia of*

    evaluation. SAGE

Starks, H., & Brown Trinidad, S. (2007). Choose your method: A comparison of

    phenomenology, discourse analysis, and grounded theory. *Qualitative Health*

    Research, 17(10), 1372-1380.

Stevahn, L., Berger, D. E., Tucker, S. A., & Rodell, A. (2020). Using the 2018 AEA Evaluator

    Competencies for effective program evaluation practice. In J. A. King (Ed.), The

    American Evaluation Association's program evaluator competencies. *New Directions for*

    Evaluation, 168, 75–97

Stevahn, L., King, J. A., Ghere, G., & Minnema, J. (2005). Establishing essential competencies

    for program evaluators. *American Journal of Evaluation, 26*(1), 43-59.

Stufflebeam, D. (2001). Evaluation models. In G.T. Henry (Ed.), Evaluation models [Special

    issue]. *New Directions for Evaluation, 89*, 7-98

Symonette, H. (2004). Walking pathways toward becoming a culturally competent evaluator:

    Boundaries, borderlands, and border crossing. In M. Thompson-Robinson, R. Hopson, &

    S. SenGupta (Eds.), Special issue: In search of cultural competence in evaluation:

    Toward principles and practice. *New Directions for Evaluation, 102*, 95-109

Thompson, S., & Thompson, N. (2018). *The critically reflective practitioner* (2nd ed.). Palgrave

Tourmen, C. (2009). Evaluators' decision making: The relationship between theory, practice, and experience. *American Journal of Evaluation, 30*(1), 7–30. https://doi.org/10.1177/1098214008327602

Usher, R., & Bryant, I. (2014). *Adult education as theory, practice, and research: The captive triangle*. Routledge

van Draanen, J. (2017). Introducing reflexivity to evaluation practice: An in-depth case study. *American Journal of Evaluation, 38*(3), 360–375. https://doi.org/10.1177/1098214016668401

Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Prentice Hall

Weiss, C. H. (2013). Rooting for evaluation: Digging into beliefs. In M.C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences* (2nd ed., pp. 130-143). SAGE Publications, Ltd.

Wilce, M., Fierro, L.A., Gill, S., Perkins, A., Kuwahara, R., Barrera-Disler, S., Orians, C., Codd, H., Castleman, A.M., Nurmagambetov, T., Anand, M. (2021). *Planting the seeds for high-quality program evaluation in public health*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Environmental Health, Division of Environmental Health Science and Practice, Asthma and Community Health Branch

Wilkin K. (2002) Exploring expert practice through reflection. *Nursing in Critical Care 7*(2), 88–93

Worthen, B. R. (1990). Program evaluation. In H. J. Walberg & G.D. Haertel (Eds.), *The international encyclopedia of educational evaluation* (pp.42-47). Oxford, England: Pergamon Press

Worthington, Roger L., & Tiffany A. Whittaker. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806-838

Yarbrough, D.B., Shula, L.M., Hopson, R.K., & Caruthers, F.A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd. ed). Corwin Press

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology, 9*(2), 79-94

Zander, P. E. (2007). Ways of knowing in nursing: The historical evolution of a concept. *Journal of Theory Construction and Testing, 11*(1), 7–11

Appendix A: Ways of Knowing in Nursing

Ways of Knowing	Dimensions				
	Critical Questions	Creative Processes	Formal Expressions	Authentication Processes	Integrated Expression in Practice
Empirics	What is this? How does it work?	Conceptualizing Structuring	Facts Models Formal description Theories Thematic descriptions Empirics research reports	Confirmation Validation	Scientific competence
Ethics	Is this right? Is this responsible?	Clarifying Exploring	Principles and codes	Dialogue Justification	Moral and ethical comportment
Personal	Do I know what I do? Do I do what I know?	Opening Centering	Personal stories Genuine self	Response Reflection	Therapeutic use of self
Aesthetics	What does this mean? How is this significant?	Envisioning Rehearing	Aesthetic criticism Works of art	Appreciation Inspiration	Transformative art(s)
Emancipatory	Who benefits? What is wrong with this picture? What are the barriers to freedom? What changes are needed?	Critiquing Imagining	Action plans Manifestos Critical analyses Visions for the future	Social equity Sustainability Empowerment Demystification	Praxis

Note. From Knowledge Development in Nursing (10th ed., p. 15) by P.L. Chinn and M.K. Kramer. Copyright 2018 by Elsevier Inc.

Appendix B: Reflexive Action Plan

Intentional Activities	Timing/Frequency	Objectives
Social Identity Map (Jacobson & Mustafa, 2019) and Positionality statement (Creswell, 2013)	Pre-data collection/once	• Support prospective reflexivity (Attia & Edge, 2017). • Sensitize myself, my advisors, the readers, and others to my motivations and interests in the study (prospective reflexivity). • Unearth and communicate preconceived notions and assumptions that influence how I perceive the research context, data, and participants. • Provide a baseline for retrospective reflexivity (Attia & Edge, 2017).
Field Notes	During data collection/written after each interview/reviewed weekly	• To capture a descriptive and reflective account of each interaction with a participant. • Descriptive notes will capture information pertaining to the context, participant, and interview process. They may lead to adjustments to procedures or protocols. • Reflective notes will focus on what was learned (the expected and unexpected), my experience and what I think accounted for that experience, including my assumptions and speculations. • New theoretical insights and the evidence to back them up.
Self-interviewing/journaling (Bolam et al., 2003)	Data collection to completion/weekly	• Maintain an ongoing narrative about the researcher's position, context, and the dynamic between the two. • Document and scrutinize the research process, patterns of thought, and formulated conclusions.

		• Unearth assumptions and assess their impact on the research. Determine what, if anything, should be done to manage assumptions moving forward.
Dyadic reflection with dissertation chair	Data collection to completion/bi-monthly	• Explore and scrutinize theory development and findings. • Review and reflect on matters pertaining to the integrity of the study. Discuss the impact of these matters on the study and any needed next steps. • Trouble-shoot issues that arise.
Dyadic reflection with critical friend	Data collection to completion/monthly	• Reflect on the research with the support of "a trusted person who asks provocative questions, provides data to be examined through another lens, offers critiques of a person's work as a friend … is an advocate for the success of that work" (Costa & Kallick, 1993, p. 50). • Explore and scrutinize theory development and findings. • Review and reflect on matters pertaining to the integrity of the study.

Appendix C: Positionality Statement

We do not write about things as they are or were or will be. We write about these things as we are.

—Peter Ives, The Whole Truth

The purpose of this statement is to introduce myself, explain my motivations for doing this study, and disclose some additional details that I feel are important for my readers to know. Before getting into all that, I should mention that my approach to writing this positionality statement was (I think) a little unorthodox. I intended to finalize my positionality statement when my research proposal was approved and use it as a reference point for ongoing reflection.

However, as the study unfolded, previous experiences and new connections between me and the research began to surface. I found myself returning to the statement to add and remove details that had become more and less relevant. This document is the final version of an ongoing dialogue that unfolded across the research process.

I decided to begin the statement with a personal story. Robert Nash talks about how stories can get us closer to knowing who we are and who others are. In this spirit, the statement begins with a personal story about an encounter with the ways of knowing in nursing, which is a connection I made between my history and the research topic during the research process.

How I Can to Know the Ways of Knowing

Well over a decade ago, my family and I gathered at a hospital for the last day of my grandmother's life. My grandma lived a long and happy life and had good health for most of it. But her health had been deteriorating, and it was clear to all involved that it was time to let Grandma go. So, the day at the hospital was about being there for Grandma and each other as she passed away.

My grandmother had been assigned to nurse as per standard hospital procedures. The first thing I noticed about the nurse was that they seemed young and probably hadn't been nursing for very long. However, they seemed knowledgeable and confident, so their age was not a concern.

Another thing that I remember about the nurse is how attentive they were and their efforts to keep Grandma comfortable. I found out later that I wasn't the only one who noticed this about the nurse. My Mom, for example, remembers how the nurse washed and styled my grandma's hair. What I remember is the nurse periodically switching out Grandma's blankets with warm ones, even near the very end when grandma was unconscious and non-responsive. It was a difficult day, but I've always felt at peace with how it unfolded.

Over the years, I have reflected on that day a lot, and usually after learning something new about dying and what's involved when a person passes from natural causes. As part of this, I have had to come to terms with my experience and things that I thought happened (e.g., a hand squeeze) that probably did not. It also led me to wonder about the reasoning behind some of the nurse's care. Obviously, they weren't intended to change the outcome. But would someone in Grandma's condition even notice a warm blanket or being repositioned?

Fast forward to the present time when I am familiar with the Ways of Knowing in Nursing. This literature discusses the therapeutic use of self, which, in nursing, is an expression of personal knowledge that is integrated with all ways of knowing (Chinn & Kramer, 2018). In the simplest terms, the therapeutic use of self is the understanding that you influence the contexts in which you are providing care and can, therefore, contribute to therapeutic processes. Contributing to therapeutic processes can take many forms, including being present, non-judgment, and doing things that will make patients feel at ease with their situation.

Admittedly, it did take some time to connect the therapeutic use of self with my hospital experience. Now I can see that it was my narrow understanding of practice, including the patients receiving care and outcomes of practice, that led me to question the nurse's actions. The nurse, despite their newness to practice, seemed to understand that caring they were caring for a family and that for their patient is therapeutic for a grieving family, which is a desirable outcome of practice.

Who am I?

This section corresponds to the social identity map I created as part of my reflexive action plan. The details included here are those identities, and related attributes, that I feel are most important for the study's audience. I am a white, middle-class, heterosexual cis woman. I enjoy many privileges and live a pretty comfortable existence free from ridicule. In fact, many of the things most meaningful (i.e., marriage, motherhood, education, and employment) stem from my privileged position.

I have been educated and socialized with western values and norms, which shape my worldview. My western orientation has influenced all aspects of the research, including my belief that practical knowledge and ways of knowing can and should be studied. That said, my interest in critical reflection and some other life changes have opened me to new values and perspectives (e.g., wanting to understand what is) and helped to alleviate some western values that I no longer see as useful (e.g., wanting to control what is).

Thinking specifically about myself in relation to the study, there are three additional qualities or characteristics to raise. First, I am an immigrant, having moved to the U.S. from Canada about 10 years ago. As with many other immigrants, I'm sure I often feel stuck in limbo somewhere between my home and my new country. Whether I'm talking to Canadians or Americans or learning about something that occurred in either country, I always feel like I'm on the fringes until I figure out how to situate myself. The silver lining is that I think this feeling of limbo has caused me to acquire a more reflective stance toward life, which includes an appreciation for reflective practice and critical reflection.

Second, I identify more as a practitioner than a researcher. After starting my doctoral studies, I immediately learned how poorly prepared I was for primary research. Adding to this was an established practitioner identity from a previous career. This is not to say I am not interested in research but that my immediate concerns tended to be about service and meeting a defined need. Moreover, I see myself as a practitioner whose work intersects with research in various ways, and I see ways of knowing as embodying that intersection. Thus, the topic of this research is connected to how I see my practice.

Last, I am a pretty pragmatic person, as I think the research area (i.e., practical knowledge) suggests. But I'm also theoretical in the sense that I like to know how things work, why they work the way they do, and to what end. I think this is part of the reason why I tend to gravitate to more descriptive research and topics that are unfamiliar to my discipline.

It also explains why I see professional practice as theoretical. I believe there is an underlying logic to how people engage in their craft that does not necessarily match with something you might read in a textbook.

Motivations for the Research

My initial motivation for this research was to help evaluators gain a deeper understanding of self and a reflective stance toward their practice. Reflection and critical reflection are often described as essential to evaluation. However, there is limited information about what these concepts mean in evaluation and how evaluators can develop and refine their reflective capabilities.

As part of my doctoral studies, I have explored the broader literature on reflective practice (Schön) and critical reflection (Mezirow, Brookfield) as well as adult learning and management (Argyris & Schon, Senge). This literature has fed my interest and helped me to understand these concepts and how they are embedded in professional practice. I have also felt humbled by the efforts taken by other disciplines to ground reflection and critical reflection into their scholarly works, educational programs, and practice.

For instance, in fields such as social work, education, and nursing, reflective practice and critical reflection are often integrated into educational programs, so students develop an understanding of them and how to connect them to practice. Additionally, these fields offer a wealth of literature that has been developed over decades. Evaluation, comparatively speaking, had taken much smaller strides despite arguments that it is essential to practice.

My explorations also exposed me to different frameworks and models intended to aid practitioners in their reflective work. I have always been a bit leery of reflective frameworks, particularly those that involve pre-established steps or questions, because I think they can lead to reflection that is impersonal and superficial. However, my feelings changed when I came across the Ways of Knowing in Nursing, which seemed better suited for productive self-exploration and learning.

As I read more about the Ways of Knowing in Nursing, I learned that it was more than a tool for reflection. It represents the epistemological and ontological foundations of nursing and the processes that nurses engage in to understand what is needed of them and how to shape their practice to meet that need. It acknowledges that practice is both personal and empirical, as well as dynamic and thoughtful. To be truthful, it was a breath of fresh air after spending years learning about evaluation checklists lists, steps, and cycles.

When I compared nursing's conceptualization of practice with evaluation, it seemed like there was room for more self-discovery. As such, the second motivation for the study was to help evaluation field develop a better understanding of its practice, including the underlying epistemology and the wisdom of practitioners.

My Assumptions

As a practitioner and student of evaluation, I am not a neutral participant in the research process. I have preconceived notions, assumptions, and 'baggage' (Finlay, 2002, p.698) related to evaluation practice and evaluator knowing based on my experiences, identities, and worldview. This study is premised on the assumption that evaluation practice involves ways of knowing and is a process of exploring, understanding, experimenting, and reflecting. Underlying this premise are numerous other related assumptions.

- I assume that ways of knowing help evaluators understand evaluation contexts and how to adapt their practice accordingly.
- I assume that some ways of knowing are unique to individuals, while others are common to the discipline, at least at some level.
- I assumed that evaluator knowing could be connected to various dimensions of our being (e.g., mind, body, spirit). These dimensions draw in critical intel needed to know or understand our and others' experiences.

These assumptions were key areas of focus in my efforts to maintain a reflexive stance throughout the study. They were commonly revisited in my field notes and when I engaged in self-interviewing. For example, I would often consider the theorists or evaluators who would disagree with my interpretations and rationalize the basis of their argument. I would also challenge myself to see the weakness of my interpretations and identify contradictions or "tunnel vision" in my thinking.

These assumptions were also points of discussion during my dyadic reflection discussions with my critical friend and advisor. The intent of these discussions, particularly those with my critical friend, were to scrutinize and critique my interpretations and examine how my identities and assumptions were influencing the study.

The Ways of Knowing in Nursing framework was also an assumption that I brought to this study. While I did not use it in my instrument development or analyses, it ultimately refined my research topic and decision to examine evaluation practice holistically and through a knowing lens. Additionally, it provided some of the language I used in the study, such as expressions and creative processes of knowing, and informed how I described some of the findings.

The Ways of Knowing in Nursing has influenced me personally as described. The research was well underway before I made the personal connection, so it's difficult to determine the extent to which it influenced my perception of professional practice or how I approach the study. However, the connection brought the concept into my reality, which is difficult to undo for the sake of research. Also, observing the ways of knowing in a new professional reinforced an existing assumption that it is a teachable concept.

My Relationship to the Research Context and Participants

The research focused on two contexts: the practice and formal education of evaluators. As an evaluation doctoral student and practitioner, I am situated in both contexts and have spent most of the last seven years of my life toggling between them. I am also related to each context differently.

As an evaluation practitioner, I shared a professional identity with the participants in Study 1. This shared identity included a unique language, competencies, theories, and social purpose. Importantly, our professional practice and, therefore, our perceptions of quality and ethical practice were guided by similar values and principles.

Still, my interpretations were likely influenced by my sense of how evaluation is best practiced. I likely gravitated towards participants who spoke readily about the importance of self-awareness and critical reflection and see knowledge generation and enlightenment as key evaluation purposes. Additionally, as a formal sponsor and user of evaluation, I was potentially more sensitive to data that could be interpreted as undermining or undervaluing the contributions of evaluation partners and clients.

Fortunately, the incorporation of the rep-grid technique into Study 1 helped to control for these biases by allowing participants to describe their practice in their own words. Additionally, I was able to provide participants with a copy of their grid for member checking immediately after their session.

Despite this professional connection, it is worth noting that there was a considerable gap in experience and practical knowledge between the participants and me. This gap contributed to a power dynamic in this research context. This dynamic may not have been a substantial concern in a more conventional, semi-structured interview format. However, I felt that I needed to tread carefully in the repertory grid interviews, where I often needed to push participants for clarification and assist them in clarifying their thinking.

Study two engaged evaluation faculty and alumni/advanced graduate students from a sample of evaluation programs in the U.S. As an advanced evaluation student, I easily related to the experiences and perspectives of the student participants. I also have my own perceptions about what an effective evaluation professor and advisor does and what aspects of teaching evaluation are and are not beneficial to students.

It is my hope that the rigor I applied to the analysis helped to control for these biases. I felt the directed content analysis approach provided some assistance in this regard because I had an existing theory to guide the early stages of my analysis. The structured nature of my analysis also provided a clear boundary for coding. Additionally, I tried to present as much evidence (i.e., raw data) as possible, without overwhelming the manuscript, so that the audience can make their own determinations about the validity of the findings.

Evaluation from Both Sides

This final section provides a bit of background on the title of the manuscript. Throughout the study, I became sensitive to various dichotomies, some rigid and others subtle, that seemed to surface. Undoubtedly, part of this had to do with the hours spent learning and implementing the repertory grid technique, which is intended to draw out bi-polar contrasts. But my sensitivity started before that and in the early stages of the research proposal when I could sense a tension

between western and eastern orientations regarding ways of knowing. Evaluation theory and practice are also often treated as dichotomous in the literature and within the field.

However, the most impactful dichotomy surfaced when I started thinking about how the reflections of the experienced practitioners compared to those of the alumni/advanced students. It reminded me of a song called "Both Sides Now," where the singer-songwriter describes their perceptions of clouds, love, and life as both a young person and someone with more years and experience under their belt. The latter is a more knowledgeable and realistic perspective, but it's also less certain because it appreciates the complexities of life.

In some ways, I feel like this research was about looking at evaluation from different angles, particularly from novice and seasoned points of view. So, the song and story behind it seemed fitting.

Appendix D: Study 1, Strand 1 Pre-Analysis

Question	My Responses
What counts as a theme?	• A prevalent pattern across the data set means that it appears at least once in each document. Some exceptions may be made to themes that appear in two documents OR themes that are particularly informative and a rationale exists as to why they appear in one document and not others. • Capture an important element of evaluator knowing, which means: a basis for knowing (i.e., something knowing is ground in), a process of knowing, or a manifestation (or expression) of knowing. • It is anticipated that themes will be hierarchical, consisting of superordinate and subordinate themes. The highest theme will be the basis of knowing, which will encompass various processes and manifestations. • The researcher will maintain some flexibility regarding what counts as a theme as per Braun & Clarke's (2008) guidance.
What type of analysis will I do?	• The researcher is interested in providing an account of what and how evaluators know what they need to know to practice evaluation and additional information about how this knowledge shapes their practice.
How will themes be identified?	• Inductively, beginning "with specific observations" and building "towards general patterns" (Patton, 2002, p. 56). I intend to code diversly for data regarding knowledge and knowing in evaluation without consideration for the ways of knowing in nursing or other details regarding practical knowing or wisdom discussed in the evaluation literature. That said, this information is in my head, and I will take strides to be explicit about what is influencing my coding. Themes will be built from these codes.
At what level will themes be identified?	• Initial coding will be semantic (explicit to the meaning articulated). My further analysis (2nd, 3rd, and so on sweeps) will allow latent-level meaning that will go "beyond the semantic content of the data…to identify or examine the underlying ideas, assumptions and conceptualizations – and ideologies – that are theorized as shaping or informing the semantic content of the data" (Braun & Clarke, 2006, p. 13). Latent level analysis is consistent with the study's constructivist paradigm (Braun & Clarke, 2006). Additionally, it will allow the researcher to communicate what may underpin the data.

What epistemological paradigm will shape my analysis?	• Constructivist paradigm which is consistent with the overall study aims.
What questions will guide my analysis?	• How does knowing or knowledge relate to this segment? • What does this mean with respect to what and how evaluators know, and how do I know that this is the case? • Who would agree and disagree with my interpretation and how, if at all, does this impact what I am seeing in the data?

Appendix E. Summary of Personal Construct Theory and the Rep-grid Technique

The rep-grid method is a primarily quantitative approach rooted in Personal Construct Theory (PCT; Kelly, 1955, 1991) that is intended to draw out how and what individuals think about a given topic. The fundamental basis of PCT is that "a person's processes are psychologically channelized by the ways in which he anticipates events" (Kelly, 2003, p. 7). This sentence encompasses two critical components of PCT: ways (contrasts) and events (elements). The term ways refers to constructs, which are bi-polar contrasts that individuals make when construing (Fransella, 2015; Jankowicz, 2004; Kelly, 1991). In this regard, constructs are not concepts but opposing possibilities that represent how an individual conceptualizes a person, object, situation, and so on within a given context (Kelly, 2003, p. 10). Kelly (2003) argues that constructs operate within a specific context or realm of one's experience. For example, one might contrast weak with strong when thinking about their friend's physical abilities or with convincing when describing arguments made in a debate.

According to PCT, individuals encompass a network of interrelated constructs, referred to as a personal construct system (Kelly, 2003). In these systems, constructs are organized hierarchically, whereby the superordinate personal constructs form the mainspring of our psyche and the subordinate the tributaries (Fransella, 2015). An individual's personal construct system is the basis upon which they anticipate and make predictions and, consequently, respond to the world around them (Fransella, 2015; Kelly, 2003). Yet, despite this influence, people are typically unaware of their personal construct system (Kelly, 2003).

The second component referenced in the statement is events, which are the elements of PCT. Elements are simply the people, objects, or events to which constructs are applied (Kelly, 2003; Fransella et al., 2004). For example, McDowall & Saunders (2018) asked participants to identify three high, medium, and low performers with whom they interact, and these individuals became the elements of the study on workplace performance. Kuipers and Grice (2009) used various tasks of the occupational therapy process as elements when examining clinical reasoning among occupational therapists.

Kelly (2003) identified eleven corollaries inherent to personal construct systems. Among these, three are of particular importance to this study. First, the construction corollary indicates that an individual "anticipates events by construing their replications" from their personal construct systems (Kelly, 2003, p. 9). In other words, individuals draw on past experiences to strategize in the present. Second, the dichotomy corollary specifies that a personal construct system consists of "a finite number of dichotomous constructs" (Kelly, 2003, p. 10). When we construe an event, we consider its likeness while simultaneously distinguishing it from something else, the latter of which clarifies meaning (Fransella et al., 2004, p. 8).

The final corollary of significance is commonality. For Kelly (2003), the motivation behind PCT was to understand the unique and personal constructs that are meaningful to individuals (individuality corollary). However, Kelly (2003) also acknowledged the potential for similarities in constructs among groups, which motivated the use of rep-grids when studying the personal construct systems of professional groups (Cornelius, 2015; Hillier, 1998; Kuipers & Grice, 2009; McDowall & Saunders, 2018).

Appendix F: Study 1, Strand 2 Participant Invitation

Greetings,

My name is Heather Codd, and I am a Ph.D. Candidate in Evaluation and Applied Research Methods at Claremont Graduate University (CGU). As part of my dissertation, I am investigating how evaluators conduct evaluations and the knowledge that underlies their practice.

You have been identified as an individual with an interesting and valuable perspective on the study topic. As such, I would like to invite you to participate in a 60-90 minute interview over Zoom (or another platform of your preference). The interview questions will draw on your perspectives of evaluation practice and how you engage in evaluation practice.

I appreciate that you are busy and routinely receive requests for your time. However, I would very much appreciate the opportunity to learn from you and would do my best to conduct the interview as efficiently as possible. Should you be interested in and willing to participate in an interview, please complete this expression of interest survey {insert link}.

Thank you very much for your time.

Sincerely,

Heather Codd, Ph.D. Candidate
Evaluation and Applied Research Methods
Claremont Graduate University
Heather.codd@cgu.edu

Appendix G: Study 1, Strand 2 Expression of Interest Survey

Thank you for your interest and willingness to participate in the study. The purpose of this survey is to assess your eligibility for the study and gain your informed consent. You will also be able to book your interview date and time. The survey may seem a little long but my intent is to limit the number of times I need to request information from you in advance of the interview.

1. To be eligible for the study, participants must meet *both* of the following two criteria (1) reside in the United States and (2) have at least 10 years of evaluation experience. Please select the option(s) that best describes you.

   ☐   I reside in the United States and have at least 10 years of evaluation experience
   ☐   I reside in the United States but have less than 10 years of evaluation experience
   ☐   I have at least 10 years of evaluation experience but do not reside in the United States

Those who select the first option will continue to question 2. Those who select the second or third options will skip to the end of the survey.

2. Great news! You are eligible for the study. This next section is the informed consent for the study and includes important information about the benefits and risks of participation and how participant's confidentiality will be maintained. Please review it carefully and indicate whether you do or do not agree to participate.

   Agreement To Participate in a Study about *Examining Evaluators Practice* (IRB # 4024)

   You are invited to participate in a research project. Volunteering will probably not benefit you directly, but you will be helping the investigator to create new knowledge that may benefit the evaluation field. If you volunteer, you will be asked to participate in an interview that involves no more risk than what a typical person experiences on a regular day Your involvement is entirely up to you. You may withdraw at any time for any reason. Please continue reading for more information about the study.

STUDY LEADERSHIP: This research project is led by Heather Codd, PhD candidate of the Claremont Graduate University, who is being supervised by Dr. Leslie Fierro.

PURPOSE: The purpose of this study is to examine the knowledge that evaluators use to conduct evaluation work and the processes they use to perceive and interpret practice situations.

ELIGIBILITY: You are eligible for this study because you are an experienced evaluation practitioner who has been recognized by your field for exemplary practice through the American Evaluation Association's (AEA) Outstanding Evaluation award program or through

recommendation from a colleague. Additionally, you have met the residency and experience criteria.

PARTICIPATION: During the study, the investigator will facilitate a structured interview. Your role will be to respond to questions about how you engage in your evaluation practice. The interview will take 90 minutes or less of your time.

RISKS OF PARTICIPATION: The risks of participating in the study are minimal and do not exceed those of everyday life. To the extent possible, the interview time will be scheduled at your best availability to minimalize any inconvenience. Additionally, you have the options to: (1) refuse to answer specific questions and (2) withdrawal from the study at any time without penalty.

BENEFITS OF PARTICIPATION: We do not expect the study to benefit you personally. However, due to the unique focus of the study and its relationship to evaluation practice, it is believed that your participation will benefit the field of evaluation. Additionally, the interview will provide you with an opportunity to reflect on your practice, which may benefit you personally.

COMPENSATION: There is no direct compensation to you for participating in this study. However, I greatly appreciate your participation.

VOLUNTARY PARTICIPATION: Your participation in this study is completely voluntary. You may stop or withdraw from the study at any time or refuse to answer any particular question for any reason without it being held against you.

CONFIDENTIALITY: To maintain your confidentiality, I will not tie your name or any other identifying information to any details included in reports, manuscript papers, presentations, or other communications resulting from the study. I may use the data collected for future research. During the interview, I will ask for your permission to record our interview for transcription purposes. If permitted, two recordings will be made: one using Zoom (primary recording), stored in Zoom's cloud, and one on a digital recording device (backup recording). Following the interview, both recordings will be downloaded/transferred to a password protected computer and deleted from their original source (i.e., Zoom's cloud, the recording device). Once the transcription is complete, all recordings will be deleted from the computer. In order to protect the confidentiality of your responses, you will be assigned a unique identifier, which will be used on all recordings and documents resulting from your participation in the study. While I believe these measures will maintain your confidentiality, it is important to acknowledge that your participation in the study may be assumed by others since recipients of the AEA Awards are publicly available.

FURTHER INFORMATION: If you have any questions or would like additional information about this study, please contact:

Heather Codd, Ph.D. Candidate

Evaluation and Applied Research Methods
Claremont Graduate University
150 E. 10th Street
Claremont, CA, 91711
heather.codd@cgu.edu

You may also contact my faculty advisor, Dr. Leslie Fierro, Senior Fellow at Claremont Graduate University at leslie.fierro@cgu.edu.

[IF APPROVED BY IRB] The CGU Institutional Review Board (IRB) has approved this project. If you have any ethical concerns about this project or about your rights as a human subject in research, you may contact the CGU IRB at (909) 607-9406 or at irb@cgu.edu. A copy of this form will be given to you if you wish to keep it.

CONSENT. To provide your consent to participate in the study, please select "I consent to participate in this study." If you do not provide your consent, please select "I do not consent to participate in this study."

☐   Yes, I consent to participate in the study.

☐   No, I do not consent to participate in the study.

Thank you for agreeing to participate in the study! I am truly grateful for your support and look forward to interviewing you.

As a final step, please click the arrow below and you will be redirected to my interview scheduling webpage. Please note, if none of the options provided work for you, please contact me at heather.codd@cgu.edu so we can work together on finding a suitable interview time.

Section One: Rep-Grid Description
• The rep-grid consists of four components (1) topic, (2) elements, (3) constructs, and (4) ratings [show where each is located in the grid]. The topic is the subject-matter of the grid; the elements are activities or components of the subject-matter; constructs are attributes individuals assign to elements; and the ratings explain how construct and elements relate.   • The topic will be supplied to you. The topic for this rep-grid is "your evaluation practice." I am interested in the knowledge that informs your practice. In particular, I'd like to learn about what you do and pay attention to when conducting an evaluation.   • You will provide the elements, constructs, and ratings.   • My role is to facilitate that process by asking you to make a series of systematic comparisons and guiding you along.   • Do you have any questions about the grid before we begin?

Section Two: Eliciting Elements	Questions/Instructions
Elicit elements	The first step is to define the elements of your practice. I'd like you to think about your typical approach to conducting an evaluation. Please list eight – 12 things you typically do when conducting an evaluation that you feel are important to your unique practice. *Researcher documents elements in the grid.*

Section Two: Eliciting Constructs & Ratings	Questions/Instructions
Elicit and rate constructs	Thanks for providing the elements of your practice. Now, I would like to learn about how you think about these different elements, in particular, how you see them as similar and different. We will use a process of systematic comparisons to do this.
Step 1: take 3 elements at random	Which two of these are the same in some way, and different from the third in terms of what you need to know in order to conduct them?"
Step 2: record construct and check understanding.	
Step 3: present the contrast on the rating scale.	Please rate each of the three elements on the scale. Please rate each of the remaining elements on the construct.
Step 4: record the ratings on grid	

Step 5: continue to elicit constructs (repeats steps 1 – 4) until the participant cannot offer any new constructs.

Seps 6: supply overall construct and ask participants to rate each element on that construct.

Please rate each element on how fundamental or not so fundamental it is to your practice.

The Study 1, Strand 2 pre-analysis was conducted before the RTA. The RTA was conducted on the data obtained from only those participants who did not complete a full rep-grid (*n*=5). Table I1 summarizes the questions and responses that resulted from the pre-analysis.

Table I1

Strand 2 Pre-Analysis Questions and Responses (n=5)

Question	My Repones
What counts as a theme?	• Themes are prevalent patterns of meaning regarding how evaluators know and construe their practice. The analysis will identify themes that are prevalent across the sub-set and within a transcript. Both themes will be advanced for further analysis (i.e., synthesis of Strand 2 results). • Similar to Strand 1, themes will capture an important element of evaluator knowing, which means: a basis for knowing (i.e., something knowing ground in), a process of knowing, or a manifestation (or expression) of knowing. • Hierarchy among themes is anticipated. The highest theme will be the basis of knowing, which will encompass various processes and manifestations.
What type of analysis will I do?	• I am interested in providing an account of what and how evaluators know what they need to know to practice evaluation and additional information about how this knowledge shapes their practice.
How will themes be identified?	• Inductively, beginning "with specific observations" and building "towards general patterns" (Patton, 2002, p. 56). I intend to code diversly for data regarding knowledge and knowing in evaluation without consideration for the ways of knowing in nursing or other details regarding practical knowing or wisdom discussed in the evaluation literature. That said, this information is in my head, and I will take strides to be explicit about what is influencing my coding. Themes will be built from these codes.
At what level will themes be identified?	• Initial coding will be semantic (explicit to the meaning articulated). My further analysis (2nd, 3rd, and so on sweeps) will allow latent-level meaning that will go "beyond the semantic content of the data…to identify or examine the underlying ideas, assumptions and conceptualizations – and ideologies – that are theorized as shaping or informing the

	semantic content of the data" (Braun & Clarke, 2006, p. 13). Latent level analysis is consistent with the study's constructivist paradigm (Braun & Clarke, 2008). Additionally, it will allow the researcher to communicate what may underpin the data.
What epistemological paradigm will shape my analysis?	• Constructivist paradigm which is consistent with the overall study aims.
What questions will guide my analysis?	• What does this tell me about this participant's practice? What does this mean for what and how they know? • How is this like or different from other codes, themes, and participants? What assumptions am I making to draw these comparisons? • How do I know my interpretation is credible? What other evidence do I have? What, other than my interpretation, could this mean?

Appendix J: Study 2 Faculty Invitations

Greetings,

My name is Heather Codd, and I am a Ph.D. candidate from Claremont Graduate University (CGU). As part of my dissertation, I am conducting research on graduate programs that concentrate on evaluation. In particular, I am interested in gaining an in-depth understanding of the types of knowledge gained by doctoral students in your program and how this knowledge prepares them for evaluation practice.

To inform the study, I am interviewing faculty who teach evaluation courses and work with students on evaluation research and projects. In this regard, I would like to invite you to participate in a 45 minute interview over Zoom (or another platform of your preference). The interview questions will draw on your approach to teaching evaluation as well as advising and mentoring students.

I appreciate that you are busy and routinely receive requests for your time. However, I would very much appreciate the opportunity to learn from you and would conduct the interview as efficiently as possible. Should you be interested in and willing to participate in an interview, please click on this link {insert link}, which we take you to an informed consent form as well as an interviewing scheduling webpage.

Thank you very much for your time. Please feel free to contact me with any questions or concerns.

Sincerely,

Heather Codd, Ph.D. Candidate
Evaluation and Applied Research Methods
Claremont Graduate University
Heather.codd@cgu.edu

Appendix K: Study 2 Alumni/Advanced Students Invitations

Greetings,

My name is Heather Codd and I am a PhD candidate in Evaluation and Applied Research Methods at Claremont Graduate University. As part of my dissertation, I am examining various graduate programs in evaluation, including {insert name of program}.

As a {graduate or current student} of the program, I would like to invite you to participate in a 60 minute interview over zoom (or another platform of your preference). The interview questions will draw on your experience in the program, including the knowledge and skills you acquired through your participation in the program.

I appreciate that you are busy and routinely receive requests for your time. However, I would very much appreciate the opportunity to learn from you and would do my best to conduct the interview as efficiently as possible. Should you be interested in and willing to participate in an interview, please schedule a time using this link: {insert Calendly link} or by replying to this email.

Thank you very much for your time. Please feel free to contact me with any questions or concerns.

Sincerely,
Heather Codd, Ph.D. Candidate
Evaluation and Applied Research Methods
Claremont Graduate University
Heather.codd@cgu.edu

Appendix L: Study 2 Faculty Expression of Interest Survey

Agreement To Participate in Research on Evaluation Educational programs (IRB # 4024)

You are invited to participate in a research project. Volunteering will probably not benefit you directly, but you will be helping the investigator to create new knowledge that may benefit the evaluation field. If you volunteer, you will be asked to participate in an interview that will take 60 minutes of your time. Volunteering for this study involves no more risk than what a typical person experiences on a regular day Your involvement is entirely up to you. You may withdraw at any time for any reason. Please continue reading for more information about the study.

STUDY LEADERSHIP: This research project is led by Heather Codd, PhD candidate of the Claremont Graduate University, who is being supervised by Dr. Leslie Fierro, Senior Fellow, Claremont Graduate University.

PURPOSE: The purpose of this study is to gain an in-depth understanding of your institution's evaluation program, including your role as a faculty member.

ELIGIBILITY: You are eligible for this study because you are a current faculty member that teaches and advises students.

PARTICIPATION: During the study, the investigator will ask you questions about your role as a faculty member as well as your approach to teaching and mentoring doctoral students. The interview will take about 60 minutes.

RISKS OF PARTICIPATION: The risks of participating in the study are minimal and do not exceed those of everyday life. To the extent possible, the interview time has been scheduled at your best availability to minimalize any inconvenience. Additionally, you have the options to: (1) refuse to answer specific questions, and (2) withdrawal from the study at any time without penalty.

BENEFITS OF PARTICIPATION: We do not expect the study to benefit you personally. However, due to the unique focus of the study and its potential relationship to evaluation practice and theory, it is believed that your participation will benefit the field of evaluation. Additionally, the interview will provide you with an opportunity to reflect on your role, which may benefit you personally.

COMPENSATION: There is no direct compensation to you for participating in this study. However, I greatly appreciate your participation.

VOLUNTARY PARTICIPATION: Your participation in this study is completely voluntary. You may stop or withdraw from the study at any time or refuse to answer any particular question for any reason without it being held against you.

CONFIDENTIALITY: Your participation in the study will be confidential. You will not be identified in any papers, books, talks, posts, or stories resulting from this study, and your name will not be tied to any of your responses. Additionally, the evaluation programs investigated in this study will not be disclosed in any communications. The data collected may be used for future research.

During the interview, I will ask for your permission to record our interview for transcription purposes. If your interview is occurring over Zoom and you permit to recording the interview, two recordings will be made: one using Zoom (primary recording), stored on the researcher's password-protected computer, and one on a digital recording device (backup recording). Following the interview, both recordings will be downloaded/transferred to a password protected computer and deleted from their original source (i.e., Zoom's cloud, the recording device). If your interview is occurring over the telephone and you permit to recording the interview, the researcher will utilize two digital recording devices to make a primary and backup recording. Following the interview, all recordings will be saved to a password-protected computer and deleted from the digital recording device(s). Once a transcript of the interview is complete, the recordings will be deleted from the computer. In order to protect the confidentiality of your responses, your data will be assigned a unique identifier, which will be used on all recordings and documents resulting from your participation in the study.

FURTHER INFORMATION: If you have any questions or would like additional information about this study, please contact:

Heather Codd, Ph.D. Candidate
Evaluation and Applied Research Methods
Claremont Graduate University
150 E. 10th Street
Claremont, CA, 91711
heather.codd@cgu.edu

You may also contact my faculty advisor at leslie.fierro@cgu.edu.

[IF APPROVED BY IRB] The CGU Institutional Review Board (IRB) has approved this project. If you have any ethical concerns about this project or about your rights as a human subject in research, you may contact the CGU IRB at (909) 607-9406 or at irb@cgu.edu. A copy of this form will be given to you if you wish to keep it.

CONSENT: To provide your consent to participate in the study, please select "I consent to participate in this study." If you do not provide your consent, please select "I do not consent to participate in this study."

☐   Yes, I consent to participate in the study.

☐   No, I do not consent to participate in the study.

Thank you for agreeing to participate in this study! I look forwarding to interviewing you.

As a final step, please click on the arrow below and you will be redirected to my interview scheduling webpage.

Please note, if none of the options provided work for you, please contact me at heather.codd@cgu.edu so we can work together on finding a suitable time.

Appendix M: Study 2 Alumni/Advanced Students Expression of Interest Survey

Agreement To Participate in Case Study Research on Evaluation Educational programs
(IRB # 4024)

You are invited to participate in a research project. Volunteering will probably not benefit you directly, but you will be helping the investigator to create new knowledge that may benefit the evaluation field. If you volunteer, you will be asked to participate in an interview that will take about 45 - 60 minutes of your time. Volunteering for this study involves no more risk than what a typical person experiences on a regular day Your involvement is entirely up to you. You may withdraw at any time for any reason. Please continue reading for more information about the study.

STUDY LEADERSHIP: This research project is led by Heather Codd, PhD candidate of the Claremont Graduate University, who is being supervised by Dr. Leslie Fierro.

PURPOSE: The purpose of this study is to gain an in-depth understanding of your experience as a graduate student.

ELIGIBILITY: You are eligible for this study because you are an alumni or current student of one of the educational programs included in the study.

PARTICIPATION: During the study, the investigator will ask you questions about your academic experience, including the types of activities in which you were or are involved as part of the program and the benefits of these activities. The interview will take about 60 minutes.

RISKS OF PARTICIPATION: The risks of participating in the study are minimal and do not exceed those of everyday life. To the extent possible, the interview time has been scheduled at your best availability to minimalize any inconvenience. Additionally, you have the options to: (1) refuse to answer specific questions, and (2) withdrawal from the study at any time without penalization

BENEFITS OF PARTICIPATION: By participating in research on evaluation, you will be contributing to the evaluation scholarship. In particular, it is hoped that the findings from this study will provide new insights on evaluation education and practice.

COMPENSATION: There is no direct compensation to you for participating in this study. However, I greatly appreciate your participation.

VOLUNTARY PARTICIPATION: Your participation in this study is completely voluntary. You may stop or withdraw from the study at any time or refuse to answer any particular question for any reason without it being held against you.

CONFIDENTIALITY: Your individual privacy will be protected in all papers, books, talks, posts, or stories resulting from this study. I may use the data we collect for future research or share it with other researchers, but I will not reveal your identity with it. I will ask for your permission to digitally record our interview for transcription purposes. The recordings will be made on a password-protected device and transferred to a password-protected computer immediately following the interview. The recording on the digital recording device will be deleted at this time. In order to protect the confidentiality of your responses, you will be assigned a unique identifier, which will be used on all recordings and documents resulting from your participation in the study. All recordings and documents will be saved on a password-protected computer and device.

FURTHER INFORMATION: If you have any questions or would like additional information about this study, please contact:

Heather Codd, Ph.D. Candidate
Evaluation and Applied Research Methods
Claremont Graduate University
150 E. 10th Street
Claremont, CA, 91711
heather.codd@cgu.edu

You may also contact my faculty advisor at leslie.fierro@cgu.edu.

[IF APPROVED BY IRB] The CGU Institutional Review Board (IRB) has approved this project. If you have any ethical concerns about this project or about your rights as a human subject in research, you may contact the CGU IRB at (909) 607-9406 or at irb@cgu.edu. A copy of this form will be given to you if you wish to keep it.

CONSENT: To provide your consent to participate in the study, please select "I consent to participate in this study." If you do not provide your consent, please select "I do not consent to participate in this study."

☐   Yes, I consent to participate in the study.
☐   No, I do not consent to participate in the study.

Thank you for agreeing to participate in this study! I look forwarding to interviewing you.

As a final step, please click on the arrow below and you will be redirected to my interview scheduling webpage.

Please note, if none of the options provided work for you, please contact me at heather.codd@cgu.edu so we can work together on finding a suitable time.

Appendix N: Study 2 Faculty Semi-Structured Interview Guide

Preliminary Topics	Sample Items
• Roles within the program	• First, I'd like to learn about your involvement with the program. • Could you tell me about your role(s) in the program, particularly those that involve students?

• Courses taught and objectives of these courses
• How courses do an do not prepare students for practice
• Other roles in the program and the objectives of those roles
	• You mentioned that one of your roles is teaching. I'd like to ask you some questions about the courses you teach. In particular, I am interested in what students are expected to gain from their course work that helps to prepare them for practice.
	• Over the last couple of years, what courses have you taught?
	   • What is the overall intent of the course?
	   • What knowledge are students expected to gain?
	   • In what ways does the course prepare students for practice? How does this occur? Probe for competent practice, ethical practice.
	   • Probe for social, self, others, field (methods and ethics).
	• You mentioned other roles that involve students. [For each role that involves students]: can you talk to me about the kind of work you do with students as part of this role?
	   • Why has this component been integrated into the program? What is the intent?
	   • What knowledge are students expected to gain?
	   • Probe for social, self, others, field (methods and ethics).
	   • How does students' involvement in this activity help prepare them for evaluation practice?
	• What aspects of the program are typically most challenging for students? How do you help them overcome this challenge?
	• Take a moment to think about the demands that evaluators face in their practice and what is

	required to conduct quality and ethical practice. With those demands and requirements in mind, in what ways, is the program doing a good job of preparing students for practice? • What are some areas in students' development that are not being addressed by the program? Why?
• Personal goals	• My final question relates to your personal goals as someone who teaches and mentors future evaluators. • What are you trying to accomplish with your students? Why is this important? • What growth or changes do you want to see in students?
• Wrap-up	• That's all the questions I have for you. Is there anything else you can tell me about your program or work that helps prepare students for evaluation practice?

Appendix O: Study 2 Alumni/Advanced Students Semi-Structured Interview Guide

Section	Sample Items
• Background information on participants	• First, I'd like to learn about your journey to your doctoral studies. • What were you doing prior to starting your graduate/doctoral studies? • Why did you decide to pursue a graduate/doctoral degree in evaluation? • What knowledge or experience did you bring to the program?

• Describing and analyzing program activities	• Now, I'd like to learn about different components of your doctoral studies and how they helped develop your evaluation practice.
	   • Let's start with your course work.
	   • What type of courses did you take?
	   • Probe for focus (evaluation, research methods).
	   • Probe for key topics covered.
	   • Probe: knowledge gained/used.
	• Probe other program components using prompts above.
	• Considering your educational experience, it sounds like your program helped to build your knowledge of X and X.
	• I'm wondering how, if at all, your educational experience developed your knowledge in other areas.
	   • Probe for examples
	• In what ways, is the program doing well in preparing students for practice?
	• What are there some areas in students' development that are not being addressed by the program? Why?

• Connections to practice	• These last few questions focus on your experience conducting evaluations • To what extent, has your doctoral program shaped how you currently engage in evaluation work? • What aspect of your practice do you see your education having the most influence?
• Wrap-up	• That's all the questions I have for you. Is there anything else you can tell me about your experience in the program that has influenced your evaluation practice?

194

Appendix P: Study 2 Summary Table

A Priori Codes	Sub-codes (program component)	Sample Excerpts
Knowing self	Self-awareness	• "I take feedback seriously from my students, so that I can model good practice as evaluators, because I think that our evaluators need to know themselves where they fit in the big scheme of things" (faculty).
	Self-awareness; positionality	• "When I came into the program, as well, I think I had done a lot of self-reflection and development of myself awareness as an individual, and that's part of what took me to the program, being willing to like, quit my career…and then what the program helped me do, or my experiences in the program, was to take that next step and think about my positionality…that was a brand new concept to me…when I got to the stuff around like paradigms and mental models, like my head exploded…and…so really unearthing my assumptions about evaluation, about research, about reality, about knowledge, about epistemological pathways …all of that was just huge for me, and fun, like it was really energizing, and hard and all of that" (alumni/advanced student).
	Self-awareness (course work)	• "In the theory course, they're required to keep a journal, and where they turn those in every week, and they have journal prompts. And so there are several prompts that are about understanding their own cultural position. Thinking about what power, privilege they do have, and how they manage it, you know, they're thinking about how people respond to and perceive them…they're meant to have a conversation with me about, you know, some of those more difficult moments and practice, but also just to learn to become reflective" (faculty).
	Professional identity (practical experience)	• "I'm feeling like an evaluator, I felt like I was able to make an impact and, you know, maybe help provide some guidance of some ways that this team can move forward…that's kind of in that moment, I started to kind of switch identities a little bit, which was a neat feeling" (alumni/advanced student).

		• "I think going through these experiences, and seeing, I remember starting off and being like a little bit unsure about what I was doing in the program, or like, what these interests were, where I kind of, like, fit in to the field, or even if I like belonged in the field at all. And so I think over time…building more confidence in that area, too" (alumni/advanced student).
Knowing others	Interpersonal skills (professional development sessions)	• "I think it's just the interpersonal dynamics that, you know, we attempt for the lab to address some of that" (faculty).
	Communication; Interpersonal skills (coursework)	• "In that class, we spend some time talking about how to use design and facilitate face to face interactions. And, you know, we use the example of getting people to try to understand results and think about their implications. So, you know, see what, what, you know, what, how, what would work for different kinds of audiences, and that kind of face to face facilitated session?" (faculty).
	Communication (coursework, practical experience)	• "I feel like that's also where I developed more of my facilitation skills. And where I really honed in on what I enjoy about this is the educative part of it. Like I enjoyed being a facilitator of other people's thinking…you're the expert in your experiences, you know about your program. I have some skills in how to help structure and scaffold thinking and in what kind of tools we can bring to capture the information you're interested in" (alumni/advanced student).
		• "I think being able to communicate effectively with stakeholders was a skill that I learned, specifically in that evaluation course, but then also in my research assistantship, with some of the projects I was on. And I think, coming into the program, that was something I didn't realize that was, I guess, an important skill to have in evaluation" (alumni/advanced student).
	Relationship building	• "Every project I work on, my students are intimately involved with the client" (faculty).
	Relationship building (practical experience)	• "I think seeing how faculty or other PhD students were able to kind of talk to clients and approach them in a certain way was something that that wasn't

necessarily taught in our courses, that was kind of helpful to see in like a real-world application" (alumni/advanced student).

Knowing the discipline	Evaluation design/procedures	• "I felt like this first course, really introduced me to what is evaluation, and some of the different methods of conducting evaluations, things to be aware of when you are in evaluations, such as identifying stakeholders, kind of basic things like that" (alumni/advanced student).
	Evaluation theory	• "We would have all kinds of questions regarding these theoretical approaches and our faculty that was running that course is actually very, very, very knowledgeable in these frameworks. And she would present questions that just made you turn upside down and be like, I never thought about it that way" (alumni/advanced student).
		• "They do a pitch competition, each of them has to do the pitch using a theory I assign them. So they all have the same RFP…it's a way for them to learn what the theory is…to practice applying it and practice selling it to a client with a real RFP" (faculty).
		• So the objectives is to give the students a broad understanding of the field's history as related to theory development and the relationship between the fields, sort of foundational theory and social theory more broadly, and concepts of program theory" (faculty).
	Research methods	• "People coming out when you're in that program, so you can have the quantitative focus, or the qualitative focus, and then there's an evaluation focus, but either way, like you're coming in to get pretty solid training on the different methods" (faculty).
		• "I would say for that first year, like my coursework was pretty methods based" (alumni/advanced student).
	Evaluation ethics (course work)	• "That was covered in all my classes. So it was sort of woven into, in their theory classes, it would be framed, often around sort of the trade offs and decisions that evaluators need to make" (alumni/advanced student).

197

	Evaluation ethics; course work; practical experience	• "So, ethics, I think, are built in like, so, we don't have a course on ethics, but we're always covering, right, what's ethical behavior? And I think that comes through discussions on what's culturally responsive and appropriate" (faculty).
Knowing the common good and equity	Role of evaluation/evaluators	• "From the decision to even evaluate to the entire process and the interests and consequences of what the evaluation will achieve or seeking to achieve. There's pretty constant questioning around who does this benefit? whose perspective does this reflect?" (faculty) • "I definitely see connections between the ways in which an equity lens or culturally responsive lens have been used in my classes, and the ways in which we are doing this work. Even in our or on my projects, I tell students, we need to have an evaluation question that's explicitly asking about equity" (faculty). • "I was upfront with the students and with our clients, particularly given the programs that we are working with this semester, that the paradigm in which I operate from with this class, and I think generally in my own work is we are doing this from an equity focus lens, from a culturally responsive lens. Because we're also working in communities where many of these clients or not the clients themselves, but the participants of these programs have experiences with being exploited. And so I'm leaving, knowing we don't want to leave, and we don't want to leave our clients and the participants of these programs worse off than when we entered. So I told that to my students up front. I told that to our clients upfront when I had individual meetings, here's how I'm thinking and operating. And it was important for me to be explicit about that" (faculty).
	Social consciousness	• "I just felt like I grew to better understand, not to be so naive about this, you know, not to be so superficial about what systems are. I've always been oriented with a system lens…but really understanding the systems and policy within the system to actually make changes about this policy…and so really, getting a feel for that has helped me develop a better idea for how things really operate" (alumni/advanced student).

		• "Our program is focused on culturally responsive and social justice-oriented evaluation. And so students come in and leave with that framework, with that viewpoint. And so they learned what it means to do that, both in the coursework and in the lab" (faculty).
Knowing in action	Managing ambiguity (practical experience; coursework)	• "What I learned was listening skills are very important and interpreting how people, what the client or what the population is saying might mean something different than what is said" (alumni/advanced student).
		• "I think more recently what has been coming up in the class, as well as in came up in the last practice class, is this concept of a clear objective sense of what is the right way to go about something and measure something or know whether it's good or not? And that evaluators don't bring with them an indisputable authority around what is good design? Or what is the right method? Or what is successful or not? That if that kind of valuing and methodological approach is debatable, then how can they argue what they're doing is good? Or how do they know if what they're doing is good? And how do they establish credibility and evidence with stakeholders and like these issues" (faculty).
		• "We have to remind them, like, I'm reminding myself always, like, we have multiple roles in the evaluation space. Like sometimes we facilitate, sometimes we're building ideas, sometimes we actually are collecting data, but there's a lot of team building and relationship building" (faculty).

	Managing complexity (practicum experience)	• "We talked about relationship building, trust, right? But in practice, it looks completely different" (alumni/advanced student). • "Very often, students will have an idea, the fact that we ask them to reflect in… course, on what issues in society they would like to be engaging with as part of their practicum experience. If their thing is, is health inequity, then often they get involved in evaluations that address that particular issue. And so there's their opportunity to understand and begin to develop more complex, nuanced understanding of that aspect of our social life and they can develop and grow" (faculty).
Expressions of knowing	Logic Models Reporting/Presenting findings	• "So in the theory class that's when you learn how to do logic models" (faculty). • "Then the second half of that class in the second semester, covered more of the nitty gritty of how you, you know, plan out your evaluation process, collecting data, reporting that data back to the stakeholders" (alumni/advanced student). • "So the goal is for them to come out having, you know, written evaluation reports having done data collection, right, having worked on a project as much as possible, for a long term that you see all aspects of evaluation. So from writing up an evaluation plan to, you know, attending meetings, presenting to clients, submitting reports, presenting at conferences" (faculty).

Appendix Q: Study 2, Strand 2 Principal Components Analysis Summary Table

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
P1	4.64	66.29	focus is on organization's work and purpose; action or strategies…focus is on my purpose and work	0.91	0.85
			about outputs and outcomes…about the process of evaluation	0.89	0.79
			information about the research questions…information about socio-cultural influences that can impact the data	-0.87	0.81
			about helping people to learn technical evaluation knowledge…about using my technical evaluation knowledge	0.85	0.89
			fundamental to my evaluation practice...not so fundamental to my evaluation practice	0.83	0.7
P2	3.34	47.67	need to understand the context and the purpose…how, understanding the logistics	-0.86	0.75
			building relationships, open conversation, local expertise…nuts and bolts, transactional and contractual	-0.92	0.87
			falls under my responsibility…clients responsibility to lead	0.78	0.61
			foreseeing the burden of evaluation and how it fits with other priorities…monitoring the evaluation burden in the moment	-0.69	0.76

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
P3.1	2.88	41.08	technical expertise; about data, data sources, and dissemination…related to values	0.94	0.91
			practical; common set of expectations (trying to understand how clients' and evaluators' expectations relate and foresee challenges...philosophical and conceptual conversation, imagining something different than what is in front of them, thinking about diversity, equity, and inclusion; challenging clients to think about who is not included but have a stake in the program (3)	0.84	0.72
			ethics involved, what needs to be considered from an ethical standpoint...who might be involved in the evaluation, how they will be involved, doesn't necessarily involve ethics	0.62	0.42
			alignment between expectations and resources...alignment between expectations and all matters related to the evaluation (ways of communicating, who are the decision makers	0.75	0.84
P 3.2	2.00	69.65	strategies to get at needs and intended use...technical specifications, what I am producing	0.68	0.61
			grounding the evaluation, the four stakes in the ground (what are we trying to do and who is that for...staying in touch, having a clear plan to make sure you are doing what you've agreed to do	0.82	0.79

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
			fundamental to my evaluation practice...not so fundamental to my evaluation practice	0.73	0.59
P7.1	3.28	36.45	what skills does the team have and need; skill-based...what are the client's questions and information needs; question-based	0.6	0.37
			Client terms; Contractual deadlines and client needs...On the evaluator's terms; are these interesting and innovative; merit beyond the single organization	0.69	0.65
			Creativity, thinking about the puzzle, aspiration, big picture thinking...experience and responsibility	0.94	0.9
P7.2	3.27	72.80	evaluator's role in enhancing this program; formative evaluation...evaluator's role in providing final information	0.79	0.7
			feel certain about, confident that it will be achieved...uncertain about what you can achieve (practical side of dissemination and will this rise to a level of interest for academic/policy community)	0.94	0.9
			fundamental to my evaluation practice...not so fundamental to my evaluation practice	0.81	0.65
P9.1	3.03	37.92	ways of communicating information and knowledge...about knowledge generation	-0.92	0.93
			call for estimations of where they are and want to go...mechanisms/tools to help facilitate the group understanding where they are and/or where they want to go (have to invite collaboration)	0.92	0.93

203

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
			revolve around the evaluator modelling certain behaviors and attitudes (invite participation, collaboration, openness, ensuring voices are heard, etc….revolve around the clients; they are taking charge of the evaluation	-0.9	0.84
P9.2	2.56	69.86	internal; helping them to take charge of their own assessment...move things outwards	0.77	0.82
			dependent on other elements...overarching, shapes everything you do; relevant to every element	0.91	0.88
			product; audit trail and respecting community knowledge...process; modeling respect as a mechanism for creating an audit trail	0.84	0.71
P10.1	2.54	31.70	Deep conversations with people; highly interactive...Internal	0.94	0.88
			interpersonal competencies; situational analysis...technical knowledge	0.89	0.8
			think with me and gain understanding; facilitate understanding...my challenge to figure out	0.78	0.67
P10.22	2.24	59.98	about action...internal, preparation	0.69	0.49
			have I got it right; validation...can I do this work; reflective	-0.8	0.71
			doing the job/evaluation...personal evaluation; looking at who I am	-0.7	0.58
P11.1	2.45	35.06	Fundamental to my evaluation practice...Not so fundamental to my evaluation practice	0.64	0.55

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
			making a judgement in the moment, spontaneous, rules are less clear...practical decisions, following the formula, rules are clear	0.82	0.74
			established; set...less established; seeking creative and news ways	-0.75	0.59
			thinking strategically about the client's needs...being tactical	0.79	0.66
P11.2	2.14	65.69	convincing clients on your credibility and value...delivering value; doing credible work	0.96	0.93
			collaboratively make sense...communicating your expertise; what I know	-0.61	0.4
			participants, respondents are the priority...client is the priority	-0.74	0.62
P12.1	4.09	51.10	gaining agreement; ensure that intended uses are agreed on…getting users to act on the results; utilization commitment	0.94	0.88
			the findings are actionable...looking for the personal factor; can these people take action	-0.87	0.88
			more conceptual; understanding the situation and its effect on the evaluation…more actionable; facilitating action among intended users	0.85	0.76
			focusing on intended use and intended users…focus on changes in context (what has changed) at every level ; monitoring	0.8	0.65
			intended users have commitment and capacity to use the evaluation…determining who needs to have that commitment and capacity	-0.76	0.65

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
P12.2	2.36	80.65	Fundamental to my evaluation practice…Not so fundamental to my evaluation practice	0.88	0.78
			people management; getting it done and used, which includes adaptation…people and purpose; who is involved and what we are trying to accomplish	-0.83	0.84
			assessment based on commitment and position…about information needs and how to gain information	0.8	0.78
P13	3.73	53.34	understanding and explaining how the aspects of the evaluation fit together…understanding what you could undertake; relates to constraints	0.9	0.98
			relates to methods that will be used for the evaluation…who's the broader audience for the publication; why would it be accepted; what is it contributing to the literature	-0.9	0.9
			Fundamental to my evaluation practice...not so fundamental to my evaluation practice	0.9	0.98
P14	3.18	39.80	Fundamental to my evaluation practice...Not so fundamental to my evaluation practice	0.9	0.81
			is it likely that the process good...is it likely the outcomes will be good	-0.68	0.47
			somebody's expectations are influences the evaluation...stopping those expectations; exiting the relationship	0.84	0.74

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
P15.1	3.24	40.49	think about best way to report the results for the clients' understanding…focused on the evaluators understanding; about getting an understanding to do a good evaluation	-0.78	0.89
			design and implementation; equitable process takes more time and, therefore, resources…outwardly focused; not about the design and implementation	0.84	0.81
			Fundamental to my evaluation practice…not so fundamental to my evaluation practice	0.79	0.64
P15.2	3.18	80.28	evaluator's responsibility…evaluators facilitate sense making; resulting information comes from participants; knowledge is generated from the audience and evaluator; local wisdom	0.9	0.83
			questions are linked to what intended users want to learn…evaluation runs smoothly	0.93	0.86
			what the financial resources will allow for…what the client resources will allow (time and inclination)	0.74	0.63
P16	2.51	73.37	I don't always have access to these groups, I need to find others that have access…I usually have access to these groups	0.87	0.93
			emphasis is on inclusion and voice…emphasis is technical and neutral process	0.84	0.71
			Fundamental to my evaluation practice…Not so fundamental to my evaluation practice	0.67	0.62

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
P17	3.75	53.59	creation of a new thing; initiate something that didn't exist…abiding by existing laws, standards, and rules	0.9	0.85
			I'm trying to get knowledge…working to get the job done	0.87	0.89
			requirement; compliance with laws and regulations; did I do what I was supposed to…voluntary; am I consistent with standards and practice in the field	-0.87	0.76
			planning…assessment	0.82	0.7
			Fundamental to my evaluation practice…Not so fundamental to my evaluation practice	0.66	0.46
P18	3.41	48.69	actively connected to aspects of facilitating engagement…the act is not as much of a collaborative process; more so setting the stage for engagement	0.86	0.74
			technical aspects of conducting the evaluation; methodology…create space for honest, inclusive, and meaningful conversation and exploration	-0.96	0.96
			making sure everyone's voice heard/involved…framed in a way that guides evaluation decision-making and in an appropriate way	0.97	0.95
			Fundamental to my evaluation practice...Not so fundamental to my evaluation practice	0.65	0.56

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
P.20	3.88	64.75	understanding what data already exists; availability and processes and approaches...get an understanding of the broader landscape (driving or not driving the work) so you can see how it could impacts the evaluation	0.91	0.93
			what data (or voices) are missing and how to tweak the plan/program...who is missing; in this case audiences that are not missing or not prioritized as they should be	0.9	0.91
			focus is on understanding what's useful and how to help the client use it...understanding if the questions lead to information that will drive decisions	-0.97	0.94
			what is the data saying and how to make meaning from it...who you want to inform with the data you are collecting and how their information needs might be different	0.96	0.93
P.21	2.78	46.37	relational (drawing connections) and energizing; relate to direction (mine and the evaluations)...what is realistic and fair; creating a mental project and making it work	0.96	0.91
			more of a calculation; playing with what's possible in order to get at what they know; adhering to their written words or something that has been prescribed...knowing is intuitive, create a simple and elegant process that captures the spirit of the work	-0.89	0.8

Participant (P) Component	Eigenvalues	Cumulative % Variance	Constructs	Loadings	Communalities
			produce an elegant solution within constraints...expansive; arrange in different ways without constraints	-0.93	0.89
P.22	2.57	36.78	Fundamental to my evaluation practice...Not so fundamental to my evaluation practice	0.82	0.92
			Exploratory phase in the evaluation; limited constraints...Working with what you already have; constraints in place	0.86	0.74
			Provide a lens of importance; help us to understand what is important...Learning within an existing frame of what's important; already defined what's important	0.89	0.81
	2.12	66.99	About building relationships that will benefit the evaluation...Specifically about analysis; drawing inferences from the data	-0.76	0.61
			Focused on the data (quality, interpretation, enough data)...More focused on how to best communicate (preparing a report, presentation)	0.96	0.92