Socialization of U.S. Doctoral-Degree Students into Evaluation Professionals: The Use of Evaluator Competencies and Experiential Learning Strategies in Selected Programs

Michelle Minear Chandrasekhar
SOCIALIZATION OF U.S. DOCTORAL-DEGREE STUDENTS INTO EVALUATION PROFESSIONALS: THE USE OF EVALUATOR COMPETENCIES AND EXPERIENTIAL LEARNING STRATEGIES IN SELECTED PROGRAMS

By

MICHELLE MINEAR CHANDRASEKHAR

A Dissertation submitted to the Department of Educational Leadership and Policy Studies in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Degree Awarded:
Spring Semester, 2012
Michelle M. Chandrasekhar defended this dissertation on March 16, 2012.

The members of the supervisory committee were:

Linda B. Schrader  
Professor Co- Directing  Dissertation

Robert A. Schwartz  
Professor Co- Directing  Dissertation

Vanessa Paz Dennen  
University Representative

Richard C. Kunkel  
Committee Member

Sande D. Milton  
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.
ACKNOWLEDGEMENTS

My family and friends "worked" on this degree with me - listening to me think out loud and encouraging me to continue. I thank them for their patience and support. I also thank Dr. Linda Schrader. It was her cheerful, unflagging support that made this possible.

This scholarly effort is the result of mentoring I received from faculty located at Florida State University, Tallahassee Community College, the University of North Carolina at Charlotte, American University, and the Georgia Institute of Technology. I found other mentors in the following professional associations: American Evaluation Association (AEA), Association for Institutional Research (AIR), Preparing Future Faculty (PFF), Professional and Organizational Development (POD), Southeastern Evaluation Association (SEA), and the Southern Association of Colleges and Schools (SACS).

James C. Smith shared his love of experimental psychology and of teaching in that discipline with me. Anita J. Harrow inspired me to do more with teaching in higher education, to go beyond teaching a course and look at the curriculum and the entire student experience. Finally, my discussions with George Walker and his work with Preparing Future Faculty and the Carnegie Initiate on the Doctorate had a significant impact in choosing this topic. The insights of these mentors and my committee into higher education, research, and teaching, were invaluable. Their passions informed and inspired me - and the conversations were a joy. It is an honor to be a part of their academic 'tree'.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Evaluators and Two Approaches to Evaluation</td>
<td>3</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>4</td>
</tr>
<tr>
<td>Conceptual Framework: Socialization and Experiential Learning</td>
<td>5</td>
</tr>
<tr>
<td>Research Questions</td>
<td>8</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>9</td>
</tr>
<tr>
<td>Assumptions of the Study</td>
<td>13</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>13</td>
</tr>
<tr>
<td>Delimitations of the Study</td>
<td>13</td>
</tr>
<tr>
<td>Significance of Study</td>
<td>14</td>
</tr>
<tr>
<td>Summary</td>
<td>15</td>
</tr>
<tr>
<td>CHAPTER 2: LITERATURE REVIEW</td>
<td>16</td>
</tr>
<tr>
<td>Mandate for Evaluation Expertise</td>
<td>16</td>
</tr>
<tr>
<td>Accountability in Education</td>
<td>18</td>
</tr>
<tr>
<td>The Decline in Number of Evaluation Degree Programs</td>
<td>22</td>
</tr>
<tr>
<td>What is an Evaluator?</td>
<td>25</td>
</tr>
<tr>
<td>Approaches to Evaluation Inquiry</td>
<td>26</td>
</tr>
<tr>
<td>Use approach.</td>
<td>27</td>
</tr>
<tr>
<td>Method approach.</td>
<td>27</td>
</tr>
</tbody>
</table>
Value approach ........................................................................................................... 29
Professionalization of Evaluators .................................................................................. 30
Program Evaluation Standards ......................................................................................... 30
AEA Guiding Principles for Evaluators ........................................................................... 32
Evaluation as a Discipline ............................................................................................... 33
Developing a taxonomy of evaluator competencies ....................................................... 35
Professional organizations as sources for evaluator skill sets ......................................... 37

Preparing Future Evaluators ......................................................................................... 39
Socialization of Students ................................................................................................. 40
Experiential Learning or Practical Training ......................................................................... 43
Evaluation Literature on Experiential Learning ............................................................... 45
Examples of experiential learning ...................................................................................... 47
  Simulations ...................................................................................................................... 48
  Role-Play ........................................................................................................................ 48
  Case studies ..................................................................................................................... 49
  Single course projects .................................................................................................. 50
  Focused practicum experiences .................................................................................. 51
  Broad practicum experiences ....................................................................................... 53

Summary ............................................................................................................................ 54

CHAPTER 3: METHODOLOGY ......................................................................................... 56
  Program artifacts .......................................................................................................... 58
  Program-Informants ...................................................................................................... 58
  Prominent evaluation theorists .................................................................................... 58
Selection of Programs ........................................................................................................ 59
Recommendations from Prominent Evaluation Theorists ................................................. 60
Criteria for Selection of Evaluation Degree Programs .................................................... 60
  First criterion for program selection .......................................................................... 61
  Second criterion for program selection ...................................................................... 61
  Third criterion for program selection ....................................................................... 62
Data Collection Process ............................................................................................................. 63

Phase I: Interviews with Four Evaluation Theorists ................................................................. 65
  Selection process. .................................................................................................................. 65
  Interview protocol .............................................................................................................. 65
  Consent ................................................................................................................................ 66

Phase II: Information Describing Six Selected Evaluation Programs ..................................... 67
  Program artifacts – website material .................................................................................... 67
  Program informant participation and consent ...................................................................... 68
  Institutional characteristics ................................................................................................. 68
  Evaluation-specific course syllabi. ....................................................................................... 69

Phase III: Interviews with Program-Informants ....................................................................... 69
  Interview protocol .............................................................................................................. 69
  Interview Questions .......................................................................................................... 70
  Program informant participation and consent ..................................................................... 70

Phase IV: Development of Program Profiles ........................................................................... 71
  Validation of program profiles. ............................................................................................ 71

Data Analysis ........................................................................................................................................ 73
  Coding Rubric Development ............................................................................................... 73
  Definitions for Eight Axial Codes ......................................................................................... 75
  1. Program mission. ............................................................................................................ 75
  2. Degree requirements .................................................................................................... 75
  3. Program size .................................................................................................................. 76
  4. Evaluation competencies .............................................................................................. 76
  5. Career foci ...................................................................................................................... 80
  6. Instructor information ................................................................................................... 81
  7. Course information ....................................................................................................... 81
  8. Instructional strategies ................................................................................................... 81

Identifying Program Practices to Socialize Students ................................................................. 83

Study Quality .................................................................................................................................. 83
  Construct Validity ................................................................................................................. 84
Research Question 4: What practices are used? ................................. 119
Iterative learning.................................................................................. 119
Hands-on experiences................................................................. 120
Extensive faculty mentoring........................................................... 123

Common Attributes of Programs ................................................. 125

Differences among Programs .......................................................... 126

Section III: Developing Evaluation Professionals ............................. 127
Dimension 1: Socialization Strategies .................................................. 128
Evaluator competencies element................................................. 128
Experiential learning element................................................... 129
Dimension 2: Individualized Career Preparation .............................. 129
Flexible coursework element..................................................... 130
Tailored experiences element..................................................... 131
Two Approaches to Support, Guide, and Enhance ......................... 133
Faculty mentoring approach....................................................... 134
Practica experiences approach..................................................... 134
Emerging Themes in Common Program Practices ......................... 135

Summary of Study Results................................................................. 136

CHAPTER 5: DISCUSSION ................................................................. 139

Creating Evaluation Professionals .................................................. 140
Promoting Professional Knowledge and Skills ................................ 142
Socializing Students into Evaluation ............................................. 144
Training for Academe and Praxis ....................................................... 145
Learning by Doing ........................................................................ 146

Identification of Program Practices ............................................... 147

Implications for Programs ............................................................. 148
Socialization Requires Active Planning ........................................ 149
Evaluation is Resource Intensive ...................................................... 149
Become Part of the Culture ............................................................... 150
LIST OF TABLES

TABLE 1. AN OVERVIEW OF DIRECTORIES OF EVALUATION DEGREE PROGRAMS IN THE UNITED STATES ........................................................................................................................................................................... 23
TABLE 2. OVERVIEW OF RESEARCH QUESTIONS AND DATA SOURCES ................................... 57
TABLE 3. SELECTION OF EVALUATION DEGREE PROGRAMS ....................................................... 62
TABLE 4. PHASES OF DATA COLLECTION ........................................................................................................... 64
TABLE 5. PROGRAM PROFILE COMPONENTS ........................................................................................................... 72
TABLE 6. AXIAL CODING STRUCTURE ................................................................................................................ 74
TABLE 7. AXIAL CODING FROM THE AEA GUIDING PRINCIPLES FOR EVALUATORS (ABBREVIATED) (2004) ........................................................................................................................................................................ 77
TABLE 8. AXIAL CODING FROM THE PROGRAM EVALUATION STANDARDS, 2E (JSCEE, 1994).. 79
TABLE 9. AXIAL CODING FOR THE IBSTPI EVALUATION COMPETENCIES (IBSTPI, 2006) .......... 80
TABLE 10. CLASSIFICATION OF INSTITUTIONS BY PROGRAM, 2010 ....................................................... 92
TABLE 11. EVALUATION PROGRAMS BY PROGRAM SIZE, 1992-2008. ....................................................... 94
TABLE 12. STUDY PROGRAM DESCRIPTORS BY SIZE ........................................................ 95
TABLE 13. CODING RUBRICS WITH AXIAL AND OPEN-ENDED CODES ................................................. 98
TABLE 14. EVALUATION COURSE SYLLABI BY PROGRAM ........................................................................... 100
TABLE 15. OVERLAP IN EVALUATOR COMPETENCY SKILL SETS BY SOURCE ........................................ 105
TABLE 16. EVALUATOR COMPETENCIES IN EVALUATION SYLLABI BY SOURCE AND PROGRAM ............................................................................................................................................................................ 110
TABLE 17. STUDENT PROGRESSION BY CAREER GOAL .............................................................................. 118
TABLE 18. COURSE SYLLABI BY INSTRUCTIONAL STRATEGIES AND BY PROGRAM ................. 121
TABLE 19. CREDIT HRS (SEMESTER OR QUARTER) PER DOCTORAL DEGREE BY PROGRAM. 130
LIST OF FIGURES

FIGURE 1. FRAMEWORK FOR UNDERSTANDING THE STRUCTURE OF DOCTORAL EVALUATION DEGREE PROGRAMS. .......................................................................................................................... 7

FIGURE 2. PERCENT OF SUBMITTED SYLLABI FOR EACH PROGRAM USING EXPERIENTIAL LEARNING: SIMULATIONS, ROLE-PLAY, SINGLE COURSE PROJECTS, AND FOCUSED PRACTICA EXPERIENCES. NOTE: PROGRAMS A, B, AND C WERE MEDIUM IN SIZE, WHILE PROGRAMS D, E, AND F WERE LARGE. ................................................................................................. 122

FIGURE 3. PRACTICES FOR DEVELOPING EVALUATION PROFESSIONALS. NOTE: SOCIALIZATION STRATEGIES INCLUDE EVALUATOR COMPETENCIES AND EXPERIENTIAL LEARNING; INDIVIDUALIZED CAREER DEVELOPMENT INCLUDES FLEXIBLE COURSEWORK. ........................................................................................................................................ 128
ABSTRACT

Recent conversations in the field of evaluation concern the number of graduate degree evaluation programs (LaVelle & Donaldson, 2010), a high job-market demand for trained evaluators (Stufflebeam, 2001), and a call to meet that demand using evaluator competencies (Stevahn, King, Ghere, & Minnema, 2005; Russ-Eft, Bober, de la Teja, Foxon, & Kosxalka, 2008). There has also been an increase in the promotion of practice-based instructional strategies for the teaching of evaluation (Alkin & Christie, 2002; Kelly & Kaczynski, 2008; Oliver, Casiraghi, Henderson, Brooks, & Muslow, 2008; Patton & Patrizi, 2005; Skolits, 2009; Trevisan, 2004). This study examined six evaluation degree programs in the United States to describe the knowledge, skills, abilities, and values reflected in the program structure (course work and student experiences).

This study utilized a purposeful sampling strategy to identify doctoral programs that had been sustained over time. Practices programs used to develop evaluation professionals were identified from material used to describe programs in terms of evaluator competencies, the program structure used to promote student achievement of those competencies, and the socialization experiences that prepared students for careers in evaluation. Data was collected from interviews with prominent theorists in the field and program-informants, program web-based documents, and evaluation-specific course syllabi.

The practices for Developing Evaluation Professionals can be described in two dimensions: Socialization of Students and Individualized Career Preparation. Critical elements include (a) use of evaluator competencies to guide and inform student learning goals, (b) use of experiential learning strategies to facilitate learning, (c) fostering flexible coursework options in designing a program of studies that meets student career goals, and (d) creating tailored practica experiences that engage students with skill sets matched to their career goals. Programs framed these dimensions with two faculty approaches important to supporting, guiding, and
enhancing the process of developing evaluation professionals: extensive faculty mentoring and practica experiences culminating in leadership roles.

Like the Carnegie Initiate on the Doctorate, discussion addresses opportunities and challenges in identifying “the desired core ingredients of an enriched form” (Golde & Walker, 2002, p. 2) of university-based evaluation training for doctoral students.
CHAPTER 1: INTRODUCTION

The demand for the evaluation of programs across disciplines and fields of work is increasing. A reflection of this demand for effectiveness and efficiency can be seen in the wording of the 1993 Government Performance and Results Act (GPRA) which requires that federal government agencies, such as the Environmental Protection Agency and the Department of Education to plan, implement, and make use of evaluation and evaluators and submit annual evaluation reports to the President and Congress. These reports are used to address a growing attitude towards accountability in spending and program implementation.

This mandate drives a growing need for evaluators with graduate degrees to provide quality evaluation services, staff academic programs, train new evaluators, and develop theory and research in the field. Recognizing this, and the inherent diversity of experiences this need implies, the American Evaluation Association (AEA) in concert with the National Science Foundation (NSF) funded an international survey of evaluation degree programs in 2002. The survey identified the characteristics of current evaluation programs and briefly described the common course offerings.

Engle and Altschuld (2003) noted a marked decline in the number of evaluation degree programs in the United States—from 44 in a 1984 survey of programs to 29 in 2002. One reason for the decline was the retirement of faculty who led these programs. Three other concerns were also identified.

The second concern was the competition from national training institutes, in-service training programs, evaluation certificates and continuing degree courses that compete for university students and professionals. As more students from other disciplines seek evaluation training “further study of the effect of alternative ways of preparing evaluators in universities is an important area of research” (Engle, Altschuld, & Kim, 2006, p. 359).

The third concern was the small number of programs that offered advanced evaluation courses in specialized topic areas. Only seven of the 29 programs in the United States offered advanced topics and seminars. The authors questioned whether
students are prepared for evaluation jobs as accountability pressures increase if they have not undertaken advanced courses. The fourth concern the authors noted was a lack of program emphasis on research and scholarship in evaluation. “Few courses offered in the programs…would suggest that evaluation scholarship is an important emphasis” (Engle et al, 2006, p. 359).

Contrasting with the declining number of degree programs, there is a growing need for trained evaluators. Stufflebeam (2001) noted an increasing number of jobs posted at the American Evaluation Association Job Bank. He claimed there are often more than 100 in a single day, of which “about 40% were seeking evaluators with doctoral degrees” (p. 447). He also cited the growth in the American Evaluation Association membership, and “the wide-ranging topics in AEA’s national meetings and journals make clear the interdisciplinary aspects of evaluation work” (p. 447).

This, in part, is likely due to the 1993 Government Performance Act, which required an evaluation component in all federally funded programs. A movement towards greater accountability in government, education, as well as in private and non-private corporations has led to a demand for context-specific evaluators. In effect, there is a need for evaluators with knowledge of evaluation methods as well as a degree of subject matter expertise in various content areas. The additional government accountability requirements require employers to seek trained evaluators who can work in different content areas. These employment areas include education, human services, foundations and corporations and are as varied as a community’s water treatment plant or organizational management in private businesses (Connor, 1986; Perloff & Rich, 1986).

The need for trained evaluators should not be surprising. Scriven (1991) described evaluation as a ‘trans-disciplinary’ field because it crosses over into or is used extensively by other fields (social sciences, health sciences, policy and public administration). Given the varied experiences and performance skills of students coming to evaluation from other fields, how do current evaluation degree programs meet the need for evaluators with a shared skill set?
Evaluators and Two Approaches to Evaluation

Alkin and Christie (2004) discuss the pervasiveness of evaluation approaches in various disciplines as being rooted in the need for: (1) social accountability and fiscal control, and measured through (2) social inquiry research practices. Evaluations serve a societal demand for accountability in contracted products or services.

The epistemological understandings in evaluation are as varied as the disciplines from which it is derived—education, psychology, public health, sociology, economics, public administration, political science, and social work (Rossi, Lipey, & Freeman, 2004). Chen's (2004) definition for an evaluation approach is equally pervasive, as “a systematic set of procedures and principles guiding evaluators . . . [in] conceptualizing problems and determining the evaluation’s focus; it guides the application of research methods for collecting and analyzing data, as well as interpretation of data” (p. 144). Despite the wide variety of contexts where evaluation appears there is agreement as to what constitutes evaluation standards and principles. The Program Evaluation Standards, 2e (Joint Committee on Standards for Educational Evaluation [JCSEE], 1981, 1994, 2011) outlines standards needed to conduct evaluations regardless of the evaluation approach. The American Evaluation Association’s (AEA) Guiding Principles for Evaluators (American Evaluation Association [AEA], 1994, 2004) provides for the role of ethics in conducting evaluations. Both publications were vetted by a broad range of people involved in evaluation studies, research, and theory development. However, agreement upon a list of specific evaluator competencies in terms of skills, abilities, and values has not yet occurred.

This increasing interest in detailing evaluation competencies led to an exploratory study on developing a taxonomy of specific evaluation competencies by King, Stevahn, Ghere, and Minnema (2001). King’s taxonomy was generated using Multi-Attribute Consensus Reaching in a group of 31 students, faculty and local evaluators. The taxonomy was revised in 2005 and now contains six categories and 69 “essential competencies for program evaluation” (Stevahn, King, Ghere, & Minnema, 2005, p. 54). Stevahn et al. (2005) state that one reason to develop and adopt a specific list of
evaluation competencies is to ensure that evaluators are trained systematically and with a comprehensive set of skills.

In a similar effort targeting evaluators for organizations, the International Board of Standards for Training, Performance, and Instruction (International Board of Standards for Training, Performance, and Instruction [IBSTPI], 2006) developed their own list of skills, knowledge and attitudes that could be expected for evaluators. IBSTPI argues that the *Program Evaluation Standards, 2e* (JSCEE, 1994) and the *American Evaluation Association’s Guiding Principles for Evaluators* (AEA, 2004) are too vague to provide guidance in training evaluators in specific skill sets. Also, Russ-Eft, Bober, de la Toja, Foxon, and Kosxalka (2008) in discussing the IBSTPI evaluators competencies argue that references to social responsibility/public welfare in these publications are not relevant, or may be counter-productive, for evaluators in organizational settings. The instructional design context of IBSTPI resulted in evaluation performance statements that were phrased as measurable knowledge, skills, and attitudes required for organizational evaluators. The result was 84 performance statements for 14 competencies categorized in four domains. The method used to generate and validate the statements made them "applicable to evaluators working in a variety of organizational settings and national and organizational cultures" (Russ-Eft, Bober, de la Toja, Foxon, & Kosxalka, 2008, p.153).

**Purpose of the Study**

The purpose of this study was to develop an understanding of how evaluation degree programs train students in terms of evaluator competencies and to use that understanding to identify practices that programs use to develop evaluation professionals. Previous surveys of evaluation degree programs only briefly described course offerings and did not provide detailed information on evaluation competencies or instructional strategies used by the programs. A demand for evaluation skills that cross disciplines is part of the reason for this study’s exploration of how current programs are promoting evaluator competencies. The *Program Evaluation Standards, 2e* (JSCEE, 1994) and the *AEA Guiding Principles for Evaluators* (AEA, 2004) were used to
compare programs because of their general adoption by the profession at the time data was collected. The IBSTPI (2006) evaluator competencies for organizational evaluation were also included because they were developed using literature and program review with an extensive validation study. Given the current call to change or enhance evaluation programs, it is important to describe the ways in which the knowledge, skills, abilities, and values implied in the evaluation standards and principles, and detailed in terms of specific evaluation competencies by IBSTPI, are fostered by the program structure (defined as coursework and student experiences).

**Conceptual Framework: Socialization and Experiential Learning**

The conceptual framework for this study used two perspectives: socialization of students into the profession and experiential learning. The broad, common goal of evaluation degree programs is to facilitate the socialization of its students into the profession. Socialization of graduate students "refers to the processes through which individuals gain the knowledge, skills, and values necessary for successful entry into a professional career requiring an advanced level of specialized knowledge and skills" (Weidman, Twale, and Stein, 2001, p.iii).

In *Re-envisioning the Ph.D.: What Concerns Do We Have*, Nyquist and Woodford (2000) share comments from students, faculty, business & industry, government, foundations, and disciplinary societies that express concerns "that [traditional] doctoral education inadequately prepares students for the other responsibilities and aspects of their careers" (p. 5). This is a concern because socialization of students for the profession of evaluation, which is context-based and often interdisciplinary, is already complex. Students in evaluation have a variety of career options. Students may choose an academic research focus or become an applied evaluator. Evaluators practice in a variety of settings that are non-profit, for-profit, or government based. They may also specialize in various content areas (health, education, public policy, psychology, etc.) Evaluators also differ in terms of the evaluation approaches they favor and their methodological strengths. Finally, evaluation
degree programs may also vary in how they socialize doctoral students in evaluation - what career paths they support and what resources they provide in the program.

Evaluation is practice-based, which made experiential learning an appropriate perspective from which to view evaluator development in university-based evaluation programs. Exploring what experiential learning strategies are currently used in evaluation degree programs was referenced as a need in the 2002 directory by Engle and Altschuld (2003). Trevisan's (2004) survey of publications on teaching in evaluation also suggested a need for more information on practical training in evaluation degree programs. Trevisan found that "none of the articles in the review were formal research studies...many of the articles did not address pedagogical theories or frameworks" (2004, p. 266). Experiential learning strategies are critical in a degree program where learning practice-based skills is important, and where many of the students come with existing skill sets from other disciplines. The assumption is that students leave these programs with an identifiable skill set and some degree of familiarity with the values, norms, and ethical behavior expectations for evaluators, regardless of the disciplinary background or skill sets with which they entered. A recognized set of evaluator competencies and a means for testing applied skills would facilitate student professional development. Experiential learning concepts help us to understand how doctoral evaluation degree programs prepare students for evaluation careers.

Because the purpose of this study was to identify practices programs in evaluation use to promote doctoral student professional development in evaluation, a mixed methodology used: (1) archival materials or program artifacts (program information and evaluation-specific course syllabi provided via website and by request) and (2) interviews via telephone. From this data, socialization practices used by programs were identified that built upon the evaluator competencies previously described in the literature and by professional associations. Doctoral degree programs were chosen for this study that had (1) an evaluation emphasis, (2) demonstrated long-term viability and (3) been associated with faculty who have made significant contributions to theory in the field of evaluation research and teaching. Figure 1. illustrates the conceptual framework for this study and is discussed in more detail in the
section below. Additional information on the sampling frame and methods are discussed in Chapter 3.

Figure 1. Framework for understanding the structure of doctoral evaluation degree programs.
As this study was largely qualitative in orientation, data was coded using clustering of codes, categorizing words into categories using investigator questions and key concepts from the qualitative data analysis framework suggested by Miles and Huberman (1984). Development of the data analysis rubric was guided in part by the following strands of interest: (1) program incorporation of evaluator competencies as they have been identified in the literature and by the discipline, (2) coursework and students experiences used for the socialization of students, and the use of experiential instructional strategies.

**Research Questions**

This descriptive study examined how evaluation doctoral degree programs (defined as programs with a strong evaluation emphasis, long term viability, and associated with major contributors to the field) train students in terms of evaluator competencies. As can be seen in Figure 1, the following questions were addressed:

1. What are the evaluation competencies, in terms of knowledge, skills, and values, proposed in the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004), the Program Evaluation Standards, 2e (JSCEE, 1994), and the International Board of Standards for Training, Performance, and Instruction Evaluator Competencies (IBSTPI, 2006)?

2. How are these competencies expressed in evaluation degree programs?

3. How are graduate evaluation degree programs structured to enable students to develop these competencies?

4. What practices support the development of student evaluation knowledge, skills, and values?

The illustration in Figure 1: Framework for Understanding the Structure of Doctoral Evaluation Degree Programs shows the research questions, information sources, and perspectives used in this study. The first research question (Q.1) identified evaluator competencies among students in doctoral evaluation degree programs could expect to develop. The Program Evaluation Standards, 2e (JSCEE, 1994) and the AEA Guiding Principles for Evaluators (AEA, 2004) imply evaluator competencies that could
be used by evaluation degree programs. The International Board of Standards for Training, Performance and Instruction (IBSTPI, 2006) provides a list of evaluator competencies for evaluators working in organizations. All three were used to identify competencies that might be used by evaluation degree programs. The second research question (Q.2) in the figure explored the how evaluator competencies are expressed by the program(s). The third research question Q.(3) provided information on how programs are structured (via coursework and student experiences) to socialize students in the profession. Question four (Q. 4) asked what practices are used by the department to support student development of evaluation knowledge, skills, and values. Data for questions two, three and four were obtained via interviews with program-informants and analysis of archival material (university web pages and evaluation-specific course syllabi). Data was analyzed using a conceptual framework (socialization of students and experiential learning) well-suited to describing the preparation of doctoral students for careers in evaluation.

**Definition of Terms**

For the purposes of this study, the following terms are used and are defined as follows:

1. **Curricula**: Programs of study required to earn a particular degree. These consist of coursework and experiences known as the curriculum. The curriculum usually includes a core set of courses and a minimum number of credit hours. Increasingly, curricula are also described by the program outcomes expected for students at each degree level.

2. **Degree programs in evaluation**: Selected programs in this study were evaluation degree programs awarding doctoral degrees as defined by the 1992, 2002, and 2008 surveys of evaluation programs (Altschuld, Engle, Cullen, Kim, & Masse, 1994; Engle et al, 2006; LaVelle & Donaldson, 2010). An evaluation degree program was defined as (1) held across time, (2) evaluation was a significant aspect of the degree, and (3) multiple courses (2 or more) were offered that were designed to teach students evaluation principles and concepts. The common definition for program is as follows: “A program consists of multiple courses, seminars, practica, offerings, and so on designed to teach what the respondent considered to be evaluation principles and concepts” (Altschuld et al., 1994, p. 72).
3. **Evaluation**: Evaluation is the process of gathering and analyzing information for the purpose of making a judgment or decision. There are many types of evaluations but they can be thought of generally as either formative (part of developing a new program or facilitating the ongoing process of improvement) or summative (providing a description of what occurred – impacts and outcomes). Examples of formative evaluations include needs assessments (who, what, where, why), implementation evaluation (proceeding as planned?), and process evaluation (describing how the process works). Examples of summative evaluations include outcome evaluation (usually targeted outcomes), impact evaluation (overall effects) and cost-benefit analysis (how efficient was the program?). The methods used can be qualitative, quantitative, or a mixture and vary widely depending upon context, purpose, and approach (such as participatory versus management-oriented models).

4. **Experiential learning**: Practical or hands-on learning can be categorized in various levels. This study used the definitions in Trevisan’s (2004) survey of practical training examples in the evaluation literature: (1) simulation, (2) role-play, (3) single course projects, and (4) practicum experiences (p. 258). Practicum experiences were further categorized, in part, using Weeks (1982) definition of short-term (one-semester) and long-term experiences (more than one-semester) but were qualified in terms of the extent to which students were involved with focused skills or the entire evaluation process.

5. **Simulation**: A description or scenario provided by the instructor along with a set of student instructions. An example would be analysis of data using NVIVO software. Simulations that run the entire course may be considered single course projects.

6. **Role-Play**: A less structured simulation in which students have a greater degree of input in how they respond to a given description or scenario. Students in role play may work through an interview process or a meeting with a client to discuss results.

7. **Single course projects**: These are student projects that form a large portion of the instruction for a given course. Assignments may be for teams or individuals, towards specific evaluation tasks or methodologies, and may incorporate external interactions (such as site visits or negotiations with a possible client). Projects result in a student generated product presented for instructor feedback. A typical example is to produce an evaluation plan for a potential client, usually in the students’ area of interest. Another example requires students to implement some aspect or task of an existing or imaginary evaluation (such as design, conducting, and analyzing focus group or interview data).

8. **Focused practicum experiences**: Based in a real context, a practicum or field experience consists of a period of supervised, practical, hands-on
training over a period of time that is possible only after completion of core methods courses. In this definition, the applied experience is beyond the time expectations possible in one course. The timeframe is usually at least one semester but sometimes more or less. The extent of student involvement in the external setting varies, sometimes limited to specific tasks and sometimes to multiple tasks. In this study, tasks assigned target specific evaluator skills: such as instrument design, particular data collection methodology and/or analysis, or report writing. Some internship experiences are paid, some are not (i.e., the student may earn a wage in addition to college credit) but all technically require the supervision of a faculty member and a field representative.

9. **Broad practicum experiences**: As with the focused practicum experience, these experiences are based in a real context and feature supervision, guidance, or mentoring from faculty or evaluation professionals. Students engage in practical, hands-on training over a period of time that is possible only after completion of core methods courses. In this definition, the applied experience is beyond the scope possible in one course or one semester. The extent of student involvement in the external setting is much greater than in the focused practicum. Students with broad practicum experiences are more intimately involved in the entire evaluation process: from first meetings with clients, to question and methodology development, data collection and analysis, and to writing the final report. Many internship experiences are paid, some are not (i.e., the student may earn a wage in addition to college credit or may earn no college credit) but all technically require the supervision of a faculty member and/or a field representative.

10. **Program directors**: Program directors are the primary faculty member responsible for the administrative aspects of an evaluation degree program. In cases where the program director was new to the position, the previous director was interviewed.

11. **Program evaluation**: Program evaluation addresses the systematic assessment of programs or activities that have a specific goal or outcome in order to answer questions about the program. These questions could be related to assessment of the goals or outcomes, effectiveness of program strategies, stakeholder involvement, etc. Program evaluation can be housed in a variety of disciplines: education, psychology, business, human services, or other social sciences. Regardless of the content area reflecting the disciplinary association, evaluation programs share common methodological and analytical skill sets along with the study of evaluation theory and approaches.

12. **Program outcomes**: Expected attitudes and skill sets or competencies that students will achieve as a result of, or by the end of, the degree program.
13. **American Evaluation Association (AEA):** “An international professional association of evaluators devoted to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation” (AEA About Us section, para. 1, n.d.). AEA published the *Guiding Principles for Evaluators* (1993, 2004) used in this study. Professional journals associated with the organization include the *American Journal of Evaluation* and *New Directions for Evaluation*, which routinely addresses issues related to the teaching of evaluation. AEA’s mission, through the journals, and the organizations conference series is to improve evaluation practices and use, and specifically fosters the professional development of new and current evaluators. In addition, AEA provides a voice for evaluation theories and approaches for users of evaluation: government, private, public, and non-profit organizations.

14. **JCSEE (JCSEE):** Serving as a major source of information on knowledge, skills, attitudes for evaluator competencies via the *Program Evaluation Standards, 2e* (JSCEE, 1994). “The Joint Committee is a coalition of major professional associations concerned with the quality of evaluation” (JCSEE About Us section, para. 1, 2012). Because the third edition was not available until 2011, well after most syllabi in this study were written, the 1994 standards were used in this study to compare what evaluator competencies were addressed by programs. The committee is accredited by the American National Standards Institute (ANSI), which also approved the program evaluation standards as recognized American National Standards.

15. **American National Standards Institute (ANSI):** Founded in 1918, this institute “oversees the creation, promulgation and use of thousands of norms and guidelines that directly impact businesses . . . ANSI is also actively engaged in accrediting programs that assess conformance to standards” (American National Standards Institute [ANSI] About ANSI section, para. 2, n.d.). ANSI is composed of individuals, companies, organizations, academic and international groups, and government agencies; it is the U.S. representative for the International Organization for Standardization (ISO).

16. **International Board of Standards for Training, Performance, and Instruction (IBSTPI):** “Is primarily involved in research and development activities that lead to the development of standards in the areas of training, performance and instruction” (IBSTPI What we do section, para. 1, 2010). IBSTPI validated their 2006 list of evaluator competencies for organizational settings using a review of literature and feedback from 450 international professionals in the field (Russ-Eft et al, 2008). The IBSTPI evaluator standards are also used to compare competencies addressed by programs.

17. **Carnegie Foundation for the Advancement of Teaching:** “Chartered in 1906 by an act of Congress, the Carnegie Foundation for the Advancement of Teaching is an independent policy and research center” (Carnegie Foundation for the Advancement of Teaching, para. 1, 1999). The Carnegie
Classification of Institutions of Higher Education was used to describe the participating programs (extent of research focus). Also, the Carnegie Initiative on the Doctorate model, which worked with disciplines in terms of restructuring programs to better prepare graduate students for careers in their discipline (Carnegie Graduate & Professional Education, para. 2, n.d.) was influential in developing this study.

18. National Center for Education Statistics (NCES): This government organization maintains a database of information available to any user; the “Higher Education Act of 1965, as amended, requires that institutions that participate in federal student aid programs report data on enrollments, program completions, graduation rates, faculty and staff, finances, institutional prices, and student financial aid” (National Center for Education Statistics [NCES] About the Integrated Postsecondary Education Data System [IPEDS] section, para. 1, n.d.). NCES and IPEDS provided additional information to describe participating programs and their institutions.

Assumptions of the Study

The assumptions in this study were:

Program directors will be motivated, based on recent conversations at AEA conferences and publications by recognized theorists, to participate in the survey.

All programs address evaluator competencies in their curriculum but may vary in terms of emphasis and in terms of strategies used.

Limitations of the Study

This study uses archival documents and interviews with program-informants and prominent theorists. The qualitative data collected cannot be used to make cause and effect statements as to actual student learning or program activities.

Delimitations of the Study

The relatively small number of programs selected to participate reflects an intentional focus towards a non-generalizable sample. Participating programs represent
purposefully chosen mini-case studies. As such, they do not reflect the range of evaluation degree or certificate programs currently extant in the United States.

**Significance of Study**

An increasing demand for skilled evaluators can be met by increasing the number of degree-based evaluation training programs. This topic has been prominent in the evaluation literature since the 2002 AEA sponsored survey of university-based evaluation degree programs in the United States. Suggestions for enhancing university-based evaluation degree programs include looking at partnerships (interdisciplinary and field-based) in preparing future evaluators, exploring the role of a taxonomy of evaluator competencies, updating the *Program Evaluation Standards, 2e* (JCSEE, 1994), and improving the use of evaluation results through capacity building. A taxonomy of evaluator competencies (IBSTPI, 2006; King et al, 2001; Stevahn et al, 2005) and a revision of the program evaluation standards (JCSEE, 2011) were already in progress at the time this study was conceived.

A think tank session at the 2006 AEA conference titled “How Best to Improve University-Based Evaluation Programs” determined that not enough was known about current programs. It was suggested that more information be gathered on what programs in the U.S. were doing, and how they were doing it. This study begins that task by addressing Engle and Altschuld’s (2003) comments from the 2002 Survey of Programs. A critical need noted by the author in personal communication (Altschuld, 2004) was to catalog the instructional methods used by evaluation degree programs to teach evaluation competencies, and how these, and other program resources, might vary. Trevisan’s 2004 review of the sparse literature on the teaching of evaluation also suggested that there might be gaps between instructional strategies used, and the theory or purpose behind them. Describing how successful programs develop students’ professional skills and socialize them into the evaluation profession is the first step in understanding the range of university-based evaluation programs.
Summary

Stufflebeam (2001) identified a need to grow the number of evaluators available to meet the increasing demand for evaluations. In their 2002 directory of university-based evaluation degree programs Engle and Altschuld (2003) suggested that the number of programs might be decreasing compared to numbers indentified in previous directories. However LaVelle and Donaldson (2010) using a web-based search protocol disagreed. Instead they suggested that the overall number of programs may be increasing, finding twice the number of small programs, 31 or 64% of surveyed programs, in 2008. This is compared to the number identified in the 2002 directory by Engle and Altschuld (2006), where 17 or 63% of surveyed programs were identified as small. Engle et al (2006) suggested that perhaps more important to the conversation on preparing future evaluators was to more clearly describe what evaluation degree programs were doing. This suggestion is supported by LaVelle and Donaldson’s (2010) comment that “we found evidence of an additional 13 programs that report training evaluators” although these did not include enough information to include them in the directory (p. 20). This was in addition to their finding that a number of other programs only offered one evaluation course (and hence did not meet the definition of a program).

In order to more fully describe what programs were doing, and to identify commonalities or differences among programs, six recommended programs in the U.S. were selected to participate. This study began with the identification of evaluator competencies identified in the literature that the six programs might use to help students develop evaluator attitudes and skills. The conceptual framework used to describe these programs explored the role of experiential learning strategies programs described using to socialize students into the profession. Finally, a qualitative data approach provided a picture of program components and strategies program-informants and evaluation theorists considered important in developing future evaluation professions.
CHAPTER 2: LITERATURE REVIEW

The 1993 Government Performance and Results Act (GPRA) resulted in a trickledown effect in terms of increasing the need for evaluation. As one example of this, the United States Department of Education discretionary grant process requires all proposals to include statements about their intended program goals, objectives, and the performance measures that will be used in annual and final reports. New grantees under the *Fund for Improvement of Postsecondary Education (FIPSE)* programs are required to submit an evaluation plan by March the year after they start funding (United States Department of Education, 2010). These mandates result in an increasing need for diversely trained evaluators at a time when degree programs in evaluation appear to be decreasing in number (Engle et al., 2006; LaVelle & Donaldson, 2010). The question that arises is how do program evaluation degree programs help to produce the broad range of discipline specific evaluators needed? This chapter provides background information for this question, by focusing upon the curricula of program evaluation—the training of quality practitioners, ensuring the continuance of the field, and the use of various instructional strategies.

**Mandate for Evaluation Expertise**

Recent national conferences of the American Evaluation Association repeatedly highlighted the increase in job market demand for evaluators (American Evaluation Associate Conference Program, 2001), and the relationship between what employers want and what students focus upon in developing evaluation skills. Practitioners and leaders in the field agree that university-based training should meet the needs of practice across disciplinary lines. However, limited experience in practical skills (such as completing literature reviews, using various evaluation methods, conducting interviews, constructing surveys, analyzing data, etc.) made it difficult for employers to find and retain new evaluators (Donaldson, 2006; Dewey, 2006). The accountability movement showed that external concerns are also interested in evaluation, whether pushed by
congressional mandates, by professional organizations that set standards for members, or by private corporations, foundations and non-profit organizations.

The clearest example of a Congressional mandate for performance-based budgeting comes from the purpose statement for the GPRA. There are five stated purposes to the 1993 act. The act is intended to: (1) reform program performance through public goal setting and assessment; (2) enhance effectiveness by focusing on the end results of programs, quality, and client satisfaction; (3) increase the level of service provided through the use of planning objectives and results reports; (4) inform congressional decision-making; and 5) increase citizen confidence in government (Office of Management & Budget – Executive Office of the President, Government Performance Results Act of 1993 [GPRA], 2006, Section 2(b)(1-5)).

Section 306 outlined the need for evaluation professionals to accomplish these purposes. Strategic plans, part (a) number (6), requires all agencies to have “a description of the program evaluations used in establishing or revising general goals and objectives, with a schedule for future program evaluations” (GPRA, 2006). In practice, GPRA requires that all executive government agencies have in place a 5-year strategic plan and an annual performance plan as part of the government’s movement toward performance-based accountability. This will require an increase in trained evaluators, given the need to provide accountability information on the wide range of government services and related disciplines covered by this act.

The Bill & Melinda Gates Foundation policies for grants provide a second example of the need for trained evaluators. They use a well worded justification for using evaluation in non-profit venues. The foundation website lists 15 Guiding Principles or Values that frame all of the non-profit programming sponsored by the foundation (Bill & Melinda Gates Foundation, 2006, para. 1). The eleventh principle states, “Delivering results with the resources we have been given [italics added] is of the utmost importance—and we seek and share information about those results” (para. 13). Implicit in this statement is the idea of using implementation evaluation (program follow-through) and outcome evaluation (achievement of targeted outcomes). Sharing results suggests sharing the formative evaluation tools used, such as needs analysis, cost-benefit analysis, personnel evaluations, and budget reviews, as well as the summative
results – the outcomes or impacts of programming. Just as a corporation evaluates its internal processes to determine if it is successful, a non-profit organization benefits from evaluating the impact of its programming, both in a formative sense (to provide ongoing change or adjustments) and in a summative sense (to determine value, best practice, or generalizability).

Performance reports with the evaluative elements of the GPRA protocols are also found in regional and national accreditation requirements for postsecondary education institutions. For example, as part of its ten-year accreditation process, the Southern Association of Colleges and Schools (SACS) requires institutions to develop specific, measurable outcomes for all program areas (academic as well as non-academic), and provide ongoing, appropriate data as well as action plans related to stated program goals. This means that the accreditation process now involves a trend toward accountability by requiring a match between goals, objectives, and the data collected to assess outcomes (Commission on Colleges: Southern Association of Colleges & Schools, 2004). While assessing the match between goals and outcomes may seem a standard component of program planning and implementation, it has not been a required skill set. But it is a skill set in demand as the reporting requirements for institutional Quality Enhancement Plans (QEP’s) spread throughout the institution’s academic and non-academic programs. The need for content matter experts with experience in evaluation perspectives to provide training and guidance has grown as institutions add evaluation to the list of faculty and staff administrative tasks.

**Accountability in Education**

The Secretary of Education's Commission on the Future of Higher Education (U.S. Department of Education, 2006) called for changes in how higher education produces accountability reports “Higher education must change from a system primarily based on reputation to one based on performance. We urge the creation of a robust culture of accountability and transparency throughout higher education.” (p.20). This culture change highlights four areas of concern (Access, Affordability, Quality and Accountability) with how the American higher education system is currently structured,
funded, and overseen. The two areas that advocate a performance-based approach are Quality and Accountability. Participation in higher education should result in some measurable level of value-added. To set standards, the commission states that the “faculty must be at the forefront of defining educational objectives for students and developing meaningful, evidence-based measures of their progress toward those goals” (U.S. Department of Education, 2006, p. 2(3))

In terms of accountability, the commission suggested that accreditation agencies have not been outcome-based and are too reliant on process data. Because accreditation polices are changing towards an outcome or performance reporting system, departments and programs are required to create objectives and measurement systems to track their progress in achieving those objectives. The accreditation process does not tie in directly to how institutions receive their funding, despite the consequences if an institution fails to be re-accredited. Accreditation is an internal system designed to guide change, even though users, the general public, and state funding agencies may be unaware of why institutions go through the accrediting process or how it might guide change in a formative way. The commission recommended that performance report information be made available in aggregate form to the general public in a readily accessible format. The information could then be used to make decisions about how institutions are funded or utilized, and would be called a performance-based accountability system.

Decision-making to determine effectiveness in higher education is more complex than as suggested in the report from Commission on the Future of Higher Education. Burke (2005) describes an accountability triangle in higher education. Each point of the triangle is pulled by a different stakeholder based upon their own assessment of what constitutes an appropriate accountability system. These three points are public, market, and academic demands and have both positive and negative impacts in evaluating higher education. They are examples of different approaches that could be taken in evaluating an institution. The public concern for an educated citizenry and workforce is exemplified by state mandates and guidelines for higher education institutions. Market demands focus upon the needs of employers and employees [students]. Academic demands can be thought of as disciplinary and administrative directives that guide
curricula as well as set policy. Both the public and market demand positions are represented in a positive light by the report from the Commission on the Future of Higher Education. But the negative aspects of political partisanship in the public demands, or an economic bias from the market demand that results in behavior like the Enron scandal, were not mentioned. Conversely, the commission's report over emphasizes reputation-building behaviors (faculty stars, selective recruitment and endowments) in the academic corner that may not contribute to institutional mission.

The positive pulls where “academic concerns . . . encourage free inquiry and discussion of ideas, beliefs, and institutions infused by openness, scholarship, and objectivity” are mentioned only briefly when the report refers to the faculty role in determining learning objectives and in developing new ideas (Burke, 2005, p. 22; U.S. Department of Education, 2006).

There is a history to the similarities in the recommendations made by the GPRA, the Bill & Melinda Gates Foundation, the accreditation agencies, and the Commission on the Future of Higher Education Report. This is because they go beyond the earlier matching spending with doing approach to accountability. The 1980’s bureaucratic model of accountability centralized governance stressed efficiency, with compliance to rules and regulations. “The need and desire for accountability presents a need for evaluation” (Alkin & Christie, 2004, p. 12). The accountability movement in the 1990’s was focused on results with an interest in the outcomes of programs rather than the processes used (Burke, 2005). These examples require a performance orientation in accountability that goes beyond accounting for spending to include outcomes, which may or may not be tied to intended goals or indicators.

However, according to Burke (2005), success standards or benchmarks are often linked primarily to funding and budgets. The No Child Left Behind Act of 2001 (NCLB Act) is an example with its link between Florida’s Department of Education budget allocations and student and school performance on the Florida Comprehensive Assessment Test (FCAT) (United States Department of Education, 2001; Florida Department of Education, 2005). The Commission on the Future of Higher Education Report refers to benchmarks to be developed by higher education or by government. The Government and Performance Results Act of 1993 allowed directors of agencies to
set annual performance plans, goals and objectives. Both the higher education commission's benchmarks and the GPRA have some component where budgets and performance are tied. Another way to view these is to consider performance reporting as more formative and performance-based accountability as more summative in nature.

As Alkin & Christie (2004) noted, evaluators straddle both perspectives frequently, moving from the collecting of information for Accountability & Control to a Social Inquiry approach depending upon the client demands. Unlike the NCLB Act, neither the accreditation agencies nor the Gates Foundation, which relies “on others to act” (Bill & Melinda Gates Foundation, 2006, para. 6), set indicators or success standards. This separation of funding from information gathering allows a degree of flexibility to respond to changes in market, public, or academic needs that occur after budgets are set. Instead, this ability to be responsive is what Alkin and Christie (2004) referred to as the Social Inquiry root in evaluation purposes, which “emanates from a concern for employing systematic and justifiable set of methods for determining accountability” (p. 12).

In its article, “The Challenge from Within: A National Commission on the University-Community Interaction to Harness the Potential of America’s Cities” (1993), the Milken Family Foundation National Education Conference stated “The goal of [university-community partnerships] is to positively affect communities—their students, business people, and organizations—and the universities which are a part of the program. The long-term aim is to revitalize the community, and consequently, reinvigorate the university’s mission” (McCroskey & Einbinder, 1994, p. 3). University degree programs offer one avenue to meet the demand for trained evaluators. The practice of evaluation offers feedback on evaluation that impacts the teaching of evaluation to those theorists responsible for degree programs.

According to Scriven, "to evaluate is to judge the [relative] worth or merit of something" (Worthen, Sanders, & Fitzpatrick, 1997, p. 5). Evaluation makes use of the same research skills found in other disciplines (qualitative and quantitative methods for data collection and analysis for example). This is one reason evaluation degree programs are located in disciplines such as education, psychology, sociology, social work and medicine (May, Fleischer 1986). Scriven believes that evaluators share a
common skill set with other disciplines and combining that skill set with content matter expertise can make evaluators more effective in their respective areas (Scriven, 1991). Altschuld (1986) challenged this in relation to training practice-based evaluators. He argued that “whether there is a field or not, the practicing evaluator needs a broader base of knowledge and skills” than that provided by additional courses in another discipline (Altschuld, 1986, p.262). His argument is that “evaluation encompasses traditional topics (measurement, formative and summative design) as well as other ones (needs assessment, personnel evaluation, multi-methods usage)” (p. 262). It is the additional topics and their integration through guided practice that develops a trained evaluator. This evaluation-specific training is critical as governments, post-secondary institutions, and corporate or non-profits who use a trained evaluator will be more effective in meeting their accountability requirements.

The Decline in Number of Evaluation Degree Programs

In 1984 both the Canadian and American evaluation societies commissioned a directory of evaluation degree programs (May, Fleischer, Scheirer, & Cox, 1986). The authors referred to a listing of 80 programs from a 1976 publication by Gephart and Potter that indicated evaluators usually earned degrees from measurement and statistics programs in the 1970’s, which numbered approximately 100 in the United States. McConnell (1982) stated that a 1980 directory by Connor, Clay, and Hill (1980) identified 67 evaluation training programs in a variety of disciplines. The 1984 directory (May et al, 1986) lists 45 programs in the United States that offered evaluation training, but this survey did not define evaluation training other than seeking “a program shall prepare students to conduct independently a program evaluation” (p. 72). The 1992 and 2002 the surveys included a more exact definition for evaluation programs where: “a program consists of multiple courses, seminars, practica, offerings, and so on designed to teach what the respondent considered to be evaluation principles and concepts” (Engle et al, 2006, p. 355). Of the 49 surveys received in the 1992 survey, only 38 met the definition of an evaluation program (Altshuld, Engle, Cullen, Kim, & Macce, 1994). In the 2002 survey there were 38 respondents but only 29 that met the definition of
evaluation program. Only 24 programs were in the United States, five were international (Engle et al., 2006). Only 10 of the 29 United States and international evaluation degree programs offered four or more classes with a predominantly evaluation focus, the rest were small programs. Yet LaVelle and Donaldson (2010) found 48 programs in the United States that met the 1992 and 2002 survey definition for an evaluation program. See Table 1 below.

### Table 1. An Overview of Directories of Evaluation Degree Programs in the United States

<table>
<thead>
<tr>
<th>Date Collected</th>
<th>Directory Data Collected</th>
<th>Number of Programs Meeting definition for Evaluation Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>~1976</td>
<td>Gerhart &amp; Potter, 1976</td>
<td>80 No known definition</td>
</tr>
<tr>
<td>1979</td>
<td>Connor, Clay, &amp; Hill, 1980</td>
<td>67 No definition</td>
</tr>
<tr>
<td>1984</td>
<td>May, Fleischer, Scheirer, &amp; Cox, 1986</td>
<td>46 (44 U.S.) “a program shall prepare students to conduct independently a program evaluation”</td>
</tr>
<tr>
<td>1992</td>
<td>Altschuld, Engle, Cullen, Kim, &amp; Macce, 1994</td>
<td>38 (U.S.) “A program consists of multiple courses [more than one], seminars, practica, offerings, and so on designed to teach what the respondent considered to be evaluation principles and concepts”</td>
</tr>
<tr>
<td>2002</td>
<td>Engle, Altschuld, &amp; Kim, 2006</td>
<td>24 (U.S.) Same as 1992</td>
</tr>
<tr>
<td>2008</td>
<td>LaVelle &amp; Donaldson, 2010</td>
<td>48 (U.S.) Same as 1992</td>
</tr>
</tbody>
</table>

Altschuld et al. (1994) suggested that the decline in the number of programs from the earlier directories to the 1992 directory was partly due to the use of a concise definition for an evaluation program (as opposed to a program in educational research, for example). Engle et al. (2006) also suggested that the apparent continued decline in the number of evaluation programs (from 38 in 1992 to 24 in 2002) was less dramatic than it initially appeared. They suggested several reasons for the lower numbers. First, the 2002 survey had a low response rate (45%), perhaps in part due to an inflated sampling frame (which included programs that had already gone defunct). Second, the
criteria may have been implemented differently, even though the same definition of an evaluation program was used in 1992 and in 2002. The web page survey by LaVelle and Donaldson (2010) supported these two ideas as they found 48 programs when they applied the definition to the data they collected. Finally, Engle et al. (2006) stated that there were an increasing number of venues for evaluation training beyond the university-based programs, such as national institutes, in-service training, and workshops offered by the national and regional evaluation organizations. They suggested that further research is needed to examine how non-degree based venues may impact university-programs, either in terms of enrollment numbers, growth in an interdisciplinary sense, or maintenance of evaluation practice standards.

One reason why the 2002 data suggested a narrowing of programs is that evaluation as a field of study has been changing as it defines itself apart from the early measurement and statistics programs. The surveys have also become more focused. In 1986 Barbara Gross Davis, editor of New Directions for Program Evaluation commented on the new attention being paid to evaluation compared to the late 1960’s. “If evaluation was taught at all, it tended to be a component of a research methods or measurement course. Within the last ten years, evaluation has come to be recognized as a legitimate professional field…that possesses its own core of knowledge, specialized concepts, and particular strategies” (p. 7). The number of evaluation-focused courses that are now offered may be a better marker than merely counting number of programs.

Engle et al. (2006) confirmed this shift in focus on evaluation as a program of studies by noting that 50% of programs in 1992 used evaluation in their titles. By 2002, 69% of programs did so. Yet the majority of programs continue to reside in the same disciplines as in 1992—education, educational psychology, or psychology, despite the growing use of evaluation in other fields. This trend was confirmed by LaVelle and Donaldson (2010) with 73% of programs using evaluation in their titles, and most programs housed in the education, educational psychology, or psychology. This traditional base of limited disciplines may be reinforcing the growth of non-university based training programs as students in other disciplines are not exposed to evaluation concepts in their degree areas.
A critical issue to the authors was the size of the university-based programs in the 2002 survey. This echoes the point made by Barbara Gross-Davies (1986) that evaluation was becoming a field of its own. In the 2002 survey evaluation programs were categorized by the number of evaluation specific courses offered (Engle et al., 2006). Typical course titles included: Evaluation of Educational Programs and Curricula, Applied Evaluation Design, Advanced Evaluation Methodology, Practicum in Evaluation, Introduction to Needs Assessment, Foundations of Policy Analysis, Cost-Benefit Analysis, and Personnel Evaluation (p. 358). Of the 29 domestic and international programs surveyed in 2002, 19 programs were small (2-3 courses), nine were medium (4-6 courses), and one program was large (7 or more). The 2008 survey by LaVelle and Donaldson (2010) found that the number of programs had increased but that proportions remained the same: 64% of programs were identified as small in 2002 and 64% in 2008, with the number of programs identified as medium combined with large programs at 37% in 2002 and 36% in 2008.

Program size data cannot account for the change in the number of degree programs as the number of evaluation-focused courses was not measured in the previous surveys. This poses a potential problem for doctoral programs. A limited number of evaluation courses (theory, practice, and methods) in a doctoral degree program would have implications for the skill set of graduates. Compared to master’s degree programs, doctoral programs extend the breadth and depth of what students learn in the discipline, usually through the number and type of courses offered as well as through internships and dissertation studies. New evaluation faculty will be competitive hires for university departments only if they have a comprehensive background in theory and methodology as well as practical experience. It is not known how many degree programs have program objectives that include preparing future faculty for scholarship, research, and teaching in university-based training programs.

What is an Evaluator?

Evaluators differ in (1) their opinions of the optimal approach to an evaluation inquiry, (2) in their employment venues, academic discipline, and evaluation preparation
experiences, and to a small extent, (3) in their view of evaluation as a discipline. These differences influence the training evaluators receive, how evaluations are designed, what questions are asked, who participates, and how evaluations are reported. Evaluations can have different goals and venues which in turn require different evaluator skill sets. But how are these different emphases - evaluation approach and preparation for evaluation careers - reflected in the curricula of evaluation degree programs? Are they reflected as well in the evaluator competencies identified in the literature? The following reviews what is available in the literature.

**Approaches to Evaluation Inquiry**

A major influence on the guidelines and standards for evaluation that have been developed, and are discussed later, are the approaches to framing an evaluation inquiry. These approaches can be varied, as was noted in Chapter 1 with Chen’s 2004 definition of an evaluation approach as “a systematic set of procedures and principles guiding evaluators . . . [in] conceptualizing problems and determining the evaluation’s focus” (p. 144). As an example of the range of conceptualizations, Rossi et al. (2004) describes evaluation procedures and principles as being found not only in the traditional disciplines of education but also in psychology, and sociology, and all of the social sciences. He cites significant contributions in evaluation methodology and perspectives from economics, political science, public administration, and public health. The danger, Rossi states, is defining the epistemological understanding of evaluation solely in terms of the traditional disciplines (education, psychology, and sociology). For Rossi, evaluation covers a broad territory. As an example of the variety of approaches in evaluation inquiry he reminds his readers that the “major large-scale evaluations of critical federal programs in the last two decades have been designed and conducted by economists” (p. 131).

Alkin and Christie (2004) agreed in general with Chen and Rossi. They submit that evaluation inquiry approaches can be categorized in terms of three primary divisions regardless of discipline: (1) Use, where the evaluator is guided by how evaluation information will be used; (2) Methods, where the primary issues relate to how
data is collected; and (3) Valuing or judging, which is guided by the placement of value on data, and the determination of criteria or characteristics of importance.

**Use approach.** Use approach to evaluation focuses on the utility of evaluation results as guiding the design for an evaluation inquiry. The purposes of an evaluation can range widely. Examples provided by Alkin and Christie (2004) include: Wholey’s information for managers and policy makers; Provus’s view of discrepancy evaluation focusing on the differences between intended and envisioned goals, and outcomes; Stufflebeam’s focus on decision-makers; Patton’s utilization-focused evaluation that emphasizes work with the real users of evaluation rather than the external mandates to conduct the evaluation; Preskill’s transformational learning where long-term utilization is achieved through capacity-building during the evaluation; and Fetterman’s empowerment evaluation, where program participants learn to evaluate themselves in the process of providing their own data. The main assumption in the use-oriented approach is that evaluation results must be actively used and that there are a variety of means to achieve that use.

Daniel Stufflebeam’s CIPP model (standing for Context, Input, Process, and Product) is an example of active use and means. This model emphasizes the decision-making aspect of evaluation related to who will use the information. One strategy is to involve stakeholders, individuals or groups who have an interest in the program being evaluated, in the design of the evaluation so that the final results are more likely to be seen as credible and useful.

**Method approach.** The second evaluation inquiry approach is oriented towards the evaluation method. This approach is concerned with the design influences resulting from various means for collecting and analyzing data. For example, the methodology determines what information will be collected, who will be contacted and how, and to what extent the evaluation information can be generalized. Alkin and Christie (2004) joke that “in the beginning, there was research. And the methods of research dominated the conduct of studies” (p. 17).

A truer beginning might have been simple observation as a source for decision making. But without understanding the underlying processes, the relationships between observations can seem magical, or just be false. Darwin’s 1859 *Origin of the Species*,
exemplified the power and limitations of the careful naturalistic observation that ruled intellectual discovery-making at the time. Predictive and causative statements are not possible without the ability to control and measure objects of interest and that control is relatively recent. Description and control of variables and accurate measurement lie at the heart of research; early social science research (which emphasized experimental and quasi-experimental designs) is an example.

Two concerns have shifted thinking on how research should be conducted in the social sciences. The first concern was an ethical one. The 1974 National Research Act officially stated that it is unethical to withhold treatment when needed, and that research subjects require informed consent. The result is that some research designs, while allowing control, fail in terms of ethics.

The second concern related to validity and has far reaching implications for design. External validity refers to the idea that the result obtained should be a fair representation of what actually occurs, and that influences outside the control of the evaluator should be taken into consideration. Internal validity refers to accuracy of measurement. Case studies are an example of giving up control by investigating more deeply, in the hope of assigning relationships that can be seen in other examples and hence extend external validity. Careful descriptions and clear logic assure internal validity.

While research designs using true randomized control/experimental groups are the standard in the sciences, the method lacks feasibility in social science research where it is not possible or is too cumbersome to conduct such research. Consequently the results are not generalizable. The early evaluation theorist Campbell (1991) continued to recommend that quasi-experimental methods (such as pre-test, intervention/treatment, post-test) be used as a means to ensure validity in evaluation studies where true randomization is not possible. Boruch (1995) used randomized field tests, where individuals are randomly assigned to treatment groups, to ensure that extraneous variables did not bias his ability to make causal attributions. Rossi and Weiss moved the continuum further in the need to explore and explain the context of the study, rather than control what cannot be controlled. Rossi emphasized designs that would have both internal and external validity (Rossi et al., 2004). Weiss (1993, 1998)
emphasized the lack of bias in evaluation designs and clear communication of questions and design constraints. Her goal in evaluations is to ensure that the validity of results could not be questioned. In all cases, methods are used that will offer the desired level of validity within the design constraints (Alkin & Christie, 2004).

**Value approach.** The third division presents evaluation approaches that explore how values are embodied in the evaluation process. Evaluator values influence a range of evaluation issues from decisions as to which questions or data should be pursued to what results are valid. Another way to describe this is how the evaluator determines the context of the evaluation. That is, the evaluator chooses key characteristics relevant to the evaluation and develops an evaluation inquiry to compare and explore those characteristics.

Alkin and Christie (2004) reference two theorists who define evaluation as being primarily concerned with making value judgments in terms of quality (good and bad based on the evaluator’s opinion). For Eisner (1979), determinations of quality depended upon the subjective opinion of subject-matter experts, whose evaluations are shared in the form of critiques. Scriven (1967) cited the familiar *Consumer Reports* articles as an example of how the evaluator uses expert judgment to decide what features merit a closer inspection, defining specific characteristics to examine using specific measurements. Stake (1991) also emphasized the role of the evaluator in making judgments, but adds the importance of an alignment between the expertise of the evaluator and with stakeholder values. House (1999) had the same approach as Stake but continued to a further level by advocating a venue for stakeholders who are often under-represented. A final example of the impact of the values approach in evaluation design is the work of Lincoln and Guba (1994), who not only included under-represented stakeholders, but consider them primary sources for placing value.

Despite these divisions, all evaluation serves a social function. Alkin and Christie (2004) conceptualized this social function as having two foundations—(a) social accountability and fiscal control, and (b) social inquiry. These societal demands can be likened to a perspective of maximizing the cost-benefit ratio, wherein, for example, programs should be held responsible or accountable for delivering contracted products or services. An accounting audit would be a very basic example. The second part of the
base, social inquiry, extends this idea to suggest that the assessment should be of good quality, a “systematic and justifiable set of methods for determining accountability” (p. 12). Whether the approach is dominantly oriented toward Use, Method, or Valuing, it still speaks to both of the foundations of evaluation. Each approach provides a range of perspectives as to the best means to achieve their critiques (of goals, implementation, and outcomes) and the extent to which they assign a degree of social efficacy.

Professionalization of Evaluators

The *Program Evaluation Standards, 2e* (JSCEE, 1994) and the *AEA Guiding Principles for Evaluators* (AEA, 2004) provide a general list and definitions for evaluator skill sets and behaviors. These are considered to be common for most evaluators and their evaluations regardless of their employment venue (corporate, non-profit, government). This list was generated in part for users of evaluation products, but is also used as a guide for practicing evaluators and in the training of evaluators. When these were developed, input was solicited from practitioners with a wide representation of groups. The representative groups were deliberately selected to be varied in nature (AEA, 2004; Sanders, 1999); the groups included full and part-time evaluators employed in corporate, government, not-for-profit, and small-business settings as well as faculty in academic programs. Some had experience as internal or external evaluators. Some had an academic background in evaluation but others had evaluation training from in-service programs, institutes or certificates. The contributors and reviewers of both the guidelines and the standards came from a diverse set of disciplines. This diversity in training and employment was part of the impetus behind the structure provided in the *AEA Guiding Principles for Evaluators* (AEA, 2004) and the *Program Evaluation Standards, 2e* (JSCEE, 1994).

*Program Evaluation Standards.* The JCSEE developed the *Program Evaluation Standards*, now in its third edition (JSCEE, 2011). The standards address the process-based skills for the conduct of evaluations. How evaluators should behave is addressed by the *AEA Guiding Principles for Evaluators* (AEA, 2004), discussed below. The JCSEE started in 1975 with representatives from the American Psychological Association, the American Educational Research Association, and the
National Council on Measurement in Education. This first group added other members to represent the venues where educational evaluations occurred and eventually included 18 professional organizations. The first set of standards for evaluating educational endeavors was published in 1981.

The first set of standards was modeled after the Certified Public Account (CPA) standards for accounting. This format allowed the group to maintain the standards as well as address new ones. The non-profit JCSEE was formed at this time (Sanders, 1999). In 1989 the American National Standards Institute (ANSI) invited the Joint Committee to apply to become an ANSI accredited organization. This was important because an ANSI approved standard is a recognized international standard and national standard. The second edition of the *Program Evaluation Standards* (JCSEE, 1994) was approved by ANSI in 1994.

The second edition of the *Program Evaluation Standards* (JCSEE, 1994) identified four competency areas or standards: Utility, Feasibility, Propriety, and Accuracy. Each standard is defined and then further detailed with specific guidelines. The third edition of the standards added a fifth competency area or standard: Evaluation Accountability (Yarbrough, Shulha, Hopson, & Caruthers, 2011). The third edition provides additional material connecting the individual standards under the five competency areas. It also includes specific references to the evaluation context and cultural competence (the idea that cultural differences must be taken into account in the design and implementation of the evaluation). An overview of the standards from the third edition including the fifth area, taken from the JCSEE (2011), is below:

**Utility Standard**
The utility standards are intended to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs.

**Feasibility Standard**
The feasibility standards are intended to increase evaluation effectiveness and efficiency.

**Propriety Standard**
The propriety standards support what is proper, fair, legal, right and just in evaluations.
Accuracy Standard
The accuracy standards are intended to increase the dependability and truthfulness of evaluation representations, propositions, and findings, especially those that support interpretations and judgments about quality.

Evaluation Accountability Standard
The evaluation accountability standards encourage adequate documentation of evaluations and a metaevaluative perspective focused on improvement and accountability for evaluation processes and products. (JCSEE Program Evaluation Standards Statements section, para. 4-8, n.d.).

It is not expected that every guideline will be applied equally in all evaluations. Each evaluation must meet the needs of the client. For example, under the Propriety Standard is a guideline for transparency and disclosure of evaluation findings and materials. There are cases where this is not allowed by the client. Similarly, while the accuracy standard is crucial for interpreting data and making recommendations, not all forms of data collection are feasible for each evaluation budget.

AEA Guiding Principles for Evaluators. AEA’s Guiding Principles for Evaluators (AEA, 2004) stated that they “are intended to guide the professional practice of evaluators, and to inform evaluation clients and the general public about the principles they can expect to be upheld by professional evaluators” (AEA Guiding Principles for Evaluators section, 2004, para. C). The guidelines were drafted in an iterative process by individuals representing various fields of practice and include assumptions embedded in various evaluation theories. An initial AEA task force and subsequent ad hoc committee drafted the guidelines and offered them for review by discussion groups at the annual American Evaluation Association Conference in 1993.

In 2002-2003 a review and revision was conducted by the AEA Ethics Committee and the AEA Board. AEA member input was solicited via an online survey, a meeting at the annual conference, and material from the American Journal of Evaluation. Additional feedback was obtained from AEA’s Diversity Committee, Building Diversity Initiative, Multi-Ethnic Issues Topical Interest Group, and past AEA presidents. It was finally approved by the board in 2004 (AEA, 2004). The AEA Guiding Principles for Evaluators (AEA, 2004) outlined expectations for evaluator behavior in five areas (non-ranked order):
A. Systematic Inquiry: Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.

B. Competence: Evaluators provide competent performance to stakeholders.

C. Integrity/Honesty: Evaluators ensure the honesty and integrity of the entire evaluation process.

D. Respect for People: Evaluators respect the security, dignity and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.

E. Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare. (AEA Resulting Principles section, 2004).

But how these expectations are to be achieved, within the context of training or skills sets in conducting an evaluation, is not specified. Similarly, the *AEA Guiding Principles for Evaluators* (AEA, 2004) are carefully prefaced as a suggested approach for producing evaluations of quality:

the primary purpose of evaluation . . . is not limited to the following: bettering products, personnel, programs, organizations, governments, consumers and the public interest; contributing to informed decision making and more enlightened change; precipitating needed change; empowering all stakeholders . . . the common ground is that evaluators aspire to construct and provide the best possible information that might bear on the value of whatever is being evaluated. The principles are intended to foster that primary aim. (AEA Resulting Principles section, 2004, para. B).

**Evaluation as a Discipline**

Scriven (2001) has suggested that evaluation is a transdiscipline. In his analogy, some fields (such as statistics and measurement) serve as support areas that cross over to the “discipline estates” - disciplines like psychology, education, and even physics. Evaluation provides tools to its customers (the disciplines) but also maintains its own operations. This service component of providing an interface with other disciplines is important to Scriven. But Scriven also felt that growth in evaluation degree programs, and in research that contributes to the field of evaluation, could be greater if evaluation were taken to be a discipline in its own right.
To assure that trained evaluators are available for multiple content areas, Scriven (2001) proposed that evaluation degree programs recruit students from other disciplines so that evaluation training is seen as a ‘trans-disciplinary’ field—one that crosses over into others. Evaluation faculty whose work encompasses theory as well as practice would enhance not only what is known about evaluation from a practical sense, but would also refine the training of skilled evaluators.

Defining skill sets and competences is one aspect of developing evaluation as a disciplinary area. However, research matching those skills with actual practitioner needs lags behind. Stufflebeam (2001) stated that a growing job market in evaluation is demanding more doctorates but that current programs, while producing quality evaluators, tend to orient those evaluators toward specific disciplines. This is producing a gap between the education and training of evaluators within a broader range of disciplines and the needs of the job market. Stufflebeam (2001) proposed four needs in expanding higher education programs in the field of evaluation, namely an increase in the:

1. Number of evaluators available discipline-wide,
2. Number of evaluators who are (trained to be) cross disciplinary,
3. Cross-disciplinary communication between graduate students, and
4. Practice-linked component in the evaluation curriculum (field or internships).

Stufflebeam’s call for an increase in a practice-linked component in evaluation curricula, beyond exposure to theory, suggests the following general question - How can we create practice-based evaluation programs that generate competent evaluators for a variety of disciplines?

Hoberman and Mailick (1994) shared Stufflebeam’s practice-linked concern by addressing educational tasks. They defined expected outcomes of professional education as “helping students acquire special competencies for diagnosing specific needs and for determining, recommending, and taking appropriate action” (p. 3). The tie to the community of practitioners is critical in that these outcomes go beyond standard academic knowledge, skills, and abilities (KSA’s) to include socialization of “students in the ‘thought processes’ of the profession and to inculcate them in the customs, ethics, and working relationships, and the behaviors expected from members of the profession”
It is this acculturation to an evaluative perspective through broad practical experience that Stufflebeam seeks to provide to students in cross-disciplinary programs. This programmatic need to include practice-linked components makes cross-disciplinary studies difficult. Stufflebeam (2001) noted that the gap between practice and theory can be bridged by incorporating experiential learning activities into the curriculum. Education in the professions (law, engineering, education, medicine) commonly utilizes internships, simulations, and other types of experiential learning activities. In these examples, professional associations also play an active role in helping to determine both the curricula and the instructional methodology, largely through their influence in certifications and board examinations, which assess practical skill. Utilizing what is known about experiential learning to enhance learning activities in evaluation could help meet the increasing demand for evaluation while providing both students and practitioners with an open dialogue about theory and practice.

Most authors agree that evaluators and clients benefit from specific standards of practice (Patton, 1990; Scriven, 1991; Worthen et al, 1997). Altschuld (1999) indicated that there is less agreement as to certifications or licensure, which imply a relationship between practitioners (usually the professional associations) and those who provide education and training for future practitioners. In some fields, they include a relationship with government (such as the medical licensure in different states). Whether a discipline or a profession, evaluation is not at the point of requiring licensure, and may never be in the way physicians are licensed because of the variety of its applications and methods (Worthen et al., 1997). But a discussion of general or core competencies or skill sets is relevant and ongoing. The question of what comprises an ideal curriculum in evaluation in terms of evaluator competencies, if it is to be considered a separate field, and how it should be taught, continues to be a topic of interest, although the number of articles devoted to this area is small (Trevisan, 2004).

**Developing a taxonomy of evaluator competencies.** King, Stevahn, Ghere, and Minnema (2001) developed a taxonomy of evaluator competencies. They utilized iterative discussions with faculty and practitioners to develop a list of skills for evaluation degree programs. The study looked more specifically at the development of program evaluation skills in order to delineate those skills, or competencies, that in a skilled
evaluator. Their goal was to start a conversation on what an ideal curriculum, or training program, might entail. Their first step was to ask if there was a consensus among evaluation professionals for a “taxonomy of essential evaluator competencies” (p. 229). To answer this question, they used a variation of MACB (multi-attribute consensus building). The focus group included diverse members of the evaluation community (age, job title, years as an evaluator, training, etc.) who were tasked with identifying and then ranking the knowledge, skills, and attitudes (KSAs) for evaluators. The resulting list of competencies (taxonomy) was proposed as the basis for a core curriculum in program evaluation.

Their competencies were grouped into four main categories with a focus on skills, not just knowledge. Two primary themes emerged. First, they drew a clear distinction between doing evaluation and doing research. None of the evaluators considered their work to be research. Second, they determined which competencies should be rated as more important and that ranking influenced which competencies were included in the taxonomy. Evaluators tended to rate those aspects directly under their responsibility rather than delegated (such as setting up the evaluation processes) as being most important. The resulting list of skills within sub-areas may have missed competencies identified by other groups (Russ-Eft et al., 2008).

The first competency category was Systematic Inquiry and included the ability to do a) research-oriented activities, b) evaluation-oriented activities, and c) those activities common to both research and evaluation. The second was Competent Evaluation Practice and included a) information management (for intended users), b) situational analysis, and c) organization and management skills. In the King et al. (2001) study, each of these areas and sub-areas, as well as those that follow, were itemized according to the ranges of agreement. The third category was General Skills for Evaluation Practice –communications skills including a) logical and critical thinking, b) written communication, c) verbal communication, d) communication, and e) computer application skills. The final category was Evaluation Professionalism and included a) knowing yourself as an evaluator, b) ethical conduct, c) knowledge of professional standards (the Program Evaluation Standards and & AEA Guiding Principles for Evaluators), d) application of professional standards, and e) professional development.
The King competencies were refined by Stevahn et al. (2005), who collected additional information from conference and session participants. The general consensus was that the categorization was complicated and was neither intuitive nor user-friendly. The revised list was linear with “six distinct competency categories, namely, (a) professional practice, (b) systematic inquiry, (c) situational analysis, (d) project management, (e) reflective practice, and (f) interpersonal competence—each of which more clearly identifies the specific competencies contained within.” (p.52).

The most recent revision by Ghere, King, Stevahn, and Minnema (2006) further clarifies the various definitions and provides a seven-point rating scale where [0-2] equals entry/novice level, [2-4] equals proficient or skilled level, and [4-6] is equal to mastery/expert:

1. Professional Practice – Fundamental norms and values of evaluation practice
2. Systematic Inquiry – Technical aspects of evaluation practice
3. Situational Analysis – Unique interests, issues, and contextual circumstances of evaluation
4. Project Management – “Nut and bolts” of evaluation work
5. Reflective Practice – One’s own evaluation expertise and need for growth
6. Interpersonal Competence – “People skills” necessary for evaluation practice

These 73 competencies are worded in terms of evaluator behaviors very similar to the 84 performance standards created by IBSTPI (Russ-Eft et al., 2008).

**Professional organizations as sources for evaluator skill sets.** While the essential competencies taxonomy continues to be refined, it is not the only list available. The Canadian Evaluation Society and IBSTPI have also detailed competencies or skills they considered crucial to the training, or instruction, of future evaluators (Canadian Evaluation Society, 2004; IBSTPI, 2006). A comparison made by Stevahn et al. (2005) indicated considerable overlap between the *Program Evaluation Standards* developed by the JCSEE in 1994 and the *AEA Guiding Principles for Evaluators* revised by AEA in 2004. The Canadian Evaluation Society addressed the same skill sets as: Understanding Program Evaluation, Building an Evaluation Framework, Improving Program Performance, and Evaluating for Results (Canadian Evaluation Society, 2004).
The International Board of Standards for Training, Performance and Instruction (IBSTPI, 2006) has a similar list of evaluation competencies divided into four domains with 14 competencies:

Professional Foundations
1. Communicate effectively in written, oral, and visual form.
2. Establish and maintain professional credibility.
3. Demonstrate effective interpersonal skills.
4. Observe ethical and legal standards.
5. Demonstrate awareness of the politics of evaluation.

Planning and Designing the Evaluation
6. Develop an effective evaluation plan.
7. Develop a management plan for the evaluation.
8. Devise data collection strategies to support the evaluation questions and design.
9. Pilot test the data collection instruments and procedures.

Implementing the Evaluation Plan
10. Collect data.
11. Analyze and interpret data.
12. Disseminate and follow-up the findings and recommendations.

Managing the Evaluation
13. Monitor the management plan.
14. Work effectively with personnel and stakeholders.

(IBSTPI Evaluator Competencies section, 2006)

The different professional associations have differences in naming structure between the domains or divisions for skill sets, but the general topics are similar. The IBSTPI competencies were relatively unique in that, as with Ghere et al.’s taxonomy of 73 competencies (2006), each IBSTPI competency is further defined in terms of measurable definitions or performance statements, a total of 84 across the four domains.

In writing about the IBSTPI competencies, Russ-Eft et al. (2008) described the process of determining core competencies as an identification of expected behaviors for internal evaluators. Part of their process included a review of other sources for evaluator competencies, such as the American Evaluation Association and the Canadian Evaluation Association. They also reviewed the evaluator taxonomy work of King et al.
(2001) and the preliminary data from the survey of evaluation programs done by Engle and Altschuld (2003). Unlike other work on evaluator competencies which did not make qualifications for where evaluations were being done, IBSTPI considered it critical to have specific evaluation competencies for specific types of evaluation specifically "internal staff or external consultants conducting evaluations in organizational settings, such as for-profit and not-for-profit organizations, military, and government agencies evaluating their own internal programs" (IBSTI, 2006, para. 1).

According to IBSTPI, the King, Stevahn, Ghere taxonomy, the Program Evaluation Standards, 2e (JCSEE, 1994) and AEA Guiding Principles for Evaluators (AEA, 2004) were lacking core competencies that are "the generic skills of an evaluator, independent of setting and organization" (IBSTPI Evaluator Competencies section, 2006, para. 2). Examples of missing core competencies ISBTPI identified are "establish and maintain professional credibility" and "monitor the management plan". However, an argument could be made that these competencies or skill sets are embedded in the more generally stated Program Evaluation Standards and the AEA Guiding Principles for Evaluators. Regardless, specific skills still tend to depend upon the needs of the evaluator (qualitative versus quantitative data), the discipline (health professions versus education), and the client (corporate versus government). Faculty teaching evaluation will therefore need competence in multiple areas of evaluation, other content areas, and other skill sets such as communication and negotiation skills (Russ-Eft et al., 2008).

Preparing Future Evaluators

There are several influences on the way in which master’s and doctoral degree programs in evaluation might be structured to help students develop professional skills as evaluators. The conceptual framework for this study uses two factors in looking at program structure, (1) socialization of students and (2) experiential learning or practical training. Socialization of students provides an overview of the behaviors and skills considered part of the acculturation to the discipline, as well as the program’s view of how to achieve or begin that acculturation. Experiential learning stresses using learning strategies where learning skill-mastery is practiced. Many of the evaluator competencies
are skill-based. Socialization and experiential learning concepts create a framework in understanding how doctoral evaluation degree programs in the United States promote student development of evaluator competencies.

Defined traditionally, “professionalization should be viewed as the transmission of content knowledge; the informing about professional norms, ethics, and the teaching of technical skills. Socialization . . . [requires] the internalization or adoption of the profession’s norms, values, and ethics” (Antony, 2002, p. 369). The term professionalization in this study refers to the means by which students acquire the knowledge, skills, and abilities (KSAs) of evaluators. Socialization is better referred to as the process by which students acquire an awareness of evaluator norms, values, and ethics or beliefs. These norms and values are also embedded in the evaluator competencies. The evaluator competencies provide direction on the evaluation components needed for socialization of evaluation students. The program structure determines how students are socialized, or made aware of evaluation values, ethics, and standards.

**Socialization of Students**

Socialization has been defined in many disciplines (sociology, anthropology, education) as the process of adopting attitudes and values appropriate to the culture Antony (2002). It is a dynamic process occurring throughout the lifetime of an individual, particularly when learning new skills, new contexts, and new cultures. Two dimensions are of interest in the socialization of evaluators: socialization to a role and socialization to a culture. Role socialization refers to the knowledge, skills, and abilities required for evaluation practice. Learning rules and requirements, standards, theories, and methodologies are all part of professional development for the role of evaluator (Sarbin & Allen, 1968) and frequently addressed in program descriptions. Cultural socialization refers to the development of norms - understanding of what it means to be an evaluator within particular contexts (Louis, 1985). Here it refers to values and attitudes about how evaluation is practiced. These are not as explicitly stated in program descriptions but are part of the individual mentoring students receive from faculty.
That socialization includes the development of a sense of ethics particular to the discipline, as well as specific skill sets, is common. For example, the medical motto “do no harm” is the basis of physicians’ ethical practice. Physicians who adopt the motto into their value system and apply it in using their skills as physicians have been socialized into the profession (at least for that value). Similarly, the *AEA Guiding Principles for Evaluators* (2004) outlines suggestions for the ethical conduct of evaluators and is often part of the professional development experiences for evaluators.

But Antony (2002) argued that socialization into a profession requires only an awareness of the professional values, ethics, and attitudes of the profession in general. Antony argued against the traditional definition of socialization where students must change their norms, values, and ethics. He points out that: (1) socialization does not require a student to replace their own values, but rather to develop an awareness of the values of the discipline; (2) the assumption of linear development (learning values progresses through stages) is incorrect; and (3) that the linear concept fails to account for gender or perceptual differences and changes in role expectations over time (beyond the degree program for example). In Antony’s view, students develop a dynamic approach to socialization into their profession that changes depending upon career goals and experiences. For example, he references discussion from the “Re-Envisioning the PhD Conference” held in Seattle in 2000. This was an initiative sponsored by the Pew Charitable Trusts as part of the discussion around the question, "How can we re-envision the Ph.D. to meet the needs of the society of the 21st Century?" (Re-Envisioning the Ph.D. About Us section, 2000, para. 4). The consensus at the conference, Antony reported, is that doctoral programs need to extend the content knowledge and practical experiences currently provided in doctoral education and to address a greater diversity in students and the careers they choose (other than research institutions). This requires a socialization approach that recognizes the greater diversity in students and their careers and in the ways in which students might be socialized.

Literature on the socialization of graduate students is not restricted to any one discipline. The 1993 national initiative on Preparing Future Faculty (PFF) is an excellent example of the cross-discipline interest in the socialization and professionalization of
graduate students. PFF was sponsored by the PEW Charitable Trusts in association with the American Association for Higher Education and the Council of Graduate Schools (Austin, 2002; Gaff, Pruitt-Logan, & Weibl, 2000). The program focused on students seeking academic careers in higher education, but its structure (mentoring, foundation knowledge combined with practice, and the explicit socialization into faculty careers at different institutions) mirrored the literature on the socialization of students.

Boyle and Boice (1998) published similar results from their survey of exemplary departments at a large research university. The departments they investigated were rated by the National Research Council as average to extremely effective in training doctoral students. Boyle and Boice identified three best practices: “collegiality, mentoring, and structure [that] help incoming graduate students adjust more readily to the culture and academic demands of the graduate department” (p. 93).

Master’s students also benefit from socialization activities in their programs. In a national study of student experiences in professional master’s programs, Conrad, Duren, and Haworth (1998) interviewed students, faculty, staff, and alumni from 47 programs in 11 fields. Their results are similar to PFF and to Boyle and Boice in that students reported three categories of beneficial experiences: meaningful learning, professional development, and leadership. However, they expand upon the ideas by including greater information on the instructional strategies used at the program-level, such as required collaborative learning activities. In other words, practice-based socialization opportunities were built into the program rather than as an add-on, as often occurred with PFF, or as a series of orientations, seminars, and advisor meetings as reported by Boyle and Boice.

Two of the three categories Conrad et al. (1998) listed are of interest here—meaningful learning experiences and professional development experiences. Meaningful learning experiences have two critical components, (1) participation in a community of learners, and (2) engagement in a critical dialogue. A community of learners fosters collegial and collaborative relationships with faculty and fellow students. They defined critical dialogue as an interactive dialogue or shared conversation that engaged students with the material, encouraged reflection, and developed connections
with faculty. Both were highly rated by students as being an important part of their experience.

In the professional development category, Conrad et al. (1998) describe doing-centered learning as "students actively and consciously practiced their profession—or craft—through engaging in hands-on teaching and learning experiences . . . both on and off-campus" (p. 69). The off-campus professional-residencies or internships were considered the most helpful. In these internships, students applied and tested classroom knowledge in real-world settings. This doing-centered learning was reflected in a second professional development aspect, integrative learning. In integrative learning, an effort is made to tie theory with practice, using experiential learning strategies: "faculty modeling; hands-on learning activities such as role-plays, case studies, simulations, field trips, and artistic performances; and bringing non-university workplace professionals into classes" (p. 70).

The successes shared in this last example highlight the important of hands-on learning for performance-based skills and socialization of students. Experiential learning strategies reveal instructional tools programs use to develop student abilities in the doing of evaluator tasks. Concurrently, socialization of students in this study refers to what goals programs have for students in terms of knowledge, skills, and abilities. Socialization in this study was used to refer both to Sarbin and Allen's (1968) role socialization (the knowledge, skills, and abilities required for evaluation practice or for an academic career) and to Louis’ (1985) view of cultural socialization (the development of norms with the understanding of expected values for practicing evaluators).

Experiential Learning or Practical Training

“Education must be conceived as a continuing reconstruction of experience; that the process and the goal of education are one and the same thing.” [italics added]

“To prepare him [the student] for the future life means to give him command of himself . . . train him that he will have the full and ready use of all his capacities; that his eye and ear and hand may be tools ready to command, that his judgment may be capable of grasping the conditions under which it has to work . . . ”
“The only true education comes through the stimulation of the child's [the student’s] powers by the demands of the social situations in which he finds himself. Through these demands he is stimulated to act . . . to emerge from his original narrowness of action and feeling . . . . Through the responses which others make to his own activities he comes to know what these mean in social terms. The value which they have is reflected back into them.” (Dewey, 1897, para.’s 40, 6 & 2).

“'Experiential education’ refers to learning activities that engage the learner directly in the phenomenon being studied” (National Society for Internships and Experiential Education [NSIEE], 1986, p.1). This instructional strategy is based on the premise that students learn more by being involved in the process. This is the point that John Dewey makes in his 1897 work, My Pedagogic Creed. In the first paragraph quoted above he states that the process of learning mirrors the goal. In evaluation that would imply that if the goal is to be a practitioner or faculty member, than the learner must practice the skills to be used. In the second paragraph, Dewey emphasizes that it is the student who must be in control as it is through his own capacities that he will later act. To continue the analogy, one cannot be an effective evaluator until one has actual experience and skills in one’s own hands. The last paragraph quoted focuses upon the relevance of the experience. Here Dewey would say that simulated experiences do not reflect reality and it is the reaction to reality, and reflection upon that reaction, which comprise true education.

Dewey’s work continues in Rogers’ (1969) presentation of two types of learning. One was based in an academic or cognitive sense; the kind of learning that is often imposed and requires memorization (which Dewey railed against). Rogers gave an example of learning that cannot be achieved except by rote, Rogers refers to sight words such as the, that, is, and it; such vocabulary words have to be memorized because these words do not have any concrete examples that can be manipulated in a physical sense.

The second kind of learning was more personal and reflects Dewey’s perspective, in that it related directly to something that the learner wants to know or be able to do. Rogers termed this kind of learning ‘experiential’ and it has three characteristics. One, the learner has to be actively involved in, and in control of, the process (in other words, responsible for his own learning). Second, learning is based on
direct, hands-on experience with real-world problems (rather than artificial simulations or role-playing). Finally, self-evaluation is required. The learner has to take a metacognitive perspective and reflect upon his experiences to determine what is truly experienced and understood. This kind of personally integrated and practical or hands-on training is widely accepted as the only valid training in many applied professions and can be seen in the clinical training programs for nurses and physicians, for example.

More recently, Merrill (2002) built his model by incorporating experiential and adult learning theory into the student-instruction focus of Gagné (1965; 1977). Merrill’s four phases of instruction are: (a) activation of prior experience, (b) demonstration of skills, (c) application of skills, and (d) integration of these skills into real-world activities (Merrill, 2002). This study used Merrill’s concept that instructional design (as revealed by the use of experiential learning activities) provides a window into the curricular intent of programs. That is, experiential learning strategies would be used (perhaps extensively) where programs have integrated the hands-on learning valued by Rogers (1969).

Evaluation Literature on Experiential Learning

In experiential learning, instructional methods that facilitate hands-on skill development are considered most appropriate. This emphasis on practice-based learning is critical to evaluator skill development. Several of the instructional strategies listed in Trevisan’s (2004) review of the literature on practical training in evaluation were rooted in experiential learning. His review included evaluation literature and the literature of other profession-oriented disciplines such as medicine, engineering, education, which supported the use of adult and experiential learning theories in the teaching of evaluation. The need for trained evaluators with content matter expertise, interdisciplinary perspectives, and practical experience is clear. But Weeks (1982) argued that experiential learning should be a preferred teaching method in evaluation because of the potential for harm that exists with incompetent practitioners.

In his study, Weeks (1982) chose to use Coleman’s (1976) model of experiential learning. This begins with an action by the student. The effects are observed, and the student must be able to understand the effects in that instance. “The final step is to
understand the general principle under which the particular instance falls, and to be able to apply the principle in new situations.” (p. 22). Weeks used Colman’s model of experiential learning to examine two evaluation-training experiences – a long-term and a short-term supervised internship and two comparison groups comprised of students who did not participate in an internship experience and students who were just entering the long-term program. To be effective, Coleman stated that the student experience required (1) student authority – the ability to make decisions and see their effects, (2) variety – the opportunity to engage in a number of activities, and (3) centrality – an active participant rather than a peripheral observer.

Weeks (1982) found these requirements applicable to evaluation since most evaluation projects involve a variety of decisions over a period of time as well as the practice of several skill sets. The extended time and involvement actually allow the generalization of principles, which is the final stage of the model. This was an important finding in the study. Students from the Year-Long Internship performed significantly better than did the other three groups. The Short-Term Internship group (only one quarter in duration) was not significantly different from the No-Internship and Incoming Student groups.

Trevisan (2002) stated that most methods for incorporating a practical component into Evaluation training use either simulations, short-term evaluation projects within a class, and the (somewhat longer) field or practicum experiences. This is consistent with Weeks’ (1982) early research on combining short and long-term training experiences. Trevisan (2002) indicated that the major emphasis in practice is geared more toward the long-term and involved evaluation models, stating that “The aforementioned short-term approaches do not prepare students for long-term evaluation work, and as a consequence, do not provide students with exposure to the kinds of evaluation projects they will likely face as professional evaluators” (p.84).

In a review of the literature on teaching of evaluation Trevisan (2004) lamented the dearth of research-based articles on instructional practices in evaluation. He summarized the articles by outlining three major issues: (1) faculty supervision of students, (2) need for resources, and (3) pedagogy and student achievement. In his summary of the 18 peer-reviewed articles published over a 38-year period, he found
that many did not make connections between theory presented in classroom contexts and the field experiences of students. Yet the importance of reflecting upon theory and its applications to the field is a critical component in both experiential and service-learning. Again, one of the issues frequently raised was the difficulty of providing both a reflective component and the faculty or mentor guidance to help students interpret their learning.

**Examples of experiential learning.** *Simulations, Role-Play, and Case Studies* are in-class exercises or homework projects that require students to work on a prepared activity. These can be group, class, or individual activities and may involve some amount of preliminary work (either lecture or homework projects for example). The work may involve working with actual data or results, or with material created specifically for the activity. *Single Course Projects, Focused Practicum and Broad Practicum Experiences* are examples of real-life experiences where students are involved in actual evaluations. What makes all of these activities experiential is that they provide the characteristics Dewey (1897) and Rogers (1969) said were important: students had to be involved in the learning process, they needed hands-on experiences, and some kind of reflective component was critical.

The impact of reflection on learning is illustrated by Stufflebeam and Wingate’s (2005) use of reflection as a learning and assessment tool. Reflecting on what was learned compared to what was already known was the method used in developing their self-assessment instrument for their three-week summer course. Their instrument was piloted with experienced faculty to determine scale end-points and wording. The instrument was then given to workshop participants as a pre-post-test instrument that asked them to reflect upon the amount of knowledge they felt they had in each of eight areas before and after the workshop experience. The results indicated that the reflective activity prior to and after the workshop allowed participants to gauge what they had learned during the workshop. Rather than a simply assessment or feedback tool reflection becomes another way to learn.

Examples of the different kinds of experiential learning that have been used, and shared in the evaluation literature, are described below.
Simulations. Simulations are a description or scenario provided by the instructor along with a set of student instructions. An example of a simulation activity (seen in introductory evaluation courses) is the creation of an evaluation plan (Trevisan, 2002). Evaluation plans typically include an outline of the evaluation process from development of questions, selection and identification of data collection and analysis methodology, design of instruments, timeline for products, and a suggested budget or cost-analysis. Evaluation plans may also include elements of other evaluation tasks – such as evaluator-client communication protocols and identifying who will receive and have access to final reports. As an experiential learning activity, introductory or beginning students in evaluation work through an evaluation plan, or any one of its components, as a simulated activity for an evaluation that may or may not occur.

Nassif and Khalil’s (2006) described using a simulation activity to help students understand concepts or skills necessary to evaluation design. In their paper, they described their attempt at debunking student perceptions that understanding the concepts of validity and reliability as they related to evaluation was too difficult. Their evaluation students were new to scientific methodology and evidenced low self-confidence in their ability to learn statistical material, particularly as it applied to an evaluation design. To simulate the concepts while using something students would be more comfortable with, students were asked to describe how they would make and evaluate baked pies. This activity enabled a discussion of validity and reliability in ways that made sense to the students. Since the idea of making and judging (or evaluating) pies was part of their reality they learned the concepts the instructors intended. Relating what they already knew to new information was exactly what Dewey (1897) said would happen if experiential learning was used as an instructional activity.

Role-Play. Role-Play is a less structured simulation in which students have a greater degree of input in how they respond to a given description or scenario. Students may work through different scenarios of either a fictional or actual experience—such as a first meeting with an evaluation client. Role-Play is typically a group activity where discussion of student perspectives and choices is

Alkin and Christie (2002) described a well-constructed series of role-playing activities threaded throughout the course that met specific course goals. “Role play is
used primarily to teach students how to initiate and focus an evaluation, as well as to develop the skills necessary to present evaluation proposals and findings to stakeholders.” (p. 214). The class made extensive use of models to ensure students understand expectations (several activities were modeled after television programs, which were demonstrated in class). The level of student involvement and complexity of the role-play increased over the course of the semester. This was designed to provide students a period of practice, to ask questions as an evaluator, to minimize anxiety over the activity, and to give students a perspective that evaluation can be an enjoyable activity. In this example, the reflective discussions with the instructor and classmates—a debriefing the role-play experience—comprised the third characteristic of experiential learning as described by Rogers (1969).

**Case studies.** Case studies are instructional activities that make use of an ambiguous scenario. The instructor introduces new material related to the sample case, or redefines the perspective students should take (such as applying a different evaluation approach or changing the methodology). Students typically work in groups, though they may have individual tasks, which may be assigned as homework prior to the class meeting. The key is that cases should not have a *one right answer*. In fact, it is the multiple possibilities from a real-world scenario that make a case useful (Patton & Patrizi, 2005). The goal is to have students work in depth with the material and to practice what they have learned. Instructors model questions and interactions. Feedback is an important component, whether it comes from the activities, is given during a class presentation or role-playing, or in a written format. Instructors may build up to using one case as a summative activity, or use smaller cases more frequently throughout the class. Case studies are an excellent model of experiential learning as they involve students working through content, incorporate some aspect of hands-on practice, and always involve the reflective component.

Patton (2005) discussed another approach to using cases. He combined three cases in order to have students compare individual differences, evaluator roles and purposes, complex relationships and institutional arrangements, controversies and politics as well as various parts in a larger picture. Students also used the larger, combined cases to work on report-writing skills, stakeholder mapping, and issues
related to ethics, standards, and evaluator-competencies. In one respect they used the material to think about the project, and in another, to think about their own learning in relation to the project. Both are characteristics of experiential learning.

Another way to look at the utility of cases in the teaching of evaluation is in terms of how cases can help evaluators avoid delusional thinking. According to Bare (2005), delusional thinking in evaluations is: (1) **Anchoring**—where a first impression continues to dominate subsequent thinking. Using a case approach forces the consideration of other, outside views. (2) **Competitor neglect**—too quick to implement an idea, in this trap the evaluator fails to consider what others may be likely to do. (3) **Exaggerating abilities and control**—an attribution bias where people tend to take credit for positive events and attribute external causes for negative events. By structuring the facts and questions into a case, students, and evaluators, have an opportunity to not only consider alternate views, but to practice the desired responses.

**Single course projects.** These are student projects that form a large portion of the instruction for a given course. Assignments may be for teams or individuals, towards specific evaluation tasks or methodologies (developing an instrument, analyzing data) and may incorporate external interactions (such as site visits or negotiations with a possible client). Projects result in a student generated product presented for instructor feedback and may over several courses result in a portfolio of examples for evaluator skills. A typical example of a final course project is an evaluation plan for a potential client.

Oliver et al. (2008) detailed how they accomplished a single course project during a compressed (5-week) introductory evaluation course. One half of the course was spent learning foundation material, often within the context of the program to be evaluated, while the practical component was completed in the second half and concluded with a final report at week 5. Because this course involved a short timeline, the students, in concert with their client (a center on campus) determined what aspect of an evaluation they would be able to complete. In this case, the students developed an assessment instrument that the center could use in future evaluations. While this example was extreme, many courses offered in a single term also have a limited
timeframe. Instructor flexibility combined with cooperative learning (involving students as a team rather than individuals) made this single course project work.

Kelly et al. (2008) used the same concept for advance evaluation students but substituted a grant proposal for the evaluation plan. “The intent of the course was for students to identify and select a social problem of interest to them, design an initiative to address the issue, and develop and submit an authentic grant proposal to fund the initiative” (Kelly et al., 2008, p. 549). Kelly’s instructional strategy is a good example of using experiential learning precepts, where a course long project makes use of information students already possess, ties the new skills to the students’ area of interest, and reflect upon appropriate strategies (or in this case, grant proposals).

**Focused practicum experiences.** Based in a real context, a practicum or field experience consists of a period of supervised, practical, hands-on training over a period of time that is possible only after completion of core methodology courses. In this definition, the applied experience is beyond the time expectations possible in one course. The timeframe is usually at least one semester but sometimes more or less. The extent of student involvement in the external setting varies, sometimes being limited to specific tasks and sometimes to multiple tasks with greater responsibility. In this study, tasks assigned target specific evaluator skills: such as instrument design, particular data collection methodology and/or analysis, or report writing. Some practica are paid internship experiences, some are not (i.e., the student may earn a wage in addition to college credit) but all technically require the supervision of a faculty member and a field representative.

Three articles provide examples of what instructors do to facilitate supervised learning projects where students go out on their own or under the guidance of a mentor to do an evaluation project or professional development activity. Each author acknowledged the need to provide real-world experiences for their students, and addressed the constraints of meeting diverse needs in too short a time and with limited opportunity for one-on-one interactions.

Brown’s (1985) model for supervising evaluation practicum and intern students built upon what was known about the developmental stages of counseling psychologists during their internship experience. For Brown, “the framework views students as
growing professionals in a field that demands interpersonal skills as well as technical skills” (p. 166). Instructor understanding of how the stages affected what students learned, how they reacted, and how their behavior might influence evaluations was an important part of mentoring in the internship. Three stages were recognized that were relevant to evaluation students and their supervisors, (1) *Naivety and Rigidity* – where the students initially were unaware of their lack of skill and depended upon external sources for information and guidance; (2) *Disequilibrium*—realizations that simple solutions were not possible and that the student lacked the skill set without some kind of effort or change; (3) *Assimilation and Integration*—where problem solving was used to generate creative solutions. Brown gave several examples of student responses in various situations and suggested how supervisors might steer students through stages that are ineffective or dysfunctional. For example, beginning students need a *facilitative* supervisor, someone who is available for consultations in an atmosphere of trust. Some supervisors may be *catalytic*, moving students toward an active rather than passive approach by setting goals or even changing evaluator roles. Both of these can help students move through stages one and two. For students having trouble with stages two and three, a *conceptual* intervention (referring to current experiences compared to different theories).

Gredler and Johnson’s (2001) article on the one-semester directed evaluation experience is another example of sharing of lessons learned with other evaluation instructors. Students were offered evaluation experiences that were guided by the faculty member and their comments on the experience were collected. In order to avoid the logistical issues of matching classroom and practica appointments, these mentoring experiences were not tied to course enrollment, but were guided by the faculty member’s available client-base. Students reported positive aspects in mentoring with faculty, that the *real-world* experience broadened knowledge in unanticipated ways (scheduling & time issues for example), and that they gained confidence as evaluators. Gredler and Johnson did not discuss developmental stages, but the student comments on mentoring and expanded learning appeared to reflect the developmental process Brown described in 1985. As an example of the facilitative role mentors play, one student experiencing difficulty in scheduling visits found that faculty mentor availability
and guidance reduced her anxiety. For other students the experience highlighted “the disconnect that can sometimes occur between theory and real-world reactions” (Gredler & Johnson, 2001, p. 100) and provide opportunities for mentoring on conceptualizing those disconnects.

In a slightly different example of a focused practicum experience, Hurley, Renger, and Brunk (2005) structured a three-semester experience where the first (Spring) semester consisted of regular course material on theory and methods. A second, unofficial, semester (Summer) was required to establish the placement and begin a needs analysis prior to doing the project in the Fall. The third semester (Fall) consisted of a short-term placement to do an evaluation project, which was billed as a professional development experience. Instructional issues occurred because of different instructors for the two graded courses, no instructor-student interaction for the summer, minimal distance-based instruction during the Spring to explain the Fall commitments, differing expectations (and assumptions) for instructor and student, and issues related to establishing credibility and buy-in in a placement situation. In this example, Brown’s (1985) developmental stages were again evident and suggested that course design (and student behavior) might benefit from a review of his results.

**Broad practicum experiences.** As with the focused practicum experience, these experiences are based in a real context and feature supervision, guidance, or mentoring from faculty or evaluation professionals. Students engage in practical, hands-on training over a period of time that is possible only after completion of core methods courses. In this definition, the applied experience is beyond the scope possible in one course or one semester. The extent of student involvement in the external setting is much greater than in the focused practicum. Students with broad practicum experiences are more intimately involved in the entire evaluation process: from first meetings with clients, to question and methodology development, data collection and analysis, and to writing the final report. Many internship experiences are paid, some are not (i.e., the student may earn a wage in addition to college credit or may earn no college credit) but all technically require the supervision of a faculty member and/or a field representative.
Levin-Rosalis and Rosenstein (2003) summarized their article with the comment that “the complexity of evaluation work requires a complex way of teaching” (p. 257). This is the premise behind long-term field experiences, or internships. In this study, they are defined as broad practica because the expectation for student involvement is much broader and may include leadership roles. Evaluation work is complex in that it involves several layers of activity, different skills, and the implementation of a timeline in order to frame questions, write the proposal, collect and analyze data, compile reports and provide communication throughout the process. As Weeks noted in 1982, few evaluation projects take place within the scope of one semester and those that do seldom allow a student to participate in all aspects of the evaluation. Yet students still need guidance and time for reflection during these intensive experiences.

In their article, Levin-Rosalis and Rosenstein (2003) began by structuring the activities they wanted the experience to encompass: theoretical and methodological skills, conceptualization of the practical, and personal and practical skills. They then structured learning activities to support each area. For example, instructional strategies included traditional lecture format, skill practice using simulations or role-playing, student-instructor reflective activities, and products such as field diaries, reports, and portfolios. The year-long experience also made extensive use of mentoring “as a tool for helping students adapt to a new role and acquire new skills in non-threatening surroundings” (p. 255). But the authors also included socialization characteristics that Anthony (2002) and Conrad et al. (1998) had espoused – professional development through discussion of values, ethics, and attitudes expected of evaluators within a context of meaningful (real-world) experiences.

Summary

Enhancement of teaching and access to training for evaluation in order to meet external and internal accountability needs has long been a part of the evaluation community dialogue. Indeed, it could be said that these were some of the same issues that helped to create the American Evaluation Association in 1985. Its conferences, workshops, and opportunities to network have been critical to the growth of evaluation.
Surveys of degree programs were published in 1986, 1994, 2006, and 2010 as interest in evaluation as a field and as a public endeavor increased. Altschuld (1981; 1991; 1995) and Stufflebeam (2001), among others, have published longitudinally on evaluation curricula and program training. Scriven (2001) articulated a need for evaluation skill sets across disciplines and the growing need for evaluation outside the field of education and social science. King et al. (2001), Stevahn et al. (2005), and Ghere et al. (2006) published their work on a taxonomy of evaluator competencies in terms of how training might be structured. The AEA Guiding Principles for Evaluators was updated in 2004 while the JCSEE approved its third edition of the Program Evaluation Standards in 2011. IBSTPI (2006) published their version of evaluator competencies for organizational evaluators. Others continue to explore ideas and strategies to enhance the teaching of evaluation in either university-based programs or in training opportunities offered elsewhere.

The ascent of higher education into the 21st century fostered a flurry of initiatives to review, assess, and redesign graduate education—Preparing Future Faculty (1994), Re-Envisioning the Ph.D. (2000), and the Carnegie Initiative on the Doctorate (2003). Each of these cross-discipline, national discussions asked faculty and programs to explore what skill sets, knowledge, and attitudes students needed to not only successfully practice their discipline, but to make contributions to society and to grow the knowledge-base of their fields. It is that concept – what do students need and how can we get them there, that frames this study. Its focus is a conceptual framework that views the teaching of evaluation in terms of the socialization of students, specifically through the use of experiential learning. In that, it is not far different from Dewey's My Pedagogic Creed on student learning just over 100 years ago (1897), when he too asked what did we need to do to help get students where they needed to be.
CHAPTER 3: METHODOLOGY

The purpose of this study was to understand how evaluation degree programs utilized evaluation competencies and program structure (defined as coursework and student experiences) to prepare doctoral students for careers in evaluation. The overarching research question in this study, *what are the goals* (the expected evaluator competencies) *behind a program’s structure of coursework and student experiences*, lies at the heart of the recent movement in accreditation of institutions and their academic programs. Before institutions can develop Quality Improvement Plans that address “the enhancement of educational quality”, programs must define measurable goals for students earning degrees (Southern Association of Colleges and Schools Commission on Colleges (SACS), 2011). This interest in relating program structure to the needs of students and the discipline was a critical component in the Carnegie Initiative on the Doctorate (CID). Part of the Professional and Graduate Education unit in the Carnegie Foundation for the Advancement of Teaching, the “CID was a five year action and research project [begun in 2003] that worked with doctoral-granting departments committed to restructuring their programs to better prepare graduates” (Carnegie Foundation, 2011).

These broader questions about academic programs are mirrored in questions raised by Engle et al (2006) in their article on the directory results, and from the discussion with the authors at the 2002 AEA Conference. Questions also were raised at the 2006 American Evaluation Association Conference session, “How Best to Improve University-Based Evaluation Programs.” A point was made during the discussion that not enough was known about current programs; more information was needed on what evaluation degree programs in the U.S. were doing, and how they were doing it. Discussion on evaluator competencies and the session title made it clear that there was interest not simply in what programs were doing, but why programs were structured as they were. This study addresses those concerns. As shown in Table 2, I identified four research questions and data sources to answer them.
<table>
<thead>
<tr>
<th>Research Question</th>
<th>Data Sources</th>
<th>Description</th>
</tr>
</thead>
</table>
| 1. What are the evaluation competencies, in terms of knowledge, skills, and values, proposed in the American Evaluation Association (AEA) Guiding Principles for Evaluators for Evaluators (AEA, 2004), the Program Evaluation Standards 3e (JSCEE, 1994), and the International Board of Standards for Training, Performance, and Instruction Evaluator Competencies (IBSTPI, 2006)? | American Evaluation Association’s Guiding Principles (AEA, 2004)              | • Guidelines for professionals  
• Information for the general public  
• A tool for the professionalization and socialization of new evaluators                                                                                                                                 |
|                                                                                  | The Program Evaluation Standards, 2e JCSEE (2011)                            | • Product of joint committee of professional associations  
• Lists standards for conducting evaluations of programs  
• Promotes involvement of program stakeholders (evaluation users)  
• A quality enhancement tool for new and current evaluators                                                                                                                                       |
|                                                                                  | IBSTPI Evaluator Competencies International Board of Standards for Training, Performance, and Instruction (2006) | • Professional competencies developed by professionals in training and performance  
• Knowledge, Skills, and Attitudes for evaluations in organizational settings  
• Training tool for evaluators                                                                                                                                                                       |
| 2. How are these competencies expressed in evaluation degree programs?           | Program Artifacts                                                            | • Web material describing programs  
• Evaluation-specific course syllabi                                                                                                                                                                          |
| 3. How are graduate evaluation degree programs structured to enable students to develop these competencies?                             | Program Informant Interviews                                                    | • Confirmed web-based materials  
• Provided background information and philosophy of program  
• Provided member-checking of program profiles (summaries)  
• Shared thoughts on best practices and future directions                                                                                                                                              |
| 4. What practices support the development of student evaluation knowledge, skills, and values?                                     | Prominent Evaluation Theorists Interviews                                      | • Provided background on early evaluation programs  
• Shared philosophy for program structure (coursework and student experiences)  
• Shared thoughts on best practices and future directions                                                                                                                                              |
The first question in this study was to engage in a literature review to answer the question “What are the evaluation competencies, in terms of knowledge, skills, and values, proposed in the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004), the Program Evaluation Standards, 2e (JSCEE, 1994) and the International Board of Standards for Training, Performance, and Instruction Evaluator Competencies (IBSTPI, 2006)?” These sources are briefly described in Table 2; additional information is provided later in this chapter.

Data sources for the other three research questions “How are evaluation competencies expressed in evaluation degree programs?”, “How are graduate evaluation degree programs structured to enable students to develop these competencies?” and “What practices support the development of student evaluation knowledge, skills, and values?” are also summarized in Table 2. Other elements examined in this review of doctoral evaluation programs included doctoral student professional development and socialization into the field. Data sources for these questions included program artifacts and interviews with program-informants and prominent theorists.

**Program artifacts.** Information collected for the six participating programs included web-based information describing each institution, program mission and goals, degree and course requirements, and student experiences. In addition evaluation-specific course syllabi were collected for each of the six programs.

**Program-Informants.** Leaders in academic programs who were knowledgeable about evaluation theory and practice served as a primary data source in this study. These leaders had a familiarity with and an understanding of the structure and function of doctoral degree evaluation programs. This meant that at a minimum they had experience as evaluators, taught or had taught courses in evaluation, and had mentored doctoral students. In addition, all informants (prominent theorists and program-informants) had extensive publications on evaluations topics. All but one was a full professor. All had served or were serving as directors of programs and/or directors of evaluation centers.

**Prominent evaluation theorists.** Questions for prominent theorists comprised two areas. One area was a request for recommendations of programs (and specific
faculty contacts) who might make appropriate study participants. The second asked about their perspectives on what evaluation programs were, and what they should be. The questions were open-ended and designed to encourage discussion along whatever lines the theorist considered most important. The interview started with an overview of the Engle et al. (2006) study and the questions it had raised. One question asked how did programs train students as evaluators – what were their instructional strategies? The other question asked about evaluator standards, goals, or competencies. Whenever possible, prompting questions were used to elicit theorist perspectives and views on learning outcomes, experiential learning, approaches to training practitioners and academics, and lessons learned.

In summary, this chapter describes the way programs were selected to participate and how the descriptions of those programs were analyzed in order to develop a classification schema to describe evaluation degree programs. This chapter describes data collection in terms of the variety of sources used - program websites, interviews with program-informants (faculty), evaluation course specific syllabi, and interviews with evaluation theorists. This chapter also describes instrumentation and development of the coding rubrics used in the analysis of collected data, and to develop the classification schema. The chapter ends with a discussion of strategies for quality in this study – and addresses issues such as pattern matching, consent, and limitations.

Selection of Programs

Current and previous evaluation directories, various articles on evaluation competencies and recent conversations in at AEA conferences have indicated an interest in knowing more about existing evaluation degree programs in the United States. Of particular interest in discussions on this topic at AEA Conferences (2002-2008) were those programs that had been consistently active, or that might be considered strong programs by students or experts in the field. The 2002 directory of evaluation programs by Engle et al. (2006) provided a potential sampling frame of 22 active evaluation doctoral degree programs in the U.S. (Three programs were no longer functioning and one provided only a master’s degree.) However, the interest in this
study was to describe what strong evaluation degree programs were doing, in some degree of detail. A blanket survey was felt to be too superficial a method for describing what programs did in terms of student professional development in evaluation.

Recommendations from Prominent Evaluation Theorists

I used a combination of snowball sampling and purposeful criterion sampling strategies to identify a sampling frame of university-based doctoral evaluation degree programs in the United States. I started the snowball process by recruiting four prominent theorists in the field of evaluation to interview. I defined prominent theorists as those who had: (1) been closely associated with a program of long-term duration (20 years or more); (2) been recognized as making major contributions to the field of evaluation; and (3) evidenced an interest in teaching and learning in evaluation through publications and discussions on the topic. These criteria, evidence of strong research and applied skills in evaluation in combination with an interest in student academic and career preparation, support the purpose of the study – to understand how evaluation degree programs prepare doctoral students for careers in evaluation. Five theorists met these criteria; however one was approached for discussions on this topic prior to study development and is not included here. The remaining four were approached and asked if interested in participating, all consented to participate.

The purpose in interviewing prominent theorists was twofold: first, to gather recommendations of programs to include in this study and second, to collect additional perspectives on critical components for training and development of doctoral students in evaluation degree programs. More about the interviews for this second purpose is detailed in the data collection section below. The four prominent theorists recommended ten evaluation degree programs for my sampling frame. Three of the ten programs had not participated in the 2002 directory (one was apparently omitted) but all were evaluated against the criteria for selection.

Criteria for Selection of Evaluation Degree Programs

I used a purposeful criterion sampling strategy to select programs for possible participation in this study. The intent was to gather information from programs that
provided a doctoral-level degree focused on evaluation and had demonstrated long term viability as a program. My assumption was that programs with a focus on evaluation at the doctoral level that had been in existence for an extended period would have more fully developed program structures and would hopefully yield more useful data on multiple career options in evaluation.

Each of the ten recommended programs were researched using web-based material on three criteria that would determine selection for participation in this study: (1) the program currently offered a doctoral degree, (2) an evaluation emphasis was available in the program, and (3) the program had demonstrated long term viability (15 years or more), preferably in association with a prominent theorist.

**First criterion for program selection.** *The program currently offered a doctoral degree for students interested in evaluation.* This criterion was chosen in order to keep the focus of the study on doctoral student preparation. Doctoral students are exposed to a greater depth of content and receive more one-on-one training than is possible for the shorter master’s degree programs. Doctoral students may choose to do research in areas related to or within evaluation, go into other academic careers, or pursue applied careers as evaluators. Opportunities for these options were investigated. However, the discipline area of the degree program was not a variable in choosing participants as long as evaluation was an emphasis in the degree. Information on Master's degrees, specializations, or post-doctoral opportunities, if offered, was not collected as these degrees did not offer the same time commitment. One recommended program did not offer a doctoral degree in evaluation and was not included in the final sample.

**Second criterion for program selection.** *An evaluation emphasis was available within the program.* This criterion used the definition for an evaluation program from the 1992, 2002, and 2008 directories of evaluation training programs (Altschuld et al., 1994; Engle & Altschuld, 2006; LaVelle & Donaldson, 2010). The definition states that programs that emphasize evaluation are those that offer “multiple courses, seminars, practica, offerings, and so on designed to teach what respondents considered to be evaluation principles and concepts” (Engle & Altschuld, 2006, p. 354). This definition required a minimum of two evaluation-specific courses for the degree (practically 2-8 credit hours). This is a small number of hours to dedicate to evaluation.
as a specific topic when it is considered against the fact that most graduate degrees require an average of 60 credit hours. Note that the definition refers specifically to evaluation principles and concepts. Methodology courses were not included in this count. Two of the programs did not offer two or more evaluation specific courses as a requirement for the degree and were not included in the final sample.

Third criterion for program selection. The program must evidence a level of long-term viability. This was determined largely by the age of the program as documented on program websites and in previous evaluation directories. Because this study investigated program structure – the coursework and student experiences required for the doctoral degree in evaluation – programs that were relatively new were not as credible in terms of the viability of their program structure. In contrast, while it was possible that some established programs might be undergoing changes in their programs (two were in that process), it was anticipated these programs would continue a large part of their historical activities and philosophies. Of the ten programs, two were relatively new at the time this study started, with five years or less as an evaluation degree program. In addition, one of the two did not require two or more evaluation courses for the degree and neither was included in the final sample. See Table 3.

<table>
<thead>
<tr>
<th>Programs</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programs in 2002 directory considered active in 2006.....................</td>
<td>22</td>
</tr>
<tr>
<td>Programs recommended as participants in this study by prominent evaluation theorists, interviewed 2008-09 ...........</td>
<td>10</td>
</tr>
<tr>
<td>Programs meeting selection criteria, 2009 ................................</td>
<td>6</td>
</tr>
<tr>
<td>Programs agreeing to participate, 2009 ..................................</td>
<td>6</td>
</tr>
</tbody>
</table>

As can be seen in Table 3, six of the ten recommended programs met the three criteria for selection and were selected to participate. Contacts for program-informants were solicited from prominent evaluation theorists during interviews. All six potential program contacts, or informants, were emailed a description of the study along with a request to participate. All six contacted programs agreed to participate. However, one
informant was unable to complete the final step in the data collection process (member-checking of program profiles). Data for that program is included based on published web documents, the informant interview, and submitted evaluation-specific course syllabi. Additional detail on this process and the programs selected is discussed in Phase II of the data collection process that follows.

Data Collection Process

The data collection process involved four phases. The first phase was begun in interviews with prominent theorists in the field of evaluation and their recommendations for programs to participate in this study. Selection of programs was done using purposeful criteria described previously. The second phase consisted of the collection of information describing the six selected programs. This included multiple sources of information: the collection of program artifacts (information describing the program goals / student activities and evaluation-specific course syllabi), collation of descriptive information for institutional characteristics of the institution and the program, and the editing of interview scripts for program-informants (based on analysis of artifacts).

The third phase was comprised of interviews with program-informants. In the fourth phase a program profile was developed and shared with each program informant. Program profiles were summaries of data pulled from program informant interviews and program artifacts. The profiles described the program mission and goals, listed the evaluation-specific courses required and/or available for the degree, and outlined any unique student experiences or program activities. Member-checking or validation of content of the program profile was a critical aspect of this phase. A second interview or email correspondence provided the opportunity to amend profiles to program informant satisfaction. See Table 4 for a brief overview, greater detail on each phase is provided below.
Table 4. Phases of data collection

<table>
<thead>
<tr>
<th>Phase</th>
<th>Data Collection Activity</th>
</tr>
</thead>
</table>
| Phase I: Interviews with Four Evaluation Theorists | A. Selection process  
  a. Leaders in field  
  b. Closely associated with education in evaluation  
  c. Published in literature on teaching / learning in evaluation  
 B. Interview process  
  a. Solicited participation  
  b. Obtained consent  
  c. Conducted interview protocol  
 C. Conducted analysis of interview transcripts for:  
  a. Status of evaluation degree programs in U.S.  
  b. Development of rubric elements  
  c. Recommendations of programs to include in study  
    i. Generated list for Phase II.  
    ii. Applied selection criteria to determine final sample set for Phase II. |
| Phase II: Information Describing Six Selected Evaluation Programs | A. Artifacts for selected programs  
  a. Web-based material describing:  
    i. Institution  
    ii. Department  
    iii. Program  
  b. Evaluation-specific course syllabi  
 B. Institutional characteristics  
  a. Carnegie Classification of Institutions of Higher Education  
  b. National Center for Education Statistics  
 C. Developed program informant interview protocol and modified questions in script as required after artifact analysis |
| Phase III: Program Informant Interviews | A. Selection process  
  a. Explained referral process  
  b. Briefly shared who made referral  
 B. Interview process  
  a. Contacted program-informants  
  b. Discussed selection process  
  c. Obtained consent  
  d. Recorded interviews |
| Phase IV: Program Profiles for Classification Schema | A. Program profiles developed  
 B. Member-checking of profile information by program-informants  
  a. Emailed copy of profile to program informant  
  b. Follow-up interview or email correspondence  
 C. Amended profiles with additional program information |
Phase I: Interviews with Four Evaluation Theorists

In 2004, George Walker, Senior Scholar for the Carnegie Initiative on the Doctorate, stressed that the purpose of doctoral training was to create "stewards of the disciplines" (Golde & Walker, 2004). These stewards were to be young researchers who would carry-on the traditions of making new contributions to the field, recognizing foundational concepts, and transforming theory and practice with teaching and learning. Boyer had identified a similar importance for graduates of doctoral programs 14 years earlier in “Scholarship Reconsidered: Priorities of the Professoriate” (Boyer, 1990). Both theorists advocated fostering specific skills doctoral degree students needed to be functioning professionals in their field. The range of skills needed by students in evaluation highlights the importance of the conversation; what do students need to learn, and how can they learn it best? The addition of interviews with prominent theorists in evaluation in this study provides data on this topic from individuals with the experienced perspective Walker and Boyer identified as critical to any field.

Selection process. As described previously, four theorists were purposively selected who had (1) been closely associated with a program of long-term duration (20 years or more); (2) been recognized as making major contributions to the field of evaluation; and (3) evidenced an interest in teaching and learning in evaluation through publications and discussions on the topic. These individuals were listed in the both 1994 and 2002 evaluation directories, and were published widely in the literature on teaching of evaluation. All four agreed to participate in the study. A fifth theorist had also met criteria but was heavily involved in the initial design and development of the study, so was not interviewed. However, private conversations with this individual were instrumental in the design of the research questions and the study questions.

Interview protocol. From February 2008 to November 2009, four theorists who met the sampling criteria were approached via email and in two cases, personally, and asked to participate in this study for a one-hour interview. All four agreed. Interviews with three theorists were conducted face-to-face with an audio record using the previously developed scripts. One theorist approached preferred to use email correspondence in lieu a telephone or face-to-face interview. In this case, questions were sent in a word document, responses were added, and the document returned. All
audio records were transcribed into Word documents. Follow-up questions using email finalized the interview content for that respondent. Copies of the email templates and interview scripts are in Appendix A and B.

Each interview had two sections. The simplest was a request for recommendations of evaluation degree programs, and the best individual to contact for that program, which might be useful to the purposes of the study. Each theorist was asked if they would recommend any doctoral degree programs in evaluation that they felt were exemplary and would fit the purpose of this study. This information was used to develop the sampling frame for the study. Theorists provided the names of ten programs and faculty contacts knowledgeable for each. Of the ten programs recommended, six met selection criteria for inclusion in the study.

The second section of the interview featured open-ended questions that made each interview unique to the perspectives and interests of each theorist. Theorists were asked to discuss their views on how evaluation programs should be structured to meet the current demands of students and the needs of clients. This included a discussion of desirable learning outcomes for evaluation practitioners as compared to evaluation researchers, the role of evaluator competencies and disciplinary areas, and whether these might differ by degree level or type of evaluation activity. Another topic addressed the role of experiential learning for students of evaluation and asked what theorists considered to be the challenges and advantages of using experiential learning in evaluation programs. In this area, each theorist emphasized the importance of meeting the individual needs of students for an iterative, growth-oriented mentor experience. Theorists were asked if they had any “lessons learned” or ideas based on their history in mentoring evaluation students and running programs or centers. Finally, they were asked if they had thoughts about the future of evaluation programs that they considered critical to student experiences in evaluation.

Consent. Audio recordings and transcripts for each interview were created using Word documents. Each audio record (or email response) included a description of the study and collected consent to participate. Although a standardized interview protocol was used with each interviewee, follow-up questions developed during the interview were used to ascertain additional information. Appendix D contains a sample
of the interview script. In order to assure confidentiality, information reported in this study was reported using random code names rather than personal identifiers.

**Phase II: Information Describing Six Selected Evaluation Programs**

The second phase of this study comprised the bulk of artifact collection. These were existing materials that described the institution, the department, and the program in terms of institutional characteristics, department and program goals and resources, and student degree requirements. Information was collected on program and course descriptions and student activities. Evaluation-specific course syllabi were collected to provide greater insight into the learning outcomes, learning and assessment strategies used in the program. Particular interest was paid to material that described the use of experiential learning strategies or addressed the professionalization and socialization of students into the field of evaluation.

**Program artifacts – website material.** During the selection process, program websites were reviewed briefly to determine if the program met the criteria for inclusion in the study. After programs were selected as possible participants in the study, information downloaded from websites included all evaluation faculty contact and website information as well as program catalogs, student handbooks and college bulletins, available course materials, and any other documents providing program and/or evaluation specific course information. This included forms and timelines or other instructional aides that were intended for current or future doctoral students in the program.

To prepare for determination of participation and save program-informants time, the scripts for the program informant interviews (collected as Phase III) were modified with questions on website locations, availability of evaluation-specific course syllabi, and accuracy of content in program information sources. The information was confirmed in Phase III and updated as required. This material was saved in HTML, PDF, and Word formats. All material was redacted for reporting purposes, that is, no institutional or individual identifiers were reported. Instead each program was identified using an abbreviated random code.
**Program informant participation and consent.** Recommended faculty informants for six selected programs were approached for participation in the study using an email message (Appendix A). Contact information used was posted online at department websites. The email message introduced the investigator, the purpose of the study, and provided an explanation for how the program was chosen (referred and met selection criteria). For programs interested in participating, a reply option (email or telephone) was provided for setting up an interview at a later date. The reply provided an opportunity for the program informant to state a preferred date and time for the interview (within a 10-day time frame).

All respondents replied to the initial contact email and agreed to participate. Consent to make audio recordings of the interview session(s) and to create transcriptions was obtained at the time of the interview. Due to unique time constraints, one program respondent used email correspondence instead of a telephone interview. A copy of the email template and the telephone script, as well reminder posts, is located in Appendix A. A generic script for interviews is provided in Appendix B.

All telephone interviews were recorded and transcribed into Word to be stored electronically. Email documents were also stored electronically in PDF or Word format. Electronic documents will be destroyed at a later date according to university guidelines. Information from program-informants reported in this study did not include personal or institutional identifiers but instead used assignment of random abbreviations or code names as identifiers.

**Institutional characteristics.** Institutional data was collected for each program from the Carnegie Foundation for the Advancement of Teaching and the College Navigator at the Institute of Education Sciences National Center for Education Statistics. The descriptors for academic institutions presented here are at a high level and traditionally used - such as public versus privately funded or comprehensive doctoral versus single doctoral degree institutions. Traditional descriptors for programs were accessed from the Carnegie Foundation (size and types of degrees awarded) and the National Center for Education Statistics (location, source of funding, and cost of tuition).
Evaluation-specific course syllabi. All evaluation-specific course syllabi that were publicly available were downloaded by the investigator from program website materials. During the program informant interview, a list of evaluation-specific courses in the program compiled from website data was shared with program-informants for confirmation. This list included syllabi that had been downloaded and a request for access to other syllabi. One program informant put participation on hold dependent upon permission from program faculty before gathering and sharing syllabi, which was granted. At the time requests for syllabi were made, and as part of the consent process, informants were assured that all materials collected, including course syllabi, would remain confidential and reported only in the aggregate.

Coding of syllabi (use of evaluation competencies, required texts, instructional and assessment strategies) was stored in an EXCEL format that facilitated comparisons between programs. All syllabi identification information was kept confidential. Non-descriptive abbreviations or code names were used and where possible, information was aggregated for reporting purposes. All files were saved as Word, HTML, or PDF formats and secured by the researcher to be destroyed at a later date according to university guidelines.

Phase III: Interviews with Program-Informants

As with THE prominent theorists, all but one program informant were senior faculty with years of experience as leaders in the field of evaluation, all were highly published, actively teaching and mentoring graduate students, and in addition to leadership roles in the evaluation program, were all leaders or had been leaders in their departments or centers. These people were uniquely able to discuss what their programs did to foster student learning and professional development and to discuss how the program had changed over time, and what changes it anticipated in the future. Their thoughts and perspectives contributed to building an understanding of how evaluation degree programs were structured.

Interview protocol. From June 2008 to December 2009, six program-informants were contacted via email and asked to participate in this study for a one-hour telephone interview with a second follow-up interview to be set up at a later date. The
follow-up interview was scheduled after program profiles were created (described in Phase IV below). These summaries of the program were shared via email. The follow-up interview provided an opportunity for program-informants to respond. Telephone interviews were recorded and transcribed to a Word format. All six agreed to participate, although one did seek clearance from other faculty in the program before participating. One was not able to do the interview by telephone, but corresponded using email.

**Interview Questions.** The initial interview was partially scripted in order to confirm the accuracy, and if needed, location, of web-based artifacts that could describe the program. Questions for program-informants were very similar to questions for prominent theorists in that one area addressed straightforward topics. In this case it was confirming material and location of program information and evaluation course-specific syllabi at the institutional website, and asking questions about the history of the program.

The second area repeated the open-ended questions used with the prominent theorists: about evaluation competencies, instructional strategies, learning outcomes, approaches to training practitioners and academics, and lessons learned. The open-ended question format allowed informants to discuss elements of graduate student professional development they considered critical or unique to their program – these were often experiences that were not formally addressed in coursework, but were part of mentoring experiences in applied settings.

Questions for the follow-up interviews were driven by informant feedback on the program profiles and any questions they raised. This process of member-checking for the final summary document was an important continuation of the data collection process, as each profile was modified by program informant feedback. This helped to ensure that the classification scheme would be based on accurate information. Copies of these instruments are available in Appendix A and B.

**Program informant participation and consent.** As discussed previously, six programs met selection criteria and contacts for each program were approached via email with a request to participate. Using the approved Institutional Review Board format, the email messages introduced the investigator, the purpose of the study, and
provided an explanation for how the program was chosen for inclusion. The templates for initial, follow-up, reminder, and thank-you email messages are in Appendix A.

All program-informants agreed to participate. Consent to make audio recordings of the interview session(s) and to create transcriptions was obtained at the time of the interview. One program respondent used email correspondence instead of a telephone interview. A generic script for interviews is provided in Appendix B.

All telephone interviews were recorded and transcribed into Word to be stored electronically. Email documents were also stored electronically in PDF or Word format. Electronic documents will be destroyed at a later date according to university guideline. Information from program-informants reported in this study did not include personal or institutional identifiers but instead used assignment of random abbreviations or code names as identifiers.

**Phase IV: Development of Program Profiles**

This final phase of data collection required an analysis of data. The program profile, or summary description for each program, was used as a source of information in identifying program practices. See Table 5 below. Because profiles represented a summary of data unique to each program, it was critical that the information be accurate. Phase IV provided a venue for member-checking or validating the data that was used in summary form as part of the research question results.

Each program in this study was described in a summary document called a "Program Profile" that used the following identified areas: Program Identifiers (name, contacts, location, degrees offered and disciplinary area), Mission Statement, Evaluation-Specific Courses (required and elective), and Unique Student Experiences. This material was summarized from program artifacts and interviews with program-informants. A copy of the template for email about the program profile, and a template of a profile, is provided in Appendix A.

**Validation of program profiles.** An important aspect to this study was the validation of the program profile by program-informants. This was done using follow-up interviews. A copy of the interview script is provided in Appendix C. Program-Informants were sent a draft of their program profile via email. If changes were noted, a request
was made to schedule a short, 15-20 minute interview. In one case this was conducted using email. Data for the program included interviews with program-informants, submitted syllabi material, and information obtained from outside sources (institutional characteristics data). All program-informants were able to validate their program profiles.

As mentioned previously, the second or follow-up conversation with program-informants provided an opportunity to amend or correct interpretations of the data as well as to add additional insights to the program profiles. Additional input from program-informants as they commented on their profiles became part of the data describing each program. The additional information also provided an opportunity to discuss insights into perceived best practices for the teaching of evaluation. The program profile components that were shared with program-informants are outlined in Table 5 below.

Table 5. Program Profile Components

<table>
<thead>
<tr>
<th>Section</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program Identifiers</strong></td>
<td>• Institution Name</td>
</tr>
<tr>
<td></td>
<td>• Program Location</td>
</tr>
<tr>
<td></td>
<td>• Program Title</td>
</tr>
<tr>
<td></td>
<td>• Degrees Offered</td>
</tr>
<tr>
<td></td>
<td>• Name and titles for Program Informant</td>
</tr>
<tr>
<td></td>
<td>• Program Size = Number of Evaluation-Specific Courses Offered</td>
</tr>
<tr>
<td></td>
<td>o Small – 2 to 3</td>
</tr>
<tr>
<td></td>
<td>o Medium – 4 to 6</td>
</tr>
<tr>
<td></td>
<td>o Large – 7 or more</td>
</tr>
<tr>
<td><strong>Program Mission Statement</strong></td>
<td>Direct quotes from website</td>
</tr>
<tr>
<td><strong>Degree Requirements Specific to Evaluation</strong></td>
<td>Name and number of units or credits for core (required) and elective courses in evaluation</td>
</tr>
<tr>
<td><strong>Unique Program Characteristics</strong></td>
<td>Student learning experiences designed to enhance skill development and socialization into the field</td>
</tr>
</tbody>
</table>

Telephone interviews and/or email correspondence with program directors focused mainly on director’s perspectives of lessons learned in running a successful
program, meeting student needs from a perspective of hands-on or practice-based learning, ideas for enhancing the program, and the role the program plays in the continuum of evaluation degree programs. Topics included (1) how the program was developed, (2) what program strategies were most important and why, (3) how the program focus come about, and (4) what lessons were learned from this program.

Data Analysis

The unit of analysis in this study was the doctoral evaluation degree program. That is, the data analysis, discussed in detail below, resulted in six descriptive products—one for each program. Analysis included both modification of predetermined coding rubrics and development of new codes until a description of the codes was finalized and applied to each of the programs. Coding is an iterative process and each program was coded multiple times as definitions, variables, and values were modified to better fit the data. This process of explanation-building helped to identify the themes and patterns that led to the development of the classification schema. Developing the classification schema comprised a secondary data analysis process, where the programs were analyzed with respect to each other. Both processes are detailed here.

Coding Rubric Development

Rubrics provide a systematic procedure for making decisions. This study used the concept of axial coding - initial or predetermined codes for variable definitions and values along categories of interest (Strauss & Corbin, 1998). These were developed using focused coding. Focused codes are developed a priori based on suggestions from the literature or from the variables in the research questions. Focused codes jumpstart the coding process by providing a platform of defined variables and values that are relevant to the research questions. Table 6 illustrates the initial codes used to organize information from the various sources of data: program artifacts, program-informants, and evaluation syllabi. These codes addressed research questions 2, 3, and 4. Note that portions of this data were anticipated to be quantitative, such as the percentage of required evaluation-specific courses to the total number of courses. These involved
<table>
<thead>
<tr>
<th>Source</th>
<th>Axial Codes</th>
<th>Indicators</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program Artifacts</strong>&lt;br&gt;Research Questions 2, 3, and 4</td>
<td>A. Program mission&lt;br&gt;B. Degree requirements&lt;br&gt;C. Program size</td>
<td>A. Mission statement&lt;br&gt;B. Evaluation core, elective, and non-eval courses&lt;br&gt;C. Number of evaluation courses offered</td>
<td>A. Clarity:&lt;br&gt;0 = No [indistinct, vague, unclear]&lt;br&gt;1 = Implied [implicit, tacit, inherent]&lt;br&gt;2 = Yes [explicit, overt, clear-cut]&lt;br&gt;B. Percentage of each (credits or course) (0-100%)&lt;br&gt;C. Size&lt;br&gt;○ Small – 2 to 3&lt;br&gt;○ Medium – 4 to 6&lt;br&gt;○ Large – 7 or more</td>
</tr>
<tr>
<td><strong>Program Informant Interviews</strong>&lt;br&gt;Research Questions 2, 3, and 4</td>
<td>A. Evaluation Competencies&lt;br&gt; i. AEA Guidelines for Evaluators (AEA, 2004)&lt;br&gt; ii. Program Eval. Standards, 3e (JSCEE, 1994)&lt;br&gt; iii. Evaluator Competencies (IBSTPI, 2006)&lt;br&gt;B. Career Foci</td>
<td>A. &amp; B. Number of competencies by type in program&lt;br&gt;B. Toolkit skills, Professional skills, Academic skills</td>
<td>A. &amp; B. Name and number of units or credits for core (required) and elective courses in evaluation and Percentage of each to total number of required competencies (0 – 100%)&lt;br&gt;B. Count of types of Career Foci offered in the program</td>
</tr>
<tr>
<td><strong>Evaluation-Specific Course Syllabi</strong>&lt;br&gt;Research Questions 2, 3, and 4</td>
<td>A. Instructor Information&lt;br&gt;B. Course Information&lt;br&gt;C. Instructional Strategies</td>
<td>A. Contact Inform.&lt;br&gt;B. Descriptions, pre-requisites, learning outcomes&lt;br&gt;C. Type of Experiential Learning</td>
<td>A. Present: Yes or No&lt;br&gt;B. Utility of syllabus&lt;br&gt;0 = No [unable to determine]&lt;br&gt;1 = Incomplete [insufficient information to describe activities/assessment]&lt;br&gt;2 = Poor construction [activities/assessment not clear]&lt;br&gt;3 = Well construct [clear activities / assessment]&lt;br&gt;C. Type of Experiential Learning&lt;br&gt;0 = Unable to determine&lt;br&gt;1 = Lecture&lt;br&gt;2 = Reading/Discussion&lt;br&gt;3 = Writing&lt;br&gt;4 = Simulations&lt;br&gt;5= Role Play&lt;br&gt;6 = Single Course Projects&lt;br&gt;7= Practicum (focused/broad)</td>
</tr>
</tbody>
</table>

*Note: Research question 1 required the identification of existing material defining evaluation competencies.*
counts and percentages. Other portions in this initial rubric were definitions developed from the literature review and previous evaluation directories.

Another useful tool for this study was an open-coding method, via a pattern-matching process. In pattern-matching, as described by Miles and Huberman (1984), variables of potential interest are defined at the beginning of a study based upon study propositions and the literature review (the axial codes outline in Table 6 and defined below) and then compared to collected data. Definitions of codes are adjusted or tweaked until they reflect patterns detected in the data. If needed, new codes are added to the process. These modified codes are discussed more fully in Chapter 4.

Definitions for Eight Axial Codes

Eight axial or beginning codes were identified and examined: Program Mission, Degree Requirements, Program Size, Evaluation Competencies, Career Foci, Instructor and Course Information and Instructional Strategies. For each program in the study, judgments about the extent to which these elements were evident, and in what formats, was conducted. Results were tabulated in a spreadsheet format to allow comparisons between programs. Changes to these axial codes are described in Chapter 4.

1. Program mission. The variable, Program Mission, referred to the stated program outcomes or student learning goals for each evaluation degree offered by the program. Evidence of this variable was found on web pages describing the program, in catalogs, or in program brochures. Because it was important to correctly identify what the program wants students to know and do, the online mission statements were validated, and often modified, through interviews with program directors. Final versions of these mission statements were approved as part of the member-checking process that validated the program profiles shared with program-informants. This helped to ensure that accuracy and currency of the statements related to program mission.

2. Degree requirements. This variable referred to the number and type of evaluation-specific courses required for the degree. It included internships or other student experiences that were required for degree completion. Evaluation-specific requirements were compared to requirements that were not evaluation-specific (content area courses for example).
3. Program size. Program size is often referred to as a function of the number of enrolled students. But this method has not been used traditionally by the evaluation directories of programs. The agreed method to describe program size refers instead to the number of evaluation-specific courses that are offered by the program. A small program offers only two to three courses. (Note that programs which offer only one course are not considered to be evaluation programs for directory purposes.) A medium program offers four to six evaluation-specific courses while a large program offers seven or more courses on evaluation-specific topics.

4. Evaluation competencies. This variable was important in answering research questions 2, 3, and 4. But it was also the information required to answer research question one (evaluation competencies). Three publications were used to define evaluation competencies and to assess this important variable: the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004), the Program Evaluation Standards, 2e (JSCEE, 1994) and the IBSTPI Evaluator Competencies (IBSTPI, 2006; Russ-Eft et al., 2008).

Program informant interviews, program website documents (learning activities or outcomes) and evaluation-specific course syllabi were examined for each program. A checklist was created to track whether or not any of the evaluation competency concepts from the three publications were included into student learning experiences. Evaluation competencies were defined as any student knowledge, skill, or attitudes the program advertised that were identified as a competency required for the practice of evaluation.

Table 7 lists the five major areas in the AEA Guiding Principles: Systematic Inquiry, Competence, Integrity and Honesty, Respect for People, and Responsibilities for General and Public Welfare. It includes an abbreviated list of the 25 components for the five major areas. Table 8 defines the four Program Evaluation Standards, 2e found in the most recent edition (JSCEE, 1994): Utility, Feasibility, Propriety, and Accuracy. This table also provides a list of 30 specific statements associated with each standard. The final source for defining evaluation competencies is detailed in Table 9. There are fourteen IBSTPI competencies described in terms of four domains.
<table>
<thead>
<tr>
<th>Code</th>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systematic Inquiry:</strong> Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.</td>
<td></td>
</tr>
</tbody>
</table>
| S 1-3 | 1. To ensure the accuracy and credibility of the evaluative information they produce, evaluators should adhere to the highest technical standards appropriate to the methods they use.  
2. Evaluators should explore with the client the shortcomings and strengths both of the various evaluation questions and the various approaches that might be used for answering those questions.  
3. Evaluators should communicate their methods and approaches accurately and in sufficient detail to allow others to understand, interpret and critique their work. |
| **Competence:** Evaluators provide competent performance to stakeholders. |
| C 1-4 | 1. Evaluators should possess (or ensure that the evaluation team possesses) the education, abilities, skills and experience appropriate to undertake the tasks proposed in the evaluation.  
2. To ensure recognition, accurate interpretation and respect for diversity, evaluators should ensure that the members of the evaluation team collectively demonstrate cultural competence.  
3. Evaluators should practice within the limits of their professional training and competence, and should decline to conduct evaluations that fall substantially outside those limits.  
4. Evaluators should continually seek to maintain and improve their competencies, to provide the highest level of performance in their evaluations. |
| **Integrity/Honesty:** Evaluators ensure the honesty and integrity of the entire evaluation process. |
| I1-7 | 1. Evaluators should negotiate honestly with clients and relevant stakeholders concerning the costs, tasks to be undertaken, limitations of methodology, scope of results likely to be obtained, and uses of data resulting from a specific evaluation. It is primarily the evaluator's responsibility to initiate discussion and clarification of these matters, not the client's.  
2. Before accepting an evaluation assignment, evaluators should disclose any roles or relationships they have that might pose a conflict of interest (or appearance of a conflict) with their role as an evaluator.  
3. Evaluators should record all changes made in the originally negotiated project plans, and the reasons why the changes were made.  
4. Evaluators should be explicit about their own, their clients', and other stakeholders' interests and values concerning the conduct and outcomes.  
5. Evaluators should not misrepresent their procedures, data or findings. Within reasonable limits, they should attempt to prevent or correct misuse of their work by others.  
6. If evaluators determine that certain procedures or activities are likely to produce misleading evaluative information or conclusions, they have the responsibility to communicate their concerns and the reasons for them.  
7. Evaluators should disclose all sources of financial support for an evaluation, and the source of the request for the evaluation. |
<table>
<thead>
<tr>
<th>Code</th>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R1-6</strong></td>
<td><strong>Respect for People: Evaluators respect the security, dignity and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.</strong></td>
</tr>
<tr>
<td>1.</td>
<td>Evaluators should seek a comprehensive understanding of the important contextual elements of the evaluation.</td>
</tr>
<tr>
<td>2.</td>
<td>Evaluators should abide by current professional ethics, standards, and regulations regarding risks, harms, and burdens that might befall those participating in the evaluation; regarding informed consent for participation in evaluation; and regarding informing participants and clients about the scope and limits of confidentiality.</td>
</tr>
<tr>
<td>3.</td>
<td>Because justified negative or critical conclusions from an evaluation must be explicitly stated, evaluations sometimes produce results that harm client or stakeholder interests. Under this circumstance, evaluators should seek to maximize the benefits and reduce any unnecessary harms that might occur, provided this will not compromise the integrity of the evaluation findings.</td>
</tr>
<tr>
<td>4.</td>
<td>Knowing that evaluations may negatively affect the interests of some stakeholders, evaluators should conduct the evaluation and communicate its results in a way that clearly respects the stakeholders' dignity and self-worth.</td>
</tr>
<tr>
<td>5.</td>
<td>Where feasible, evaluators should attempt to foster social equity in evaluation, so that those who give to the evaluation may benefit in return.</td>
</tr>
<tr>
<td>6.</td>
<td>Evaluators have the responsibility to understand and respect differences among participants, such as differences in their culture, religion, gender, disability, age, sexual orientation and ethnicity, and to account for potential implications of these differences when planning, conducting, analyzing, and reporting evaluations.</td>
</tr>
<tr>
<td><strong>W1-5</strong></td>
<td><strong>Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.</strong></td>
</tr>
<tr>
<td>1.</td>
<td>When planning and reporting evaluations, evaluators should include relevant perspectives and interests of the full range of stakeholders.</td>
</tr>
<tr>
<td>2.</td>
<td>Evaluators should consider not only the immediate operations and outcomes of whatever is being evaluated, but also its broad assumptions, implications and potential side effects.</td>
</tr>
<tr>
<td>3.</td>
<td>Freedom of information is essential in a democracy. Evaluators should allow all relevant stakeholders access to evaluative information in forms that respect people and honor promises of confidentiality.</td>
</tr>
<tr>
<td>4.</td>
<td>Evaluators should maintain a balance between client needs and other needs.</td>
</tr>
<tr>
<td>5.</td>
<td>Evaluators have obligations that encompass the public interest and good. These obligations are especially important when evaluators are supported by publicly-generated funds; but clear threats to the public good should never be ignored in any evaluation.</td>
</tr>
</tbody>
</table>
Table 8. Axial Coding from the *Program Evaluation Standards, 2e* (JSCEE, 1994)

<table>
<thead>
<tr>
<th>Code</th>
<th>Section</th>
<th>Standard</th>
</tr>
</thead>
</table>
| U1-7 | **Utility Standards**  | 1. Stakeholder Identification  
|      | The utility standards are intended to increase  
|      | the extent to which program stakeholders find  
|      | evaluation processes and products valuable  
|      | in meeting their needs.  |
| F1-3 | **Feasibility Standards** | 1. Practical Procedures  
|      | The feasibility standards are intended to  
|      | increase evaluation effectiveness and  
|      | efficiency.  |
| P1-8 | **Propriety Standards** | 1. Service Orientation  
|      | The propriety standards support what is  
|      | proper, fair, legal, right and just in  
|      | evaluations.  |
| A1-12| **Accuracy Standards** | 1. Program Documentation  
|      | The accuracy standards are intended to  
|      | increase the dependability and truthfulness of  
|      | evaluation representations, propositions, and  
|      | findings, especially those that support  
|      | interpretations and judgments about quality.  |
|      | 2. Context Analysis  
|      | 3. Described Purposed & Procedures  
|      | 4. Defensible Information Sources  
|      | 5. Valid Information  
|      | 6. Reliable Information  
|      | 7. Systemic Information  
|      | 8. Analysis of Quantitative Information  
|      | 9. Analysis of Qualitative Information  
|      | 10. Justified Conclusions  
|      | 11. Impartial Reporting  
|      | 12. Metaevaluation  |
Table 9. Axial Coding for the IBSTPI Evaluation Competencies (IBSTPI, 2006)

<table>
<thead>
<tr>
<th>Code</th>
<th>Domain</th>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF</td>
<td>Professional Foundations</td>
<td>1. Communications</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Professional credibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Interpersonal skills</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Ethical and legal standards</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Politics of evaluation</td>
</tr>
<tr>
<td>PD</td>
<td>Planning and Designing the Evaluation</td>
<td>6. Evaluation planning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7. Evaluation management planning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8. Data collection strategies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9. Pilot testing</td>
</tr>
<tr>
<td>IE</td>
<td>Implementing the Evaluation Plan</td>
<td>10. Data collection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11. Data analysis and interpretation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12. Dissemination</td>
</tr>
<tr>
<td>ME</td>
<td>Managing the Evaluation</td>
<td>13. Monitoring the evaluation plan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14. Working effectively with personnel and stakeholders</td>
</tr>
</tbody>
</table>

5. Career foci. This variable addressed statements the program made about the career options available for students. It also included the different areas of emphasis available to doctoral degree students within a department or program. Three types of Career Foci were identified: Preparation for Academic Careers, Preparation for Praxis, and Disciplinary Foci or Subfields. This information was found in web pages describing the program, in catalogs, or in program brochures. As with the program mission, interviews with program-informants were used to determine to what extent these foci were utilized by each program.

All programs were housed in departments at research intensive institutions. Therefore, each department required preparation for an academic career as research faculty within the departmental discipline. However, evaluation degrees included a strong applied component or preparation for praxis. This option allowed students to choose an applied career as evaluation practitioners by emphasizing activities in this
area. In some departments, additional areas of emphasis within the discipline (specific topics or subfields) were either optional, or required as part of the students' program of studies (elective or core courses). The extent to which this was an option in a program was defined as the amount of coursework in subfields or related fields that students could choose to include in their program of studies.

6. Instructor information. Syllabi from the evaluation-specific courses were examined to determine if information on contacting the instructor was provided. The intent here was to include an additional measure that would indicate the detail used for the various syllabi.

7. Course information. An aspect of any course syllabus is information describing the nature of the course - how it fits into the curriculum (pre-requisites) and intended student learning outcomes. Evaluation-specific course syllabi were examined to determine the extent that this information was made available.

8. Instructional strategies. This variable refers largely to data collected from the evaluation-specific course syllabi for courses required to complete the doctoral evaluation degree offered by the program. Non-evaluation specific competencies (measurement and analysis courses) tended to be common across programs and these courses were not included. However, all syllabi vary in terms of detail (the extent they document what occurs in a course), and between professors and/or terms. This ambiguity in the available data was potentially problematic—as experiential learning strategies comprised a primary interest in this study. Of particular interest was how these strategies were used by programs to achieve evaluator competencies. Given these possible issues, and to ensure an accurate account of what programs were doing, program-informants were asked during the program profile validation process, or when syllabi were requested, to add comments or caveats that might impact data analysis in this coding rubric.

Note that three traditional instructional strategies (lecture, reading / discussion, and writing) have been included in addition to the experiential learning strategies of interest - the four types of experiential learning strategies suggested by Trevisan (2004) in his survey of research literature on teaching in evaluation. These were elaborated
upon using the more descriptive identification for the types of internships suggested by Weeks (1982).

Instructional strategies or learning activities were either explicitly stated, or could be inferred, from descriptions of student activities and assessments in each program’s evaluation-specific course syllabi. The seven instructional strategies below were identified for each individual course and aggregated overall for each program:

*Lectures* were class sessions that featured instructor or student provided material for a given course topic.

*Reading / Discussion* sessions in a course featured assigned reading on course topics with instructor or student-led discussion on the material. Guidelines for discussion may or may not be included.

*Writing* assignments were another featured instructional activity that is also, at times, a form of student learning assessment. Writing was coded as a strategy when syllabi indicated that the process of writing for a topic was intended to develop performance-based skills.

*Simulations* were defined as an instructional activity (with assignments, directions, and rules for interfacing with other materials) that was modeled after a real experience. Simulations were artificial recreations, often stylized; of a real-life event or activity and included case studies, role-playing, and data analysis.

*Role Play* was defined as a more complex form of simulation; role plays used real programs and incorporated on the spot student participation. Instructors use role play as an attempt to mimic reality, while providing a vehicle for additional instruction, practice, and feedback. Students may have conducted interviews and interacted with program stakeholders as part of developing an evaluation plan, collecting real or fictitious data, analyzing results, and preparing a report. Role play scenarios did not include completion of an actual evaluation.

*Single Course Projects* were framed in the context of a single course. Students worked in teams or individually on a real or simulated evaluation project, or an aspect of an existing project. Instructor guidance and feedback student performance were emphasized. The goal was to have students work in depth with the material and to practice what they had learned in previous courses.
Practicum (short or long term) were evaluation projects that provided real-world evaluation experiences for students, either on their own, or with guidance from an instructor over the course of one academic term, sometimes more. By definition, short-term practicum experiences were limited in terms of depth or breadth of student involvement in a given evaluation project. Long-term practica were real-world experiences that typically encompassed multiple aspects of an evaluation project, if not the full scope.

Identifying Program Practices to Socialize Students

The method to identify common program practices was suggested by Miles and Huberman (1984). A meta-matrix was used to structure operational definitions of the eight variables described above to create a series of “master charts assembling descriptive data…in a standard format”, which would identify patterns between and among the six programs participating in this study (Miles & Huberman, 1984, p. 152). Coding the data in this way involved identifying themes in the matrix that were: (1) descriptive, interpretive, or explanatory; (2) addressed the research questions; and (3) suggested key concepts or strands from the conceptual framework, all of which will be described in Chapter 4.

Study Quality

Construct validity (Yin, 200(3) and evidence of pattern-matching (Miles & Huberman, 1984) are two criteria for judging the qualitative research designs. These, along with the process of consent, are addressed below as they influenced the design of the study. Establishing and maintaining a strong design was needed in order to promote accuracy in data analysis at the program level. Program level data analysis was the first step in development of the classification schema, discussed in Chapter 4.
**Construct Validity**

Construct validity refers to the accuracy of the measures used in a study. Two tactics were used to enhance construct validity: data triangulation by using multiple data sources and methods, and member checking.

The first tactic was *using multiple sources of data and a mixed methodology*. Data was collected from a) documents providing program and course descriptions on websites and in program catalogues, b) evaluation-specific course syllabi, and c) semi-structured interviews with program-informants. Using multiple sources, particular when one source was an interview format with a reliable program informant, made it possible to note where data conclusions were supported by multiple sources. For example, program website documents included quantitative information on the number of required courses for the degree. This information was confirmed by qualitative data from program-informants during interviews (and subsequent member-checking) during the process of collecting evaluation-specific course syllabi.

I developed a spreadsheet and codebook for each of the variables in the data analysis to “allow(s) an external observer...to follow the derivation of any evidence, ranging from initial research questions to ultimate case study conclusions” (Yin, 2003, p.105). The program database established a common set of variables and values and helped to document data across programs. It included definitions for codes (variables and where appropriate, values). Different versions and field notes documented when those definitions changed.

The pattern-matching process used to interpret data resulted in frequent changes in coding as each program was analyzed or when different types of data sources were included. The final coding rubrics where then applied to data collected for each program. As an example, evaluation-specific course syllabi were coded multiple times, documents in the database featured highlighting for sources as well as value or variable categorization.

The second tactic in construct validity is to *validate study results*. This was done through member-checking, which is asking program-informants to review a draft version of material such as the program profiles and provide feedback. Member-checking was critical where programs were in the process of updating website materials. For example,
in some cases, inaccurate information was posted by web developers needing to fill space in a newly designed website—before faculty could provide appropriate content. In other cases, program reorganizations, faculty sabbaticals, retirement, or local crises impacted the availability or accuracy of program data on websites or in evaluation-specific course syllabi.

**Pattern-Matching**

In this study, pattern-matching logic was used as a technique for addressing whether inferences were consistent with the data. Pattern-matching defines variables at the beginning of a study based upon study propositions from the literature review or pilot study results and compares cases that match or contrast with the definitions in the rubric. Initial coding rubric variable definitions and values were changed until they reflected the data collected, regardless of the data source (website materials, program informant interviews, or evaluation-specific course syllabi).

Detailed coding rubrics provided consistent definitions that were applied to all data for each program. The iterative process of coding each program provided an ongoing test for consistency. For example, this study used pattern matching across the 6 cases to develop final coding rubrics. One of the modified variables discussed in Chapter 4, Socialization of Students into Evaluation, was developed in order to incorporate new data on leadership development activities. This changed the variable from a generic description of degree requirements to an elaboration of the purposes in those activities (promotion and development of evaluation skills).

**Consent**

Although the process of consent to participate was presented as part of the data collection process, it is addressed in further detail below as it related to reporting of findings. Significant portions of this study relied upon qualitative data. This type of data requires that findings are shared in a rich narrative format, often as direct quotations. As this study utilized a small number of programs, theorists and program-informants in the study results were potentially identifiable. The consent section below describes how participant anonymity was protected.
The Human Subjects Committee at Florida State University required that all email and telephone correspondence provided subjects an accurate description of the study’s purpose and potential harm or benefits. An opportunity to decline participation, and contact information for supervising faculty and the university’s Institutional Review Board, was included in the materials. This information was shared early in the contact interview, with the program informant repeating a statement of informed consent for the record, according to university guidelines.

All contacted respondents for interviews and selected programs agreed to participate. Security of data and anonymity was discussed with each respondent using the protocol approved by the Institutional Review Board at Florida State University. Institutional Review Board procedures for the use of human subjects (see Appendix D), including obtaining and documenting informed consent (see Appendix B), were followed. A copy of the informed consent letter is available in Appendix E. Interviews were conducted by telephone and by email, whichever the respondent preferred. The material was transcribed with the electronic data kept in a secure location.

This study was designed to allow respondents (senior theorists, program-informants, and syllabi creators) to participate anonymously, and all but one chose to remain anonymous. Senior theorists and program-informants were not individually identified in this study. Qualitative results in the form of quotations were referenced by an arbitrary letter designation (Senior Theorist Q, R, S, or T). Program-Informants constituted the individuals representing each program who were interviewed, and were denoted by the following arbitrary letters (Respondent J, K, L, X, Y, or Z). The six programs were also given an arbitrary letter designation not associated with other respondent letters (A, B, C, D, E, and F). The same program level designations were used for other faculty at a given program who may "speak" through text from syllabi or program materials.

Limitations

The limitations in this study were related to the design that makes the study interesting – it relies upon qualitatively based word descriptions to build a vision of what
a few doctoral degree programs in evaluation are doing to educate and socialize new evaluators. It does not address what all programs are doing, how newer programs are designed, or even the number of programs currently active. The study does not explore student perspectives, or faculty perspectives beyond their syllabi and what was shared in the program informant interviews. But it can still promote a discussion of evaluation competencies and the socialization process for students among those interested in the unique perspectives presented here.

Summary

Evaluators have been active in updating the latest directory of evaluation degree programs with six published records in slightly over a thirty-year period. True to form the last three have endeavored to maintain discrete definitions so that comparisons can be made between decades. We are after all, evaluators at heart and we want good data if we are to make suggestions or recommendation. But this study is a different kind of directory. Rather than surveying programs extant and reporting various discrete counts, this study has attempted to more fully describe how six vital programs function in terms of preparing future evaluators and research faculty in the field. Purposeful sampling and high participation rates resulted in data collected from four prominent theorists with a long history as stewards of the discipline. Included here are the thoughts and perspectives of six strong faculty members and the structure, the course work and student experiences, in the doctoral degree programs they shepherd.

Eleven evaluation faculty devoted time to this project in the form of interviews and mentoring. Other faculty shared their creative work in terms of 47 different evaluation-specific course syllabi. Programs, departments, and institutions have also devoted resources in providing web-based information about programs and faculty (no small thing if one remembers what a chore it was to collect program brochures only to hear that they were already outdated, with faculty having retired, switched focus, or stopped offering that interesting series of courses!) What data we can collect and distribute in terms of institutional characteristics is just as amazing – most of it available literally at your fingertips 24 hours, 7 days a week, 52 weeks a year.
The information is thick with context; the issues raised are each worth a lively discussion. Because of the nature of graduate degree programs, which straddle a period of decades linked by faculty careers and students' programs of study, this study reaches beyond the last three attempts to produce a directory accurate at a single point in time. Instead the participants in this study have shared a bit of the past along with their predictions for the future.
CHAPTER 4: RESULTS

The study describes the program structure of six doctoral evaluation degree programs in the United States. Prominent theorists in the field of evaluation recommended ten programs for review. Six programs met the selection criteria created by the researcher. Informants from all six programs agreed to participate. The focus of the study was to examine how doctoral programs in evaluation provide a degree with an evaluation emphasis, evidence long term viability, and create close associations with prominent contributors to the field, as well as professionalize students by teaching critical evaluator competencies.

This study began by identifying existing sources detailing evaluator competencies to answer its first research question. The identified competencies were then used in an analysis of program artifacts and informant interviews to answer the remaining research questions for each program:

1. What are the evaluation competencies, in terms of knowledge, skills, and values, proposed in the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004), the Program Evaluation Standards, 2e (JSCEE, 1994), and the International Board of Standards for Training, Performance, and Instruction Evaluator Competencies (IBSTPI, 2006)?

2. How are these competencies expressed in evaluation degree programs?

3. How are graduate evaluation degree programs structured to enable students to develop these competencies?

4. What practices support the development of student evaluation knowledge, skills, and values?

Finally, this material was used to develop a response to the 2006 AEA conference think tank session titled “How Best to Improve University-Based Evaluation Programs”. Individuals at the session determined that before improvements could be made to the field, we first needed to know what current evaluation degree programs in the U.S. were doing. What skills did they consider were needed for evaluators, for research in evaluation, and how were programs structured to help students develop
those skills? The response this study offers is the identification of practices used by programs to socialize students as evaluation professionals. The description of practices I developed provides an outline of common practices and concerns each program in this study identified as critical to the development of future evaluators and, future evaluation faculty.

In the first section of this chapter the six programs are described in terms of traditional classification characteristics as an answer to “who were the participants?” This is followed by a section providing notes on coding, followed by a section on the results across the six programs, for the remaining research questions. The third section describes practices programs use to socialize students as evaluation professionals. These practices to socialize students are composed of two dimensions defining four critical elements and framed by two approaches, developed from emergent themes in the data. This representation describes how programs are structured to better understand evaluation degree programs. The chapter ends with a brief summary.

Section I: Describing the Six Programs

Programs in this study were described at two levels. The first is specific description of each program compiled from program web sites and program-informant interviews. In this study, this second description was called a “program profile”. The second is a general description of where the program is housed – what kind of institution, what department, where in the United States is it located, what is the funding source, faculty FTE’s, etc. However, the main goal in this study was not to compare and contrast the six programs per se. Instead, the main goal was to identify the characteristics and components of their programming, to understand how this subset of doctoral degree programs facilitate the socialization of students through the vehicle of evaluator competencies and experiential learning. This was accomplished by using pattern matching and analysis of trends after analyzing all programs in terms of the research questions. These practices, in terms of identified characteristics and components identified, are discussed in Section III.
Program Profiles

Program profiles were summaries of data pulled from program informant interviews and program artifacts. They were important because they succinctly described what programs did to develop evaluation professionals. The profiles listed the program mission and goals, listed the evaluation-specific courses required and/or available for the degree, and outlined any unique student experiences or program activities. A program profile was created for each program, however, to protect confidentiality of informants, data were aggregated as noted in the tables throughout this chapter. The components of the program profiles are found in Chapter 3 (see Table 5 specifically). See Appendix A for an example of the template.

The primary purpose in creating program profiles was member-checking or validation of content of the program profile. Even after reviewing web-based material describing the program and then discussing it with the program-informant interviews, programs could be too complex to understand for an outsider. Member-checking via a second program-informant interview was accomplished for all six of the programs.

Four program profiles did not need to be changed but for two programs, sharing of program profiles (the descriptive summary of each program) as part of the member-checking process resulted in changes based on program-informer comments. The complexity in program activities was evident in answering the research questions in Section II of this chapter.

Traditional Ways of Describing Programs

Quantitative information to describe programs and institutions was collected from program informant interviews, some website documents, and from institutional classification websites: the Carnegie Foundation and the National Center for Education Statistics (NCES). Information collected included institutional descriptors: size (number of undergraduate or graduate students), public or private, setting (city size), average cost of attendance, level of research activity, and type of graduate instructional programs. Program descriptors collected included: year program started, enrollment and graduation rates, number of evaluation faculty, number of required hours for
degree, and the number of course hours for any required cognate or specialty area/certificates. Traditional indicators or characteristics of institutions are frequently used to compare departments and institutions and therefore are provided here.

As can be seen in Table 10, the major difference between the six programs institutions was location (small cities versus large cities) and size (residential-large versus non-residential-large) with an even split on each variable. A large institution has at least 10,000 undergraduate students (Carnegie Foundation, 2010). A review of institutional profiles developed by the Carnegie Foundation in 2010 ranked each institution in this study as a comprehensive doctoral program in research universities with high research activity. This is not shown in the table. Additional profile information shown indicated five of the six programs were offered at public institutions; one program was offered at a private institution which increased its cost as compared to the public institutions.

Table 10. Classification of Institutions by Program, 2010

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Variable</th>
<th>Medium Size</th>
<th>Large Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A  B  C</td>
<td>D  E  F</td>
</tr>
<tr>
<td>Sizea</td>
<td>Large four-year, primarily residential</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large four-year, primarily non-residential</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Exclusively graduate / professional</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Locationb</td>
<td>City: Small</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>City: Large</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Suburb: Large</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding Sourceb</td>
<td>Public Private Non-Profit</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
| Cost of Tuitionb | $10,658-$10,367 $25,694-$21,143       | $12,534-$10,367 $25,800-$21,143 $13,495-$10,367 $20,593-$21,143 $35,368 $10,367-$35,368 $13,495-$21,143 $20,593-

The College Navigator (National Center for Education Statistics, 2010) listed three programs in small cities, two programs in large cities, and one program in a large suburban location. Information on graduate student tuition and fees showed a
difference: from $10-13,000 for the in-state students attending public institutions ($20-26,000 out of state) to $35,000 for students attending a private institution. Additional descriptive information that has been part of the published information for the various evaluation directories (such as number of full-time faculty, number of students graduating) was collected from the above national organizations, program artifacts, and program-informants. But there are problems with using this data to make comparisons between programs as is discussed below.

**Evaluation directory descriptions.** In the evaluation training program directories, programs are described by type of degree offered (doctoral, masters, specialist), academic location (discipline), and program size (number of faculty, number of students, number of graduates, number of evaluation-specific courses). All programs in this study offered doctoral degrees as a condition for selection; other degree information was not considered. Academic location information was collected and while some programs had multiple affiliations, the majority were housed in education departments: four of six (67%) of participating programs were located in education. This percentage is consistent with the major academic location determined in the 1984, 2002 and 2008 directories showing 54%, 62%, and 60% of programs respectively located in education (May et al., 1986; Engle et al., 2006; LaVelle & Donaldson, 2010). The 1992 survey indicated that 85% of programs were located in education, educational psychology, or psychology (Altschuld et al, 1994). This is the same percentage found by LaVelle and Donaldson in their 2008 survey. For this study, other academic locations included psychology (n=1, 17%) and a multiple or interdisciplinary program (n=1, 17%) which suggest that academic location is consistent in the study compared to previous directories.

Table 11 shows the trend in the 1992, 2002, and 2008 directories in terms of distribution by program size. Program size in evaluation has been, since 1992, described in terms of the number of evaluation-specific courses offered, in part because it was difficult to identify the number of students within evaluation as compared with the department (Engle et al., 2006, p.357). More on that issue is provided below. LaVelle and Donaldson (2010) also used the number of evaluation-specific courses as a measure of program size for their internet-based directory in 2008. The percentages of
Table 11. Evaluation Programs by Program Size\textsuperscript{a}, 1992-2008.

<table>
<thead>
<tr>
<th>Program size\textsuperscript{a}</th>
<th>1992 Directory</th>
<th>2002 Directory</th>
<th>2008 Directory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number</td>
<td>49</td>
<td>29</td>
<td>47</td>
</tr>
<tr>
<td>Small</td>
<td>31%</td>
<td>65%</td>
<td>64%</td>
</tr>
<tr>
<td>Medium</td>
<td>27%</td>
<td>31%</td>
<td>30%</td>
</tr>
<tr>
<td>Large</td>
<td>43%</td>
<td>3%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Note: 1984 Directory did not include program size information.
\textsuperscript{a} Where small = 2-3 courses, medium = 4-6 courses, and large = 7 or more courses.

programs of a given size identified in the most recent directories are consistent. However in this study, three programs were identified as large (7 or more courses) and three programs were of medium size (4-6 courses). There were no small (2-3 courses) programs referred by prominent theorists who meet selection criteria.

While the 1986 directory did not collect information in terms of number of evaluation-specific courses offered, it contained survey responses for either the number of faculty participating in the program, or the FTEs dedicated to the program (full-time-equivalents) (May et al., 1986). This mix of definitions made comparisons on program size impossible. Similarly, in the 1992 directory survey responses did not "distinguish between the number of faculty involved, and the FTEs dedicated to evaluation training" (Altschuld et al., 1994, p. 74). The 2002 survey requested but did not publish information on the number of students admitted and graduated, and on faculty FTEs (Engle et al., 2006). The 2008 directory was compiled using web-based material and did not include information on the number of students or faculty (LaVelle & Donaldson, 2010).

Departments are frequently described in terms of the number of enrolled graduate students as an indicator of size. But as mentioned previously, this has not been the case for most of the evaluation directories. The 1984 directory of evaluation training program gave the information on number of students enrolled and graduated from programs - a median number of doctoral students admitted per year as three and a median of ten graduates over the previous five years across the programs (May et al., 1986). An attempt was made to collect this information for this study but the number of
students enrolled and graduated was difficult to separate from data for other degrees, and other specializations, in the department. See Table 12.

Table 12. Study Program Descriptors by Size\textsuperscript{a}.

<table>
<thead>
<tr>
<th>Traditional Size Indicators</th>
<th>Small, N=0</th>
<th>Medium, N=3</th>
<th>Large, N=3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty teaching evaluation-specific courses</td>
<td>N/A</td>
<td>2 - 5</td>
<td>4 - 8</td>
</tr>
<tr>
<td>Students currently enrolled</td>
<td>N/A</td>
<td>8 - 21</td>
<td>17 - 26</td>
</tr>
<tr>
<td>Doctoral degrees earned 2005-2009</td>
<td>N/A</td>
<td>? - 13</td>
<td>4 - ?</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Where small = 2-3 courses, medium = 4-6 courses, and large = 7 or more courses.

Across the six programs in this study, the total number of students currently enrolled ranged from eight to twenty-six. The number of students graduating with an evaluation emphasis between 2005 and 2009 ranged from 4 - 13 but this range also reflects compromised results (it is not certain that graduation information does not include master degree students in at least one program, hence the question marks). The number of faculty in the three programs of medium size (4-6 evaluation-specific courses) ranged between two and five. The number of faculty in three large programs (7 or more evaluation-specific courses) ranged between four and eight. So functionally the only reliable information for program size was the number of evaluation-specific courses offered.

One reason for the difficulty in collecting student enrollment and graduation counts is that this data is not usually tracked and reported at the program or specialization level. Institutions use the Classification of Instructional Programs (CIP) code from the National Center for Educational Statistics, which does indicate college, department, and specialization (NCES, 2010). In practice institutions report enrollment and graduation numbers by CIP codes, usually collapsed by college and/or department. Evaluation programs tend to be specializations or tracks within the degrees offered by departments. CIP codes that refer to this sub-level of information, particularly for
graduate programs, are not currently required material for reports submitted to the Integrated Postsecondary Education Data System (IPEDS) at the National Center for Education Statistics. The result is that I had to rely upon information the departments collect informally. Changes in department structure, program names, and faculty made it difficult to trace information that was not readily available at the specialization level. In an additional caveat, program respondents indicated that some students maintain program status (CIP code) elsewhere, in addition to mentoring and taking coursework in evaluation. This made it very difficult to determine the number of students specializing in doctoral evaluation degree.

In summary, the six programs in this study were fairly similar to each other (size, location in a large research intensive university). Other descriptors were marginally different (funding source, discipline, number of evaluation-specific courses). As will be seen in the next section, programs did not differ greatly in terms of the research questions. This suggests that this subset of programs is at least partially representative of the 40-odd institutions that house doctoral degrees in evaluation. But the programs in this study do differ in one major way. According to previous and the most current directory of evaluation degree programs at least 64% of programs overall are small in size, e.g. having less than 4 evaluation-specific courses. In this study, there were no small programs; the split was even between medium (4-6 evaluation courses) and large (7 or more evaluation courses) sizes. These cases were purposefully selected for their rich history and strong (in terms of longevity) programs. The characteristics and components that describe them must be considered within the context that may be representative of only 30% of current evaluation degree programs.

Notes on Study Results

In order to better understand the results, this section of notes discusses the expansion of codes required during the coding of data and an overview of results for two of the data sources used (program-informant interviews and evaluation-specific course syllabi). Axial coding modifications were done as part of fitting codes to data to ensure valid results. These notes explain qualifications in the way data was collected
and analyzed. The overview of source results also helps to clarify how individual programs were analyzed in terms of defining characteristics and components of programs. These were used to identify patterns in practices programs use to produce future evaluation professionals.

**From Axial to Open-Ended Codes**

As described in Chapter 3, this study began using pre-determined or axial codes (Strauss & Corbin, 1998). These focused codes were developed through a review of the literature and in response to anticipated variables identified through the research questions. Not all of the original or axial focused codes worked as originally defined. As can be seen in below in Table 13, some modification to the codes described previously in Chapter 3 was required. These open-ended codes were developed using the pattern-matching process described by Miles and Huberman (1984). In pattern-matching codes are changed to fit emerging trends in data. With six programs to be analyzed, this meant a frequent series of coding, re-defining, and then re-coding until all codes were determined to be relevant to the data. For example, in answering the research questions, the counts and percentages of courses for degree requirements were not as relevant as a listing of the evaluation-specific courses (see B. under the Codes column, in italics). The percentage of evaluation courses compared to other required courses in each program was dropped in favor of a simple count of the number of evaluation courses offered.

As another example, note Classification Characteristics, D under the Codes column was added. Discussed later in the chapter, these were used to describe program participants. A smaller change was made in the definition for instructional strategies (see G, under the Definitions column) which were changed to a checklist instead of a count. While identifying the seven levels of experiential learning worked well, it was difficult to determine, based on syllabi alone, the extent to which experiential learning was applied across the program (since the unit of analysis is the program level and not at the course level). Most syllabi did not contain sufficient detail to make a definitive judgment. However, explanation-building using key word analysis of syllabi,
<table>
<thead>
<tr>
<th>Data Source</th>
<th>Codes</th>
<th>Indicators</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program Artifacts (including Evaluation-specific Course Syllabi) and Program Informant Interviews</td>
<td>A. Program Mission</td>
<td>A. Mission statement</td>
<td>A. Word analysis</td>
</tr>
<tr>
<td></td>
<td>B. Evaluation-Specific Courses</td>
<td>B. Syllabi, program descriptors</td>
<td>B. Required or elective courses w/ evaluation as main topic</td>
</tr>
<tr>
<td></td>
<td>C. Program Size</td>
<td>C. Syllabi, program descriptors</td>
<td>C. Number of evaluation-specific courses offered</td>
</tr>
<tr>
<td></td>
<td>D. Classification Characteristics</td>
<td>D. Program descriptions, Professional organization data sites</td>
<td>D. Size (number of students, faculty), Degree Type, Location, Funding Source, Cost of Tuition</td>
</tr>
<tr>
<td></td>
<td>F. Career Foci: o Preparation for Academic Careers o Preparation for Praxis o Disciplinary foci or subfields</td>
<td>F. Student course activities and internships</td>
<td>F. Word/category analysis (academic, praxis, subfields)</td>
</tr>
<tr>
<td></td>
<td>G. Instructional Strategies</td>
<td>G. Course descriptions, learning outcomes, assessments, learning activities, interviews</td>
<td>G. Present: Y/N o Lecture o Reading/Discussion o Writing o Simulations o Role Play o Single Course Projects o Focused Practica</td>
</tr>
<tr>
<td></td>
<td>H. Socialization of Students</td>
<td>H. Internships, capstone experiences, development of leadership opportunities in evaluation, interviews</td>
<td>H. Key word analysis</td>
</tr>
</tbody>
</table>

Note: Variables and values in *italics* represent open codes that were not included in the original coding structure.

Program artifacts, and interviews to create a checklist of experiential learning strategies across the entire program made this feasible. But it also revealed that the seventh level of experiential learning, practica (focused or broad), had no examples in course syllabi that fit the definition of a broad practica experiences. The levels were therefore adjusted to include only focused practica experiences.
Finally, socialization of students (see H. under codes in Table 13) is another example of explanation-building to identify new themes and patterns. This new theme came out vaguely in the first interview as a specific goal for internships and other student experiences. This early identification of a code not originally anticipated made it possible to specifically address in later program informant interviews and follow-up questions during member-checking of program profiles.

Course Syllabi

Where possible, a list of evaluation-specific course syllabi was identified for each program from program artifacts, usually web-based material. If located and accessible, copies of syllabi were downloaded. In each interview program-informants readily agreed to share all evaluation-specific course syllabi they could access. However, not all syllabi were available (not every class is taught every semester) and program-informants did not always have access to syllabi for courses they did not teach. An additional caveat in using syllabi as an indicator of what happens in class is that many syllabi are used as simple logistical tools. One syllabus didn’t list any information about required readings other than where to purchase a course packet from a copy center. Information about teaching strategies and tools was more clearly determined from the program informant interviews. One program informant even stated:

The syllabi – it doesn’t reflect the dynamism of the class. It is sort of like read this, do this, but that doesn’t say too much about what happens in the class.— Respondent X

The total number of syllabi for evaluation-specific courses (47 for all six programs) that were either required or offered for the doctoral degree was requested from all programs. It was not possible to obtain a copy of all syllabi, however 76% (N=35) of syllabi for all courses identified as evaluation-specific in each of the six programs were obtained. See Table 14. Two programs made 100% of their syllabi available; one program included an extra syllabus for a different topic in a special topics course. A second program required students to repeat a special topics course at least twice for a total of six required courses. Each program was able to contribute 50% or more of the syllabi for evaluation-specific courses available in their program.
Table 14. Evaluation Course Syllabi by Program

<table>
<thead>
<tr>
<th>Item</th>
<th>Medium Size</th>
<th></th>
<th>Large Size</th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Evaluation-Specific Courses in Program</td>
<td>6(^a)</td>
<td>6</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>9</td>
<td>47</td>
</tr>
<tr>
<td>Evaluation-Specific Courses Syllabi Provided</td>
<td>2</td>
<td>7(^b)</td>
<td>3</td>
<td>14</td>
<td>5</td>
<td>5</td>
<td>35(^b)</td>
</tr>
<tr>
<td>Submitted Syllabi (%)</td>
<td>92%</td>
<td>100%</td>
<td>50%</td>
<td>100%</td>
<td>63%</td>
<td>55%</td>
<td>76%</td>
</tr>
</tbody>
</table>

\(^a\) Two core + 1 special topics course – students required to repeat special topics at least twice.

\(^b\) Two copies of a special topics course were provided, the second was not counted in the total.

Syllabi were also requested from prominent theorists when interviewed but none were able to provide copies (either they did not use syllabi or were no longer teaching). One program respondent and all of the theorists interviewed indicated that their syllabi would not provide information about instructional strategies as these were simply logistic documents identifying instructor contact and meeting time/location information. Syllabi submitted were screened for conformity to the definition of "evaluation-specific" content, rather than content that was methodological, e.g. quantitative methods or diagnostic, e.g. evaluation of special needs students. No syllabi were rejected. However, it was noted that syllabi in two large and one medium-sized programs were more likely to have combined topics – evaluation and something else—for example, mixed methods in evaluation and applied research or skills applicable to other content areas, e.g. consulting theory and practice.

All programs had some degree of flexibility in meeting perceived needs of students through the use of special topics courses. As an example of this and the upcoming revision of the *Program Evaluation Standards, 2e* (JCSEE, 1994), two programs provided special topics courses or activities. In one, the instructor construct[ed] a three week exercise at the end where I am giving them the draft materials on the new standards and the old standards and having them do an analysis of the changes and we are having a class discussion on what does it mean that these changes have taken place.—Respondent X
Interviews

Interviews were conducted with all six program-informants and four prominent theorists in evaluation. Prominent theorists were interviewed as part of the program selection process. Then program-informants were interviewed after a review of each program’s web-based material. In one case, the program-informant most knowledgeable about the program (and specifically recommended) was the previous program director. In all other cases program-informants were also program directors. In two cases, the department and its programs were in the process of reorganization, resulting in a redesign of the program websites. In one case, the program-informant was unaware that a computer technician had used inaccurate material as a placeholder. This error made the initial review of the material with subsequent member-checking during the interview very important for validity of web-based material.

Interviews were also used to supplement understanding of the issues facing evaluation degree programs in the context of department, colleges, and institutions. While each interview included open-ended questions that allowed informants to address issues of their concern, it was helpful to ask program-informants specific questions that further helped to describe programs in terms of goals, learning outcomes, instructional strategies, use of experiential learning techniques and evaluation competencies, and unique student experiences in the program. The interviews resulted in a richer description of the issues in professional development of evaluators and of program structure than was evident in program artifacts or evaluation course syllabi alone.

Section II: Research Question Results

The research questions were developed to guide a study that was largely qualitatively designed. They provided the foci for interviews and data analysis and were part of the conceptual framework for the study – to look at socialization of students into the field of evaluation across the programs using evaluator competencies. As the framework indicated that socialization of students relied primarily on experiential learning, how this instructional approach was being used by all six programs was also
examined. Results, aggregated and non-aggregated across programs, for each question are presented below.

**Research Question 1: What are Evaluation Competencies?**

The first research question in this study asked: What are the evaluation competencies, in terms of knowledge, skills, and values, proposed in the *American Evaluation Association (AEA) Guiding Principles for Evaluators* (AEA, 2004), the *Program Evaluation Standards, 2e* (JSCEE, 1994), and the International Board of Standards for Training, Performance, and Instruction Evaluator Competencies (IBSTPI, 2006)? Essentially this question asks, “What do evaluation students need to know and be able to do?”

Initially this question was simply stated in terms of identifying any evaluator competencies. But a review of the literature suggested three organized sources – the AEA Guiding Principles for Evaluators (2004), the Program Evaluation Standards, 2e (JCSEE, 1994), and the IBSTPI Evaluator Competencies (IBSTPI, 2006). These versions have a degree of standardization; indeed, one is a national set of standards from an accredited program. The process used to develop them (recruiting professionals nationally and internationally from diverse backgrounds and experiences to create and edit them) made these choices credible – and more importantly, likely to be used in evaluation degree programs. While the specific competencies for each source were described in the coding sheets presented in Chapter 3 (see Tables 7, 8, and 9), each is described below. A comparison of the three follows after.

**Guiding principles.** The *AEA Guiding Principles for Evaluators* (2004) are a general list of values in five categories (Systematic Inquiry, Competence, Integrity/Honesty, Respect for People, and Responsibilities for General and Public Welfare) intended to help evaluators make ethical decisions. They provide a useful tool for discussion of evaluation principles and concepts by professionals and the general public. The structure they outline for designing and conducting evaluations can be modified to suit conditions as such, they are not prescriptive.

**Evaluation standards.** The *Program Evaluation Standards, 2e* (JSCEE, 1994) are framed in four categories (Utility, Feasibility, Propriety, and Accuracy) and provide a
checklist for evaluation tasks. Like the *AEA Guiding Principles for Evaluators* (AEA, 2004), the standards are not defined as a measurable listing of knowledge, skills, and abilities that could be agreed upon as specific evaluator competencies. Instead, they are described in general ways that can be interpreted to fit the situation. However, they are descriptive enough that it was possible to identify examples of instructional activities that fostered the skills. The standards particularly promote the involvement of all stakeholders in an evaluation.

**Evaluator competencies.** Although perhaps not as well known as another behavior-based list of competencies in the evaluation literature (King et al., 2001; Stevahn et al., 2005; Ghere et al., 2006), the International Board of Standards for Training, Performance, and Instruction (IBSTPI, 2006) developed its own competencies list to address the needs of evaluators in organized settings. Comprising four domains (Professional Foundations, Planning and Designing the Evaluation, Implementing the Evaluation Plan, and Managing the Evaluation) and 14 standards the IBSTPI competencies have been vetted nationally. Most importantly, they are framed as a training tool, in clearly measurable terms. Eight-four performance standards refer to specific evaluator knowledge, skills and attitudes. But the IBSTPI group is very careful to state that these competencies were developed to meet the needs of organizations using program evaluation. They recommend that evaluators for other venues consider the *AEA Guiding Principles for Evaluators* (2004) and the *Program Evaluation Standards, 2e* (JSCEE, 1994).

It should be noted that not all participants were in agreement on the concept that evaluator competencies were a key part in understanding what programs were doing. The point one theorist made is that the evaluator skill set is wide and individualized and reliance upon an established set of competencies would not guarantee students were exposed to all skills they might need. He went on to point out:

I see competencies used to mean a relative [sic] common (across people) small set of quantitative variables useful in comparing people or change in people. I see such competencies as often oversimplifying the educational process and narrowing the curricula. I think educators should be very reluctant to identify
competency outcomes or standards that students should obtain as a minimum or should pursue as an indication of purpose or success.—Respondent X

In this theorist’s view, an exhaustive list of performance standards would be counter-productive – too narrow to meet the range of skills and content areas evaluators use. This perspective is seen clearly in the research question material below. All program-informants stressed the importance of flexibility in programming in order to meet student needs for evaluator skills sets within the context of their own career goals. The discussion on the practices programs used later in this chapter also indicates a need for flexibility in the choice of skills sets. But the concept of identifying common evaluator competencies made a good starting place to make comparisons across programs, and provided a platform for discussion with program-informants.

Comparing competency sources. A comparison of the three sources was conducted using the 84 performance standards from the IBSTPI Evaluator Competencies (IBSTPI, 2006). Each standard or area in the AEA Guiding Principles for Evaluators (2004) and the Program Evaluation Standards, 2e (JCSEE, 1994) was compared to and matched with the 84 IBSTPI performance standards.

Looked at in detail using the 84 IBSTPI performance standards, 60 or 71%, are matched to material in either the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004) or the Program Evaluation Standards, 2e (JSCEE, 1994). In addition, 21 or 25% of the 84 IBSTPI performance standards match to both the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004) and the Program Evaluation Standards, 2e (JSCEE, 1994) while only 24, or 29%, of IBSTPI performance standards do not match either.

As can be seen in Table 15, when the competency sources are compared and collapsed to the highest level (Principles, Attributes and Domains, respectively) there is an almost complete overlap in evaluator skill sets. Even where gaps exist as illustrated below, all basic skill sets are covered.

For example, the first IBSTPI competency (Communicate effectively in written, oral, and visual form) contained 7 performance standards. The 7th, write clearly and concisely, matched to the third example of the first guiding principle (Systematic Inquiry)
Table 15. Overlap in Evaluator Competency Skill Sets by Source

<table>
<thead>
<tr>
<th>IBSTPI Evaluator Competencies</th>
<th>AEA Guiding Principles</th>
<th>Program Evaluation Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>System. Inquiry</td>
<td>Competence</td>
</tr>
<tr>
<td>Professional Foundations</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Planning and Designing the Evaluation Plan</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Managing the Evaluation</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

* AEA Guiding Principles for Evaluators (AEA, 2004), Program Evaluation Standards 2e, (JCSEE, 1994) and IBSTPI Evaluator Competencies (IBSPTI, 2006)

as well as to the 5th examples in two categories (Utility & Propriety) of the program evaluation standards. Two other IBSTPI performance standards also had three-way matches: the 33rd performance standard for the 6th competency, Identify the evaluation focus and key questions to be answered, and the 50th performance standard for the 8th competency, Construct reliable and valid instruments. However the 2nd performance standard, use active listening skills, from the first IBSTPI competency listed above, did not match to any of the guiding principles or program evaluation standards.

Only three examples from the *Evaluation Association (AEA) Guiding Principles for Evaluators* (AEA, 2004) did not match any of the IBSTPI competencies. These principles refer to stakeholder concepts (reducing harm to stakeholders from negative results, promoting social justice, and a full democratic disclosure of findings to all stakeholders). The majority of guiding principles 25 components can be matched to 35, or 42%, of IBSTPI’s 84 performance standards. All 30 of the *Program Evaluation Standards, 2e* (JSCEE, 1994) matched to one or more of the IBSTPI performance standards, matching 43 or 51%, of IBSTPI’s 84 performance standards.

Only one of the 14 IBSTPI Evaluator competencies had no matches for any of its three performance standards. The 9th competency, Pilot test the data collection instruments and procedures, is one of four competencies in the second domain, Planning and designing the evaluation.
While no other IBSTPI competency was completely unmatched, two others were notable for the number of performance standards that were unmatched. The 3rd competency, Demonstrate effective interpersonal skills, had no matches to four of the five performance standards. For example, Use consulting skills to clarify issues and Monitor and respond to the dynamics of groups and teams had no concrete matches to the guiding principles or the program evaluation standards. The 7th competency, Development a management plan for the evaluation, had nine performance standards, six of which had no matches. For example, Identify internal and external personnel requirements, Determine training needs of personnel, and Allocate personnel and resources to support the plan did not make direct matches. It should be noted again that when rolled up to higher levels, all basic skills were addressed by each source.

**Research Question 2: How are competencies expressed in programs?**

The second research question in this study asked: How are these competencies expressed in evaluation degree programs?

Evaluation competencies as identified by the *AEA Guiding Principles for Evaluators* (AEA, 2004), the *Program Evaluation Standards, 2e* (JSCEE, 1994), and the IBSTPI Evaluator Competencies (IBSTPI, 2006) are expressed directly in web-based material describing the programs, in evaluation-specific course syllabi, and in program-informant interviews.

**Web-based material.** Web site information on each program was accessed and downloaded. Page locations, currency, and general content were confirmed during interviews with program-informants. Web material included: program and faculty information pages, graduate student handbooks, brochures or sample documents, course progression and course descriptions, graduate student policy statements, and job aides (such as program of studies forms). This material contained lists of graduate skill sets, requirements for portfolios and products required for the degree, course descriptions, and program goals and mission statements.

For all programs, web materials addressed the five categories of the *AEA Guiding Principles for Evaluators* (2004): Systematic Inquiry, Competence, Integrity/Honesty, Respect for People, and Responsibilities for General and Public

The following quotes from program web site material are examples of statements describing what students will learn or do which align with evaluator competencies:

Each student’s curriculum will be designed around his or her specific needs and interests, as well as a set of core competencies required for program completion [such as] Social, Political, and Cultural Context of Evaluation (psychology of evaluation, politics of evaluation, “kill the messenger,” stakeholder analysis, diversity and multicultural issues) —Web P

Evaluation studies students build upon a foundation of knowledge in evaluation theory and practice supplemented with coursework in such areas as organizational development, education systems, and conflict resolution. Our graduates leave with a portfolio filled with the tools of the evaluation trade—qualitative and quantitative inquiry methods, communication skills and computer database analysis experience—gleaned not only from time in the classroom but also from internships and collaboration with evaluation professionals in real-world settings—Web M

[Student Handbook requires] maintaining professional standards of conduct in all your activities . . . the major principles are confidentiality of privileged information, not overreaching or overstating your competence and knowledge or that of the discipline, fulfillment of obligations and commitments, reasonable self-control, protecting clients from risk, jeopardy, deception, or embarrassment, and general professional demeanor—Web X

[This] course explores prevalent evaluation theories, systems for categorizing these theories, and the process of theory development in educational evaluation. [Another] course includes assessment methodologies appropriate for evaluation problems, writing evaluation proposals, developing program monitoring procedures, selecting appropriate evaluation design
strategies, coping with ethical considerations in evaluation, framing the decision context, and reporting evaluation results—Web G

Society’s organizations need evaluation to identify and assign priorities to unmet needs; assess progress and identify areas requiring improvement; assess costs and seek ways to make services more efficient and cost-effective, document and assess outcome, provide credible reports to accrediting/oversight bodies—Web P

A common element in most of the evaluations is working with programs to develop a "logic model" or "theory of action." Not only do these theoretical models provide clarity regarding program processes, but also serve as a frame to help guide data collection and analysis—Web G

Graduates leave with a portfolio filled with evidence of their expertise with the tools of the evaluation trade—qualitative and quantitative inquiry methods, communication skills, and computer database analysis experience—Web M

This program prepares students for top leadership positions in evaluation. All Ph.D. students are required to complete a portfolio. The portfolio is expected to represent a cohesive set of experiences that balance the student’s training in their area of specialization—Web X

The Ph.D program prepares you for leadership positions in evaluation in diverse settings. Some careers settings, and job titles are: Educational Evaluation, Cluster, Multi-site and Multi-lever Evaluation, Collaborative, Participatory, and Empowerment Evaluation—Web O

Competence in evaluation importantly rests on methodological competence that is both broad and deep. Because good evaluation practice routinely includes a mix of qualitative and quantitative methods, evaluators need basic competence in multiple social science methodologies. This includes understanding of large-scale experimental approaches to social inquiry, basic skills in survey methods, and firm grounding in interpretive, qualitative approaches to social inquiry. Thoughtful evaluation practice also rests on conceptual knowledge of the field, as represented in the evolution and
diversification of evaluation theory, and experiential familiarity with the field practice of evaluation, as best captured in an actual evaluation project—Web N

A wide variety of quantitative and qualitative research methods are employed, as needed. Depending upon the evaluation questions, data are collected using interviews, case studies, surveys, or document analysis. In cases where program effect needs to be estimated, the group will employ quasi-experimental research techniques—Web G

**Evaluation-specific course syllabi.** All programs used evaluator competencies in evaluation course syllabi. Competencies were identified first through required reading and second through document analysis of course / student learning objectives, student learning and assessment activities, and course topics.

**Required reading.** Evaluation course syllabi show that the *AEA Guiding Principles for Evaluators* (AEA, 2004) was utilized as required reading by 67% (n=4) of the programs in the study. One program not showing the guidelines as required reading on syllabi listed “Disciplinary Guidelines” instead. The other program without a reading requirement on syllabi submitted only a very small number of syllabi and indicated that other course syllabi usually referred solely to room location and meeting times. Of the total number of submitted syllabi, 7 of 36, or 19%, listed the guiding principles as required reading. In contrast, 100% of programs had syllabi that listed the *Program Evaluation Standards, 2e* (JSCEE, 1994) as a required text. The *Program Evaluation Standards, 2e* (JCSEE, 2010) are addressed as required reading for 100% of the programs in this study. The standards are also listed as required reading in 7 course syllabi, or 19% of all submitted syllabi. It should be noted that the only other required reading material that was common across programs were articles from the “American Journal of Evaluation” and “New Direction in Evaluation”.

**Syllabi document analysis.** In addition to required reading, evaluator competencies were also identified in the course objectives, student learning outcomes statements, learning activities, assessment materials, and course topics. All programs had syllabi evidence for evaluator competencies, from each of the three sources: the *AEA Guiding Principles for Evaluators* (AEA, 2004), the *Program Evaluation Standards, 2e* (JSCEE, 1994), and the IBSTPI Evaluator Competencies (IBSTPI, 2006).
As can be seen this summary of the data in Table 16, each program had multiple examples of evaluator competencies in its evaluation course syllabi. The greatest percentages of competencies identified in each program were for the IBSTPI Evaluator Competencies. These were written in terms of short performance standards or behaviors that were easily matched to the dense text of syllabi. In comparison, the more generally worded guiding principles and program evaluation standards were harder to match. The guiding principles were the second greatest percentage, overall and in four of the six programs. However, each program differed in the number of syllabi provided.

Table 16. Evaluator Competencies in Evaluation Syllabi by Source and Program

<table>
<thead>
<tr>
<th>Competencies</th>
<th>Medium Size</th>
<th>Large Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>GP^a</td>
<td>9</td>
<td>23%</td>
</tr>
<tr>
<td>Systematic Inq</td>
<td>2</td>
<td>22%</td>
</tr>
<tr>
<td>Competence</td>
<td>1</td>
<td>11%</td>
</tr>
<tr>
<td>Integrity</td>
<td>2</td>
<td>22%</td>
</tr>
<tr>
<td>Respect</td>
<td>2</td>
<td>22%</td>
</tr>
<tr>
<td>Public Good</td>
<td>2</td>
<td>22%</td>
</tr>
<tr>
<td>PES^b</td>
<td>6</td>
<td>15%</td>
</tr>
<tr>
<td>Utility</td>
<td>2</td>
<td>33%</td>
</tr>
<tr>
<td>Feasibility</td>
<td>1</td>
<td>17%</td>
</tr>
<tr>
<td>Propriety</td>
<td>1</td>
<td>17%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>2</td>
<td>33%</td>
</tr>
<tr>
<td>IBSTPI^c</td>
<td>25</td>
<td>63%</td>
</tr>
<tr>
<td>Prof Found.</td>
<td>9</td>
<td>36%</td>
</tr>
<tr>
<td>Plan / Design</td>
<td>6</td>
<td>24%</td>
</tr>
<tr>
<td>Implement</td>
<td>6</td>
<td>24%</td>
</tr>
<tr>
<td>Manage</td>
<td>4</td>
<td>16%</td>
</tr>
<tr>
<td>Totals</td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>

^a AEA Guiding Principles for Evaluators (AEA, 2004), ^b Program Evaluation Standards 2e, (JCSEE, 1994) and ^c IBSTPI Evaluator Competencies (IBSPTI, 2006)

(2 to 14), as well as the amount of detail provided in each. For example, program A’s two syllabi provided 40 counts of evaluator competencies, nine, or 23% linked to the guiding principles; six, or 15% to the program evaluation standards; and 25, or 63%
linked to the IBSTPI competencies. In contrast, program C’s three syllabi yielded only 28 counts of evaluator competencies, while program B had seven syllabi with a total of 82 competencies. At 14 syllabi, program D had double the number of program B, but had a similar number of identified competencies – 80. Programs E and F each had five syllabi, and a similar number of identified competencies – 29 and 30 respectively. Another way to compare detail in syllabi is to consider that the three medium size programs included 28, 40, and 82 competencies in 3, 2, and 7 course syllabi. Large size programs had similar numbers with 29, 30, and 80 competencies in 5, 5, and 14 course syllabi.

Differences between programs could also be seen in a closer inspection of the principles, attributes, and domains respectively for the guiding principles, the program evaluation standards, and the IBSTPI competencies. As an example, within the IBSTPI evaluator competencies, the domain for Managing the evaluation was the smallest percentage of identified competencies in all six programs. But the medium sized programs had a greatest percentage of identified competencies in Professional foundations while the large programs had more competencies identified in Planning and designing the evaluation.

Below are excerpts of data identified in syllabi. As an example, the 6th IBSTPI competency refers to Evaluation Planning and is part of the second domain, Planning and designing the evaluation. A hit in one syllabus as the second day’s topic, an overview of evaluation and its tasks (Syllabus O1). In another syllabus for a different program, evaluation planning comprises a series of topics slated for the second through sixth weeks of the term (Syllabus N2). Other examples of programs’ syllabi list course outcomes that include the following:

The categories of the Program Evaluation Standards and understand how to apply the Program Evaluation Standards. —Syllabus X4.
Describe the standards and ethical practices of evaluators.—Syllabus X8.
Students will [know] what meta-evaluation is and why it is important.—Syllabus M3.
Determine effective communicating and reporting methods for disseminating evaluation information . . . developing sensitivities to ethical and political issues involved in evaluation—Syllabus J4.

Evaluation practice is constrained by many factors, including its applied nature, its political nature, and the interests of different stakeholders/audiences . . . that ethical issues affect the evaluator’s role . . . . [and] design ways [in the report] to encourage people to use the results.—Syllabus M4

At other times the reference in the syllabus is indirect as when Syllabus M5 refers to a course focus of "analysis of problems and issues in evaluation" and proceeds to outline an intensive discussion of data collection methods.

No syllabi listed interpersonal skills per se as a course goal but specific examples were addressed. For example, when students were asked to identify and understand:

The issue in establishing and maintaining client relationships.—Syllabus X6

Negotiating evaluation contracts or agreements with clients.—Syllabus N2

Demonstrate leadership skills in various roles . . . involve [others to] . . . build consensus, communicate, and resolve conflicts—Syllabus O1

**Program-Informant interviews.** All program-informants spoke of using evaluator competencies as instructional tools for introducing new evaluators to the skill sets and continuing practice in practicum and internship discussions. Informants referred to working with students personally and discussing experiences in the components of doing an evaluation – from developing questions, identifying a framework, instruments, communication skills, logistics of communication & budgeting, to report writing. An example using IBSPTI evaluator competencies occurred when program-informants talked about the need to give students increasing levels of responsibility for running the real-world evaluations via evaluation management planning and monitoring the evaluation plan.

But program-informants in all programs also talked about using competencies for problem-solving, particularly in terms of interacting with clients. These soft skill sets or interpersonal communication skills, often referenced as *other skills*, was frequently mentioned by program respondents as something developed during hands-on
evaluation activities. This is an important competency to note because interpersonal skill development, unlike data collection and analysis, or planning the evaluation, was often embedded into internship activities or listed as a specific topic on a syllabus. Program-informants indicated that feedback and discussion for interpersonal skill development was conducted during debriefing seminars where students met to discuss their work and progress or as part of the faculty-student mentoring process. Below are examples of comments from program-informant. They suggest a range of competencies the program addresses (doing an evaluation) while including competencies in “soft skills”, e.g. communication and client interactions:

Writing clearly and concisely for both academic and non-academic audiences...communication and interpersonal skills which are vitally important in evaluation.—Respondent Y

We actually do an evaluation . . . we socialize students to be evaluators in class [discussions]. We pay a lot of attention to what was it like to meet with the clients, what did we talk about, what were some of the red flags at that conversation, why were they red flags, did you feel like you prepared well enough for that meeting, why would evaluators have that kind of meeting etc. So we actually do an evaluation. We go through it, but we spend most of our time on the communicative and relational dimensions of evaluation—Respondent L

We will talk about it . . . what's happening, and why are you doing this, what problems have you encountered, and what are possible ways you might get around it and so forth and so on—Respondent X

[If you don't do evaluations] you don't interact with people, you don't have to develop communication skills. [Research] is a very different world . . . . [Regarding learning evaluator skills] most of it can't be learnt through coursework. A lot of it is learnt by doing it. Sort of trial and error.—Respondent P

In this research question, evaluator competencies were useful as guides to developing ethical attitudes as well as content knowledge and skills that comprise the socialization process in developing evaluation professionals. Most competencies were easily identified, regardless of their source. They were expressed throughout the programs—in web-based materials describing the program activities, in evaluation-
specific course syllabi as part of course descriptions, outcomes, and activities, and in the faculty-student mentor experience that formed a key aspect of internship or practica experiences. The latter area was critical as it provided a backdrop for discussion and practice of the soft skill sets, e.g. communication, interpersonal skills, and program management. How programs got students to that point is answered in the last two research questions.

Research Question 3: How are programs structured?

The third research question asked: How are graduate evaluation degree programs structured to enable students to develop these competencies?

Program structure refers to the curriculum – the coursework and non-coursework experiences all students are required to complete for the degree. All programs used evaluation-specific courses, practica experiences, and faculty mentoring and therefore had a similar program structure. Program profiles were similar in terms of the courses required, and the evaluator experiences provided. All programs were willing to accommodate student career interests in either academic or applied venues, although it was more or less difficult to follow an academic path in some programs. Critically, all programs indicated that doctoral students required time intensive mentoring during multiple hands-on evaluation opportunities.

Evaluation-specific courses. Using the program size indicator based on the number of evaluation-specific courses offered that was first used by Altschuld et al (1994), three of the participating programs were medium in size, each having 6-courses. One of these used a single course ID with a required, repeated enrollment (at least four times), content for which was based on students’ current experiences. Three programs were large in size and offered 8, 9, and 14-courses. Program-informants indicated that evaluation courses tended to draw majors as well as non-majors and could include master’s and doctoral degree students. One program-informant said:

>This class brings to many students’ attention to this whole social practice of evaluation. And these students say, oh I am so glad I bumped into this. I really want to do this as part of my doctoral program.—Respondent L
All programs used a developmental model in their three types of evaluation-specific courses: introductory, advanced, and special topics. Introductory courses described the entire evaluation process in a brief overview. Students start with basic skill sets and work their up. In the introductory level students take core courses in methodology, measurement, statistics, analysis, and instrument development. These tend to be fairly standard courses outside the field of evaluation (being applicable to many other disciplines). An introductory or foundation course in evaluation includes an overview of the field and includes the basics of evaluation design and planning. It builds upon the core courses in terms of data collection and analysis and runs students through developing a sample evaluation plan for a real or imagined client.

Advanced courses also covered the entire evaluation process, but tended to go into more depth on various aspects, e.g. data collection and analysis techniques, report writing. At the advanced level students spend more time on evaluation theory, alternative approaches, and practicing additional evaluator skill sets through data collection and analysis opportunities. Students widen their experience in hands-on evaluation through role-play, case studies, and single course projects.

A third course type found in programs that offered more courses tends to focus on special topics, often blended within the content area, such as Consulting Theory and Practice, Data-Informed Decision Making, Evaluation in Secondary Education, or Human Resource Education. These appear to focus on evaluator skill areas that may not get enough attention in the introductory and advanced courses and are offered only occasionally, e.g. Needs Assessment and Cost-Benefit Analysis. Evaluation program directories (Engle et. Al, 2006; LaVelle & Donaldson, 2010) provide a more comprehensive list. Rarely do special topics courses address soft skills sets, e.g. communication, interpersonal skills, and program management. These tend to be found informally in the student-faculty mentoring that anchors hands-on practice opportunities in internships and practica. Internships and practica do not appear to follow the standard course format, which leads to the second structure in evaluation programs.

**Practica experiences.** The second program structure all programs used as a formal part of the curriculum was multiple internship or practica experiences. Internships and practica experiences provide students opportunities to practice real-life evaluation
skills within a facilitated or guided context. Practica experiences start by building skill sets. Students are gradually involved more intimately across the entire evaluation process from question development to report writing and delivery until students take on leadership roles in conducting and managing evaluations.

Many practica are one or two terms in length. A longer period of time—several semesters, sometimes up to a year could be devoted to participating in actual evaluations, usually through employment at an evaluation center. When possible, these experiences allowed students an opportunity to reflect on practice as they move toward greater responsibility and developed new skills. However, broad practica experiences (a year or more) as a required learning vehicle were not observed in this study. In practice, all programs encouraged or required students to seek multiple internship experiences to expand not only skill sets, but experience in multiple evaluations. For example:

This typically involves taking a series of increasingly challenging roles [and] projects as the student progresses through his or her degree—Web P

Multiple internships with evaluation professionals—Profile M

I see that as one of my responsibilities, as a faculty member, to get evaluation projects as vehicles for students to get experience in being an evaluator. . . I really think it is quite a responsibility that we have as faculty because this is a practice that you only learn by doing it. It can't just sit in a classroom.—Respondent L

Practical experience opportunities feature significant student-faculty mentoring, which forms the reflective characteristic required in experiential learning (Rogers, 1969). For example:

[The Res. Apprenticeship course] they always take it with me…for the full time they are here not just one course…this really is the primary means of instruction for evaluation. It is a mentorship thing—Respondent X

The guided aspect of the practica experience was part of the third program structure.

**Faculty mentoring.** All program-informants indicated that mentoring students comprised a significant amount of their time. Mentoring included helping students to determine skill sets and evaluation experiences that would facilitate career choices – usually in various applied evaluation venues but also for academic career. Mentoring
was also a featured aspect of the learning process for students engaged in practica experiences. All of the programs used a debriefing or discussion component to practica experiences that allowed faculty and students to reflect on the process.

Program-informants said they did their best to design student experiences to meet their career goals, although this was difficult in some cases. Students enter evaluation already straddling two worlds – their home discipline or content area (psychology, education, social work, policy studies, health science) and the new world of evaluation. Some students incorporate additional content areas or minors. These students, more so than in other doctoral programs, enter programs with a variety of career options and need careful guidance from experienced faculty. One program-informant said:

The majority of the students who graduate from our evaluation program want to be evaluators, work as evaluators. They work for the government somewhere, state or national government, occasionally a school district evaluator. Some get some expertise in assessment and work for an assessment company. Some might work for a mid-size or large firm. Some might work in a research evaluation center in a university, but the majority will be practitioners rather than academics. There aren’t very many jobs as academics in evaluation. And many of them love the field. They just love to do it. So that is their first choice.—Respondent L

Many students choose to go the route of an applied evaluation practitioners. They may work in government, for-profit, non-profit, corporate, or education venues within their content area. For a progression of student activities, see Table 17. Others will be academic practitioners and seek a traditional academic career as a faculty member in a department or program similar to what they are experiencing as a graduate student (which are highly competitive in the United States). Note that the differences in student progression are minor; adding a reference to teaching skills and then strategies related to successful academic job searches (where a research plan in the content area would be important). According to three of the program-informants, many academic practitioners are international students who will go home to a less competitive job market. Still others will stay in higher education as 12-month administrative or technical
### Table 17. Student Progression by Career Goal.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Activities for Applied Practitioner</th>
<th>Activities for Academic Practitioner</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>• Introduction to Evaluation</td>
<td>• Introduction to Evaluation</td>
</tr>
<tr>
<td></td>
<td>• Core Courses in:</td>
<td>• Core Courses in:</td>
</tr>
<tr>
<td></td>
<td>o Research Methodology</td>
<td>o Research Methodology</td>
</tr>
<tr>
<td></td>
<td>o Data Collection &amp; Analysis</td>
<td>o Data Collection &amp; Analysis</td>
</tr>
<tr>
<td></td>
<td>o Content Area or Discipline</td>
<td>o Content Area or Discipline</td>
</tr>
<tr>
<td></td>
<td>• Participation in professional</td>
<td>• Participation in professional</td>
</tr>
<tr>
<td></td>
<td>conferences</td>
<td>conferences</td>
</tr>
<tr>
<td></td>
<td>• Participation in professional</td>
<td>• Participation in professional</td>
</tr>
<tr>
<td></td>
<td>conferences</td>
<td>conferences</td>
</tr>
<tr>
<td></td>
<td>• Opportunities to publish, write</td>
<td>• Opportunities to publish, write</td>
</tr>
<tr>
<td></td>
<td>grants / RFPs</td>
<td>grants / RFPs</td>
</tr>
<tr>
<td></td>
<td>• Start Dissertation Process</td>
<td>• Start Dissertation Process</td>
</tr>
<tr>
<td>2</td>
<td>• Advanced Evaluation Course</td>
<td>• Advanced Evaluation Course</td>
</tr>
<tr>
<td></td>
<td>• Elective or Special Topics Courses in Evaluation and / or Content Area</td>
<td>• Elective or Special Topics Courses in Evaluation and / or Content Area</td>
</tr>
<tr>
<td></td>
<td>• Practical Experiences:</td>
<td>• Practical Experiences:</td>
</tr>
<tr>
<td></td>
<td>o Single Course Projects</td>
<td>o Single Course Projects</td>
</tr>
<tr>
<td></td>
<td>o Focused Practica Experiences</td>
<td>o Focused Practica Experiences</td>
</tr>
<tr>
<td></td>
<td>o Faculty Mentoring</td>
<td>o Faculty Mentoring</td>
</tr>
<tr>
<td></td>
<td>• Participation in professional</td>
<td>• Participation in professional</td>
</tr>
<tr>
<td></td>
<td>conferences</td>
<td>conferences</td>
</tr>
<tr>
<td></td>
<td>• Opportunities to publish, write</td>
<td>• Opportunities to publish, write</td>
</tr>
<tr>
<td></td>
<td>grants / RFPs</td>
<td>grants / RFPs</td>
</tr>
<tr>
<td></td>
<td>• Start Dissertation Process</td>
<td>• Start Dissertation Process</td>
</tr>
<tr>
<td></td>
<td>• Teaching Assistant (TA)</td>
<td>• Teaching Assistant (TA)</td>
</tr>
<tr>
<td>3</td>
<td>• Focused or Broad Practica</td>
<td>• Focused or Broad Practica</td>
</tr>
<tr>
<td></td>
<td>Experiences</td>
<td>Experiences</td>
</tr>
<tr>
<td></td>
<td>• Faculty Mentoring</td>
<td>• Faculty Mentoring, includes</td>
</tr>
<tr>
<td></td>
<td>• Participation in professional</td>
<td>academic culture</td>
</tr>
<tr>
<td></td>
<td>conferences, <strong>publish</strong>, write</td>
<td>• Participation in professional</td>
</tr>
<tr>
<td></td>
<td>grants / RFPs</td>
<td>conferences, <strong>publish</strong>, write</td>
</tr>
<tr>
<td></td>
<td>• Dissertation</td>
<td>grants / RFPs</td>
</tr>
<tr>
<td></td>
<td>• TA or guest lecturer</td>
<td>• TA or guest lecturer</td>
</tr>
<tr>
<td>4</td>
<td>• Networking for job search</td>
<td>• Strategies for Academic Job Search</td>
</tr>
</tbody>
</table>

*Note: Student activities in *italics* are examples of differences by career goal.*

Professionals. Program-informants put it this way:

*Academics in evaluation are very few and far between. So we give our students a pretty thorough grounding, and when they apply and when I interview them about the potential career paths that they might go down.—Respondent Y*

*[Students now] are very interdisciplinary and 50% are international . . . 75% want evaluation practice [careers] and about 40 – 50% do institutional research or academic [venues].—Respondent K*
These practica experiences culminate in leadership roles in conducting actual evaluations all the while reflecting on their experiences with faculty mentors. All programs adopted these three primary program structures. Surprisingly, the structure of the programs does not change in response to the variety of career interests and content areas where students plan to apply evaluator skills. Changes that do occur tend to be in elective course selections, specific course activities, and the focused practica experiences, all of which are often the result of faculty mentoring. One program-informant said:

I like to work side by side. I like to go out into the field. I find that really fun. So we do quite a bit of evaluation . . . . It would be like socializing a sociologist—

Respondent L

**Research Question 4: What practices are used?**

Finally, the last research question asked: What practices support the development of student evaluation knowledge, skills, and values?

Program practices are the approaches, polices, and supports that programs use to facilitate student learning. All programs in this study used (1) iterative learning, (2) hands-on experiences, and (3) extensive faculty mentoring.

**Iterative learning.** John Dewey (1897) said that experiential learning is practicing the skills the student wants to be able to use or know - until they can be done independently. All programs included stages in student learning in both courses and practica experiences. Practice or iteration is a feature of the introductory and advanced evaluation courses every program included in its curriculum. These courses tended to repeat the entire evaluation process in greater depth, providing students of varying core skill sets opportunities for repetition and practice of new skills. This is important because relating what is already known to the new material creates involvement and is one characteristic of experiential learning. (Rogers, 1969)

Introductory and advanced courses would often feature the same topic pattern. This provided students with multiple exposures to the evaluation cycle, e.g. needs assessment, question development, data collection and analysis, and report dissemination. For example, introductory and advanced courses tended to have similar
activities structured around the concept of an evaluation plan. The description for an introductory course stated:

An introductory course in program evaluation; planning an evaluation study, collecting and analyzing information, reporting results; overview of the field of program evaluation.—Syllabus M4

Below is the description provided for an advanced course:

In this course, we will examine various evaluation designs, data collection methods, data analysis techniques and strategies for communicating and reporting evaluation processes and findings.—Syllabus F8

This iterative approach in courses provides opportunities for students to practice the same skill sets to greater depth and increasing variety, e.g. using other evaluation designs or approaches.

Practica experiences were also iterative. Students progressed from individual skill development within an evaluation context to assuming leadership roles that required understanding the entire process. Here is an example from web-based material:

Field Experiences in Evaluation--progressive experiences developing evaluation skills under the supervision of evaluation faculty—Web N

Including practica experiences as required experiences is another way to give students practice with the process of evaluation—moving from an introduction to evaluation and creating an evaluation plan, to discussing design options and implementing various tasks in the plan in later experiences.

**Hands-on experiences.** All programs referred to hands-on student experiences as an important part of the curriculum, either in evaluation course syllabi or in program-informant interviews. Faculty syllabi and program-informant interviews show references to the use of simulations, role-playing, single course projects, and focused practica experiences to provide practice in, and discussion of evaluation theories, concepts, and skills. Experiential learning activities include first, opportunities to practice actual skills and second, discussion and reflection on the experience. Both are characteristics in experiential learning (Rogers, 1969)
All programs used hands-on or experiential learning strategies (simulations, role-playing, small course projects) in course syllabi. See Table 18 for the number of syllabi addressing each type of learning by program. Note that the number of syllabi available from each program varied and that multiple counts (more than one activity per syllabus) are possible.

Table 18. Course Syllabi\(^a\) by Instructional Strategies and by Program.

<table>
<thead>
<tr>
<th>Type of Experiential Learning</th>
<th>Medium Size</th>
<th>Large Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Number of Syllabi Reviewed</td>
<td>2</td>
<td>7 (^b)</td>
</tr>
<tr>
<td>Simulations</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Role Play</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Single Course Projects</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

\(^a\) Most syllabi utilized multiple instructional strategies, which resulted in dual counts. \(^b\) Special topics course included two syllabi for the same course for a total of 35 different course syllabi.

It is difficult to make too many generalizations about experiential learning in the evaluation course syllabi obtained. Not all syllabi were available, and not all syllabi accurately portrayed what actually happened in classes. Some programs, and faculty, appeared to favor some types of experiential learning over others but information on personal preferences or program mandates was not collected. For example, two of the three large programs used simulations (a description or scenario provided by the instructor along with a set of student instructions) in all of the syllabi submitted while four programs used simulation as an instructional strategy in 50% or less of syllabi provided. While simulations were mentioned in the largest number of syllabi overall, 21 or 60%, role-playing was mentioned in 11, or 31% of syllabi and single course projects were mentioned in 7 or 20% of syllabi. Three programs, two large and one medium used all three types of experiential learning found in syllabi while two medium and one large program used only two. The differences may have more to do with instructor preferences and time demands than with curriculum. But information at that level was not collected.
Figure 3 further illustrates the distribution of types of experiential learning described in the six programs by including two syllabi that referred to practica experiences. Only simulations were used in syllabi for every program while role-play was used in five of the programs. Role play (a less structured simulation in which students have a greater degree of input in how they respond to a given description or scenario) was used less than simulations. One program used it heavily in the syllabi and mentioned it specifically during the program-informant interview "syllabi – it doesn't reflect the dynamism of the class... I do a lot of role playing"—Respondent X; two programs used it in 60% of syllabi and two, only 33% or less, and one not at all.

Figure 2. Percent of submitted syllabi for each program using experiential learning: Simulations, Role-Play, Single Course Projects, and Focused Practica Experiences. Note: Programs A, B, and C were medium in size, while programs D, E, and F were large.

Four programs had syllabi that referred to single course projects. These are student projects that form a large portion of the instruction for a given course. Assignments may be for teams or individuals, towards specific evaluation tasks or methodologies, and may incorporate external interactions (such as site visits). Projects result in a student generated product presented for instructor feedback. Tasks were
often broken down over the course of the course with discussion in class as one of the learning strategies. One respondent described the course project as:

We go through it, but we spend most of our time on the communicative and relational dimensions of evaluation...we go through a sequence of tasks, we talk about those.—Respondent L

Another said:

Courses are always working with data and written presentations . . . lots of writing . . . [because this is] how they develop skills—Respondent K

Only two programs had syllabi that referred to focused practica experiences. No program provided syllabi for broad practica experiences. But this is not surprising. Practica or internship experiences, if they involve credit hours, are typically faculty-student understandings. Syllabi are more useful in sharing course organizational structure and policies with multiple students. In addition, while evaluation-specific course syllabi may not have been used, program-informant interviews indicated that these focused and broad practica experiences (as defined in this study) were used by all programs. In fact they played a prominent role in how programs socialized students into evaluation.

**Extensive faculty mentoring.** Faculty serve as mentors in courses, internships, and other student experiences (such as attending conference). Mentoring provides students with feedback on their acquisition of evaluation knowledge, skills, and attitudes. Guided mentoring provides students with an opportunity to reflect on their learning, and, according to the literature on socialization, to internalize what it means to be a professional evaluator. Theorist interviews indicated that mentoring was a key faculty task:

It is important for doctoral students in evaluation to work closely with experienced faculty members, young people with vitality (including other students) and to participate in evaluation activities.—Senior Theorist R

Well, if you really believe that you should individualize programs for your students, and you can pick your students, which I was able to do, you look at what they have had and you look at what their experience is, and what their orientation and interest is, then sure you want to make sure that their proper
confident measurement and design and evaluation theory and some of the other areas.—Senior Theorist S

Mentoring, like learning specific evaluator skills, involved a developmental approach. New students don’t know what the expectations are and students changing roles, from doer to leader, aren’t always sure what their next step might entail. Students new to evaluation may not understand the process well enough to understand their experience; they rely on an experienced mentor who understands not only the context of the project, but the developmental stage of the student. For example:

You never put a brand new student in charge of something. They have to learn. More senior students take a lot more leadership.—Respondent L

Projects [and leadership roles] increase in responsibility and scope—Respondent K

One reason to start students at the bottom is to help them to understand that there is a bottom, which is where they need be in order to develop their own skills. Using a leadership development perspective implied that students will move from follower to leader. It also meant that students were not sent off raw to come back when finished with a task. They were monitored from the very beginning. As they developed skills they moved up in responsibility, gathering feedback on their performance all along the way. This expert feedback was important, as one program respondent said:

Not everyone can output. I have definitely learned that. There are those who are really, really effective and those that are mediocre.—Respondent Y

Another spoke of structuring for mentoring:

[Leadership opportunities] build on a foundation of knowledge [from courses and move to] one-on-one mentoring.—Respondent Z

Respondents reflected on the breadth of experiential activities faculty provide. The quote below summarizes how program-informants viewed their role as mentors:

The major challenge is that the competencies for an evaluator are tremendous. Most of it can’t be learned through coursework. A lot of it is learned by doing it. Sort of trial and error . . . [for example, if] you don’t interact with people, you don’t have to develop communication skills.—Respondent Y
In summary, one of the purposes of this study was to identify unique attributes or features, if any, within the selected programs that might contribute to their status as an exemplary program. The strong emphasis on the use of evaluator competencies, a common program structure featuring evaluation courses, practica experiences and an emphasis on guidance and mentoring of students are clearly important components in socializing students into evaluation. Socialization referred to any required student or faculty activities, beyond the traditional coursework and research requirements of doctoral programs, which promoted skill development specifically for evaluation career options. These include internships (paid or unpaid), capstone experiences, portfolios, evaluation projects, experience writing papers and reports, and leadership development as evaluators. This information was found on website materials describing the advantages of enrollment in a given program and in evaluation course syllabi. However, these components were most notable in the topics raised by program-informants and senior theorists.

A key aspect in these program practices is that it results in specific challenges for faculty mentors, as the individualized student experiences in courses and practica can be time intensive. Program-informants indicated that these tasks were facilitated when programs were affiliated with an evaluation center, when faculty had evaluation projects themselves or applied connections, and finally, when other faculty in the program, department, college, or university were also engaged in evaluation. These resources provide multiple opportunities and skill ranges in evaluation experiences available to students throughout their program of studies.

Common Attributes of Programs

Programs were more similar than they were different. Program commonalities include the use of:

1. An iterative curriculum (introductory and advanced courses) using experiential learning strategies and skills practice throughout the student experience.
2. A full range of evaluator competencies including soft skills sets, e.g. communication, interpersonal skills, and program management skills as identified by the
American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004),
the Program Evaluation Standards, 2e (JCSEE, 1994) and the IBSTPI Evaluator
Competencies (IBSTPI, 2006).

3. Extensive faculty mentoring in coursework and practica experiences tailored
to student career goals.

4. Student progression to leadership roles in real-world evaluation experiences,
regardless of career orientation (academic or practice-oriented).

Differences among Programs

The primary difference between programs were small differences in (1) program
characteristics: geographical location, program college, annual tuition rates, and
number of courses and (2) hands-on or experiential learning in programs syllabi. These
differences were slight and did not appear to limit or affect student options, as described
by program curricula.

First, program characteristics refer to information that describes the institution in
terms of funding, size, type, etc. All of the programs were at large research intensive
universities but half of the programs in this study are located in institutions housed in
large cities. The other half are in small cities. Three programs were housed in the
College of Education, one in Social Sciences, and one is interdisciplinary (including Arts
& Sciences, Education, Engineering and Applied Sciences, and Health & Human
Services). For five of the publically funded programs annual tuition in 2009 averaged
around $10-thousand for in-state and around $21-thousand for out-of-state. One
program was privately funded and averaged tuition around $35-thousand. Half of the
programs were medium in size, having between four and six evaluation courses. Half
were large in size, having seven or more evaluation courses. The number of faculty at
medium programs ranged from 2 to 5 while the number of faculty are large programs
ranged from 4 to 8.

Second, programs showed small differences in the number of type of hands-on
or experiential learning activities identified in evaluation course syllabi. But there are
several caveats to be considered: (1) not all programs were able to provide syllabi for all
courses, while some programs provided additional copies, (2) it is not clear to what extent syllabi reflect actual course activities, and (3) syllabi are largely the result of faculty preferences and may not fully reflect actual program practice.

All programs used simulations as an experiential learning activity but varied in the degree to which this was identified in submitted syllabi. The percentage of syllabi for each program using this activity ranged from 36% - 100% of obtained syllabi for the 6 programs. Five of six programs had submitted syllabi that identified the use of role-playing as an experiential learning strategy, with the percentage of program syllabi between 14% - 100% for the five programs. Four of six programs had syllabi that identified the use of single course projects for a range of 29% - 40% of obtained syllabi. All programs used focused practica experiences but only two programs specified this in syllabi they submitted.

**Section III: Developing Evaluation Professionals**

In this study I wanted to understand how a selected group of programs socialized doctoral students as evaluation professionals. "Developing Evaluation Professionals" is a description of the common practices that emerged from the characteristics and components identified in the analysis of the six program participants as they relate to the research questions and the conceptual framework. The practices identified were developed through analysis of data using pattern-matching and explanation-building techniques to identify emergent themes.

Two dimensions in developing evaluation professionals stood out from the data—"Socialization Strategies" and "Individualized Career Preparation". Each dimension is comprised of critical or defining elements framed by two approaches for achieving those elements—the use of "Practica Experiences" and intensive "Faculty Mentoring". See Figure 3 for an illustration. The approaches form an important dimension in their own right as they support, guide, and enhance the program elements. The program elements are unique to each student. They comprise the scaffolding of the two dimensions that lead, *for each student*, to development as an evaluation professional.
within evaluation and development of career goals that include evaluation. The first dimension, Socialization Strategies, has two critical elements—Evaluator Competencies and Experiential Learning. The second dimension, Individualized Career Preparation, also has two critical elements—Flexible Coursework and Tailored Experiences. The dimensions, approaches, and elements are discussed in more detail below.

**Dimension 1: Socialization Strategies**

Conrad (1998) described socialization of students in terms of (1) meaningful learning experiences, (2) professional development experiences and (3) leadership skills. Meaningful learning experiences had two components, (1) participation in a community of learners, and (2) engagement in a critical dialogue. These ideas from the literature on socialization of students were supported by the data in this study.

**Evaluator competencies element.** Students benefited from the guidance and perspective of evaluation competencies as they were expressed in syllabi and student learning experiences. Evaluation competencies from each of the six programs were found mentioned in program-informant interview transcripts and predominantly, in the course outcomes and required reading listed on the evaluation-specific course syllabi.
The competencies provided students with a frame of reference for the skill sets they were developing, which is one of Conrad’s (1998) components for socialization.

The developmental model for evaluation coursework identified in research question 3 as part of the structure of programs also makes use of evaluator competencies. New learners use competencies as a guide to doing evaluations. Competencies remain useful as a checklist for repetition and practice of skills sets combined with reflection by more experienced students. Soft skills sets, e.g. communication, interpersonal skills, and program management skills have not been as prominently identified as “evaluator” competencies (since they are used by other fields) but remain an area of need for many students. Clearly stated competencies for these skills it would be easier to implement and track progress in these areas.

**Experiential learning element.** Conrad’s second requirement for socialization was professional development. Professional development can be thought of in terms of the skill sets needed by practicing professionals and the methods used to obtain them; the “doing centered learning” Conrad (1998) described. Experiential learning is a foundation for socialization. Dewey (1897) advocated hands-on learning for performance skills on real-life topics, student involvement in the material—relating it to what they already knew, and some form of reflection, preferably guided for more naïve learners. The experiential learning strategies described as part of the research questions indicate that programs are already using experiential learning to socialize students. Experiential learning is well utilized by programs as a tool for developing evaluation professions. It makes an appropriate partner to the evaluation competencies in the domain of socialization of students.

**Dimension 2: Individualized Career Preparation**

The primary goal of doctoral degree programs in evaluation is to socialize, or prepare, students for their chosen career goals. But these career goals can vary on a number of fronts, most notably in terms of preparation for the traditional doctorate career—academic placement, typically a research department, and the applied career of an evaluator. Evaluation is an applied field and many students intend a career in evaluation practice. In addition, students may come to evaluation from a number of
disciplinary backgrounds. This increases the need individualized programming. Individualization can be accomplished in two ways, flexible coursework and tailored practica experiences.

**Flexible coursework element.** Coursework may not differ greatly for doctoral students in evaluation in terms of an academic or applied career. But it can differ in terms of the venue and content areas that interest students – for example students interested in government applications may be interested in empowerment evaluations or evaluations that are more organizationally focused. Flexibility in choosing elective and core courses enabled students to develop a program of studies that met their needs and interests, which is particularly important for those students coming into programs from fields that did not have a research skill background.

As can be seen in Table 19, the doctoral program of studies was more intensive in some programs, but all programs had a degree of flexibility when it came to choosing

<table>
<thead>
<tr>
<th>Degree Reqs.</th>
<th>Medium Size</th>
<th>Large Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of credit hrs required</td>
<td>A: 60+</td>
<td>D: 72</td>
</tr>
<tr>
<td></td>
<td>B*: 30-60</td>
<td>E: 62+</td>
</tr>
<tr>
<td></td>
<td>C: 56 - 61</td>
<td>F: 63</td>
</tr>
<tr>
<td>Degree area, hrs</td>
<td>student choice</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>unknown</td>
<td>6+</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Evaluation-specific course hrs</td>
<td>8-24</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>29-40</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>Methodology course hrs</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>unknown</td>
<td>13-18</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Cognate or specialty hrs</td>
<td>12 (tailored to interests)</td>
<td>via area specialization</td>
</tr>
<tr>
<td></td>
<td>student choice</td>
<td>18-21</td>
</tr>
<tr>
<td></td>
<td>12+</td>
<td>18</td>
</tr>
<tr>
<td>Minimum no. eval. specific courses required</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

*a Courses not specified by department, subject to student and advisor approval based on career interests, needs*
courses. It is obvious from the table that the doctoral program devotes only a small percentage of courses to evaluation-specific courses. Note that the non-shaded rows of the table refer to credit hours while only the shaded row refers to number of courses.

Programs ranged from a minimum requirement of 13% of program hours for required evaluation-specific course hrs in program “A” to a possible 65% in program “E”. Elective, cognates, and choices in methodology courses enabled students to design a program of studies to meet their needs and interests. Programs also adopted unique strategies to meet the needs of students for guided feedback. The number of evaluation-specific courses required by a program was six in a medium size program. In program “A” the six course experiences were actually three (3) different courses. Students enrolled four terms in the equivalent of a special topics course in evaluation that was tailored to their skill level at the time. This meant that while the number of actual courses appeared small, students were actively involved in different evaluation discussions and tasks each term they were enrolled.

The largest number of evaluation-specific course hours required was twenty-nine, from a large program that allowed students to tailor program requirements to suit career interests. Program “B” had the most freedom for student choice but each program offered a degree of flexibility in credit hours and course choices. This allowed the program to promote acquisition of additional skills or specializations within method or content area courses to suit student career goals. Individualization for each student seemed to be the primary goal for each of these programs - aptly stated in this quote:

No two of my students ever had the same program in terms of exactly what courses they took with which professors . . . you look at what they have had and you look at what their experience is, and what their orientation and interest is—Senior Theorist Q

Tailored experiences element. Doctoral evaluation degree programs offered a variety of professional development experiences tailored to student needs and interests. Two students may have similar goals but enter the program with different research backgrounds. Each program becomes tailored to the needs of the student:

They are all different. Each student has got a different agenda and a different comfort level of not wanting to be a faculty member or wanting to get in the field
or working in a large consulting firm or research firm. They are different—Respondent X

The Ph.D. program is designed to prepare you as an evaluator, assessment specialists, or researcher for leadership and teaching positions in schools, non-school organizations, institutions of higher education, and government—Respondent O

It varies completely by field. It may be international aid and development foundation positions, it may be federal agency positions, it may be international agency positions and for-profit efforts. It may be governmental positions in recipient countries. It may be [teaching at a university] but that is a small proportion of it. It may be training departments in large commercial enterprises where they can just transform the training job or initial training and such. So it varies a lot with the field. It may be training of technocrats in evaluation skills for some industrial operations, for example in nuclear energy—Senior Theorist S

In evaluation, program-informants indicated that many of the students enter with the intention of leaving for a practice or applied career doing evaluations. This applied career focus is common for evaluation students. But I also asked programs an open-ended question on what they did to socialize students for academic careers. Five of the institutions indicated that coursework requirements and practica experiences in the program were tailored to career goals. Three of the programs incorporated specific courses or internships for teaching in an academic setting. For example, a student with an academic career interest may do a lot of co-teaching with the faculty mentor. The last stated that it had a good selection of experiences and courses to meet students’ needs, except that there was "not enough preparation for teaching"—Respondent K, simply because there were no undergraduate courses to provide full responsibility.

Academic publications are still frequently done by discipline area and program-informants stress that publications are important for students considering academic careers as well as applied practice. Interviews revealed tension in this topic because of the pull of students seeking applied careers. One prominent theorist said that the purpose of a doctoral degree in evaluation was to:
Transform the field that [students] are working in . . . by learning what the field's weaknesses are . . . then they propose a dissertation ritual and transform the field by [addressing] those weaknesses . . . I am only interested in what needs to be done in the field. The business of finding a job is completely secondary.—Senior Theorist S

The implication appears to be that doctoral students are expected to contribute to the discipline and that one doesn't need a doctor of philosophy in order to be trained as an evaluator and get a job. A program respondent acknowledged that:

Academics in evaluation are very few and far between . . . academia is not the route if you are interested in this discipline. It is not going to be that you are going to get an assistant professorship somewhere within the next few years. It is just logically not going to happen . . . getting an academic position in a department or college of education with a specialty in evaluation is a different issue. And that is when you do need the subject matter expertise.—Respondent Y

Another program respondent indicated that there has been field growth and expansion, but not in just one domain and it wasn't called evaluation. It’s:

Very interdisciplinary now - 50% of students are international! - so the program tries to find bridges and collaboration—Respondent K.

Two other program respondents also reported a large international student population for whom academic placements outside the United States are more available.

**Two Approaches to Support, Guide, and Enhance**

Identifying the common practices in developing evaluation professions involved not only looking at patterns, but finding paths that led to student socialization into evaluation. The two approaches, faculty mentoring and practica experiences, are the paths that lead through and between the dimensions and their elements; they could almost be considered the primary motivators for students traveling toward a career as an evaluation professional. As would be expected, the characteristics and components in the structure of doctoral programs interact. The socialization dimension is informed by how evaluation competencies and experiential learning combine to socialize students in the knowledge, skills, and attitudes, or culture, in evaluation.
The two elements associated with the individualized career preparation dimension are relevant to socialization. Coursework is part of evaluation competencies and experiential learning, just as tailored experiential learning activities echo evaluator competencies within the context of experiential learning. The separation reflects the juxtaposition of meeting individual needs for personal career goals with the needs of the discipline and society for producing capable leaders in academic and applied evaluation settings. It also makes it easier to tease out strategies to enhance the process.

**Faculty mentoring approach.** Faculty mentoring was not a separate task assigned to a weekly afternoon or to the Fall and Spring advising week for registration. Faculty mentored students expressing an interest in the program, students enrolled in their courses, graduate assistants working on their contracts and grants or attending conferences, as well as master’s and doctoral students as chair and committee members. Helping students move toward their own level of leadership or independence in evaluation was an important task.

A second reason to use a leadership development approach is to not only develop skills, but to become socialized into the field through interactions with professionals. The gradual movement to leadership gave responsibility to the new evaluator, and also introduced that evaluator to the people who will become peers:

Before [students] graduated, I would try to find a big national project in evaluation that they would actually direct, because I wanted them, when they left to have demonstrated to their future colleagues what they could do and to work with them and to know them. When she graduated, the people in the leadership of the field knew who she was and she had demonstrated what she could do, and I try to do that with every one of my students. And so, take them up the career ladder and find them a project that they could lead before they left.—Senior Theorist Q

**Practica experiences approach.** Hands-on, relevant, engaged, and reflective, practica experiences are the deeply experiential learning activities. Each program in this study emphasized guided (mentored) experiences in conducting evaluations, as many as possible and if possible, in the students career interests. During interviews respondents spoke frequently about the need to involve students in real world evaluations however possible. Opportunities for skill development were emphasized
early in the student’s program, often through coursework. What is not visible is that each program in this study pushed internships and was either affiliated with an evaluation center, or had faculty with strong ties to their own evaluation practice. One respondent wanted to expand a student’s experience:

I have a new student who came in this year and I had an assistantship for her, but I really wanted her to get additional practical experience working with different people. So I asked a colleague if she had space in one of her projects and she said “oh sure”. Because there is that kind of understanding that people [students] have to get experience.—Respondent L

While others said:

For example, funding students to do evaluation work was important:
I have funding, bring in funding and all of my students are employed doing evaluation work in the field—Respondent X

Behind the program strategies for developing evaluation professionals is the idea of increasing responsibility with increasing skills. This is an accepted part of the experiential learning perspective to promote practice in applied and academic settings. But it is actually a critical component in the socialization of students into the field of evaluation, and is part of the development model in the structure of programs. Consider the following comment from a prominent theorist:

I believe in a career ladder for my students, when a student came in I started them out doing the dirtiest [scut] work there are in evaluation first . . . they would start at the bottom, and then when they had proved that they could work hard, do what they didn’t want to do, then I would move them up to another level where they might be a specialist on a project.—Senior Theorist Q

In sum, practica experiences provided support to students as they progressed through their program, they provided an opportunity for regular mentoring, and they enhanced the learning experience.

**Emerging Themes in Common Program Practices**

Two dimensions Socialization of Students and Individualize Career Preparation, are used by programs and four critical elements, (1) use of evaluator competencies and
experiential learning, flexibility of coursework, and student learning experiences tailored to career goals. These are closely guided and supported by two approaches used by the faculty, faculty mentoring and practica experiences. All programs are designed to socialize doctoral-degree students into evaluation professionals using individualized career preparation facilitated by extensive faculty mentoring and practica experiences that culminate into leadership roles.

**Summary of Study Results**

Six programs that met the established study criteria for doctoral degree and evaluation emphasis, sustained programming, and association with a prominent evaluation theorist, agreed to participate in this study. The intent was to understand how they socialized students—to describe how they differed and to identify programming strategies that were held in common. From that understanding practices were identified that doctoral degree programs used to socialize students into evaluation. The illustration in Figure 3 on developing evaluation professionals outlines these program practices.

*Developing Evaluation Professionals* describes how these programs are structured in terms of two dimensions, four elements, and two approaches that emerged from an analysis of the data. Applying the conceptual framework back to the research question results suggested four areas where programs are more similar than they are different: (1) an iterative curriculum using experiential learning; (2) use of a full range of evaluator competencies, including “soft” skill sets; (3) application of extensive faculty mentoring, and (4) students progression to leadership roles in real-world evaluation for both academic and practice-oriented careers.

First, it was noted that evaluation-specific courses, whether there were four or fourteen, tended to cover the entire evaluation process and in doing so, many of the evaluator competencies, varying slightly in terms of emphasis or context. Introductory and advanced courses both tended to address the entire evaluation cycle. Each course was dynamic, and iterative, based on the experience level and interests of the students enrolled at any given time. This iteration and connection to student interests and
knowledge is part of the experiential learning process and was also seen in practica experiences.

Most programs use experiential learning methods in evaluation-specific course syllabi moderately, spending a significant amount of time on lecture, reading, discussion, and writing activities. Hands-on experiences tended to target specific evaluation tasks, as it was difficult to conduct an entire evaluation study within the time constraints of an academic term. The greatest use of experiential learning was reserved for internships and faculty-student mentor activities. Yet these practica experiences and the accompanying mentor opportunities were an important part of socializing students as professional evaluators. Perhaps the informal nature and relatively low presence of these “courses”—few programs had descriptive documents for student-mentor expectations during practica, erroneously implies that these experiences are not as critical as program-informants insisted they were during interviews.

Second, evaluation programs in this study followed a common set of evaluator competencies within the coursework and internships offered. All but one program framed its evaluation coursework within the department or discipline course requirements. The interdisciplinary program used its own list evaluator competencies and left the discipline to be added integrated separately. “Soft” skill sets, e.g. communication, interpersonal skills, and program management skills were added. Depending upon the discipline and the organization of the program, some of these soft skill sets could be found in course format, although this was not the norm (for example, a course in cost analysis or theory and practice of consulting). Framing these non-evaluation skill sets as measurable evaluation competencies would likely facilitate student learning and conversation on their experiences.

Third, student preparation for careers in evaluation required intensive and individualized mentoring from experienced faculty evaluators. The skills of an evaluator were skills largely held in common, with some differences in theoretical perspective or preferred methodologies. The intensive, individualized mentoring was required in order to adjust for differences in student career goals and initial experiences or research background. Students received guided feedback from hands-on experiences tailored to their needs. Content area was considered important, but a secondary aspect was the
venue where evaluations would be conducted and where some skills and understandings might be more desirable that others. In addition, students were expected to be active participants in tailoring the program experiences to their career goals. Evaluation internships and programs were therefore much more faculty-student interactive when it came to determining student experiences.

Fourth, practical evaluator skills acquired as part of a long-term developmental process culminated in the ability to design and conduct evaluations, but also in the skills and experience to play leadership roles. Leadership involved the administrative and management aspects of conducting an evaluation as well as a certain amount of peer recognition, but it incorporated something more. Evaluations take place in context. The context might be within a discipline or within a style or perspective of evaluation. In order to function in the context, evaluators need to be aware of the issues that may impact the evaluation, as this is one way in which evaluators help guide the development of questions that will provide desirable information. The ability to take a leadership role is important in order to provide information that facilitates accurate decision-making in these different contexts. The ability to lead may also be an indicator of socialization into the profession.
CHAPTER 5: DISCUSSION

This study was a response to concerns that university-based evaluation degree programs might be decreasing, that the demand for trained evaluators in an increasingly larger number of disciplines might not be met by existing degree programs, and that as an applied field, evaluators may not be receiving the kind and amount of hands-on training and experience needed to be successful. While the field has conducted a fairly regular compilation of evaluation degree programs (May et al., 1984; Altschuld et al., 1994; Engle et al., 2006; LaVelle & Donaldson, 2010), detailed information about the structure of the programs and how students were socialized into the profession has not been available.

LaVelle and Donaldson (2010) answered the first concern by stating that their results show that the number of evaluation degree programs is not decreasing. They utilized a web survey to identify every advertised evaluation degree program or specialization and identified 48 programs with curricular information. This conclusion was in comparison to the 27 programs identified by Engle et al. (2006). LaVelle and Donaldson suggest that "university programs do not appear to be in decline, and appear to be training evaluators for the growing demands of evaluation practice" (2010, p. 20).

This statement may be debatable given that all graduate programs tend to differ to some extent. With no agreed upon competencies or universal standards for evaluators (such as the board certifications required for other applied professions--education, engineering, law, medicine, nursing), it is not currently possible to identify programs that may produce evaluators who are NOT meeting the growing demands for trained evaluators in a larger number of fields. There is little evidence as to how programs promote student acquisition of evaluation skills and socialize them into evaluation careers. However, LaVelle and Donaldson make a good point that a next step is "to examine which programs have been consistently training evaluators the longest, and to learn what characteristics have contributed to their longevity" (2010, p. 21). This study attempted to address that next step.
Creating Evaluation Professionals

Experiential learning is a tool for socialization of students (Anthony, 2002; Boyle & Boice, 1998; Conrad et al., 1998; Gaff, Pruitt-Logan, & Weibl, 2000; Louis, 1985; Sarbin & Allen, 1968; Weidman et al., 2001). Socialization involves acquiring the knowledge, skills, and abilities for the role (Sarbin & Allen, 1968), e.g. in the range of careers for evaluation professionals as well as the values, attitudes, and ethical standards found in the culture (Louis, 1985). Experiential learning advocates state that students learn more by doing and that this is particularly important for performance-based skills, e.g. evaluation (Coleman, 1976; Conrad et al., 1998; Dewey, 1897; NSIEE, 1986; Rogers, 1969). Building on this premise, Merrill (2002) incorporated experiential and adult learning theories into his model for designing performance-based instruction. He recommended the following instructional phases to maximize experiential learning: (a) activation of prior experience, (b) demonstration of skills, (c) application of skills, and (d) integration of these skills into real-world activities.

Use of evaluator competencies, experiential learning, and faculty mentoring are important aspects in practices programs use to develop student evaluation knowledge, skills, and attitudes in the variety of roles and culture for evaluation professionals. Adding the use of practical experiences and individualization of student learning goals to the other practices identified from this study for developing evaluation professionals, provides a mirror of Merrill’s design.

The purpose of this study was to understand how programs in evaluation are structured to promote doctoral student professional development in evaluation. Six programs with sustained longevity were selected and examined. The program characteristics of these programs have not been defined previously. Simple counts of graduate students enrolled, graduated, and number of faculty or FTEs generated provided in previous directories did not give a clear description of what program do to develop evaluation professionals.

If a professional is someone who has obtained knowledge and skills needed to practice in a discipline, then socialization is the process of acquiring the culture (values and standards of ethical behavior) of the individuals who constitute the profession.
Anthony, 2002). Therefore this study adopted a conceptual framework using socialization and experiential learning to interpret the four research questions listed below:

1. What are the evaluation competencies, in terms of knowledge, skills, and values, proposed in the American Evaluation Association (AEA) Guiding Principles for Evaluators (AEA, 2004), the Program Evaluation Standards, 2e (JSCEE, 1994), and the International Board of Standards for Training, Performance, and Instruction evaluator competencies (IBSTPI, 2006)?

2. How are these competencies expressed in evaluation degree programs?

3. How are graduate evaluation degree programs structured to enable students to develop these competencies?

4. What practices support the development of student evaluation knowledge, skills, and values?

This study began by identifying a list of prominent theorists in the field of evaluation who had exhibited an interest in the education of doctoral students via their publications, and who had been closely associated with well-known evaluation centers. These theorists were interviewed for their perspectives on doctoral student education in evaluation as well as their views on instructional strategies and the use of evaluator competencies for training evaluators. The theorists were asked to recommend programs that might help me to understand what programs do to prepare future evaluation professionals. Ten programs were recommended, six met the study’s criteria for participation: a) the program currently offered a doctoral degree, b) an evaluation emphasis was available in the program, and c) the program had demonstrated long term viability (15 years or more), preferably in association with a prominent theorist. All agreed to participate.

The study was largely qualitative in design, relying on a deep description of what six university-based programs in evaluation in the United States were doing to promote student acquisition of evaluation skills and socialize them into evaluation careers. This was in contrast to the sweeping but relatively superficial information describing evaluation degree programs traditionally provided by the evaluation directories (May et al., 1984; Altschuld et al., 1994; Engle et al., 2006; LaVelle & Donaldson, 2010).
Data collection included the interviews with prominent theorists and material collected from the six programs: program artifacts (web-based materials and evaluation-specific course syllabi) and interviews of program-informants. Member-checking using program-informant review of program profiles (summaries of program information on the doctoral degree program) was accomplished all six programs. Analysis was at the program level using axial codes developed from the literature and research questions (Strauss & Corbin, 1998) and open-ended codes developed using pattern-matching (Miles & Huberman, 1984). Themes and patterns identified in the data were used to answer the research questions within the context of the conceptual framework: socialization of students and experiential learning. Patterns or commonalities of practice among programs were developed as a secondary data analysis process, using a meta-matrix of patterns identified between and among the six programs participating in this study and the theorist interviews. The themes and questions below were identified analyzing the data from this study.

**Promoting Professional Knowledge and Skills**

Programs used evaluator skills sets described by professional organizations: the *AEA Guiding Principles for Evaluators* (AEA, 2004), the *Program Evaluation Standards, 2e* (ICSEE, 1994) and IBSTPI Evaluator Competencies (JCSEE, 2006).

Programs provided students with coursework and internships or practica experiences that used an iterative model to introduce students to required competencies (the knowledge, skills, and abilities of evaluators) as well as theory and approaches in evaluation. The evaluator competencies addressed the knowledge, skills, and abilities necessary to practice evaluation in general. But "soft" skill sets, e.g. communication, interpersonal skills, and program management skills, were not as clearly defined. Vaguely part of evaluator competencies or informally addressed were other skill sets identified as important, e.g. less common evaluator skills (needs assessment, cost benefit analyses), and skills related to other types of careers (research faculty, administrative professionals). Framing diverse competencies as a checklist of options would be one means for facilitating student professional development in evaluation degree programs.
Programs shared a common strategy - students were provided with multiple opportunities to practice and gain feedback on their performance in courses and practical experiences. This iterative process was used to promote development of skills sets needed for student career goals as evaluation professionals but also to provide practice in areas related to student career goals. Course activities and practica experiences were tailored to student interests in terms of evaluation approaches, content areas, and venue (government, corporate, non-profit, education).

Even so, the current evaluator competency lists are not exhaustive in terms of the full range of skills for professional developers. As one senior theorist noted, “it can be difficult to identify every possible skill an evaluation professional might need.” In acknowledgement of that diversity, the AEA Guiding Principles for Evaluators (AEA, 2004) and the Program Evaluation Standards, 2e (JSCEE, 1994) were worded in a general format in order to promote a flexible set of guidelines. They were not intended to identify all possible evaluator behaviors and contexts as these cannot be determined in advance. This is particularly true for the program evaluation standards, as they were intended to be applicable as standards for evaluation of education in the U.S. but also to a range of disciplines (education, human sciences, and other disciplinary programs), with potential interest internationally.

In contrast, the epistemological understanding behind the development of evaluator competencies is very different for the International Board of Standards for Training, Performance, and Instruction [IBSTPI]. Their 84 competencies are written in terms of performance standards—explicit descriptions of behavior coached in measurable terms. These performance standards reflect an instructional design perspective where assessment of learning is tied directly back to the original learning outcomes or goals. This same connection to training can be seen in the Stevahn et al. modification of the original version of the King et al. (2001) taxonomy of evaluator competencies into 73 performance standards, in the “Essential Competencies for Program Evaluators Self-Assessment” (Ghere et al., 2006; King et al., 2001; Stevahn et al., 2005).

This direct link between behavioral practice and original training goals is not delineated in the AEA Guiding Principles for Evaluators (AEA, 2004) or in the Program
Evaluation Standards, 2e (JSCEE, 1994). This “lack of definition” gave them a greater flexibility to meet the variety of evaluation ways, means, and venues. However, the current models for credentialing or certification in applied fields using tests would require the narrow definitions used in the performance standards of the IBSTPI and essential competencies. Given the wide variety of evaluator skills sets, theories, and arenas for practice, ideas for credentialing in the future might need to use a unique approach.

Socializing Students into Evaluation

This wide variety of needs in graduate student programming was noted by Anthony (2002). He summarized data from the 2002 Re-Envisioning the Ph.D. initiative as a recommendation for doctoral programs to adopt a socialization of students perspective. He said that greater diversity of students meant that greater career options would need to be addressed by programs. This view of greater career options was supported by program-informant interviews. While many students sought a career as an applied evaluation practitioner, others had more complicated goals.

Programs in this study utilized resources common to the profession, the AEA Guiding Principles for Evaluators (AEA, 2004), the Program Evaluation Standards, 2e (JSCEE, 1994) and the IBSTPI Evaluator Competencies (IBSTPI, 2006). Each of these begins and then guides a conversation of what it means to do evaluation, to be an evaluator. But the programs in this study went further to immerse students into the culture of evaluators from the very beginning of their program. New students attended and later presented at professional conferences, they interacted with noted theorists and authors of publications they were required to read. Students participated in evaluations, wrote extensively, interacted with more experienced students, and were closely paired with a faculty mentor or mentors, all of whom modeled the culture they are entering. Guided discussions on issues, challenges, and questions in the field pervade student coursework and internships or practica. Students were exposed to the skills they needed to be evaluators, but also to the questions, issues, and perspectives in evaluation. At another level, the program structure facilitated integration of a student’s content area and career goals into a framework of evaluation.
Training for Academe and Praxis

Program informants indicated that all students need enough expertise in evaluation to eventually assume leadership roles. But students who wanted an academic career had an added layer of both skill development and acculturation. In other words, they prepared to do both. In every program, faculty were practicing evaluators who also did research, taught classes, and mentored graduate students within two worlds – that of evaluation and that of their research discipline. Students who were interested in evaluation as a field of research were introduced to evaluation as taking place within multiple contexts or disciplines, something of which they must also master. This was the content area that housed the degree. In the end, it is the ‘housing’ discipline that recruits and selects its faculty—usually with the view that evaluation is not another professional role, but instead an approach to doing research. Research and publication experiences did not differ dramatically for students with academic career goal, but teaching experiences were emphasized to a greater degree.

Program-informants shared that to enter an academic career in evaluation is to wear two hats - one that denotes a traditional career in academic research and another as a practicing evaluator. The two hats or roles have professional skills that overlap, but adding professional evaluation to the researcher role increased demands upon faculty. This increased demand was seen in the need to find, provide, and guide multiple practica experiences, each meeting individualized goals for students. Programs reported finding support in juggling these two professional roles through administrators and colleagues who recognized that evaluation was more than a different approach to doing research.

The confusion between evaluation and research may be an artifact of an outdated higher education policy. Research encompasses clinical, applied, and basic research topics; in that sense evaluation can be an example of academic research. Currently students seek, and are admitted to, specific degree programs for both goals. For example, a doctoral student in the Educational Policy and Evaluation program within the department of Educational Leadership and Policy Studies inside the College of Education could intend an academic or applied career. The National Center for
Education Statistics would use the same classification for the student for either career goal.

The differences in career goals do lead to differences in student experiences, which is seen clearly in the program practices. Would it be more appropriate to offer a doctorate of education for applied students and a doctorate of philosophy for students seeking academic careers? How do these degrees compare to students who pursue only master’s degrees or certificates? How does the concept of majoring in “A” program similarly bind students and the faculty teaching those open enrollment courses? What happens to students “attracted” by the idea of using evaluation in their own content area? Could the concept of a degree of flexibility in evaluator competencies lead to greater flexibility in programs of higher education, particularly where students have identified applied careers?

**Learning by Doing**

Internships and practica experiences are used because programs make the assumption that hands-on experience is critical for applied practice. This is the case in evaluation so what role does experiential learning play in the structure of evaluation degree programs?

In this study, experiential learning was the lynchpin in the structure of programs. Experiential learning was used in evaluation courses in all programs, formed the major learning strategy in practica or internship experiences, and was featured part of faculty mentoring (with its emphasis on reflection and problem solving). While a good deal of content, and some skills, can be garnered through passive learning, evaluation at its foundation is an applied degree. Evaluators practice, they perform. Programs emphasized progressive performance by providing real-world examples and scenarios in coursework and through extensive, demanding internship requirements that developed leadership skills. Students who assumed leadership roles were practicing as independent evaluators and could be largely said to have been socialized into the profession. Program informants indicated an emphasis on writing for publication and attending conferences, activities that provided students a means for entering a community of evaluation practitioners.
Identification of Program Practices

The purpose of this study was to develop an understanding of how doctoral evaluation degree programs promote student acquisition of evaluator skills and socialize them into evaluation careers. Developing Evaluation Professionals is an apt title for the practices that were identified to socialize students and give them practice with evaluator competencies using experiential learning strategies. The practices identified encompass the range of career options, disciplines, approaches, and venues where evaluation occurs. It also provides a vision or perspective for structuring other applied programs. Taken in context with the recommendations in the next section of this chapter, these may be useful to other university-based degree programs.

The two dimensions, Socialization of Students and Individualized Career Preparation described what the programs in this study were doing to prepare future evaluation professionals. The dimensions confirm what Weidman and Stein (2003) said about doctoral education, “a central purpose . . . is the socialization of individuals into the cognitive and affective dimensions of social roles related to the practice of learned occupations” (p. 642). Past initiatives on graduate professional development support the need for programs to plan for socialization of students into the full roles and culture of their future careers (Carnegie Initiative on the Doctorate, 2003; Re-Envisioning the Ph.D., 2000; Preparing Future Faculty, 1993, Preparation for the Professions Program [PPP], 1999).

The dimensions and their critical elements highlight the strengths and commonalities of participating programs: (a) use of evaluator competencies to guide and inform student learning goals, (b) use of experiential learning strategies to facilitate learning, (c) fostering flexible coursework options in designing a program of studies that meets student career goals, and (d) creating tailored experiences in practical experiences that engage students with skill sets matched to their career goals. Conrad et al. (1998) identified a similar list in his three categories of beneficial experiences: meaningful learning, professional development, and leadership.

Weidman et al. 2001 said that the, “relationship of faculty to students, or practioners to students, should be interactive, collaborative, open, and mutually
evaluative” (p. 98). This statement is important because two practices or approaches were identified that are important to supporting, guiding, and enhancing the process of developing evaluation professionals: extensive faculty mentoring and progressive practica experiences. These practices were buttressed by a faculty commitment to producing evaluators who are not only trained, but experienced. Engaging students through intensive faculty mentoring opportunities created a venue for reflection-based discussion that helped to cement the skill sets and further socialized students into the evaluation profession. Incorporating a developmental process in practica experiences helped students to move into leadership roles.

**Implications for Programs**

Only when the evaluator has insight into the interests and motivations of other actors in the system, into the roles that he himself is consciously or inadvertently playing, the obstacles and opportunities that impinge upon the evaluative effort, and the limitations and possibilities for putting the results of evaluation to work---only with sensitivity to the politics of evaluation research-can the evaluator be as creative and strategically useful as he should be. (Weiss, 1993, p. 94)

Originally publishing in 1973, Carol Weiss was referring to the importance of understanding "Where Politics and Evaluation Research Meet" because understanding the social context was important if an evaluator was going to be successful. The idea of obtaining insight into obstacles as well as motivations is relevant here in much the way detail. Identifying practices that lead to sustained programs in evaluation involved more than describing what evaluator competencies or instructional strategies were used, although that is a significant aspect. The most important aspect identified for developing evaluation professionals is actively planning for the socialization of students. A second important aspect in these programs was that evaluation programs are resource intensive. Third, programs in this study emphasized the need to fit into the culture of the institution.
Socialization Requires Active Planning

First, all doctoral programs seek to socialize students into their discipline. Addressing that concept systematically and purposefully has been the purpose behind national initiatives sponsored by the PEW Charitable Trusts and the Congress of Graduate schools – Preparing Future Faculty and the Re-Envisioning the Ph.D (Austin, 2002; Gaff, Pruitt-Logan, & Weibl, 2000, Nyquist & Woodard, 2000), in articles like Boyle and Boice’s “Best practices for enculturation: Collegiality, mentoring, and structure” (1999), and the Professional and Graduate Education studies from the Carnegie Foundation for the Advancement of Teaching—The Carnegie Initiative on the Doctorate (Golde & Walker, 2006, p. 2), the Prepararion for the Professions Programs, and the Carnegie Project on the Education Doctorate (CID, 2011). To prepare to apply for participation in the Carnegie Initiative on the Doctorate, George Walker, in personal communication (Walker, 2003) asked faculty in departments to create white papers describing what doctoral students in their department should be able to know and do in order to be future stewards of the discipline. The white papers were used to discuss how faculty wanted to redesign their programs to socialize students more effectively. This study takes a step in that direction.

Evaluation programs exist within the context of their home departments, and from there, within colleges and schools. Each content area or discipline has its own cultural understandings or perspectives of what is needed or valued. As the interest in the literature of socialization practices in higher education shows, departmental practices can have impacts on students (Conrad et al., 1998; Gardner, 2011; Hoberman & Malick, 1994; McCroskey, 1998; Rosensteing, 1968; Sarbin & Allen, 1968; Weidman et al., 2001; Weidmand & Stein, 2003). A career as a faculty member represents a cultural perspective too. Evaluation is a culture that students must be actively socialized into, whether their career interests are in academic or applied settings.

Evaluation is Resource Intensive

A second important aspect is that evaluation programs are resource intensive. Programs with longevity, at least those in this study, have resources. These resources include external funding from evaluation projects, support for internal evaluations,
faculty, and students. Each program in this study was affiliated with an evaluation center of some kind or had extensive connections to active evaluation projects. This increased the availability of practical experiences for students, as well as the number of mentors. Centers included staff that had faculty appointments but critically, the centers also included staff who were not housed in departments. It can be difficult to maintain and establish centers at universities:

Institutional support waxes and wanes, stability [through] administrative support . . . resources actually . . . was very important”.—Respondent K

This respondent concluded that when administrators are aware of what evaluation can do internally and externally, particularly across disciplines, it is easier to get help and be more competitive.

Resources also include faculty who can teach evaluation courses and provide additional mentoring for students. This is important not only in reaching out to new students, but in handling the heavy time demands experiential learning strategies and extensive mentoring make on faculty:

[We] have been very successful in attracting and keeping prominent scholars in the field. That is a big part of it . . . and the college . . . has a recognition and respect for evaluation as a practice . . .There are faculty elsewhere in the college who also have some expertise in evaluation—Respondent L

**Become Part of the Culture**

Programs in this study manage this balancing act of running an intensive program as a faculty member by inserting evaluation into the culture of their institution. They understand the role the department plays in the college and the college's role for the institution and knowing that, they do everything they can to make the program fit the larger goals of the institution. Respondent L expresses again how important this fit to the college is in terms of finding resources:

It becomes more of an accepted dimension of what we do in the college . . . I think the legitimizing of this in the organization is very important. . . fostering some kind of community of practitioners and scholarship working hard to develop a sense of group . . . is hard to do.
It is clear here that fitting in is not sufficient, creating a group sense is important. Increasing awareness also means being more competitive for institutional support – and that translates into attracting grants & evaluation projects, recruiting faculty who wear both hats (evaluator and academic), and attracting quality graduate students. These issues were not explored in this study per-se, but program-informants indicated a need for the support of departments, colleges, and institutions if they were to grow and be sustained. They reported that having interests that crossed departments and colleges increased connections and made it more likely that evaluation had a presence on campus. This also tended to increase the number of students in the program. From there, involving students and other faculty in evaluation projects both on and off campus also helped to increase awareness.

Attracting students is worth addressing briefly. The promotion of evaluation degree programs is one route to meeting the demand for evaluators as it increases the number of students pursuing evaluation degrees. But this requires actively recruiting students and informing them about programs. Program-informants indicated that it can be difficult to do this if there is competition for students in their original content areas. One method that programs in this study used for attracting students was to promote evaluation courses that supported or augmented other content areas. This method attracted involved students who may never have discovered evaluation previously. Comments from respondents in this study suggested that making those connections to other departments and faculty is critical:

[In the basic] series we get students from all over the school and elsewhere . . .
[In our second series] we are picking up students from Psych, from Management, from Social Welfare, from all over campus . . . [Compared to] what I have seen in other universities, [our institution] is just tremendous in terms of the acceptance of our students by basically any department on campus—Respondent X

**Areas for Future Research**

This study used a purposeful sampling technique to understand the practices six doctoral evaluation degree programs that have had a sustained presence over time
used to socialize students as evaluation professionals. While providing rich detail, this
design limits what can be said about programs in general. Two areas for further research
stand out: (1) a better understanding of the practices used by the many programs with
less than four evaluation-specific courses and (2) more information on the resource
networks that support evaluation faculty and students.

First, as LaVelle and Donaldson (2010) point out, there were 48 U.S. programs in
their directory, 31 of which were small (2-3 evaluation-specific courses). Perhaps one
reason 64% of evaluation degree programs in the United States are small is that the
program structure tends to require only three formal courses: introductory, advanced,
and special topics. A larger number of courses is not needed if programs practice an
iterative approach: an introductory course followed by an advanced course (which may
be repeated). What practices do other programs use? What patterns might emerge that
would tell us more about how evaluation degree programs are evolving and which
components are truly critical to long term success?

Second, faculty in this study talked about the importance of their connections to
other evaluation professionals and faculty, to other departments, and to students in
other departments. In describing programs, a network map for each program may be
more enlightening than any of the more pedantic descriptors such as number of
students, cost of tuition, or even program size. It appears that three different courses
combined with strong practical experiences may be large enough. Again, that would be
particularly helpful given that the majority of programs identified in recent directories are
considered small (Engle et al., 2006; LaVelle & Donaldson, 2010).

Previous directories describing evaluation degree programs were limited to
descriptive counts of activities and services. This study focused instead on the
processes used in programs to socialize doctoral-degree students. The results were
surprising – programs were much more similar than they were different. Future studies
will likely look for different processes in different types of programs and degrees. But a
utility perspective would suggest we look as well to the impact university-based and
other venues have in developing evaluation professionals. This process and impact
focus becomes increasingly important as evaluation becomes more mainstream,
attracting more students outside of education and psychology.
APPENDIX A: TEMPLATES

Templates (Email and telephone scripts follow the same format)

Initial Email or Telephone Contact
Contact Interview Reminder Email or Telephone
Program Profile Material Email or Telephone
Program Profile Feedback Interview Reminder Email or Telephone; Sample
Program Profile Template.
Program Profile Feedback Thank you Email or Telephone

1. Initial Email or Telephone Contact
Subject line: University-based Evaluation degree programs—your view?
Dear Dr. ____,
You have shown an interest in the evolution and structure of university-based evaluation degree programs and are associated with one of the most successful programs. As one of the exemplary leaders in the teaching of evaluation I would like to request your knowledge and help in understanding your program. I am also interested in making contacts to other individuals and programs that would, in your opinion, be representative of the variety of programs in the U.S.

I am a program evaluation doctoral candidate at Florida State University working with Dr. Linda Schrader. The intent of my dissertation is to develop a classification system to describe evaluation degree programs in the United States in terms of student learning outcomes, experiential learning activities, and career focus for the program.

Last year at the AEA conference there was a meeting titled “How Best to Improve University-Based Evaluation Programs.” A point was raised that not enough was known about current programs; more information was requested on what evaluation degree programs in the U.S. were doing, and how they were doing it. My dissertation research is relevant to that discussion.
I would like to speak with you briefly to explain the study intent and answer any questions. Demands on your time will be limited. If you are interested please reply to this email with the information below that would be convenient for you:

- Date [before date+10?]
- Best time(s) to talk
- Preferred telephone number (I have ###-###-####)
  (You can reply “Not available” and I won’t send a follow-up email or telephone.)
- I can also be reached by phone at ###-###-#### (cell).
- I look forward to speaking with you!

Michelle Chandrasekhar

2. Contact Interview Reminder Email or Telephone
Subject line: University-based Evaluation degree programs—Interview?
Dear Dr. ______,
This is a reminder that we have an appointment in two days to speak about your program and my study. I will call on:
  - Day, Month, Date at 00:00 am/pm at the following number:
    - ###-###-####
  - I anticipate that this call will take approximately 15-20 minutes of your time.
  - If you need to contact me earlier, I can be reached by email, or by phone at ###-###-#### (cell).
  - I look forward to speaking with you on Day -

Michelle Chandrasekhar

3. Contact Interview Thank you Email or Telephone
Subject line: Phone meeting on Day, Date
Dear Dr. ______,
Thank you for speaking with me on Day. I appreciate the time you spent with me to discuss my study. [If not participating, end with: I am sorry you are not available to participate but I look forward to talking with you again at some other time. + signature]
  - Here are my notes from our conversation:
  - I can access program information here [give URL]
You plan to send me x,y,z via [email] OR I will send you an pre-paid envelope for # pounds of material (x,y,z) to be sent to me by US Media Post. Your address is:

123 Some Road
Sometown, XY ######

You suggested that I contact
Person A at this institution
Person B at that institution
Others?

As we discussed, I hope to have a program profile for you to review by Month.

Thank you again for your help - if you need to contact me please email or call anytime

[###-###-####] (cell).

Michelle Chandrasekhar

4. Program Profile Material Email or Telephone

Subject line: Your Program Profile (draft)!

Dear Dr. ______,

Attached please find a draft version of your program profile. [Or this is to inform you that a hard copy of the draft version of your program profile has been sent to the address below:]

Your feedback on the profile information, as well as the process used to develop it, are very important to this study. As the classification tool cannot be designed until all the profiles are finalized, would it be possible for us to talk before date+10days?

When you receive your materials, would you reply with your preferred date and time:

Day, Month, Date at 00:00 am/pm at the following number:
###-###-####

I anticipate that this call will take approximately 15-20 minutes of your time. If you would like to send me an electronic version of your feedback, that would be wonderful. We can look at it together as we talk. Once we have an appointment, I'll send you a reminder 2 days in advance.
If you need to contact me earlier, I can be reached by email, or by phone at ###-###-#### (cell).

I look forward to your thoughts, comments!

Michelle Chandrasekhar

**Program profile template.**

Name of Institution

Degree housing: College, School, Department, Division, and program

*Name of degree*

Name of Program Director

**Mission:** Statement describing the mission of the department, program, or track where the degree is administrated.

**Doctoral Degree Requirements**

- Breakdown by program’s categories (theory, practice, electives, research, internship, etc.)
- Number of credit hours for each course

**Unique Program Characteristics:** Experiences, requirements that are unique to evaluation students.

5. Program Profile Feedback Interview Reminder Email or Telephone

Subject line: Program Profile Feedback Interview Reminder

Dear Dr.______,

This is a reminder that we have an appointment in two days to speak about your thoughts on the program profile I shared with you. I will call on:

Day, Month, Date at 00:00 am/pm at the following number:

###-###-####

I anticipate that this call will take approximately 15-20 minutes of your time.
If you need to contact me earlier, I can be reached by email, or by phone at ###-###-#### (cell).

I look forward to speaking with you on Day -

Michelle Chandrasekhar

6. Program Profile Feedback Thank you Email or Telephone
Subject line: Profile Feedback on Day, Date

Dear Dr. ______,

Thank you for sharing your thoughts and comments on the program profile with me on Day. I appreciate the time you are spending on this study!

To be sure I have not missed anything, below are my notes from our conversation. If I have anything incorrect, please let me know. Otherwise I’ll consider it good to go!

One point
Second point
Third point
And so on..

Thank you again for you help – I hope to have the classification tool for you sometime in Month.

Have a great (semester, holiday, etc.)!

Michelle Chandrasekhar
APPENDIX B: INTERVIEW SCRIPTS

Interview Script for Prominent Theorists

Good Morning/Afternoon/Evening Dr. ____________.

My name is Michelle Chandrasekhar. I am a graduate student under the direction of Dr. Robert Schwartz and Dr. Linda Schrader in the Department of Educational Leadership & Policy Studies, College of Education at Florida State University.

Thank you for taking time to speak with me.

I would like to clarify how this interview will be conducted. If you have any questions concerning the research study or this interview, please ask me now or at any point during the interview or if you would like to contact Dr. Linda Schrader first, we can reschedule. Her contact information is (###-###-####) or 123@mailer.fsu.edu.

As we discussed in our email messages, I am conducting interviews of Program Directors from four well-known evaluation degree programs that have maintained longevity in their programs and are considered leaders in the field as part of a dissertation study.

I would like to clarify how this interview will be conducted. If you have any questions concerning the research study or this interview, please ask me now or at any point during the interview or if you would like to contact Dr. Linda Schrader first, we can reschedule. Her contact information is (850)644-8780 or lschrade@mailer.fsu.edu.

- For the purposes of insuring accurate representation of our discussion today, I will be recording our conversation on a audio-recording device.
- While there are no foreseeable risks to participating or not participating in this interview, the information gathered may be of benefit to evaluation degree other program directors and to the field as a whole.
- This interview contains open-ended and short answer questions and is estimated to take approximately 60 minutes of your time.
- Participation in this interview is voluntary. You may choose not to participate or to end the interview at any time with no penalty.
- The results of this interview may be published, but results will not identify respondents or their institutions.
- Information obtained during the course of this interview will remain confidential, to the extent allowed by law.
- Any records (tape, digital, hardcopy) will be kept in a locked filing cabinet. Only the researcher will have access to these recordings, which will be destroyed September 1, 2011.

If you consent to participate, Florida State University’s IRB has requested that you please say for the record:

“I understand that I will be (tape recorded) by the researcher. These tapes will be kept by the researcher in a locked filing cabinet. I understand that only the researcher will have access to these tapes and that they will September 1, 2011.”

As I mentioned in setting up this interview with you, one of the suggestions from the think tank session at the 2006 AEA conference [American Evaluation Association]: “How Best to Improve University-Based Evaluation Programs” was to gather more information on what programs in the U.S. were doing, and how they were doing it.
I am interested in following up on Engle and Altschuld’s 2002 Survey of Programs. One point they made was that it wasn’t clear, given the variety of disciplines and programs where evaluation was taught, what instructional methods were used by evaluation degree programs to teach evaluation competencies, or how these, and other program resources, might vary.

In this time I would like to first provide a rationale for the study and an opportunity to ask questions.

The purpose of this study is to create a classification system that describes evaluation degree programs in the United States in terms of evaluator competencies, experiential learning activities, and career focus for the program. These three dimensions will allow an exploration of how and why programs are structured differently. This schema will help students to determine which programs meet their career goals. Program faculty will be able to use the information to market the program to prospective students as well as to other departments within their institution. A classification schema will also help other programs and faculty interested in the training and teaching of evaluation to determine what next steps might be needed to enhance programs and meet the demand for new evaluators in the United States.

1. What would you consider to be desirable learning outcomes for:
   a. evaluation practitioners and
   b. faculty of evaluation programs?

   **Probe Questions:**
   - How do these meet the needs of theory, of practice?
   - How might these desirable outcomes differ as a function of type of evaluation activity? (different levels of work? Internal/external? Independent Consultant vs academic researcher?)
   - Would they differ, and to what extent, for various types of degrees? Certificates?
   - What implications for program learning outcomes do you see for combination degree programs—for example, Florida State University is developing a doctoral degree that combines Educational Policy and Evaluation?

2. Currently there are two approaches to the training and development of new evaluators—degree or certificate-based training that takes place in academic institutions and training that is provided by institutes, centers, and professional organizations, like AEA’s workshops.

   What are you thoughts these approaches from a practice-based or experiential learning perspective? By experiential learning, I mean “learning activities that engage the learner directly in the phenomenon being studied.” as defined by the National Society for Internships and Experiential Education [NSIEE].

   **Probe Questions:**
   - In terms of learning outcomes: advantages / disadvantages?
   - What do you consider to be the issues in implementation? Possible or effective solutions?
   - What are the implications for theory/practice from a quality perspective in these different training approaches?

3. Thinking about what you would recommend to other evaluation program directors, or to members of the field as a whole, how would you describe your “Lessons Learned” for designing and implementing a university-based evaluation training program?
Probe Questions:

- What organizational resources do you consider critical for a successful program? (space, staff, budget, connections to other disciplines, local communities of practice, technology, etc.)
- Are there other recommendations or Next Steps, that you have for how the field might move forward with these ideas?

Interview Script for Program-Informants

Good Morning/Afternoon/Evening Dr. __________.

My name is Michelle Chandrasekhar. I am a graduate student under the direction of Dr. Robert Schwartz and Dr. Linda Schrader in the Department of Educational Leadership & Policy Studies, College of Education at Florida State University.

Thank you for taking time to speak with me. [Insert names of exemplar(s) who recommend Referred programs and a short line as to the reasoning they gave for including X program.]

I would like to clarify how this interview will be conducted. If you have any questions concerning the research study or this interview, please ask me now or at any point during the interview or if you would like to contact Dr. Linda Schrader first, we can reschedule. Her contact information is (###-###-#### )or 123@mailer.fsu.edu.

- For the purposes of insuring accurate representation of our discussion today, I would like to record our conversation on a audio-recording device.
- I am conducting this study of learning activities and experiences in evaluation degree programs in the United States as part of a dissertation study. The information gathered will hopefully be of benefit to evaluation degree program directors and to the field as a whole. There are no foreseeable risks to participating or not participating in this survey.
- Participation in this survey is voluntary but if you participate you will be helping existing and potential program evaluation programs become informed of current practices in higher Education.

As I mentioned in setting up this interview with you, one of the suggestions from the think tank session at the 2006 AEA conference [American Evaluation Association]: “How Best to Improve University-Based Evaluation Programs” was to gather more information on what programs in the U.S. were doing, and how they were doing it.

I am interested in following up on Engle and Altschuld’s 2002 Survey of Programs. One point they made was that it wasn’t clear, given the variety of disciplines and programs where evaluation was taught, what instructional methods were used by evaluation degree programs to teach evaluation competencies, or how these, and other program resources, might vary.

In this time I would like to first provide a rationale for the study and an opportunity to ask questions.

The purpose of this study is to create a classification system that describes evaluation
degree programs in the United States in terms of evaluator competencies, experiential learning activities, and career focus for the program. These three dimensions will allow an exploration of how and why programs are structured differently. This schema will help students to determine which programs meet their career goals. Program faculty will be able to use the information to market the program to prospective students as well as to other departments within their institution. A classification schema will also help other programs and faculty interested in the training and teaching of evaluation to determine what next steps might be needed to enhance programs and meet the demand for new evaluators in the United States.

I would like to create a program profile that describes your program and then compare/contrast it with other program profiles. To do this, I would like to collect from you (or a URL I could download) documentation that describes your program and, most important, copies of the syllabi for the evaluation-specific courses offered by the program.

[In looking at your website I've already found some material.]

So the major data collection for this study is going to be me doing content analysis of program artifacts. But it is critical that my fact collecting and interpretations be validated and that other voices contribute to the development of the classification tool. That means that I would also ask you for feedback on a draft version of the program profile and on the final classification tool as well. None of the 3 phone conversations should take more than 20 minutes.

Finally, today I'd like your advice in selecting other programs and their key informants that would be representative of the variety of evaluation degree programs in the United States.

Do you have any questions so far? [answer any questions, do the consent bit if appropriate.]

If you don't have other questions, I'd like to record the consent phrase our IRB committee requires. Since this is an audio tape, I'll read it to you and at the end, if you would repeat a confirmation phrase that will meet the requirements.

• The results of this interview may be published, but results will not identify respondents or their institutions if you choose not to be included.

• Information obtained during the course of this interview will remain confidential, to the extent allowed by law.

• Any records (tape, digital, hardcopy) will be kept in a locked filing cabinet. Only the researcher will have access to these recordings, which will be destroyed November 1, 20XX.

*If you consent to participate, Florida State University's IRB has requested that you please say for the record:*

“I understand that I will be (tape recorded) by the researcher. These tapes will be kept by the researcher in a locked filing cabinet. I understand that only the researcher will have access to these tapes and that they will November 1, 20XX.”

-----

Thank you!

Now, in terms of program descriptors and syllabi, I've gone over the website for you program described at [give department and college, url]. Here is what I could find that talks about the
program:
1.
2.
3.

Do you have any other materials that would help me to understand your program? Would it be possible for you to send them to me electronically? If not, if you'll estimate their weight, I can send a postage page envelope for you to slip hard copies to me by US media post! Thank you –

I also was able [not] to find the following syllabi:
1.
2.

For this study it's important to review the evaluation-specific course syllabi. That is, I'd like to have copies of all the course syllabi, core and elective, for the classes your program offers that are specific to evaluation concepts and principles. For example, I'd like to see an "Evaluating Education Policies and Programs" course but wouldn't need to see the "Educational Research Methods" course that majors also take.

Would you be able to get those for me in the next two weeks or so? What format is best for you – can you send it electronically or do you anticipate a mailing envelope?

Thank you –

-----

Now, in this interview I would like to explore how successful programs such as yours develop students' professional skills and socialize them into the evaluation profession. So I thought we might talk about your program's structure (program-level goals, instructional activities, curriculum), the background on how it came to be the way it is now, and what suggestions or advice you might have for other programs based on "lessons learned".

I do have some questions to get us started, but I am most interested in what you have to say – so please feel free to speak your thoughts at any time. Do you have any questions before we begin?

4. To begin, can you tell me how the degree program was initially developed [within the home department/college/school]?
   Probe Questions:
   - Was there a critical component(s) in making this program successful? I mean, something that was needed for the program to get off the ground and stay viable?
   - I've noticed that some evaluation degree programs appear to be more of a particular discipline and quite a bit less evaluation. Can you talk about how your program developed (and has maintained) its identity?
   - Others based on response?

5. One of the issues for disciplines with applied or practice oriented paths, like social work, nursing, engineering, or evaluation is achieving a balance between the need to teach students theory as well as practice. How does your program tackle that issue? How is it different, if it is, for different degrees? (Master's, Doctorate, Certificate)
   Probe Questions:
Given your response, what aspects of your program structure do you consider to be areas of strength in achieving this goal of balance/specificity?

- How are they sustained? I mean, are they supported by administration, the faculty, a strong demand for graduates?
- What implications for program learning outcomes do you see for combination degree programs—for example, Florida State University is developing a doctoral degree that combines Educational Policy and Evaluation?

6. Currently there are two approaches to the training and development of new evaluators—degree or certificate-based training that takes place in academic institutions and training that is provided by institutes, centers, and professional organizations, like AEA’s workshops.

   What are you thoughts these approaches from a practice-based or experiential learning perspective? By experiential learning, I mean “learning activities that engage the learner directly in the phenomenon being studied.” as defined by the National Society for Internships and Experiential Education [NSIEE].

   Probe Questions:
   - In terms of learning outcomes: advantages / disadvantages?
   - What do you consider to be the issues in implementation? Possible or effective solutions?
   - What are the implications for theory/practice from a quality perspective in these different training approaches?

7. Thinking about what you would recommend to other evaluation program directors, or to members of the field as a whole, how would you describe your “Lessons Learned” for designing and implementing a university-based evaluation training program?

   Probe Questions:
   - What organizational resources do you consider critical for a successful program? (space, staff, budget, connections to other disciplines, local communities of practice, technology, etc.)
   - Are there other recommendations or Next Steps, that you have for how the field might move forward with these ideas?

Do you have any other questions for me? I have time but I don’t want to impose on you.

Ok - I plan to send you an email recapping what we’ve talked about today. I can be reached that way, or by the number in my email.

I enjoyed talking with you today and will look forward to delving into your program materials –

Good bye
APPENDIX C: PROGRAM PROFILE FEEDBACK SCRIPT

Good Morning/Afternoon/Evening Dr. __________,

Thank you for speaking with me again. Today we planned to talk about the program profile with the goal of making sure I have correct information in my final version.

Let’s start with corrections or amendments you found and perhaps we can end with some other thoughts or insights going through this process might have suggested to you?

Do you have any comments on the overview section?

Possible Probe Questions:

• Would this be an accurate description for prospective students?
• How well would this describe the program to administrators at your institution?
• Note questions based on responses--

Next let’s look at the Program Mission section, any amendments, explanations?

Possible Probe Questions:

• Is there anything else that’s part of the program mission that isn’t captured here?
• I’m very interested in the process used for accreditation of universities, given the way that process is changing, has your mission description changed?
• Note questions based on responses--

8. What about the Degree Requirements section, was that straightforward or were there aspects that I missed or that could be clarified?

Possible Probe Questions:

• What implications do you see for combination degree programs, where the program focus is shared with other disciplines, or for interdisciplinary degrees? What is your program like in this regard?
• Note questions based on responses--
9. The final area I looked at was career focus. [insert comment about how that task went in general] Is this accurate?

   Possible Probe Questions:
   - Has this focus changed much over time, by degree? How much of the focus is a product of student interest?
   - Evaluation is very practice-based. Are there other experiences for students in your program for any particular career focus?
   - Note questions based on responses--

10. Any other comments or ideas about this program profile?

   Possible Probe Questions:
   - Please tell me more about what you do to give students practical experience with evaluations?
   - Are there other recommendations or “Next Steps” that you would recommend?
   - Note questions based on responses--

   If that’s all for now, I’ll summarize this and send it to you in an email [hardcopy if that was preferred]

   I’ll used your comments to adjust my program profile to their final form and get started on the Classification Tool. At this time, I estimate I might have something for you in Month [or am not sure yet]

   I enjoyed talking with you today, let me know if you have any other thoughts!

   Good bye
RE-APPROVAL MEMORANDUM

Date: 5/14/2010

To: Michelle Chandrasekhar

Address: Tallahassee, FL
Dept.: EDUCATIONAL LEADERSHIP

From: Thomas L. Jacobson, Chair

Re: Re-approval of Use of Human subjects in Research
An Analysis of Evaluator Competencies in Selected Evaluation Degree Programs in the United States

Your request to continue the research project listed above involving human subjects has been approved by the Human Subjects Committee. If your project has not been completed by 5/11/2011, you must request a renewal of approval for continuation of the project. As a courtesy, a renewal notice will be sent to you prior to your expiration date; however, it is your responsibility as the Principal Investigator to timely request renewal of your approval from the committee.

If you submitted a proposed consent form with your renewal request, the approved stamped consent form is attached to this re-approval notice. Only the stamped version of the consent form may be used in recruiting of research subjects. You are reminded that any change in
protocol for this project must be reviewed and approved by the Committee prior to implementation of the proposed change in the protocol. A protocol change/amendment form is required to be submitted for approval by the Committee. In addition, federal regulations require that the Principal Investigator promptly report in writing, any unanticipated problems or adverse events involving risks to research subjects or others.

By copy of this memorandum, the Chair of your department and/or your major professor are reminded of their responsibility for being informed concerning research projects involving human subjects in their department. They are advised to review the protocols as often as necessary to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

Cc:
HSC No. 2010.4421
APPENDIX E: INFORMED CONSENT LETTER

Interview Consent—Script

<AFTER initial greetings…>

My name is Michelle Chandrasekhar. I am a graduate student under the direction of Robert Schwartz and Dr. Linda Schrader in the Department of Education Policy Studies, College of Education at Florida State University.

As we discussed in our email messages, I am conducting interviews of Ph.D. graduates from four well-known evaluation degree programs that have been defined as having maintained longevity in their programs and are associated with leaders in the field of evaluation. This interview is part of my dissertation study.

I would like to clarify how this interview will be conducted. If you have any questions concerning the research study or this interview, please ask me now or at the end of your interview or if you would like to contact Dr. Linda Schrader first, we can do that as well. Her contact information is (850)644-8780 or lschrade@mailer.fsu.edu.

- For the purposes of insuring accurate representation of our discussion, we will be recording our conversation on an audio-recording device. Transcripts will be composed and expunged of identifying information, will be made and stored digitally.

- While there are no foreseeable risks to participating or not participating in this interview, the information gathered may be of benefit to evaluation program directors and to the field as a whole.

- This interview contains open-ended and short answer questions and it will take approximately 45 minutes of your time.

- Participation in this interview is voluntary. You may choose not to participate and end the interview at any time with no penalty.

- The results of this interview may be published, but results will not identify respondents or their institutions.

- Information obtained during the course of this interview will remain confidential to the extent allowed by law.

- Any records (tape, digital, hardcopy) will be kept in a locked filing cabinet until researcher will have access to these recordings, which will be destroyed by 2012.

If you consent to participate, Florida State University’s IRB has requested
REFERENCES


Patton, M. Q. (2005). Diverse and creative uses of cases for teaching. In M.Q. Patton & P. Patrizi (Eds.) *New Directions for Program Evaluation, 105* (pp. 91-100). Wilmington, DE: Wiley Periodicals, Inc.


BIOGRAPHICAL SKETCH

Michelle M. Chandrasekhar

EDUCATION

Doctoral Candidate
Florida State University, Department of Educational Leadership and Policy Studies, Program in Educational Policy and Evaluation, Tallahassee, FL.
Dissertation Title: Socialization of U.S. Doctoral-Degree Students into Evaluation Professionals: The Use of Evaluator Competencies and Experiential Learning Strategies in Selected Programs. Committee Chairs: Drs. Linda Schrader and Robert Schwartz

M.S., Psychology
Florida State University, Department of Psychology, Program in Neuroscience, Tallahassee, FL.
Thesis Title: Strain Differences in Amiloride Suppression of Chorda Tympani Nerve Responses to NaCl and KCl in Sprague-Dawley and Fischer 344 Rats. Research Director: Dr. Robert J. Contreras.

College Teaching Certificate
Florida State University, Department of Educational Leadership, Program in Higher Education, Tallahassee, FL.

B.S., Psychology
University of North Carolina at Charlotte, Charlotte, NC.
Graduated with honors, magna cum laude.

WORK EXPERIENCE

4/11 – 10/11 Research Associate
Supervisor: Dr. Linda Schrader, Associate Professor, Department of Educational Leadership & Policy Studies, Florida State University.
Needs Assessment Commissioned by the FL Division of Blind Services as required by U.S. Department of Education, Rehabilitation Services Administration (RSA) Involved with literature reviews, method design, Institutional Review Board application, question and instrument design, data collection and analysis (focus group interviews, artifact analysis, data-base reports in Excel, NVIVO), and report generation.

11/06 – 1/10 Research Assistant
Supervisor: Dr. Barbara Gill, Director for Educational Research, Academic Affairs, Tallahassee Community College
Data analysis (DataWare house, SAS, SPSS, Excel), report generation, consultations, and instrument design and data collection for the Academic Affairs Office, the Lumina Foundation Achieving the Dream grant and institutional effectiveness strategic initiatives for accreditation (Southern Association of Colleges and Schools - SACs).

10/06 – 5/07 Graduate Research Assistant
Supervisor: Dr. Linda Schrader, Assistant Professor, Department of Educational Leadership & Policy Studies, Florida State University.
Assist with program evaluation and policy studies program degree coordination (Edit information materials, field email inquiries, research other program curricula, maintain committee minutes, etc.)

1/06 – 8/06  **Graduate Research Assistant**  
*Supervisor: Dr. Linda Schrader, Assistant Professor, Department of Educational Leadership & Policy Studies, Florida State University.*  
*Evaluation commissioned for the Howard Hughes Computational Biology Project, Florida State University.*  
Conduct annual evaluation of program as required by funding agency (survey sampled population & analyze data, focus group data collection & analysis, report generation).

1/01 – 3/06  **Coordinator, Academic Programs: Program for Instructional Excellence, Preparing Future Faculty**  
*Supervisor: Dr. Walt Wager, Director, Center for Teaching & Learning, Academic & Professional Program Services, Florida State University.*  
Program development and implementation of the university’s central instructional support for graduate student teaching assistants (Program for Instructional Excellence – PIE) via conferences, workshops, consultations, and the PIE Associate (discipline-based TAs) Program. Also responsible for the Preparing Future Faculty Program (PFF) which provides professional development opportunities for graduate students and postdoctoral fellows interested in academic careers.

- Provided Consultations & workshops for graduate students and departments for individual and department-wide teaching enhancement including special populations [such as performance skills (labs) and international TAs]
- Organizational responsibility for the university-wide Fall & Spring Teaching Conferences for graduate students [~500 annual participants from 67 departments & programs], including program development, selection of presenters, conference evaluation, etc. Conference includes a PIE Teaching Certificate developed in 2001 as a reflective activity to enhance conference participant experiences using an instructional design model.
- Recruited, trained [f2f and online], and provided mentoring for the PIE Teaching Associate Program, annually, 9-13 departments & their student are chosen to participate in the development of programs to enhance teaching and professional development programs in the discipline. Associates also trained as conference presenters and facilitators to act as university-wide Peer Mentors
- Administered the University’s annual Outstanding TA Award – advertising, committee training, resource development, recipient selection, event management for the awards celebration, etc.
- Created TA-Supervisor support services (BlackBoard-Resource site) and consulted with individual TAs as requested by department. Developed department based workshops as well as university-wide workshops (open to all instructors at FSU)
- Responsible for Preparing Future Faculty Program for graduate students in academic careers including resource development (via BlackBoard), providing workshops, recruitment, tracking and mentoring/consultations with PFF Fellows, program marketing, developing partner relationships (partner institutions and other FSU organizations such as the Career Center), and preparation of external reports.
- In addition to supervision of 15 PIE Associates, supervised two 0.5 FTE staff (1 A&P, 1 graduate student) in administrative tasks related to evaluation and implementation of conferences and workshops, preparation of reports, support for unit presenters, web-based resources, and department/university communication

- Resource development and information management responsibilities included oversight of university-wide email communication for PIE & PFF, preparation of flyers, brochures, and web materials, enhancement of current evaluation and administrative protocols [evaluation of unit goals, participant tracking via a mark-sense workshop evaluation form, report generation, campus communication/correspondence, file management, liaison with Graduate Studies Office, Career Center, & International Center.]

10/99 – 1/01 Coordinator, Academic Support Services: Biological Science Academic Advisor
Supervisors: Dr. Bruce Janasiewicz, Associate Dean, Division of Undergraduate Studies and Dr. Robert Reeves, Associate Chairman, Undergraduate Studies, Department of Biological Sciences, Florida State University.
Professional academic advising on department, college, and university degree requirements for undergraduate majors and prospective majors in the Department of Biological Science. Developed materials for advisees as well as tracking methods for the advising office. Supported other professional activities (Major fair, Career fair, Preview, Orientation, First Contact – freshman publication, Satellite Advisor Training Manual & Mentoring, etc.) through the home office of Undergraduate Studies. Accessed and updated student data on the Northwest Regional Database (FSMM, ZEDA – SASS) and tailoring the database displays to the needs of the department (also course equivalencies in transfer evaluation of credit). Responsible for the maintenance and content of the Undergraduate Studies portion of the Biological Science Department’s website, collection and processing of student major data.

5/98 – 8/99 Coordinator, Academic Programs: Assistant Director, 0.5 FTE
Supervisor: Dr. Lavon Gappa-Levi, Director, Program for Instructional Excellence (PIE), The Graduate School, Florida State University.
Member of the development and implementation team for university-wide and in-house Graduate Student Teaching Assistant Workshops, and sessions offered in the Fall/Spring College Teaching Conferences. Resource contact to FSU Graduate Coordinators across departments and Preparing Future Faculty (PFF) participants (includes other faculty in other institutions). Responsible for delivering workshops, evaluating workshop participant feedback, and the supervision and maintenance of office listservs, website, computer hardware and software, and desktop publishing needs (newsletters, programs, information brochures).

7/97 - 5/99 Other Personal Services: Preparing Future Faculty Coordinator, 0.25 FTE
Supervisor: Dr. Lavon Gappa-Levin, Director, Program for Instructional Excellence (PIE), The Graduate School, Florida State University
Contacted partner institution faculty and administrators, & communicated with FSU faculty and Fellows, organized Central Activities, oversaw budget reports, and developed activities to further PFF goals to enhance graduate student understanding of faculty roles in teaching, service, and advising related to post-secondary education.

8/91 - 5/91 Animal Surgical Facility Maintenance. 0.25 FTE
Supervisor; Stan Warmath, Department of Psychology, Program in Neuroscience, Florida State University.
Clean and supply (including sterilization of instruments and supplies, preparation and sterilization of Ringer's lactate) department's animal surgical facility. Maintained facility at standards for human surgery facilities (except for positive airflow).

8/90 - 12/97 **Graduate Research**
Supervisor: Dr. Robert Contreras, Department of Psychology, Program in Neuroscience, Florida State University.
Research on taste electro-physiology of chorda tympani nerves in rats. Preliminary study in cultural effects on capsaicin ingestion in humans.

5/89 - 7/90 **Executive Director, Durham County Group Home for Autistic Adults, Inc**
Supervisor(s): Board: Durham County Group Home for Autistic Adults, Inc
Responsible for four full-time, and two half-time staff, and five residents in all aspects of hiring, training, supervision, and group home management, including state and federal reports, payroll, and workman’s compensation. Instrumental in the procurement of ICF-MR and HUD grants for the development of a second 6-bed home.

**TEACHING EXPERIENCE**
5/06-7/06 **Introduction to Program Evaluation: EDF 5461 – 3 cr. hr., Department of Educational Leadership & Policy Studies, Florida State University.** Graduate-level core course for the Program Evaluation Certificate and various degree programs (~15 students). Teaching Assistant to Dr. Linda Schrader, updates to BlackBoard site, preparing class materials, attendance and minor student assessment. Responsible for delivering classes: *Drafting Evaluation Design & Logic Plans, Drafting a Management Plan, and Designing Quant/Qual. Research Designs and Instruments.*

Fall 00, 01,03 **First Year Experience (FYE): AMS 1363 – 1 cr. hr. Dean of Students (Student Affairs), Florida State University.** Instructor of Record for one section of 19-22 freshman first-year-experience, introduction to Florida State University and academe. Developed class discussions and activities for academic policies and procedures, campus life, personal responsibility issues, and career development. Course includes mentoring with a Peer Instructor (undergraduate student) and in 2001, a co-instructor from the Office of Financial Aid.

1/96 – 12/99 **Sensation & Perception: EXP 3202 – 3 cr. hr., Department of Psychology, Florida State University.** Instructor of Record for six semesters of Sensation and Perception, an upper-level majors only course (~30-75 students per section). Fully responsible for choice of text, preparing lesson plans and learning activities, assessment and evaluation of students, and determination of final grades. Responsible for supervising the concurrent graduate student lab instructors (3 semesters). One section taught at the Panama City Campus.

9/95 - 12/95 **Sensation & Perception Lab: EXP 3202L – 1 cr. hr., Department of Psychology, Florida State University.** Instructor of Record, responsible for preparation for the lab experience and student assessment.

9/93 - 5/95 **General Psychology: PSY 2012 – 3 cr. hr., Department of Psychology, Florida State University.** Instructor of Record for large lecture, liberal studies course (~200 students per section). Fully responsible for preparing lesson plans and
learning activities, assessment and evaluation of students, and determination of final grades.

5/93 - 8/93 **Lab Development, Department of Psychology, Florida State University.** Assistant to faculty in designing learning activities and equipment for the upper level Conditioning and Learning laboratory.

5/91 - 4/93 **Conditioning & Learning Lab: EXP 3422L – 1 cr. hr, Department of Psychology, Florida State University.** Co-developed and taught the upper level Conditioning and Learning Lab for a total of six semesters. Responsible for preparation for the lab experience and student assessment.

1/87 - 5/88 **Introduction to Psychology Lab Instructor, Department of Psychology, University of North Carolina, Charlotte.** Undergraduate instructor for Introductory Psychology 101-L Computer Simulation laboratory.

8/86 - 12/86 **Teaching Assistant, Department of Psychology, University of North Carolina, Charlotte.** Graded papers, and performed data analysis for Operant Conditioning & Behavior Modification faculty instructor.

**CONFERENCE PRESENTATIONS & RESEARCH PUBLICATION***

2008


2007

2006


2005

2004

2003
Session Presentation: Using Teaching Awards to Develop Graduate Students’ Vision of Teaching, Professional & Organizational Development Network 28th Annual Conference: Visions, Metaphors & Images, October 9th, 2003, Denver, CO.


2002
Session Presentation: Planning a Program for Teaching Enhancement: a Research-Based Model or Riding the Tide: A Systematic Approach to Improving Instruction, Professional and Organizational Development Network in Higher Education (POD) Annual Conference, October 10-13, 2002, Atlanta, GA. Co-Presenter: Dr. Walt Wager.

1996-1993


Panelist: Graduate Student Professional Development through Preparing Future Faculty, with Dr. Jim Slevin as Moderator, National Association of Graduate and Professional Students Annual Conference, October 1995, Miami, Florida.


Session Presentation: An Integration of Three Approaches to an Undergraduate Animal Conditioning and Learning Laboratory, Michelle Minear and Dr. Michael Rashotte, Florida State University, 16th Annual National Institute on the Teaching of Psychology, January 2-5, 1994, St. Petersburg, Florida.

TEACHING ENHANCEMENT WORKSHOPS PRESENTED

1998-2005 workshops for Teaching Assistants & Faculty through Program for Instructional Excellence Fall & Spring College Teaching Conferences, and Fall/Spring Workshops, Instructional Development Services, Office for Distributed & Distance Learning, Florida State University and its merged unit, the Center for Teaching & Learning in Academic & Professional Program Services.

Planning to Teach: instructional philosophies, planning instruction, teaching students to be collaborative, helping students stay honest, writing instructional goals & learning outcomes, principles of effective practice in teaching, using reflection to improve teaching, syllabus creation, and a 12-hr workshop series on instructional technology from a pedagogical perspective.

Teaching Techniques: lecturing and presentation skills, use of media in the classroom, increasing communication with email & WebBoard, active learning, interacting & communicating with students, incorporating research into the classroom, cultural differences (for international TAs).

Assessing Learning: grading issues, ethics in academe, evaluation of portfolios, classroom management, the role of liberal studies courses, university protocols, preparing for SACS—faculty and administrator training for the 2004 SACS Accreditation process.

Professional Development: professional development planning, research plans & statements, academic job search skills, time management, academic interviews, using videotape & observation to enhance instruction, and professional portfolios.

PROFESSIONAL AFFLIATIONS

American Evaluation Association (AEA), Student Member: 2003-current
Association for Institutional Research (AIR), Student Member: 2006-current
Professional and Organizational Development Network in Higher Education (POD), Member: 2001-2006
Southeast Evaluation Association (SEA), Member: 2003-current