

ABSTRACT

BOONYAPRAKOB, KORNSIRI. A Comparative Analysis of Longitudinal Studies of College Students' Intellectual Development. (Under the direction of Dr. Alan J. Reiman.)

Research literature on thinking indicates an association between thinking, behavior and judgment. Thus, it is critical that education support students' growth in ways that they can independently think, behave, and make judgments appropriately when confronting complexity in life. However, there is a need for more scientific and evidence-based approaches for fostering undergraduate students' intellectual growth.

The purpose of the research is to enlarge the literature evidence of how college students think, learn, know, and make judgments about ill-structured problems; and how education can help them develop intellectually.

This dissertation begins with a summary of connections among four cognitive-structural theories: Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reflection Model (1992), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt and Schroder's Conceptual Development Theory (1961). Second, two four-year longitudinal studies are described for undergraduate teacher education students at a university in southeastern, United States. These longitudinal studies report changes in conceptual development of the students as a result of attending the Deliberative Psychological and Professional Education program (DPPE). Third, through a comparative analysis, the study reports and compares effect sizes of the two teacher education studies and eight one-year to four-year longitudinal undergraduate college studies that are based on the four theories. Moreover,

it describes features of the programs that foster intellectual development among college undergraduate students. Finally implications are drawn for researchers, practitioners, and policymakers.

The analysis of the four theories suggests that the theories are connected in terms of their assumptions about the developmental process of cognitive structures and their assumptions about knowledge or information or problems.

Findings from the two studies suggest that design principles for the DPPE did foster the conceptual development of college students in teacher education. After participating in the program for four years, their conceptual level changed in positive directions. Effect sizes were very large (+1.07) for the first study and moderate (+.59) for the second study. These findings could have effective assessment of dispositions in teacher education as outlined by national accrediting bodies such as the National Council for the Accreditation of Teacher Education (NCATE). Comparative analysis across the ten studies indicates that college education positively fosters the students' intellectual development. As well, the magnitude of development, as indicated by effect sizes in the majority of the studies was found to be moderate (+.54, +.57, +.59, and +.60), large (+.73), and very large (+1.07, +1.34, and +1.43).

Six of the ten studies explored gender issue and found no statistically significant difference in trend and magnitude of development between female and male students. Findings among the studies suggest that programs can foster students' intellectual development when deliberately designed and implemented. Three studies examined relationships between students' major and intellectual development and found no

statistically significant relationships between them. No conclusive findings were found among four studies that investigated age and life experience in relation to intellectual development. Findings suggest an interpretation that students' intellectual growth occurs most consistently as a response to education. Two studies examined the relation between students' conceptual level in relation to their persistence and dropout and found no relation between conceptual level and the rate of graduation or dropout.

With the analysis of deliberative program features, the researcher suggests that features of programs designed to foster intellectual development should derive from an understanding of how students learn, what they should learn, and how to teach them.

**A COMPARATIVE ANALYSIS OF LONGITUDINAL STUDIES OF
COLLEGE STUDENTS' INTELLECTUAL DEVELOPMENT**

by

KORNSIRI BOONYAPRAKOB

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

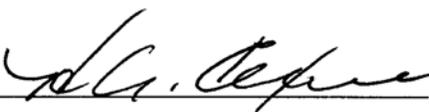
DEPARTMENT OF CURRICULUM AND INSTRUCTION

Raleigh

2002

APPROVED BY:



Chair of Advisory Committee

This dissertation is dedicated to my parents,

Tawee Sareepukkana and Ampond Sareepukkana,

and

all who highly value education.

BIOGRAPHY

Kornsiri Boonyaparakob was born in Bangkok, Thailand. She is a second of three children born to Tawee and Ampond Sareepukkana. She earned her B.A. degree majoring in English (1986) from the Faculty of Arts, Silpakorn University, Nakorn Pathom, Thailand; and her M.A. in Applied Linguistics (1991) from the Faculty of Sciences, Mahidol Univeristy, Bangkok, Thailand.

Between 1986 and 1995 she served as a civil servant in a position of program officer at the Language Institute, Testing and Training Service Division, Department of Technical and Economic Cooperation (DTEC) in Bangkok. During that time, she was granted a scholarship from the New Zealand government to study for a Diploma in Teaching of English as a Second Language (1992), at Victoria University of Wellington, New Zealand. Since 1996, she has served as a faculty member of the Department of Foreign Languages, Faculty of Science, Mahidol Univeristy. In 1999, she was granted another scholarship by the Royal Thai government to study for her doctoral degree in Curriculum and Instruction at the College of Education, North Carolina State University, USA. With the support from the Department of Foreign Languages, Mahidol University, she was granted a leave for her doctoral study and completed the degree in 2002.

ACKNOWLEDGEMENTS

During the process of writing this dissertation, I have been going through a series of assimilations and accommodations. The completion of this dissertation represents a significant “rite of the passage” in my life. However, without support from all concerned individuals, I would have never reached this point. Following are individuals and organizations whose support I deeply appreciate.

My advisory committee members include my chair, Dr. Alan J. Reiman, who is my major academic advisor. He not only models excellence as a professional educator, but he also possesses all the best dispositions of a great teacher. Without his guidance and support, the gap of my “zone of proximal development” would have been too great for successful completion of the dissertation. My appreciation also extends to his wife, Evelyn, who takes all possible occasions to provide my family and I with opportunities for food and fellowship in their home. Dr. Christopher M. Anson is a teacher whom I perceive as a “one-of-a-kind” person. His inspiring and insightful teaching provokes my enthusiasm to take a different stance in pursuing my teaching profession as an English teacher. Being in his class is an eye-opening experience. Dr. Ellen S. Vasu has engaged me in an active process of learning about “practical significance” of “statistical practice.” Her guidance has made me a better researcher. Dr. Herbert A. Exum, whose generous and peaceful personality has soothed my anxiety during the process of working on my dissertation, will never fully appreciate his support. Finally, Dr. Lois E. Huffman has shared care, empathy, support and understanding. She has inspired me to reach my

destination. She is “the one” who helped me “discover” my talent and ability to use alternative forms of teaching materials.

Although many names are not listed here, I would like to express my appreciation to all my teachers who supported me throughout my academic journey. I have always appreciated the patience, support and efforts of all my teachers.

A special thanks must be extended to all supporting organizations and staffs. The Department of Foreign Language, Mahidol University, Bangkok, Thailand granted me a temporary leave for my study, and the Royal Thai government provided essential financial support. As well, my appreciation extends to all my colleagues from both organizations for providing me with supportive services throughout the period of my study. My gratitude also goes to all support staff at North Carolina State University. This includes the staff in the Department of Curriculum and Instruction, College of Education, staff at the Learning Resource Library and at computer laboratories, and staff at D.H. Hill library.

My friends’ encouragement, love, sincerity, support, trust and belief in me has been invaluable. Their support during this difficult time made me feel connected and supported. I feel thankful to them all.

My father and my mother, Tawee and Ampond Sareepukkana, nurtured in me a high value of education, hard work, honesty, respect, trust, and choice to have a good life. Without them, I could have never lived my life up to this stage and become who I

am. My appreciation has never been adequate for their love, support and what they have done for me. My sister, Nawapond Sareepukkana, is always there for me, unconditionally. My brother, Tawal Sareepukkana and his family always cheer me up and uplift my confidence in getting through this academic passage. Last but not least, I feel thankful to my beloved husband, Ukadej, for his love and patience to share his life-long journey with me, and my precious daughter, Patarasorn, for her true love and faith in me. She has been my most important inspiration on this academic trip.

TABLE OF CONTENTS

	Page
LIST OF TABLES. x
LIST OF FIGURES..xi
CHAPTER I INTRODUCTION..1
Intellectual Development as an Aim of Higher Education.2
Statement of the Problem.6
Relationships among Intellectual Development Theories8
Longitudinal Research and Intellectual Development.. . . .	10
Deliberative Educational Programs and Intellectual Development.	11
Significance of the Study.	11
Research Questions.	13
Definitions of Terms.	15
Summary.. . . .	18
CHAPTER II REVIEW OF THE LITERATURE.	20
Importance of Intellectual Development in Higher Education and Teacher Education	22
Intellectual Development Theories and Teacher Education.	23
Relationship between Thinking and Teaching Behaviors.. . . .	27
Attempts to Change Teachers' Ways of Thinking	31
Historical Context for Research on Intellectual Development in Higher Education	36
Review of the Four Selected Theories	45
Scope of the Review.. . . .	45
Perry's Scheme of Intellectual and Ethical Development.. . . .	47
Research Validity and Reliability.	51
Research Studies Based on Perry's Scheme.	53
Baxter Magolda's Epistemological Reflection Model.. . . .	63
Research Validity and Reliability.	69
Research Studies Based on Epistemological Reflection Model.	69
King and Kitchener's Reflective Judgment Model.	73
Research Validity and Reliability.	78
Research Studies Based on Reflective Judgment Model.	79
Harvey, Hunt and Schroder's Conceptual Development Theory.. . . .	89
Research Validity and Reliability.	92
Studies Based on Conceptual Development Theory.	94
Review of Related Measurement Instruments	98
Measure of Intellectual Development	100
Measure of Epistemological Reflection.	102

TABLE OF CONTENTS

	Page
Reflective Judgment Interview.	104
Paragraph Completion Method.	106
A Proposed Connection Among the Four Theories	112
Review of Research on a Deliberative Psychological and Professional Education Program (DPPE).	119
CHAPTER III RESEARCH METHODOLOGY.	124
Research Design.	125
Methodology.	126
Research Participants.	127
Research Variables.	127
Measurement Instruments.	128
Criteria for Research Sample Selection.	128
Scheme for Data Encoding.	130
Scheme for Analysis of Program Characteristic	131
Data Analysis and Procedures	132
Research Question One.	132
Research Question Two.	132
Research Question Three.	133
Research Question four.	134
Threat of Validity.	135
Ethics and Human Relations.	136
CHAPTER IV RESULTS AND DISCUSSION.	137
Research Questions Two: Two Convenient Longitudinal Samples.	139
Participants and the Intervention.	139
Research Methodology.	141
Instrumentation.	142
Findings: Conceptual Change in Teacher Education Students in the Two Longitudinal Samples.	143
Discussion of Findings.. . . .	150
Summary of Two Convenient Samples.	152
Research Question Three: Relationships Across Longitudinal Sample Studies.	152
Part One: Quantitative Findings from Comparative Analysis Across the Ten Studies	157
Trends of Development.	158
Magnitude of Intellectual Change.	163

TABLE OF CONTENTS

	Page
Part Two: Qualitative Findings from Comparative Analysis Across the Ten Studies.	168
Gender Differences.	169
Types of Programs.	172
Students' Major Fields of Study.. . . .	176
Age and Life Experience.	177
Academic Achievement.	179
Summary of Findings from Research Analysis.	181
Research Question Four: Distinctive Features of Programs Designed to Foster Students' Intellectual Development.	182
Part One: Description of Programs in Samples.	183
Deliberative Psychological and Professional Education Program (DPPE).	183
Accent on Development Abstract Process of Thought (ADAPT).188	188
Ability-Based Learning Program.	190
Part Two: Analysis of Three Intervention Programs.	195
How Do the Students Learn?.. . . .	196
What Should the Students Learn?.	201
How Should Teachers Teach College Students..	204
Chapter Summary.	207
CHAPTER V SUMMARY, CONCLUSION AND IMPLICATIONS.	210
Summary.. . . .	210
Conclusions.	212
Common Assumptions and Connections Across Four Theories of Intellectual Development.	212
Findings from the Two Convenient Samples and Their Relation to Findings from Other Studies.	215
Features of Educational Programs Supportive of Students' Intellectual Development.	218
Implications of Findings	221
Theoretical Implications.	221
Research Implications.	226
Future Directions.	227
Theory.	227
Research.	229
Applications.	230
REFERENCES	233

LIST OF TABLES

	Page
Table 4.1 Conceptual level of pre-service teacher education students (Cohort 1988-1992) as indicated by Hunt's Paragraph Completion Method144
Table 4.2 Conceptual level of pre-service teacher education students (Cohort 1989-1993) as indicated by Hunt's Paragraph Completion Method146
Table 4.3 Proportion of students at different conceptual levels at freshmen and senior years148
Table 4.4 Proportion of students CL development based on directional trends149
Table 4.5 Students' intellectual stages at the beginning and at the end of each study as indicated by mean scores from different instruments.160
Table 4.6 Directions of students' intellectual development.161
Table 4.7 Magnitude of change in students' intellectual development.164

LIST OF FIGURES

	Page
Figure 2.1 Women's Ways of Knowing	41
Figure 2.2 Perry's Scheme of Intellectual and Ethical Development	49-50
Figure 2.3 Epistemological Reflection Model	66
Figure 2.4 Summary of Reflective Judgment Stages	77
Figure 2.5 Characteristics of position, stage, and level of development as indicated by scores of different instruments	110-111
Figure 2.6 Common features among four selected theories.	117
Figure 3.1 Scheme for data encoding.	130
Figure 3.2 Scheme for program analysis.	131

CHAPTER I

INTRODUCTION

Increasingly, higher education is a key for careers in the professions. Students pursuing education for a professional career expect to acquire knowledge about subject matter since it is generally believed to be a significant foundation for development of competence and professional performance (Rillero, 1998; Sanders & Morris, 2000). More importantly, scholars suggest that in addition to subject-matter knowledge, higher education, especially at the post-secondary level, should also foster general intellectual or cognitive competencies and skills (Erlich, 2000). This is due to the fact that unlike subject-matter knowledge, which can be forgotten unless frequently used in daily life or work after graduation, these intellectual or cognitive competencies and skills remain with graduates (Pascarella & Terenzini, 1991). Along with the acquisition of subject-matter knowledge and academic skills, the fostering of intellectual or cognitive competencies is widely viewed as an important focus in higher education (Baxter Magolda, 1987a, 1992a; Borko & Michalec, 1997; Flowers, Pascarella, & Pierson, 2000; Flowers, Osterlind, Pascarella, & Pierson, 2001; Kroll, 1992; Lazere, 1987; McMinn, 2001; Morgan, Morgan, Foster, & Kolbert, 2000; Ortiz, 1995; Pancer & Hunsberger, 2000; Pascarella, Truckenmiller, Nora, Terenzini, Edison, & Hagedorn, 1999; Reed & Kromrey, 2001; Swick, 1991; and Wubbels, & Korthagen, 1990).

Intellectual Development as an Aim of Higher Education

According to Pascarella and Terenzini (1991), intellectual or cognitive competencies and skills are known under such terms as critical thinking (Glaser, 1985; Lipman, 1988) reflective judgment (King & Kitchener, 1994), and conceptual complexity (Harvey, Hunt, & Schroder, 1961; Miller, 1981) to name just a few. They also note that these cognitive competencies and skills are believed to:

permit individuals to process and utilize new information; communicate effectively; reason objectively and draw objective conclusions from various types of data; evaluate new ideas and techniques efficiently; become more objective about beliefs, attitudes, and values; evaluate arguments and claims critically; and make reasonable decisions in the face of imperfect information. (Pascarella & Terenzini, 1991,p.114)

Higher education, in general, and teacher education, in particular, have been trying to create educated persons capable of weighing evidence related to complex real world problems (Astin & Astin, 2000; Kish & Sheehan, 1997; Pascarella, Truckenmiller, Nora, Terenzini, Edison, & Hagedorn, 1999; Royce, 1998; Thorndike, et al., 1989). In fact, in their synthesis of twenty years of research in United States, Pascarella and Terenzini (1991) found that college students make statistically significant gains both in factual knowledge and in general cognitive and intellectual skills as well as value, attitudinal, psychosocial, and moral dimensions. Such change is fostered in college environments that maximize students' subject-matter knowledge and intellectual growth. Another factor that leads to such a change is the intellectual, cultural, and interpersonal diversity and opportunities that are concentrated in college environments (Smith et al., 1997;

Tatum, Calhoun, Brown, & Ayvazian, 2000). Other sources of change include institutional characteristics, major course of study, and types of instruction. Flowers, Osterlind, Pascarella, and Pierson (2001) remarked that over the past 30 years, the studies in postsecondary education focused on what students learn during the undergraduate years, and the majority of these studies reported “statistically significant and sometimes sizable indications of growth or change” (p.565) in knowledge and cognitive development.

One of the rationales underlying a renewed commitment to both content knowledge and cognitive development in higher education is related to the fact that American society today faces increasingly complex problems. Astin (1999) observes:

There is mounting evidence that the quality of civic life and engagement in the United States has been eroding in recent years. This list of problems is a long one: shaky race relations, growing economic disparities and inequities, excessive materialism, decaying inner cities, a deteriorating infrastructure, a weakening public school system, and irresponsible mass media, declining civic engagement, and the increasing ineffectiveness of government, to name just a few (p.8).

Society now needs a new kind of citizen, one who is actively engaged in making a positive difference in the society and who serves as an effective social change agent. Thus, higher education now is being expected to play a role to “empower students, by helping them develop those special talents and attitudes that will enable them to become effective social change agents” (Astin & Astin, 2000 p.2). Baxter Magolda (1999) observes that since life in 21st-century is becoming increasingly complex, college graduates should be provided with the experience of transformation from “reliance on authority to complex ways of making meaning in which they are able to integrate

multiple perspectives and make informed judgments” (p.333). In other words, before taking in any particular information or knowledge available from any claimed authorities, students should learn how to form judgments about the information or knowledge through a process of reflective thinking, analysis and comparing of information or knowledge available from different sources.

Kegan (1994), a noted scholar of conceptual and interpersonal development at Harvard University, notes the increasingly complex mental demands of modern life, observing that conceptual and interpersonal development must become key features of educational programs. The expectation that young adults experience their emotions as inner psychological states is also a demand for the subordinating or integrating of the simple, categorical self (“I’m mad at my roommate.”) into a more complex context that relates to others (“I see why my roommate is upset, but I need to make sure he understands my views as well.”). In effect, Kegan (1994) finds that there is a hidden curriculum in higher education. This hidden curriculum is life, experience, new roles, and relationships, and students must be able to “identify their inner motivations, hold onto emotional conflict internally, be psychologically self-reflective, and have a capacity for insight” (Kegan, 1994, p. 27).

In the last hundred years, we have succeeded in recognizing the qualitative distinctions between the mind of the child and the mind of the adolescent. However, it still remains for us to discover how young adults and older adults change conceptually, interpersonally, and morally. The college experience as the beginning of adulthood is not an end point. Rather, it is a vast evolutionary expanse encompassing a variety of

capacities of the mind and heart. The analytical exploration of these capacities to meet the challenges and complexities of modern life continues to be a critical need. As well, it is essential to design and study curriculum that can promote development of these conceptual and interpersonal capacities of the young adult mind. If college graduates are to be successful and to be civically and morally engaged in today's pluralistic world, education needs to help them develop both knowledge and those intellectual or cognitive competencies and skills that are requisite for such engagement.

This dissertation examines an aspect of intellectual or cognitive competency in college students. Since the terms intellectual or cognitive competency have a broad meaning both theoretically and practically, this dissertation's scope of the terms is limited to its use according to four particular constructs. These constructs, which are explored in detail in chapter 2, include Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reflection Model (1992b), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt and Schroder's Conceptual Development theory (1961). The major aims of this dissertation are, first, to make connections among these four theories through an intensive analysis and comparison in terms of theoretical features and assumptions. Second, it aims report data from two convenient four-year longitudinal samples of teacher education students that were conducted at North Carolina State University and compare these studies to other longitudinal studies of undergraduate college students' intellectual development based on the four theories. Finally, the study describes and analyzes a Deliberative Psychological and Professional Education program (DPPE) intentionally designed to affect students'

conceptual change. This program was implemented in the two convenient longitudinal samples at North Carolina State University. Two other programs within the selected studies are also described and analyzed in order to provide insights into beneficial features of programs designed for fostering intellectual development.

Statement of the Problem

In today's pluralistic world, people are always confronting complex problems, some of which are regarded as ill-structured as opposed to well-structured problems. As pointed out by King and Kitchener (1994), any problem is ill structured if it cannot be described with a high degree of completeness or certainty. In fact, it may even be difficult to determine when a solution has been reached. Determining which political parties to choose, or judging the adequacy of different arguments for and against genetic engineering are just two examples of ill-structured problems. In contrast, the well-structured problem can be described with a high degree of completeness and has a high degree of certainty. Arranging a chronological order of processes for a scientific experiment in a laboratory is an example of a well-structured problem. The structure of problems is related to such factors as the availability of multiple perspectives, the amount of information from varieties of sources that is included and the varieties of forms of interconnections to other problems. As well, solving complex problems requires people to be aware of, open to, and to take into consideration the vast array of social factors such as religious beliefs, economic and social status, racial differences, ethnicities, and sexual orientation.

It is critical that college students are able to distinguish, detect, and filter information based on reliability, accuracy, truth, evidence and opinions among the vast arrays of information just mentioned. In addition, it is critical that they form judgments on what courses of action they should take and what commitments they should make. These judgments should be based on the preponderance of evidence relative to problems or contexts of the world they are confronting. Certainly, this is true for students in education. After all, they are responsible for teaching the younger generation to be thoughtful and civically engaged citizens. Teachers' judgments, behaviors and actions, and commitments about what to teach, how to teach, how to help children think and learn, for example, greatly influence the future of a nation.

As noted in the introduction, in addition to subject-matter knowledge, higher education is focusing on students' intellectual development. Faculty and administrators are interested in how to equip graduates with tools to confront complex real-world problems and take responsibility for actions based on their judgments in order to make changes in societies. With such an aim for education, research has been conducted to understand how college students think, learn and come to know. However, with the development of varieties of research constructs and theories in the arena of intellectual development, there is a concomitant need to examine empirical trends among the array of research based on these theories of intellectual development. In addition, there is a need for better understanding of the trends or common characteristics of deliberative educational programs designed to promote developmental change.

This study, consequently, aims to explore connections among some selected theories, analyze and compare research evidence from longitudinal studies, as well as examine characteristics of educational programs designed to promote intellectual change. Research questions of this dissertation are now elaborated.

Relationships Among Intellectual Development Theories

Different terms and definitions have been used regarding constructs and theories of college students' intellectual development. As noted by Pascarella and Terenzini (1991), "general intellectual or cognitive competencies and skills go by a number of different names" such as "reasoning skills, critical thinking, intellectual flexibility, reflective judgment, cognitive complexity, and so on" (p.115). Each of these terms has its own scope, limitation, and focus. Each theory also employs different measurement instruments. Some efforts to synthesize these theories have been done by Hofer and Pintrich (1997). Relating intellectual development to the term epistemology, Hofer and Pintrich note that epistemological theories or theories concerning beliefs about knowledge involve Perry's Scheme of Intellectual and Ethical Development (1970), women's ways of knowing (Belenky, Clinchy, Goldberger, & Tarule, 1997), Epistemological Reflection Model (Baxter Magolda, 1992b), Reflective Judgment Model (King and Kitchener, 1994), argumentative reasoning (Kuhn, 1991), and epistemological beliefs (Schommer, 1990). Hofer and Pintrich (1997) suggest that in all research based on these epistemological theories, "there is very little agreement on the actual construct under study and the dimensions it encompasses" (p.89). In addition, these studies neither address the issue of whether epistemological beliefs are domain specific or how such

beliefs might connect to disciplinary beliefs, nor make any linkages to other constructs in cognition and motivation. Furthermore, the research makes no attempts to conceptually integrate the early Piagetian-framed developmental work on epistemological beliefs to other cognitive approaches such as theory of mind or conceptual change. In their analysis of research and theories on change of traditional-aged students in college, Pascarella and Terenzini (1991) include Perry's Scheme of Intellectual and Ethical Development (1970), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt, and Schroder's (1961) Conceptual Development Theory under cognitive-structural theories. These theories concentrate on "the cognitive structure individuals construct in order to give meaning to their worlds" (p.27), and suggest students' thinking and reasoning about issues. For example, the conceptual development theory focuses on the structure of experience that affects changes in concepts due to the mediation between environmental stimuli and individuals' responses to the stimuli. In other words, it explains stages of development of concepts which occur as a result of individuals' exposure to new environment and the requirement of environment for individuals' responses to it. The theory "is relevant only to the stages of development of concepts that generate the learned responses" (p. 93), not what is being learned.

Evidence indicates that both epistemological development and cognitive-structural theories are connected. In this dissertation, Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reasoning Model (1992b), King and Kitchener's Reflective Judgment Model (1994) and Harvey, Hunt and Schroder's Conceptual Development Theory (1961) are explored for a better

understanding of potential linkages between them in order to nourish future research in the field.

Longitudinal Research and Intellectual Development

As a result of the lack of linkage among intellectual development theories, research studies rarely link their results to other promising studies of intellectual development. The core of the intellectual constructs, however, centers on describing the process of personal development. Many researchers of intellectual development have recommended longitudinal studies (Brabeck, 1984; Schmidt, 1983). This is due to the fact that evidence from comprehensive longitudinal studies could contribute to better understanding of trends in personal intellectual development. In other words, longitudinal studies allow opportunity for researchers to follow up the same individuals in terms of their change over time in the predicted direction (Brabeck, 1984). Although longitudinal research is known to be difficult and time-consuming to implement, it is “essential for exploring problems in human development” (Gall, Borg, & Gall, 1996, p. 376). King and Kitchener (1994) noted that longitudinal data are “essential for the validation of any developmental model because they indicate whether people change in a manner consistent with the description and sequence proposed by the model. Further, they allow a stricter test of the impact of educational experiences than do cross-sectional studies (at least among the particular cohorts tested)” (p.126).

Examination of longitudinal studies on intellectual development of college students in teacher education and other professional disciplines has never been explored.

This dissertation intends to fill this gap in the literature.

Deliberative Educational Programs and Intellectual Development

Included as research samples for this dissertation are two convenient four-year longitudinal studies of teacher education students carried out at North Carolina State University. A Deliberative Psychological and Professional Education program (DPPE) for teacher education students is intentionally designed to affect students' conceptual change. It is important either to understand trends or to establish common characteristics of programs that could foster or hinder students' development. Examination of characteristics of educational intervention can guide further development of programs and curriculum, not only in teacher education, but also other academic disciplines. Two other studies selected for comparative analysis in this dissertation also investigate students' development as a result of intervention programs. These programs are compared to abstract common characteristics for a better understanding of factors and program elements that promote intellectual development.

Significance of the Study

As just mentioned in the problem statement, there is a critical need for better understanding of how college students think, learn, come to know, as well as to make judgment about complex real world problems. As well, there is a need to explore and link available theories related to intellectual development in order to better understand them. There is also a need to establish a body of evidence to support those theories. Furthermore, there is a need to understand the effects of deliberative educational

programs on undergraduate students' development. This study explores both program design and the outcomes of learning process. This area has received very little attention. This is true across a broad spectrum of professional preparation programs including teacher education (Franklin, 1995; Hunsaker & Johnston, 1992; Middleton, 1999; Placek & Smyth, 1995; Sprinthall & Scott, 1989). This study will respond to this critical gap in the scientific literature.

This research is significant for four reasons. First, the identification of common assumptions across selected theories of intellectual development proposed by this research could serve as a foundation for further inquiring into intellectual development. Second, it reports the results of two convenient samples of four-year longitudinal studies (1988-1992 and 1989-1993) of a Deliberative Psychological and Professional Education program (DPPE) for teacher education students. The two convenient longitudinal samples will add evidence to existing research in related academic disciplines. In addition, the comparative analysis across longitudinal research studies in college students' intellectual development will provide insights into better understanding of quantitative and qualitative development of the students. The quantitative comparison of effect sizes among studies, for example, will provide information about the trends and magnitude of students' intellectual development. Finally, the examination of common characteristics in deliberative educational programs could enhance curriculum design in higher education aimed at fostering undergraduate students' conceptual development. Implications will also be drawn for higher education and teacher education.

Research Questions

In accordance with these stated objectives, this dissertation aims to answer the following four research questions:

1. What are common assumptions and connections across selected theories of intellectual development?
2. What are findings from the convenient samples of the two four-year longitudinal studies (1988-1992 and 1989-1993) that incorporated the Deliberative Psychological and Professional Education program (DPPE) for teacher education students?
3. What are the relationships between the findings from the two convenient samples and other longitudinal samples based on the connected theories of intellectual development?
4. What are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

As noted earlier, there is a lack of linkage among varieties of constructs and theories of intellectual development. The first research question is supposed to provide information on common assumptions and connections across selected theories of conceptual development. In doing so, four selected theories including Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reasoning Model (1992b), King and Kitchener's Reflective Judgment Model (1994) and Harvey, Hunt and Schroder's Conceptual Development Theory (1961) are explored in

detail and compared. The selection of these theories is based on the criticism about the lack of linkage among theories as remarked by Hofer and Pintrich (1997). In addition, it is based on the connections among these theories that the researcher perceives from primary review of related literature. Details are discussed in chapter two.

To understand the empirical trends among the array of research based on intellectual development theories, the second and third questions are posed. To answer these questions, results of two convenient four-year longitudinal samples that examined conceptual development of teacher education students at North Carolina State University are reported. Then, these research results are compared to results of other selected one-year to four-year longitudinal research studies based on the four selected theories. The comparative analysis includes both quantitative and qualitative aspects across all the study samples.

Question four is raised to address the need for better understanding of trends or common characteristics of deliberative educational programs designed to promote intellectual change. The question is answered through the analysis of characteristics of the deliberative educational programs.

Some terms are used throughout this dissertation. It is important to provide definitions of these terms at the onset. The next section provides definitions of terms that have particular relevance to this dissertation.

Definitions of Terms

Following are key terms employed in this research.

Change and Development

The terms change and development are distinguished in some literature (Pascarella and Terenzini; 1991). In this research, both terms are used interchangeably. As for Pascarella and Terenzini, change is regarded as a descriptive, value-free term, which refers to “alterations that occur over time in students’ internal cognitive or affective characteristics” (p. 16). Change accounts for either quantitative or qualitative regressions and progressions. Development is characterized as “changes in an organism that are ‘systematic’, [organized, and] successive... and are thought to serve an adaptive function, i.e., to enhance survival” (Lerner, 1986, p.41). Other characteristics of development include physical maturation, environmental influences and interaction of the individual and the environment as causes of interpersonal changes. In addition, by means of differentiation and integration, there is a presumption of growth, or the potential for growth toward maturity, which is “to be valued and pursued as a desirable psychological and educational end, perhaps even as a moral end” (Pascarella & Terenzini, 1991, p. 16).

Reflective Thinking

As mentioned previously, higher education is increasingly interested in fostering a kind of reasoning in which the student accepts the inevitable uncertainty of knowing while forming judgments in which they are “reasonably” certain. How is such reasoning

similar to John Dewey's notion of reflective thinking? How is such reasoning similar to more contemporary constructs of critical thinking? Why is such reasoning important?

John Dewey (1933), a noted educational scholar and philosopher at the turn of the century, provided an important foundation to our understanding of reflective thinking. He observed that true reflective thinking is engaged only after there is a recognition or awareness that a problem exists. Such problems, in Dewey's estimation, involve controversy or doubt. Dewey noted that these kinds of problems require a person to weigh evidence before forming a judgment, what he called a reflective judgment.

In addition, Dewey (1933) observed that the problem solver who is engaged in reflective judgment must evaluate the potential solutions to the problem in light of the preponderance of evidence or existing information that is available. Typically, such information is incomplete. Often, it may be unverifiable. Consequently, other criteria must be used. How coherent is the argument? What is the goodness of fit between the complex problem and the proposed solutions? How does the solution weigh the competing concerns of persons affected by the problem? Such criteria are formed as person reasons about the problem. As King and Kitchener (1994) note, "reflective thinking requires the continual evaluation of beliefs, assumptions, and hypotheses against existing data and against other plausible explanations of the data. The resulting judgments are offered as reasonable integrations or syntheses of opposing points of view"(p.7). In summary, reflective thinking is involved when there is an awareness that a problem exists or when there is uncertainty about the solution. A problem solver then

needs to open to consider multiple perspectives, weigh evidence, evaluate beliefs and assumptions, and form hypotheses against existing data and other plausible explanations.

Conceptual Thinking

Definitions of conceptual thinking, such as that provided by Hunt (1961), argue that conceptual thinking is a process of inquiry or problem solving that 1) shows increasing tolerance for ambiguity, 2) relies on criteria, and 3) is sensitive to interpersonal contexts. This view of conceptual thinking acknowledges that epistemic assumptions or assumptions about knowledge play a central role in forming judgments about complex problems. How is conceptual thinking different from critical thinking? Lipman (1988, p.39) suggests that critical thinking is “skillful, responsible thinking that facilitates good judgment because it (1) relies on criteria, (2) is self-correcting, and (3) is sensitive to context.” However, such definitions of critical thinking tend to focus on inductive and deductive logic skills. While the critical thinking approach shares some of the properties of reflective thinking and conceptual thinking, it is more limited in its approach because it tends to consist primarily of a set of skills or principles that one can apply in order to more effectively solve a problem. However, it fails to acknowledge the importance of epistemic assumptions about knowledge that not all problems can be solved by the application of the set of skills. This assumption is suggested in the work of Dewey, King and Kitchener, and Hunt. As well, these epistemic assumptions are fundamentally or qualitatively different for children, adolescents, and adults. In fact, it is only in adulthood that individuals hold epistemic assumptions that lead to optimal reflective thinking (King and Kitchener, 1994).

Epistemological Beliefs or Theories

Hofer (2002) refers to personal epistemology, or epistemological beliefs or theories as:

Ways the individual develops conceptions of knowledge and knowing and utilizes them in developing understanding of the world. This includes beliefs about the definition of knowledge, how knowledge is constructed, how knowledge is evaluated, where knowledge resides, and how knowing occurs. (p.4)

More terms related to the four theories are discussed in the review in chapter two.

Summary

In short, this research is a comparative analysis of one-year to four-year longitudinal studies of college student's intellectual development based on four selected theories: Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reasoning Model (1992b), King and Kitchener's Reflective Judgment Model (1994) and Harvey, Hunt and Schroder's Conceptual Development Theory (1961). The need for the study is to supplement and coordinate findings from the existing research for better understanding of how students develop intellectually.

Research questions include:

1. What are common assumptions and connections across selected theories of intellectual development?
2. What are findings from the convenient samples of the two four-year longitudinal studies (1988-1992 and 1989-1993) that incorporated the

Deliberative Psychological and Professional Education program (DPPE) for teacher education students?

3. What are the relationships between the findings from the two convenient samples and other longitudinal samples based on the connected theories of intellectual development?
4. What are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

Chapter two provides a review of the literature related to the research questions. Thus, chapter two first addresses the importance of intellectual development in higher education and teacher education. Second, it provides historical context for research on intellectual development in higher education. Third, it reviews four selected theories including Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reflection Model (1992b), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt, and Schroder's Conceptual Development Theory (1961). The review includes the description of the theories, validity and reliability issues, and literature on research findings based on these theories. In addition, it reviews measurement instruments used to assess theoretical assumptions. Fourth, the chapter proposes connections among these theories. Finally, since the two convenient samples employed a Deliberative Psychological and Professional Education program (DPPE) intentionally designed to foster intellectual development, relevant research literature on the program is reviewed.

CHAPTER II

REVIEW OF THE LITERATURE

As noted in chapter one, this dissertation aims to examine research studies on college students' intellectual development from longitudinal studies in order to better understand how higher education could help effectively develop their graduates' developmental growth.

This dissertation attempts to demonstrate some connections among four selected theories of intellectual development through an exploration and identification of common assumptions among them. Then, it reports the results of two convenient samples of four-year longitudinal studies (1988-1992, and 1989-1993) of a Deliberative Psychological and Professional Education program (DPPE) designed to monitor development of teacher education students at North Carolina State University. The two convenient sample studies were carried out based on one of the four theories. It also compares the results of these two convenient samples to eight selected longitudinal studies in undergraduate higher education based on the four theories. Finally, it describes and analyzes distinctive features of education programs designed to foster students' intellectual development in undergraduate programs.

The scope of this dissertation is limited to longitudinal studies in undergraduate higher education. It emphasizes how thinking, learning, and knowing in late adolescence and young adulthood change over one to four years. The general presentation of this dissertation is grounded in the context of teacher education because the two convenient

samples are studies in teacher education. However, the scope of this dissertation is generalized to undergraduate higher education. This dissertation raises four research questions:

1. What are common assumptions and connections across selected theories of intellectual development?
2. What are findings from the convenient samples of the two four-year longitudinal studies (1988-1992 and 1989-1993) that incorporated the Deliberative Psychological and Professional Education program (DPPE) for teacher education students?
3. What are the relationships between the findings from the two convenient samples and other longitudinal samples based on the connected theories of intellectual development?
4. What are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

Chapter two presents review of related literature following this outline.

1. Importance of intellectual development in higher education and teacher education.
2. Historical contexts for research on intellectual development in higher education.
3. Description of four selected theories including Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological

Reflection Model (1992b), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt, and Schroder's Conceptual Development Theory (1961). The review includes the description of the theories, validity and reliability issues, and literature on research findings based on these theories. In addition, it reviews measurement instruments based on these theories.

4. Proposed connections among the four theories.
5. Related literature on the Deliberative Psychological and Professional Education program (DPPE).

Importance of Intellectual Development in Higher Education and Teacher Education

Why is research on intellectual development important to higher education and teacher education? Today USA is in critical need of teachers who are competent in instruction and subject-matter knowledge while also serving as an agent of social justice. Hallinan and Khmelkov (2001) report a ten-year evaluation of proposed changes to teacher programs in terms of levels of training and structure of the teaching profession since 1986. They note that the proposed reform movement aims to “improve the quality of teacher education by preparing teachers to work in the rapidly changing social environment of contemporary schools and to meet the complex learning needs of students” (p.177). This aim is due to societal change, which has directly affected students and teachers. Such change has led to many problems in students' lives that leave them ill-prepared to be responsible for learning in school. As well, due to the lack of teachers at present and a crisis in the next decade, when “more than two million teachers

must be hired” (Hallinan & Khmelkov, 2001, p.177), there has been a decrease in standards for hiring teachers to meet the shortage. Most importantly, they note that many pre-service and in-service teacher training programs fail to prepare teachers to meet such dramatic societal change.

Today’s society critically needs teachers who are able to solve multiple kinds of complex problems in their teaching. It is, consequently, necessary that teacher education should serve to promote student intellectual development. Why is such development important? How does it affect teachers’ behaviors and classroom practice? The following section provides answers to these questions.

Intellectual Development Theories and Teacher Education

As mentioned in chapter one, general intellectual or cognitive competencies and skills are known under different names such as reasoning skills, critical thinking, and reflective judgment. It is, consequently, sensible to refer to the history of related theories and practices. To begin, it is worth mentioning the theory of reflection and reflective thinking introduced by John Dewey (1933) since it is generally accepted that his theory is still influential in education today (Berman & Davix-Berman, 1995; Starnes, 1999; and Woodhouse & Knapp, 2000).

Historically, Stanley (1998) observes that although the notion of reflection had been first outlined by Dewey (1933), it was not until the 80s that “the concept and practice of reflective teaching gained credence and undergone widespread discussion in Western education systems” (p. 584). As well, Hatton and Smith (1995) remark that

Dewey is acknowledged as “a key originator in the twentieth century of the concept of reflection” (p. 33). Dewey believed the type of thinking that should be an aim of education and teacher education was reflective thinking. Dewey (1933) distinguished reflective thinking from simple rote thinking. He observed that reflective thinking involves: “1) a state of doubt, hesitation, perplexity, mental difficulty, in which thinking originates, and 2) an act of searching, hunting, inquiring, to find material that will resolve the doubt, settle, and dispose of the perplexity” (Dewey, 1933, p.12). Problems that stimulate reflective thinking are those that perplex and challenge persons to realize the uncertainty of belief, that demand the persons to solve them, and that the solutions inquire the persons to search for evidence. Dewey also noted that “No one can tell another person in any definite way how he should think, any more than how he ought to breathe or to have his blood circulated” (p.1). This is generally acceptable. Thinking happens at all times in our heads but no one knows what is going on in another’s head. Dewey, however, believed that the person can optimize the effectiveness of the operations of their mental process namely thought. In order to do so, the person needs to have a will to make change for better, to understand the true nature of ways of thinking as well as to understand rationales underlying how one way of thinking is better than others.

This notion is reflected through his statement:

The person who understands what the better ways of thinking are and why they are better can, if he will, change his own personal ways until they become more effective; until, that is to say, they do better the work that thinking can do and that other mental operations cannot do so well.

(Dewey, 1933, p. 1)

Why did Dewey posit that reflective thinking be the primary aim of teacher preparation programs? According to Goodman (1984), Dewey “warned that if programs emphasized only technical expertise and failed to help students understand the relationship between theory and practice, the growth of future teachers would be stunted and the education of children thus impaired” (p.9). In addition, Hatcher and Bringle (1997) referred to Dewey’s awareness that experiences would result in learning only if they are reflected upon. They further elaborated that Dewey regarded experiences as a source of controversy, and “if the controversy is not reflected upon, it can be a misleading, even harmful experience, which produces a lack of sensitivity and responsiveness in the learner” (p. 153).

Others agree with Dewey and built upon his notion of reflection. David Kolb (1981), for example, included reflection as an element of the experiential learning cycle because he considered reflection as a link between the concrete and abstract. Reiman and Thies-Sprinthall (1998) noted that in 1983, Sprinthall and Thies-Sprinthall incorporated reflection as an essential condition among five interacting conditions, which foster adult psychological growth. Reiman and Thies-Sprinthall (1993) developed “guided” reflection to help promote the development of mentor teachers. Hatton and Smith’s (1995) literature review of research on reflection in teacher preparation programs revealed that “reflection” and “critical reflection” have been increasingly claimed as a goal in many teacher preparation programs. Referring to Zeichner’s (1983) idea, Wubbles and Korthagen (1990) noted that reflection is regarded as “a key concept within the inquiry-oriented paradigm of teacher education” (p.19).

The concepts of reflection, reflective thinking, and reflective practice have been broadly applied for research and classroom situations in teacher education (Call, 2000; Day 1993; Griffiths & Tann, 1992; Matthew & Jessel, 1998; Mewborn, 1999; Mills & Satterthwait, 2000; Norton, 1997; Pultorak, 1993; Raines & Shadiow, 1995; Stewart, 1994; Stuessy & Naizer, 1996; Woodward, 1998) and in higher education (Knight, 1996; Kroll, 1992).

Theories that are selected to focus in this dissertation are related to the concept of reflective thinking that Dewey (1933) introduced. For example, in the critique of the study by Perry and his associates (1970/1999), Perry noted that their study drew its assumptions and method from contemporary philosophical and psychological theories and procedures. As for philosophical context, the study shares the assumptions of modern contextualistic pragmatism including ideas from Dewey. This assumption is related to purpose as a control or guide for thinking. In their creation of the Reflective Judgment Model, King and Kitchener (1994) relied on the concept of reflective thinking as posited by Dewey (1933, 1938). King and Kitchener emphasized that only adults and not children hold the epistemic assumptions that allow for true reflective thinking. In addition, individuals summon reflective thinking when “there is awareness of a real problem or when there is uncertainty about a solution” (p.8). Relating to the theories about reflection and reflective thinking, conceptual development theory is one that clusters similar concepts and elements of thinking. As elaborated by Hunt (1977), conceptual level (CL) is “a characteristic based on a developmental personality theory

(Harvey, Hunt, & Schroder, 1961) that describes persons on a developmental hierarchy of increasing conceptual complexity, self-responsibility, and independence” (p.78).

The development of the characteristics as described by reflective thinking and related concepts developed and called using different terms such as reflective judgment and conceptual complexity development are clearly those that college graduates need to possess to successfully survive in today’s world. Let us now turn to evidence illustrating connections between conceptual complexity, reflective thinking and behaviors, particularly teachers’ instructional behaviors.

Relationship between Thinking and Teaching Behaviors

Like any other academic profession, teachers’ performance is predicated on a thoughtful and well-designed professional preparation program. The preparation should also be based on solid ground of theories. Reiman (1999a) notes that there are two contradictory views on the importance of preparation for the teaching profession. On one hand, there are calls for more rigorous preparation programs. On the other hand, preparation for teachers is regarded as unnecessary.

Formal teacher education in the eyes of some of its kindest critics is a lumbering giant, unresponsive to the emerging needs of a knowledge-based economy. Less generous cynics claim that the nation’s growing need for competent new teachers can only be provided by an influx of well-meaning liberal arts graduates untainted by schools of education. Teaching comes naturally, all it really requires is kindness and missionary zeal.

(Reiman, 1999a, p.94)

He notes, however, the extensive research of Jeannie Oakes (1985), which shows “there is ample evidence that untrained and kind teaching can lead to significant pedagogical harm” (p.94). Reiman (1999a) argues for a robust theory for teacher education and the inclusion of a general conception of human growth as a needed theoretical framework for teacher education. He suggests that this theoretical framework includes at least three cognitive-developmental domains—conceptual, ego, and moral. This is based on the rationale and research findings that “teaching performance is a function of complex intellectual processes, and that cognitive-developmental theory could provide a framework for systematically implementing and assessing professional programs” (p.102). Sprinthall, Reiman and Thies-Sprinthall (1996) confirm a large (and growing) body of research evidence that consistently indicates that various domains of cognitive development do predict behavior in complex human helping situations such as teaching. This raises the question: “What are the relationships between teaching performance and conceptual level?”

Miller (1981) analyzed over 60 studies that used Hunt’s version of the Paragraph Completion approach (PCM) and Schroder’s version (PCT) to measure cognitive-stage complexity. Comparing individuals at high conceptual level (CL) to those at low CL, he found that individuals with high CL show significantly more prejudice reduction, more empathy, more internal locus of control, longer decision latencies, greater use of nondirective styles and autonomy, more interdependent styles, and superior communication and information processing skills. A few studies that were included in the analysis examined teachers’ and counselors’ CL. Evidence from these studies

suggests that teachers and counselors at different CL display consistent differences in teaching styles, and display different levels of empathy. Furthermore, evidence indicates a positive relationship between CL and flexibility and adaptability of teaching style.

Another case in point, which shows the relationship between teachers' thinking and their teaching behaviors, is a study applying Perry's Scheme of Intellectual and Ethical Development (1970). Anson (1989) conducted exploratory research on writing teachers' response styles and their ways of knowing. He found writing teachers to respond differently to students' writing. Such differences in terms of the content of responses to student essays were found to clearly reflect the different epistemological assumptions categorized in Perry's scheme. In addition, each teacher's response style was found to be consistent regardless of the differences among students' essays that represented different ways of students' thinking. The consistency reflects teachers' assumptions about knowledge of the world and about learning to write. In other words, it reflects teachers' positions in Perry's scheme. Among sixteen writing teachers who participated in the study, three-fourths were categorized as dualistic responders, the rest were categorized as relativist responders and reflective responders. Dualistic teachers' responses emphasized surface mechanic and grammar and on writing aspects that can be objectively considered right or wrong. The tone of comments reflects teachers' authority in making judgment based on standards for correct and incorrect ways to write an assignment.

As for relativist responders, Anson (1989) describes:

In contrast to dualistic responders, these teachers seemed entirely unconcerned with giving the students anything more than a casual reaction, as if this is the only kind of response that can have any validity in a world where judgment is always in the eye of the beholder (p. 349).

Although comments from these teachers require student writers to inquire more about the content of their writing, the tone of the comments imply that the teachers themselves were in another realm. Suggestions from the relativist responders identify their personal preference and whatever decision to make about the change or revision in the writing depends on students' own judgment.

Reflective responders are those teachers who emphasize giving feedback on ideas and textual decisions, as well as providing their personal reaction. Reflective teachers suggest alternatives for revision. These suggestions, however, do not dictate what the students have to do for their revision. Recommendations of teachers for any particular alternative of revision are expressed in connection to general readers rather than teachers' personal preferences. In addition, the content of reflective responses clearly inform of writers' responsibility to rethink their essays through exploration of the writers' worldview prior to making decision for revising the essays. The comments usually encourage students' enquiry. Comparing the three positions of responders, Anson remarks that "a reflective style provides comments that were simultaneously tentative and goals-driven" (p.353). While reflective responders suggest alternatives for a potentially better text, they encourage student writers to weigh those alternatives and make decisions of revision on their own. The ultimate goal for reflective responders to provide such

responses is for the student writers to understand that the revision task is a part of the learning process for them.

The examples described indicate the relationship between teachers at different levels of thinking and their behaviors. These research findings and others bring about ideas that teachers should be prepared in a way that their thinking is optimally developed so that they can teach more effectively. The next sections provide evidence of attempts that have been made to develop teachers' ways of thinking.

Attempts to Change Teachers' Ways of Thinking

Realizing that some ways of thinking are better than others and that ways of thinking are associated with behaviors, teacher educators have tried to help college students in teacher education think in ways that are generally considered better than others or even best. As a result of attempts to understand how persons think and what ways of thinking are better than others, theories on development of reasoning and judgment have been grounded based on data collected from different groups of individuals (Baxter Magolda, 1992b; Belenky, Clinchy, Goldberger, & Tarule, 1997; Harvey, Hunt, & Schroder, 1961; King and Kitchener, 1994; Pascarella & Terenzini, 1991; Perry, 1970; Piaget, 1975).

Although these theories take different focuses, their common characteristic is that they try to distinguish ways of thinking, and propose factors that support or hinder more complex ways of thinking. An existing theory such as Piaget's theory about the development of intellectual reasoning and the equilibration of cognitive structures as a

way of learning and thinking, is explained based on observation from behaviors and verbal and non-verbal expressions of research participants at different points in life (Piaget, 1970, 1975).

One of the major factors that influence ways of thinking, as posited by many theories, is associated with external stimuli or environment. Examples include Lev Vygotsky's notion that social, cultural and historical forces affect development (Davydov, 1995). For a person to become who he or she is, his or her development takes place within the boundary of environment of social, cultural and historical realm where they belong. In other words, cognitive skills and patterns of thinking of individuals are the products of interactions between individuals and the history of the society in which the individuals are reared. Vygotsky explained that lower or elementary mental functions are natural mental abilities that are genetically inherited but higher mental functions develop through social interaction, being socially or culturally mediated (Goldfarb, 2000). Consequently, persons exposed to different environmental stimuli think, learn, and come to know differently. Another example includes Harvey, Hunt and Schroder's theory of conceptual systems and environmental fit (1961), and Hunt's (1971) matching models. According to this theory, people at different levels of conceptual complexity develop differently in different environmental stimuli. On one hand, a person at a lower level conceptual complexity thinks in a concrete way. In order for the person to become an abstract thinker, the person needs structured knowledge and guidance to a certain point before he or she can become an independent thinker and learner. On the other hand, a person at higher level of conceptual complexity thinks in an abstract way. If the person is

provided with a lot of concrete structured knowledge, there would be no challenge for him or her to develop their level of thinking.

Research and theories on how persons think, learn and know have many implications for education. In teacher education, interventions need to be designed and implemented in teacher preparation programs to foster ‘better’ ways of thinking and ‘better’ performance of teachers. Research studies serve as evidence of attempts to foster teachers to develop conceptually.

A case in point is an experimental study by Dieker and Monda-Amaya (1997). They examined the impact of training that integrated the introduction of two designed frameworks: effective instruction and problem solving. Examining their students’ reflection through the analysis and scoring of written journal entries of seven full-time students as research participants, the researchers found that the training changed all but one student in ways they solve problems. Findings suggest that professionals in teacher preparation can possibly influence reflective thinking of pre-service educators.

Another example is a qualitative study by Tillema (1997). With teacher educators from several institutions, Tillema developed course materials that explicitly focus on the incorporation of student-teacher beliefs to foster changes in student teachers’ beliefs and practice-teaching performance. From the analysis of pretest and posttest scores of 63 student teacher participants, the researcher concluded that although the program was not found to effect “a lasting belief change”, it did change student teachers’ performance. The small change in beliefs was explained to be related to the length of the program.

Referring to Zeichner's (1987) overview of the effects of instructional strategies aiming to promote reflective teaching, Wubbels and Korthagen (1990) noted "there has been little or no empirical support for the claim that programs designed to promote reflective teaching produce better teachers, or indeed teachers who display any special characteristics at all" (p.16). Using a quasi-experimental research design, they conducted a research to investigate the effects of a program designed to promote reflective teaching on prospective mathematic teachers and compare with those of a more traditional subject-matter-oriented program. Data were collected from two groups of graduates (between 1977-1986): one was an experimental group composed of 37 teachers from Dutch pre-service teacher education program at a teacher's college in Utrecht; and the other was a control group composed of 36 teachers from a college with a program that focused the importance of a subject-matter. The core elements of the reflective program included training for prospective teachers to reflect on their experiences as teachers as well as to consciously aspire an awareness of their own professional development. A spiral ALACT model, which was composed of five phases, was used for cycles of reflection. The ALACT stands for Action, Looking back on the action, Awareness of essential aspects, Creating alternative methods of action, and Trial. The research participants were encouraged to take actions through participation not only in classroom contexts but also among student teacher peers. The program has two noted features: the reflection that was not limited only to the pedagogical component of the program but a recurring principle in the mathematics component; and the reflection cycles that were emphasized even before students began their practical teaching. Participants in the experimental group were encouraged to also reflect on ways in which they help or cooperate with

others. As well, they learned to reflect on their thinking process, feelings, attitudes, and action in everyday relationships with others. Applying a battery of measures, including the teacher's disposition towards reflection, the quality of the interpersonal relationships with students, the adequacy of the teacher's perception of these relationships, the inclination towards innovation and job satisfaction, Wubbels and Korthagen found no difference between the two groups in terms of their reflective attitude. Regarding the student perceptions of the teacher-student relationship, the graduates of a program designed to promote reflective teaching have a better relationship with their students than do graduates of programs that did not stress this goal. In addition, teachers who graduated more than 2 years before from the college designed to promote reflective teaching were found to have a more adequate self-perception than the teachers in the control group with more than 2 years' experience. The other finding was that teachers in the experimental group were found to be more satisfied with their work in the long term. Wubbels and Korthagen concluded the program designed to promote reflective teaching yielded both encouraging and disappointing results for advocates of this type of program. They explained that while teachers in the experimental group were found to have better relationships with their students than did their colleagues, and higher job satisfaction and the greater adequacy of their perception of the teacher-student relationship in the longer term, they were neither more inclined to reflect nor did they display a greater inclination towards innovation.

These research studies indicate attempts at the implementation of programs that utilized deliberative designs to affect teachers' intellectual development. As stated in the

chapter one, challenges are growing exponentially for education and teacher education. A nation needs graduates who not only master particular sets of knowledge in particular fields of study, but who also develop intellectually. Current and future generations are facing more and more problems due to dramatic societal changes. Education needs to help prepare students for such challenges. Graduates of teacher education are a particularly important group of people who will influence the thoughts, learning and knowing process of children—the future of a country.

The next section describes the historical context for research on intellectual development in higher education.

Historical Context for Research on Intellectual Development in Higher Education

With the belief that active situations bring about problems, which further stimulate reflective thoughts, and that action is the process of verifying thoughts before they finally pass over into knowledge, Dewey started the Laboratory School associated with the University of Chicago in 1896 (Dewey, 1896; Kliebard, 1995). The Laboratory School, which was regarded as a miniature community, requires the students' participation and contribution in the community. Students were placed in a role of community members to interact with social environment where they lived. Confronted with situations where problems are posed, students are stimulated to think reflectively. The purpose of Dewey's Laboratory School was to control the process of leading the child from present interests to an intellectual command of the modern world. Dewey's

Laboratory School, even though it was established only in the elementary level, was one of the first studies illustrating the conditions for learning that support intellectual development, controlling for types of thinking, learning and knowing of students. Kliebard (1995) noted that what Dewey sought to effect through the curriculum of the experimental school was intellectual development.

As for education at higher levels, specifically in college, William G. Perry and his associates at the Bureau of Study Counsel (Harvard University) have been credited with pioneering the investigation of changes in thinking processes and views about knowledge in college undergraduate students (Perry, 1970/1999). Perry and his associates examined how students interpreted their lives in natural settings of college attendance for four years at Liberal Arts College at Harvard and Radcliffe. In 1953, the staff of the Bureau of Study Counsel undertook to document the experience of undergraduates in Harvard and Radcliffe over their four years of college. In daily counseling with students focusing on their concerns about academic work, Perry and his associates had the impression that students applied a variety of ways in responding to relativism. While some students applied multiple frames of references to respond to problems, others failed to recognize the existence of multiplicity of ideas but referred only to a single frame of reference to respond to problems. In addition to the variety of the ways of thinking, individual students were found to differently assimilate and accommodate their intellectual scheme based on such the variety of ways of thinking. In other words, the way each individual student thought differently affected the ways each individual student learned, which further affected intellectual and social development. Concomitant to multiplicity of life

in college, final examinations at Harvard at the intervals from 1900-1960 required students to consider multiple frames of reference or relativism. Multiplicity of thinking during the college years was brought about by not only courses, including History, Government, English Literature, and Foreign Literatures, but also the students daily life with peers, and diversity of students in terms of race, ethnicity, socioeconomic status and geographical areas where they were from. Perry and his associates conducted this research because they believed there was a need for education to prepare college students to confront pluralism and multiplicity of the world. This is due to the fact that there had been an increase in societal changes caused by population mobility and technology advancement of mass media. They wondered how college students understand the nature and origins of knowledge, of value and of responsibility. The research aimed to provide a description of the students' experience in terms of "its characteristic stages and its major points of choice between fragmentation and integration, alienation and involvement (p.7)."

Analyzing the data from interviews with their students, Perry and his associates formulated an "intellectual and ethical development" scheme. Perry's and his associates' research was limited to the participation of subjects who were students volunteering in a single college during the years 1954 to 1968 and whose majority was male. Nonetheless, it is now regarded as a benchmark study for understanding the ways of thinking of late adolescence and young adulthood. The scheme has been a foundation of other research on college students (Austin, 2000; Baxter Magolda & Porterfield, 1985; Erwin, 1983; Freeman, 1994; Johnson, 1999; Kubinski, 1999; Stonewater, Stonewater, & Hadley,

1986; Sweeney, 1996; Torosyan, 2000). Based on Perry's scheme, attempts have been made to extend and create other developmental models (Baxter Magolda, 1992b; Belenky et al., 1997; King & Kitchener, 1994), which include more diverse populations.

Belenky et al. (1997), being aware of the domination of males' voice during Perry's theory building stage, conducted qualitative research in which participants were women with diverse educational backgrounds. Referring to Perry's later research assessing women's development based on the formulated scheme, Belenky and associates noted that Perry found women to "conform with the patterns that had been observed in the male data" (Belenky et al., 1997, p. 9). Considering that the use of the Perry's scheme could obscure themes that might be more prominent among women and themes to which attention had not been paid, they traced a progression of women as their research participants. They collected data by interviewing a total of 135 women from a variety of age, education, and family backgrounds and from different geographical areas. Ninety of these women were students enrolling in one of six academic institutions, each of which represented different educational philosophy and composition of student body. Belenky and associates described another 45 woman participants from "invisible colleges" or family agencies that deal with clients who need information and assistance with parenting. The interviews were composed of nine sections of questions, in which embedded questions were devised for assessing Perry's epistemological positions to elicit information on the woman's assumptions about the nature of truth, knowledge, and authority, and questions that were developed by Gilligan (1977) and Kohlberg (1969) to elicit moral orientation or stage. An intensive interview/case study approach was

employed. The two-hour to five-hour interviews were taped recorded and transcribed. This brought about more than five thousand pages of text. In the coding and scoring process, the interview data were separately analyzed according to the theoretical and empirical work of Perry, Kohlberg, and Gilligan. In the analyses of data elicited by questions based on Perry's epistemological positions, Belenky and associates used Perry's scheme of classifying the women's ways of thinking about nature of truth, knowledge, and authority. They found that the data did not fit well to the scheme.

Perry and associates posited that each position is an advance over the last and persons developed in the continuum of one end, dualistic, to the other, commitment within relativism. According to Perry, students gradually progress from dualistic position and become increasingly aware of diversity of opinion and the multiple perspectives of others. They finally grow to a relativistic position, when they consciously and actively learn to analyze and evaluate knowledge. In contrast to Perry's scheme, which focuses on linear sequence in development, Belenky and associates leave the question of whether or not the way women know and view the world develops in a linear sequence or has stage like quality. They grouped women's perspectives on knowing into five major epistemological categories: silence, received knowledge, subjective knowledge, procedural knowledge, and constructed knowledge. The explanation of the five major epistemological categories is presented in figure 2.1 (Adapted from Belenky et al., 1997, P. 15)

Figure 2.1. Women's ways of knowing

Epistemological Categories	Descriptive Characteristics of Women's Perspectives
1. Silence	Knowledge only comes from external authority.
2. Received knowledge	Knowledge can be received and reproduced from that derived from external authority.
3. Subjective knowledge	Truth and knowledge is regarded as subjective and its consciousness depends on personal views and intuitions.
4. Procedural knowledge	Knowledge can be obtained, transmitted, and communicate through the process of learning and applying objective procedures.
5. Constructed knowledge	All knowledge is contextual and can be constructed, and it can be viewed as both subjective and objective.

In terms of the scope of the categories of women's perspectives on ways of thinking about nature of truth, knowledge, and authority, Belenky and associates noted that their categories are not fixed, exhaustive, or universal. Others might organize their observations about women differently. In addition, the categories are abstract and cannot adequately capture the complexities and uniqueness of an individual woman's thought and life. Similar categories might also be found in men's thinking.

Baxter Malgoda (1992b), seeing a gap in Perry's and Belenky's research in terms of gender, explored college students at a small Midwestern liberal arts college. Research participants included 101 students randomly selected from the entering class of 1986. Data was collected by interviewing students using unstructured interviews, which were believed to allow the emergence of students' frame of reference; and a short-essay questionnaire called the Measure of Epistemological Reflection (MER). The MER,

which was designed to assess epistemological development, is composed of open-ended questions addressing the same domains as the interviews. The purpose of using the questionnaire was to provide students with opportunities to write about their ways of knowing and to test against data obtained from the interviews. Analyzing data collected from both tools, Baxter Magolda finally came up with an Epistemological Reflection Model (1992b). Her model is composed of four domains including absolute knowing, transitional knowing, independent knowing, and contextual knowing (Baxter Magolda, 1992b, p.30).

Inspired by the description of Perry's Scheme of Intellectual and Ethical Development at a graduate seminar on college student development organized at the University of Minnesota, others studies involving different aspects of the intellect, and Dewey's work on reflective thinking (1933), King and Kitchener (1994) began research on a Reflective Judgment Model. The model is composed of seven distinct sets of assumptions underlying people's reasoning as well as their concepts of justification. King and Kitchener abstracted this model based on data of interviews of more than seventeen hundred people, ranging from fourteen-year-old high school students to retirees over the age of sixty-five. The creation of the model is partly based on Perry's Scheme of Intellectual and Ethical Development. However, when compared to Perry's Scheme, Reflective Judgment Model has an explicit focus on "the relationship between epistemology and judgment at each point in the developmental sequence" (King and Kitchener, 1994, p.38). Perry's scheme has such a focus only in the early positions but

not in the later positions. In addition, King and Kitchener hypothesized that “people hold epistemological positions beyond relativism” (p.38) based on data of their interviews.

All the studies described so far have become major foundations for other studies (Baxter Magolda & Porterfield, 1985; Erwin, 1983; Knefelkamp, 1974; Martin, Silva, Newman, & Thayer, 1994; Mentkowski et al., 2000; Roger, Mentkowski, Hart, & Minik, 2001; Samson, 2001; Stonewater, Stonewater, & Hadley, 1986; Taylor, 1983). The number of research studies based on these leading studies has increased (Hofer & Pintrich, 1997). Due to their view that few connections among studies based on these theories have been made, Hofer and Pintrich compiled theories, research studies and perspectives in a book, Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing (2002). They explained that the studies by Perry (1970), Baxter Magolda (1992b), and Belenky and associates (1997) aim to describe the process of college students’ interpretation of their educational experiences. However, studies by King and Kitchener (1994), Kitchener and King (1981), Kitchener, King, Wood, and Davison (1989), Kitchener, Lynch, Fischer, and Wood (1993), aim to understand the influence of epistemological assumptions on thinking and reasoning processes with the focus on reflective judgment. Hofer and Pintrich (1997), however, categorized all these studies under epistemological development theories. Some of the theories just mentioned have also been categorized under cognitive-structural theories by Pascarella and Terenzini (1991).

In 1991, there was a publication of a book entitled How College Affects Students. The authors, Pascarella and Terenzini synthesized approximately 2,600 research studies

of the late sixties through the seventies and eighties. The analysis of theories and models of students change in college includes two general families of theories: “developmental” model and “college impact” model. The former, which is influenced by psychological stage theories, focuses on the outcomes or nature of student change such as identity formation, moral or cognitive development. The latter emphasizes on the sources of change i.e. the environmental or sociological origins of student change. Under the developmental theories of student change, Pascarella and Terenzini (1991) included cognitive-structural theories, which are most influenced by Jean Piaget. The cognitive-structural theories attempt to examine process of change with the concentration on construction of cognitive structures by individuals in order to make meaning of the world. Among theories categorized under cognitive-structural theories are William Perry’s Scheme of Intellectual and Ethical Development (1970), King and Kitchener’s Reflective Judgment Model (1994), Harvey, Hunt, and Schroder’s Conceptual Development Theory (1961). It is worth noting that while Hofer and Pintrich (1997) did not include Harvey, Hunt, and Schroder’s Conceptual Level Theory in their review of epistemological development theories, Pascarella and Terenzini (1991) do include Conceptual Level Theory in their review of cognitive structural theories.

There are also other research studies that are based on earlier work of Perry (1970), Belenky and associates (1997), Baxter Magolda (1992b), and King and Kitchener (1994). Some of those studies, however, do not assume developmental trends and they are not within the scope of this dissertation.

So far, the importance of intellectual development and reflective thinking in higher education and teacher education has been addressed. As well, historical context for research on intellectual development has been provided. The next section reviews in detail the four selected theories that this dissertation focuses on.

Review of the Four Selected Theories

This section begins with the scope of the review. Then, it describes selected theories including Perry's Scheme of Intellectual and Ethical Development (1970), Baxter-Magolda's Epistemological Reflection Model (1992b), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt and Schroder's Conceptual Development Theory (1961). Along with the description of the theories, are reviews of assumptions, validity and reliability issues, and some research findings. This section finally reviews related measurement instruments employed to assess characteristics of development as described by the four theories.

Scope of the Review

The methodological focus of this dissertation is both quantitative and qualitative. It compares findings of ten studies based on the four theories. The comparative analysis takes into considerations both quantitative findings and qualitative findings. Along with the four selected theories, a number of assessment instruments have been developed to assess aspects that each theory focuses. Illustrations of these are a Checklist of Educational View (CLEV) (Perry, 1970), Measure of Epistemological Reasoning (MER) (Taylor, M.B., 1983), Measure of Intellectual Development (MID) (Knefelkamp, 1974),

Paragraph Completion Method (PCM) (Hunt, Butler, Noy, & Rosser, 1977), Reflective Judgment Interview (RJI) (King and Kitchener, 1994), Scale of Intellectual Development (SID) (Erwin, 1983), and Learning Environment Preferences (LEP) (Moore, 1987). These instruments, however, provide different types of quantitative data. Consequently, research studies to be included in the review are limited to those that employ assessment instruments that provide compatible quantitative data. In addition, those assessment instruments should have some research evidence to back up their equivalence in terms of whether they assess the same elements of theories. There should also be sound evidence of reliability and validity issues of those measurements.

King and Kitchener's Reflective Judgment Model (1994) uses the Reflective Judgment Interview (RJI), and Harvey, Hunt and Schroder's Conceptual Development Theory (1961) uses the Paragraph Completion Method (PCM) as their measurement tools. Both measurement tools provide quantitative data. Regarding research studies based on Perry's Scheme of Intellectual and Ethical Development, however, there are studies that employed tools providing both qualitative and quantitative research. Thus, only studies that employed tools that provide quantitative data are included in the comparative analysis. The measurement instruments relevant to Perry's theory are Measure of Epistemological Reasoning (MER) (Taylor, 1983), and Measure of Intellectual Development (MID) (Knefelkamp, 1974). Research studies to be included in the comparative analysis are limited to those that employ these four instruments: RJI, PCM, MER, and MID.

Perry's Scheme of Intellectual and Ethical Development

Between 1954 and 1968, Perry and his associates (1970) conducted a research study at Liberal Arts College at Harvard and Radcliffe. The research aimed to provide a description of the students experience in terms of “its characteristic stages and its major points of choice between fragmentation and integration, alienation and involvement (p.7).” Initially, Perry and his associates started exploring the variety in students’ response to the impact of intellectual and moral relativism from a random sample of 313 freshmen enrolling in the fall 1954 at Harvard College and Radcliffe. The students completed A Checklist of Educational Views or CLEV in the fall 1954 and again in the spring 1955. Based on the scores of CLEV, 55 students were sent invitations to participate in the research. Thirty-one of those being invited volunteered to participate in the interviews to share their college experience. The research participants included those freshmen who had scored at the extreme of dualistic and contingent thinking, some from the mean, and some who had notably changed their scores from fall to spring. The follow up interviews of the same participants were conducted in late May and June of each year. Final data included 17 complete four-year records.

During the interviews, participants were asked to respond to open-ended questions about what seemed salient to their own experience. Questions such as “Would you like to say what has stood out for you during the year?” and “As you speak of that, do any particular instances come to mind?” brought about a variety in the form and content of students’ responses. From such data, Perry could detect a common sequence of challenges of each individual student from experiencing his or her academic work,

social life of college, and extracurricular activities or employment. This sequence of challenges was perceived as a representation of a coherent development in the forms that support the intellectual growth, experiencing of values, and interpretation of the world. With this finding, Perry and his associates extended the purpose of the research and applied experimental and descriptive procedures.

With the more structured research procedures, Perry and his associates randomly sampled freshmen from the Class of 1962 and 1963 and finally obtained 366 interviews including 67 complete four-year reports of students' experience. From all the data, Perry and his associates formulated the developmental scheme before finally testing the validity of the formulated scheme. This scheme, which is famous as "Perry's Scheme of Intellectual and Ethical Development (1970)", is composed of nine Positions of development and three Positions of deflection namely Temporizing, Escape, and Retreat.

Perry elaborated that students pass through nine major positions in their intellectual and ethical development. Generally, they move from simplistic, categorical view of knowledge and values to a complex pluralistic perspective. The nine positions can also be grouped into three general categories: Dualism, relativism, and commitment. Dualistic students assume that all information can be classified as either absolutely right or absolutely wrong and that ambiguity is an unnatural phenomenon. The development from position one to position three represents an increase in tolerance of uncertainty. Relativistic students view knowledge as relative. Notions of right and wrong have meaning only in context and uncertainty becomes legitimate. Students at position five have a potential to think abstractly and to analyze their own thoughts. During the

positions seven, eight, and nine, while the students' assumption about knowledge remain relativistic; they gradually accept the responsibility of the world outside themselves. Such responsibility is reflected through their commitment to establish their identity. By choosing a particular career, an orientation to studies, a religious-value stance, the students also seek to find a position to interact with the world outside of themselves.

Following is the brief outline of the scheme (adapted from Perry, 1970/1999 and 1981/1997).

Figure 2.2. Perry's Scheme of Intellectual and Ethical Development

Position	Mainline of Development
Position 1: Basic Duality	Students at this position perceive the world from a dualistic viewpoint. They divide the world applying the sense of familiarity and lack of familiarity between Authority*-right-we as opposed to illegitimate-wrong others. Knowledge is absolute with either right or wrong answers known to Authority. Obedience to Authority and hard work are regarded as important elements of morality and personal responsibility.
Position 2: Multiplicity Pre-legitimate	Students at this position acknowledge the existence of diversity of opinion and uncertainty of knowledge. The source of diversity and uncertainty are perceived as derived either from unqualified Authority or as an intentionally established situation for students to learn to find the existing absolute right or wrong answers by themselves.
Position 3: Multiplicity Subordinate	Students at this position accept diversity and uncertainty as legitimate. However, they perceive that such diversity and uncertainty temporarily exist only because the Authority still hasn't discovered the absolute answer.
Position 4a Multiplicity Correlate	Students at this position highly consider the value of uncertainty of knowledge and diversity of opinion. They regard knowledge as dependent on personal and individual subjective opinions and not dependent on Authority.
Position 4b Relativist Subordinate	Students at this position discover qualitative contextual relativistic reasoning as a special requirement from Authority for them to engage.

Figure 2.2. Perry's Scheme of Intellectual and Ethical Development (Cont.)

Position	Mainline of Development
Position 5: Relativist Correlate, Competing or Diffuse	Students at this position perceive all knowledge and values as contextual and relativistic. Accordingly, dualistic point of view is perceived as a special case in context.
Position 6: Commitment Foreseen	Students at this position perceive the importance of personal Commitment** in a relativistic world.
Position 7: Initial Commitment	Students at this position decide on his own responsibility in life and initiate some Commitments.
Position 8: Orientation in Implications of Commitment	Students at this position experience their initial Commitments and raise questions about their responsibility in a relativistic world.
Position 9: Developing Commitment(s)	Students at this position validate their identity through experiencing multiple responsibilities and decide to pursue their life through some forms of personal Commitment
Alternatives to Growth	
Temporizing	Students delay or pause to grow over a full academic year in order to explore the implications of, or consolidate the position they currently attained.
Escape	Students regress from higher positions back to the previous position. This take place either between dualism and relativism or relativism and commitment to relativism
Retreat	Students decide to limit themselves to progressing to the next position to avoid complexities and responsibilities.

From glossary in Perry (1970/1999)

*Authority is explained as “the possessors of the right answers in the Absolute, or the mediators of same; or the false or unfair pretenders to the right answers in the Absolute.”

**Commitment referred to an affirmation of personal values or choice in Relativism. A conscious act or realization of identity and responsibility. A process of orientation of self in a relative world. The word Commitment is reserved for the integrative, affirmative function, as distinct from 1) commitment to and unquestioned or unexamined belief, plan, or value, or 2) commitment to negativistic alienation or dissociation.

Perry also found that the development did not occur in a linear sequence for all students. During the development process, there are some delays, deflections and regressions. Since development takes place as a result of interaction between individuals and environmental conditions, such a condition as college environment can either support or hinder the students' growth.

Research Validity and Reliability

Perry and Associates' research was validated by Cooperative Research Contract SAE-8973, Project No. 5-0825. The Final Report for the Project of this study included detailed analysis of the student samples and a study of the Checklist of Educational Views (CLEV). In the fall 1963, Perry and Associates assembled a group of six judges who were graduate students in English and in Comparative Literature. The validation process began with all the judges studies all the materials containing a copy of the chart of Development, a sample interview protocol, and a manual that described detail of the study. Then, the judges did the rating of the correspondence or the matching of the data against the scheme. The data for the rating included the four-year protocols, rating of single interviews, rating of excerpts, and rating of condensed four-year reports.

Regarding the four-year protocols, complete, unedited transcripts of four-year sequences of the interviews with 20 students were presented to the judges, one student's set at a time. Ten of the students were selected at random from the sample of the Class of 1958, and ten from the sample of the Class of 1962 and 1963. Each judge was required to rate independently of the other judges and rated all of the four interviews in a set. This

rating procedure was conducted to test the null hypothesis that there would be an agreement in matching interviews with Positions on the Chart at a level of agreement not exceeding that attributable to chance. The mean estimated reliability of the mean rating for individual interviews for each of the four years was found to be +0.966, +0.875, +0.872, and 0.916 respectively. The probability of these levels of agreement occurring by chance was found to be less than .0005 for the lowest. The reliabilities of the mean ratings of the four interviews of individual students were reported to range between .815 and .978. With this statistical range, Perry and associates concluded that the judges were able to agree reliably in relating the scheme to the report of all students in the sample.

In addition to the four-year protocols, the judges were asked to rate single interviews to check for the possibility of the influence of the judges' knowledge of the student's year on their estimation of degree of student's development. Findings confirmed that, without knowledge about student's year, their agreement about the student's Position on the Chart remained at the level reached with the rating of the first data set. The judge were also required to excerpted from other interviews 40 statements which seemed similar in character to those which the judges had noted. This procedure was conducted to examine the kinds of statements in the interviews, which had been noted by the judges as contributing most significantly to their rating of interviews. The agreement of the judges in rating excerpts demonstrated that the various structurings that constitute the scheme can be illustrated in excerpts. Another procedure that was employed to validate the scheme was to ask three judges to rated the reports in complete

form and other three to rate the reports in condensed form. Comparison of the results showed that the condensed version still represented data for rating purposes.

Research Studies Based on Perry's Scheme

After the creation of Perry's Scheme of Intellectual and Ethical Development, other researchers conducted studies based on the scheme. Among these studies, some aimed to develop a less time-consuming and less labor-intensive assessment system than Perry's interview. Knefelkamp and Widick were the first to develop the Measure of Intellectual Development or MID in 1974. Following the study by Perry and Associates, other research employed MID (Mentkowski, et al., 2000; and Moore, 1989). This instrument is still in use for research studies to collect data on students' views of multicultural diversity in the curriculum and in campus life (Perry, 1970/1999).

Taylor (1983) developed the Measure of Epistemological Reflection (MER) (Taylor, 1983; Baxter Magolda & Porterfield, 1985). Erwin (1983) developed the Scale of Intellectual Development (SID) in order to measure Perry's Scheme of Intellectual and Ethical Development. The SID was administered to 3,321 college freshmen. Factor analysis revealed four factors: Dualism, Relativism, Commitment, and Empathy. To test the construct validity of the SID, the researcher administered an identity scale and a perceived-self questionnaire to in conjunction with the SID to subsamples of 300 and 2,707 students. Although the SID scores were found to correlate with the two instruments, the researcher recommended additional validation of the instrument. Moore (1989) developed the Learning Environment Preferences (LEP) as an instrument to assess

Perry's model of intellectual and ethical development in college students. A study was conducted to evaluate construct validity of the LEP with 725 undergraduate students as participants. Data indicated that the LEP accurately measures cognitive complexity as described by Perry's scheme. The group means on the LEP's Cognitive Complexity Index across classification including freshmen through seniors suggested the expected measure of development. Findings led to the conclusion that the LEP is a reliable and valid objective measure of the Perry's scheme. It is easily administered and scored. As well, it is an inexpensive measurement.

Some studies examined the correlation among instruments designed to measure Perry's scheme. Taylor (1983) argued that available measurement techniques to assess Perry's Scheme of Intellectual and Ethical Development did not adequately address issues related to assessing cognitive development. This was due to the problematic factors including the fact that it is difficult to separate strands of development because all kinds of development occur simultaneously. In addition, there is a possibility for people to exhibit behaviors representative of more than one developmental level. Moreover, signs that appear in earlier stages reappear later in more complex forms. Taylor administered the Measure of Intellectual Development (MID) and the Measure of Epistemological Reflection (MER) to 155 college students. Data obtained from the instruments and the follow-up probes provided adequate data to elaborate on the respondent's thinking and to reveal the reasoning structure being used; this structure is the unit of analysis for rating. Preliminary percentage agreement between expert raters on the first 84 protocols was 60% exact position agreement.

In addition to studies that concern the development of instruments to measure Perry's scheme, concepts of the scheme have been applied across fields of studies and among diverse research participants with different focuses and purposes. Following are some examples of the research based on Perry's scheme.

Examples of studies that apply the concept of Perry's scheme to improve writing instructions and course design include the studies by Baer (1988), Shapiro (1985) and Irwin (1996). Baer (1988) conducted a study to investigate the effects on students' cognitive development of a freshman composition course design based on William Perry's theory of epistemological development. The course was composed of an integration of reading, writing, and discussion in order to increase students' awareness of ambiguities, uncertainties, and complexities. Six students who were participants were required the students to keep journals and write four essays. Their work was monitored through interviews and their writing was examined and categorized according to a set of criteria adapted from Perry's Scheme. These categories included dualism, multiplicity pre-legitimate, multiplicity subordinate, multiplicity correlate, relativism correlate, and commitment foreseen. Findings indicated that the cognitive development of the students increased two or more levels.

Shapiro (1985) investigated the relationship between the intellectual maturity of college students and evidence of rhetorical maturity in their writing in order to find out why some students write better than others. Data included 70 student essays. The essays were evaluated on three independent measures including Perry's scale of intellectual development, Diederich's scale of writing competence, and a measure of audience

awareness based on the writer's constructed context. Regarding that language abilities such as vocabulary and syntactic maturity cannot account for all the differences between good and poor writing, the researcher focused on the students' rhetorical maturity with respect to writing competence and context. Participants represented both traditional and older undergraduate and graduate students across a broad spectrum of majors and were assumed to have comparable general language abilities and necessary preliminary writing tools including spelling, vocabulary, and syntactic options. The researcher hypothesized that students who had moved through the more complex positions on Perry's model of intellectual development would have internalized the need to provide necessary and appropriate context in their writing. Findings suggested that aspects of intellectual development described by Perry's theory were significantly related to the quality of student writing. In addition, levels of cognitive development among college students had a statistically significant relationship to both writing competence and constructed context.

To test a hypothesis that students' ability to think critically depends on the students' developmental levels, Irwin (1996) conducted research to examine the relationship between students' performance in critical thinking ability and their developmental levels. Students' persuasive or argumentative essays were used to represent a student's critical thinking abilities. Scores from Measure of Epistemological Reflection (MER) (Baxter-Magolda and Porterfield, 1985) were used as indicator of students' intellectual development as described by Perry's scheme. Participants were volunteer students working in a university computer-writing laboratory. They were randomly assigned to an experimental and control groups. The experimental group

students used the Hermes program and treated with some traditional instruction. The Hermes program is a computer-assisted tutorial that was designed to improve students' rhetorical and dialectical writing. The control group students received the traditional online writing instruction (Linda Flower's nine steps for composing). Data were collected from student/computer interactions and the resultant essays. A holistic scoring termed Primary Trait Scoring was developed specifically for this project and was used for scoring all the essays from both groups. This scoring scheme had been tested for inter-rater reliability of the scoring procedure. The inter-rater reliability was .88. Results suggested that students in the experimental group wrote significantly more dialectical and rhetorically effective essays than those in the control group. Irwin concluded that students' performance, as determined by both the writing of the essays and the ability to benefit from the instruction provided by Hermes, was related to intellectual developmental level, as assessed by the MER. In addition, the degree of relationship was found to be greater between developmental level and performance than it was for either of educational level and performance or age and performance. Irwin noted that the results of this study are insightful for elaborating adult development theory and assessment. This is because the results indicated the effectiveness of the Hermes computer-assisted instruction designed based on the developmental level. As well, the results indicated the validity of the MER as a measure of intellectual development.

Perry's scheme is also applied to help with career planning and improvement. Moore (1983a) examined a one-credit career-planning course for undecided undergraduates at Maryland. Measure of Intellectual Development (MID) was used to

assess the intellectual dimension of Perry's scheme. The students enrolled in the course took the MID as pre- and post- tests. Findings indicated a positive change in cognitive complexity of over 40 percent of the participants. Seniors were found to dramatically increase in their cognitive complexity compared to other groups. The researchers noted that this finding suggested that the seniors were most able to respond to the challenge of the course. In addition to MID, the Myers Briggs Type Indicator was employed as an instrument to elicit students' personality type and the process by which they perceive and judge information along four dimensions: Introversion-Extraversion, Sensing-Intuiting, Thinking-Feeling, and Judging-Perceiving. Comparisons of students' MID and MBTI scores showed that there seems to be a strong tendency for Intuitive, particularly Intuitive/Perceiving types, to be found more frequently at higher levels of cognitive complexity, while Sensors and Judgers tend to be found less often at those same levels. Moore also noted that the analysis of the stage/style interactions in cognitive development showed obvious overlap between the two frameworks.

Sweeney (1996) examined the epistemological perspectives and cognitive development as described by Perry's Scheme of Intellectual and Ethical Development of 318 full time nurse faculty teaching in associate, baccalaureate, master's and doctoral degree programs. The researcher employed two types of instruments: a simple test that the researcher designed to measure faculty epistemological perspective regarding general knowledge, nursing knowledge, and nursing knowledge the faculty reported using when teaching students in the classroom; and Learning Environment Preferences (LEP). Data collected from the simple test showed that about 50% of faculty selected the highest

Perry Position (Relativism). At this position, the faculty regarded knowledge as context dependent and relative, experts as sources of knowledge in specific realms, and experts and learners alike created knowledge. As for teaching, 22% of faculty reported that they used lower epistemological perspectives of nursing knowledge in their teaching than those, which they held for themselves. The LEP was used to measure faculty cognitive development, a concept related to epistemological perspective in the Perry Scheme. Data suggested that 99% of faculty scored below Relativism in cognitive development. With a multiple regression analysis performed on the demographic data gathered, the researcher found that level of formal education, meta-cognition, and divergent thinking were positive correlates of cognitive development. The researcher expected higher scores of cognitive development from faculty because of their high level of formal education. With such findings, the researcher hypothesized that habits of meta-cognition and divergent thinking in the learner are not cultivated by nursing education. The researcher also noted that a less-than-Relativistic perspective on the part of the majority of faculty in the discipline may perpetuate the current knowledge base by not encouraging thinking in learners as well as persons with advanced degrees.

Perry's scheme is also applied in teacher education and social justice education. An example of application in teacher education includes a study by Tharp and Lovell (1995). They investigated patterns of teacher thought about student reasoning and learning. Participants were 23 preservice teachers. Their responses to "dilemma of practice about equity" were categorized as stages of development. Findings suggested that the preservice teachers' stages of development were significantly correlated with the

positions identified in Perry's scheme. Another example shows the application of the scheme in social justice education. Adams and Zhou-McGovern (1994) conducted a cognitive and sociomoral developmental studies of 165 college undergraduate students enrolled in spring 1990 at a large Northeastern public research university. The research aims to apply cognitive developmental theory to social justice education. These students enrolled for a "diversity course" undergraduate course on social diversity and social justice. Significant course effects ($p < .001$) were found for MER overall. Although age was not found to be significant, gender effects ($p < .05$) were found for the MER overall and component scores; and college class effects were found for one MER component score.

Some studies apply the Perry's scheme to examine gender difference. For example, Alishio and Schilling (1983) examined gender differences in intellectual and ego development among 31 females and 29 males. As for intellectual development, they applied Perry's scheme of Intellectual and Ethical Development to look into occupational choice, interpersonal relationships, and sexual identity. In terms of ego development, they applied Loevinger's ego development theory to investigate religious choice. Measurement instruments included Perry's unstructured interviews and Loevinger's sentence completion test (SCT). The interview transcriptions and the SCT were rated blind. Results of structural analyses showed no significant differences between gender either overall or by each area of development. Multiple and step-wise regression analyses, however, suggested difference in patterns. While male identity processes were found to consistently focus on occupational issues, female identity processes were found

to focus on interpersonal and sexual issues. Content analyses also suggested that women focused their interpersonal development on issues of trust while men focus upon issues of rightness.

There are also studies that apply Perry's scheme to college students from different cultural backgrounds. Examples include a study that examines African American students and a study that examined the generalizability of the scheme across culture and nations. Applying William Perry's cognitive theory of intellectual and ethical development, Johnson (2000) compared cognitive complexity of students on the basis of their cultural background and academic class, as well as the influence of gender and socioeconomic status. The aim of the study was to explore if Perry's theory would be valid for culturally diverse students. A cross-sectional design was employed. Participants were 1,248 students including entering freshmen and graduating seniors attending a predominantly White, public urban university. Cognitive Complexity Index (CCI) and Learning Environment Preferences (LEP) (Moore, 1987) were used to measure and identify students' cognitive development. The researcher reported both quantitative and qualitative findings. As for the quantitative findings, results of the analyses of the cross-sectional samples indicated evidence of significant differences in cognitive complexity as defined and measured by the LEP between cohorts of African American and White students. Considering freshmen, CCI scores indicated that White freshmen averaged higher than African American students on the LEP when gender and socioeconomic status were controlled. As for seniors, similar differences in cognitive development were found. However, when gender and socioeconomic status were

controlled, no significant differences were found between senior cohorts. Controlling for gender and socioeconomic status, cross-sectional analysis of the interaction between culture and academic class status indicated no significant differences in cognitive development. In terms of qualitative findings, responses were comparatively analyzed to determine if any themes, which represent cognitive development, were developed along cultural and academic levels. Interview responses were found to generally support Perry's (1970) scheme. However, the researcher noted that analyses of responses revealed that not all subjects understood and interpreted the questions in the same manner. In addition, themes that emerged within interviewee responses suggested that students may have either similar or distinct worldviews depending on their culture, academic class status, or socioeconomic level. Cultural differences were also found to exist around themes of learning orientation and perceptions of authority. Johnson concluded that although Perry's scheme provides a framework and description of the routes for intellectual potential, his study indicated inadequacies of the Perry scheme to assess the cognitive complexity of African American students.

Zhang (1999) created the Zhang Cognitive Development Inventory (ZCDI) (Zhang, 1995) based on Perry's theory of intellectual and ethical development and employed it to examine Chinese students' cognitive development. This study on cross-cultural generalizability of Perry's theory included three groups of college students as research participants. There were a group of 152 US students aged between 17 and 29, and two groups of People's Republic of China students. A group of Chinese students was composed of 808 students and the other was composed of 503 students. They were

between 17-26 years of age. Analysis of data indicated differences in college students' cognitive-developmental pattern. Zhang concluded that Perry's scheme is not universal and students' cognitive-developmental pattern varied as a function of different cultural and educational systems.

So far, research evidence validates the existence of intellectual and ethical development as described by Perry's Scheme of Intellectual and Ethical development. It is worth noting again that different instruments have been used for studies concerning development as described by Perry's scheme. The following section examines the Epistemological Reflection Model (Baxter Magolda, 1992b).

Baxter Magolda's Epistemological Reflection Model

Baxter Magolda (1992b) explored college students at a small midwestern liberal arts college. Her research paid particular attention to gender similarities and differences in students' ways of knowing. Research participants included 101 students of comparative number of both males and females randomly selected from the entering class of 1986. They had diverse academic majors but were homogeneous in socio-economic status and race. The participants were paid five dollars for their participation each year. She collected data by interviewing students using unstructured interviews, which were believed to allow the emergence of students' frame of reference; and a short-essay questionnaire called the Measure of Epistemological Reflection (MER). Since the study was longitudinal in design to allow tracing the development of students' ways of

knowing throughout the college experience, the participants were interviewed every year as well as completing the MER questionnaires.

As for the interview, the number of participants were 101, 95, 86, 80, and 70 in the first, second, third, fourth, and fifth years respectively with the balance of gender representation. The first year interview addressed six areas of epistemological development. It included the role of the learner, instructor, peers, and evaluation in learning, the nature of knowledge, and decision-making. The researcher uses open-ended questions to elicit information from the participants with an intentional avoidance to lead them within the domain of development under study. The first year data were analyzed and used as a basis for modification of the second year interview that directly addressed the nature of knowledge and that involved the participants to share their experience based on both from in and out of class. The second year data was further used to modify the third year interview, when follow-up questions were added in each area as well as direct attention was paid on out-of-class learning experience. The interview focused on the most significant experience, classroom learning, out-of-class learning, and to decision making. The same interview guide was used in the fourth year interview. The fifth year interview was also conducted among participants who had graduated. As for the interview, participants were asked to reflect on learning during the senior year with the focus on most and least effective learning, experiences. In addition, they were asked to reflect on the way they made decision, on their perceptions of learning in their current positions as graduates involving their lives in different environments. Furthermore, they were asked to reflect on the overall college experience and its impact on their lives. The

interviews were conducted in the fall of each year. Except the interviews with graduates in the fifth year that were conducted by telephone, other interviews were conducted face to face.

The Measure of Epistemological Reflection or MER is a short-essay questionnaire designed to assess epistemological development addressing the same six domains as the interviews. The purpose of using the open-ended questionnaire was to provide students with opportunities to write about their ways of knowing and to test against data obtained from the interviews. The focus of the questionnaire is on how people think and why they hold a particular view, not what they believe. After each interview, participants were given the MER and asked to wait at least two weeks before they completed the questionnaire. All participants returned the MER questionnaire in the first year. The rate of respondents, who returned questionnaire the following years, was 77 of 95, 64 of 86, 45 of 80, and 34 of 70 respectively. For data coding, reasons the respondents gave for their perspectives were sorted into themes called reasoning structures and the MER responses were assigned to theoretical description of Perry's Positions 1-5. The overall epistemological levels were derived from the empirical data rather than the initial theoretical description.

Analyzing data collected from both tools, Baxter Magolda finally came up with a gender-related pattern in knowing within her Epistemological Reflection Model (1992b). Figure 2.3 presented an overview of the model (adapted from Baxter Magolda, 1992b).

Figure 2.3. Epistemological Reflection Model

	Ways of Knowing			
	Absolute Knowing	Transitional Knowing	Independent Knowing	Contextual Knowing
Nature of knowledge	Knowledge is certain. Differences of opinion are differences in degrees of detail.	Some knowledge is uncertain. Uncertainty is a result of the answers being unknown	Knowledge is assumed mostly uncertain. One has one's own beliefs.	Knowledge is uncertain and can be judged as better depending on contexts and evidence.
Role of learners	Learners obtain knowledge from instructors. Learners acquire and remember information.	In uncertain areas, learning is more complex. Learners should pay attention for understanding rather than on memorizing information.	Learners can hold their own opinions and the opinions are considered equally valid. They should create their own views and listen to others.	Learners should exchange and compare perspectives, think through problems, and integrate and apply knowledge in context.
Role of peers	Peers share materials and explain their learning to help with the process of acquisition	Peers should discuss to expand ideas and do active hands-on activities	Peers should share views which serve as sources of knowledge	Peers are expected to make worthwhile contributions.
Role of instructor	Instructors have all answers. Instructors should ensure that they transfer knowledge to students.	Instructor should employ methods that focus on understanding and application of knowledge	Instructor should promote independent thinking by providing contexts for students to explore and exchange knowledge.	Instructor should promote application of knowledge in context, evaluative discussion of perspectives, and opportunities for students and teacher to critique each other.
Role of learning evaluation	Students reproduce from their memory knowledge that the instructors have provided to them.	Students' understanding should be measured rather than memory	Evaluation should reward students who think independently and not penalize ones for holding views that are different from those of authorities.	Competence should be accurately measured in context. Evaluation should be a process in which student and instructor work together toward a goal and measure their progress.
Gender-related patterns	1.Receiving knowledge pattern (used more by women) 2. Mastery knowledge pattern (used more by men)	1.Interpersonal pattern (used more by women) 2.Impersonal pattern (used more by men)	1.Interindividual pattern (used more by women). 2.Individual (used more by men)	(Aspects of patterns used infrequently earlier became increasingly apparent in more complex ways of knowing.)

Baxter Magolda noted that this gender-reported pattern “stemmed from reasoning structures in which the percentage of women and men using those structures differed by approximately 33 percent” (Baxter Magolda, 1992b, p.408). Baxter Magolda elaborated that similar assumptions about knowing emerged from the data and were categorized as “ways of knowing”. Based on the epistemic assumptions, there were four qualitative ways of knowing: absolute knowing, transitional knowing, independent knowing and contextual knowing. Each way of knowing leads to particular expectations of learners, peers, instructors in learning setting, understanding of how knowledge should be evaluated, and how educational decisions are made. Within each way of knowing, there are patterns that are characterized by different justifications of epistemic assumptions. Each pattern was employed with different frequency by different genders.

Within absolute knowing, they are receiving and mastering knowledge. Receiving-pattern students acquire knowledge through listening and recording information, do not expect to interact with instructor, and prefer to interpret knowledge by themselves than by consulting instructors. Peers are expected to help by listening and asking questions. Evaluation is valued as opportunities for demonstration of knowledge. Mastery-pattern students show their interest to the instructor and expect interchanges with teachers. They expect authority to resolve differences in knowledge claims. Peers are expected to be partners in arguing and challenging each other to master the material. Evaluation is valued as a mean to improve mastery of knowledge. While women tended to use the receiving pattern more often, men tended to use mastery pattern more often.

Within Transitional knowing, there are interpersonal and impersonal patterns. Interpersonal-pattern students acquire knowledge through collecting ideas from others, expect interaction with peers both to express themselves and to learn different ideas. They tend to keep relationship with instructor to enhance self-expression. These students value evaluation that take into account individual differences and make judgments on uncertainty based on their personal judgment. Impersonal-pattern students preferred to be forced to think, to debate in order to exchange their ideas with instructors and peers. They expect their instructors to challenged them. They value evaluation that is fair and practical. They resolve uncertainty based on logic and research. Women tend to use the interpersonal pattern more often while men tend to use the impersonal pattern more.

Within independent knowing, there are interindividual and individual patterns. Women used interindividual pattern more than men and men used individual pattern more than women. While interindividual-pattern students focus on both their own thinking and engaging the views of others, the individual-pattern students focus more on their own independent thinking. The interindividual-pattern students want to share view with peers and expect instructors to promote exchange ideas. They view evaluation as a process that requires cooperation between student and instructor. The individual-pattern students emphasize thinking for themselves and expect peers to think independently. They expect instructors to provide opportunities for students to define their own learning goals and view evaluation as based on independent thinking. As for the contextual knowing, Baxter Magolda noted that this way of knowing was evident in only 12 percent of the interviews.

Research Validity and Reliability

Baxter Magolda (1992b) collected data through interviews and students' writing short responses to short-essay questionnaire, Measure of Epistemological Reflection (MER). Thus, there are both qualitative and quantitative data. The qualitative data were processed using Lincoln and Guba's (1985) qualitative counterparts to evaluate credibility and transferability. The credibility of the data is supported by three of Lincoln and Guba's techniques. Member checking technique was also used to check for the accuracy of researcher's interpretation of data from research participants. Two independent readers also helped process the data. As for the quantitative data, Baxter Magolda and Porterfield (1985) reported that the accuracy of the MER in accessing epistemological level was supported by consistent significant differences across levels of education and a correlation of .93 with the interviews (Baxter Magolda, 1987b). Interrater reliability on MER is supported by a .80 correlation (N =752) and by interrater agreement ranging from .70 to .80.

Research Studies Based on Epistemological Reflection Model

Baxter Magolda expands the described longitudinal researches based on the ways of knowing with different focus. Following are some examples of extension of data analysis with different focuses.

Based on the data of a four-year longitudinal study of college students' intellectual development during college, an analysis of the data on the impact of cocurricular experiences, Baxter Magolda (1992a) found that factors that contributed to

students' intellectual development included students' peer relationships, organizational involvement, living arrangements, employment, and exchange experiences.

With an analysis of the interviews with 101 college students' perceptions of their experiences throughout the 4 years of college, students were found to have different epistemologies. Some students, for example, unquestionably accepted knowledge as certain and it was transferred to them by instructors, others believed that they could create their own knowledge. Baxter Magolda (1992c) found that results suggest teaching strategies and evaluation techniques connect with students' varying epistemologies.

During the last phase of a seven-year longitudinal study, there were 53 students. Baxter Magolda (1993a) examined the roles of impersonal or abstract and relational or connected modes of knowing in the experiences of young adults. An analysis of data drawn from the last phase of the study showed that contextual knowing is characterized by a complex interplay between relational and impersonal modes of knowing. The emergence of this intertwining was found to be clearest in contextual knowers. This took place particularly in situations where the balance of the two was not present. Such situations led to lack of ability to connect with others or to maintain the separate self. Educational implications from this knowledge that Baxter Magolda noted included the need for education to focus on advancing both modes of knowing and their integration. In addition, effective education should emphasize the construction of knowledge, the importance of the knower in what is known, and the value of others' perspectives. Moreover, education should offer direct practice with both types of knowledge.

In Baxter Magolda (1995), an analysis of the data from the postcollege phase of a seven-year longitudinal study of 101 college students suggested that the gender-related patterns of relational and impersonal knowing that these students demonstrated during college were integrated into postcollege experience in work, education, and personal environments.

In an analysis of 25 post-baccalaureate students' interviews that focus on how students arrived at contextual knowing through experiential learning experiences, Baxter Magolda (1993b) found five major themes to emerge. The experiences described offer specific examples of how to connect knowing to students' experience that could be used in undergraduate and graduate settings to promote contextual knowing. These theme include the value students have for opportunities to think and explore for themselves, to struggle with ideas, and to formulate and support their own opinion. Another theme is concerned with their interest to connect their beliefs with their own lives and identities. In addition, the students regard learning environments as a central facilitating means for them to connect their knowledge and experience. The support of such connection by learning environments is related to the roles of teachers and students as equal partners who mutually respect each other. Another theme is related to the role of peers, who are viewed as equal partners whose sharing and exchange of perspectives enhances exploration of choices of beliefs. The analysis of the data also suggested the interrelationships among these themes.

There are also studies based on Baxter Magolda's Epistemological Reflection Models by other researchers. There are studies that attempted to identify students' stages

of knowing as described by the model and to examine relationships between the stages and other variables. Atkin (1996) conducted research to identify a dominant learning stage of students applying learning stages as described by Baxter Magolda's Epistemological Reflection Model (1992b), to investigate the relationships between the stage and some demographic variables, to evaluate students' satisfaction levels, and to demonstrate application to specific counseling situations. These demographic variables included gender, age, placement in English as a Second Language (ESL) course and curriculum. Participants included 721 students from a midsize public community college. Data were obtained from questionnaires. Responses were subjected to factor analysis. Findings suggested that the students using English as a second language were most likely to be absolute knowers. No significant differences in dominant learning stage were found to be attributable to curriculum. In addition, older students were characterized by more schooling and experiential learning, and students expressed an overall satisfaction with the college. Atkin concluded that knowledge of students' dominant learning stage might be useful in the process of curriculum planning or support services. In another study, Atkin (1998) tried to identify college students' dominant learning stages. Participants included 699 alumni of a Midwestern community college. These alumni had complete at least 30 credits hours transferable to baccalaureate degrees from other institutions. Data were obtained from a survey that incorporates three of the four dominant learning stages: absolute knowing, transitional knowing, and independent knowing. Prior to the survey, these participants had been last enrolled at the institution 1 or 3 years. About four-fifths of them had completed transfers to baccalaureate-granting institutions. Factors analysis was performed. Results indicated that more male than

female alumni were further along the continuum of preferred learning stages. Regarding age, alumni who had first enrolled at the institution at age 24 or older had progressed further on the continuum than had younger alumni. In terms of their satisfaction, among men, those at the middle, transitional, stage were less likely to rate the institution as helpful than those at either of the other two stages. The association between completion of bachelor's degree and the dominant learning stage was found only among women.

Another study examined the applicability of the model to research participants with different cultural background. Shaw (2001) conducted a qualitative study to examine the applicability of Baxter Magolda's Epistemological Reflection Model to assess ways of knowing of Black and Latino students. Data were collected using semi-structured interviews and the Measure of Epistemological Reflection (MER) questionnaire. Ways of knowing of research participants who were Black and Latino students at a large southeastern urban institution was compared to those of students in the original study by Baxter Magolda (1992b). Themes that emerged from the data analysis were found to depict important issues in the learning process for Black and Latino college students.

Compared to studies based on Perry's scheme, only few studies have been done applying the Epistemological Reflection Model. Following section describes King and Kitchener's Reflective Judgment Model (1994).

King and Kitchener's Reflective Judgment Model

Based on the idea of reflective judgment as initially introduced by Dewey, King and Kitchener (1994) spent more than fifteen years exploring how people decide what

they believed about ill-structured problems. They came up with a Reflective Judgment Model, which differs from critical thinking process. King and Kitchener described that the major differences include the epistemic assumptions about knowledge and the structure of problems. As for the matter of epistemic assumptions, they explain that critical thinkers have no awareness of the existence of uncertainty. Critical thinkers only assume the existence of absolute truth based on authority and do not recognize that situations can truly be problematic. For them, simple application of appropriately logical formulas and valid arguments can solve any kinds of problems. These thinkers see no need to validate or further generate solutions to problems based on the assumption that there are authorities that ultimately know absolute answers. Reflective thinkers, however, are fully aware of “the limits of knowing, the certainty of knowing, and the criteria for knowing” (Kitchener, 1983, p. 222). In other words, they are fully aware of the existence of problematic situations, where neither consensus solutions would ever be reached nor absolute solutions would ever be found. King and Kitchener argue that this knowledge or awareness about these assumptions of knowledge is a core element of reflective thinking and only adults hold this kind of epistemic assumptions. The other difference between critical thinking and reflective thinking is concerned with the structure of problems. Only ill-structured problems require reflective thinking, but not well-structured problems. Problem structure is composed of two elements: “the degree to which a problem can be described completely”; and “the certainty with which a solution can be identified as true or correct” (King & Kitchener, 1994, p. 10). Those problems, which can be described with a high degree of completeness and can be solved with a high degree of certainty, are well-structured problems. Usually, experts agree on correct

solutions of such problems. Ill-structured problems, on the other hand, are those that can neither be described with a high degree of completeness nor be resolved with a high degree of certainty. Disagreement among experts on the best solutions usually exists despite the fact that some of such problems can be considered solved.

The Reflective Judgment Model describes a developmental progression in reasoning from childhood to late adulthood. It describes how persons come to understand their process of knowing and how they correspondingly justify their beliefs when confronting ill-structured problems. King and Kitchener use the term “epistemic cognition” to explain the ability of individuals to evaluate knowledge claims and to explain and defend their points of views on controversial issues. The “reflective judgments” is described as the ultimate outcome of epistemic cognition development. According to the model, persons could be described as having pre-reflective thinking, quasi-reflective thinking or reflective thinking. Persons with pre-reflective thinking do not differentiate between well- and ill-structured problems. They view all problems as having a high degree of certainty and completeness. They do not acknowledge real uncertainty and assume that knowledge can only be perfectly right and certain. Such knowledge can be obtained either by direct, personal observation or from authority. Persons with quasi-reflective thinking, even though they assume the existence of ill-structured problems and uncertainty, have difficulty dealing with ambiguity of such problems and making judgments. Persons with reflective thinking assume that knowledge can be constructed based on inquiries and information from varieties of sources of information. These three developmental progression in reasoning is described

based on assumptions about knowledge and how it is acquired. King and Kitchener use the term “stage” to identify each set of assumptions that has its own logical coherency. There are 7 stages of development in the Reflective judgment model that they created. King and Kitchener regard the Reflective Judgment Model as a kind of stage model because firstly it clearly explains the logical relationships between the components of each stage that reflect an underlying organized structure, and secondly, it indicates the qualitative differences between these sets of assumptions as well as documents sequential changes in the emergence of these assumptions. Figure 2.4 shows the detail description of Reflective Judgment Stages adapted from King and Kitchener (1994).

In terms of research, King and Kitchener argue that the focus of research on the Reflective judgment Model is different from others research in epistemology in two major characteristics. First, most other models employ Piaget’s concept of concrete and formal operation to characterize the general form of adolescent and adult thinking. They regard the thinking beyond formal reasoning as a mature reasoning. King and Kitchener regard such thinking as logical reasoning, which is different from reflective thinking. For them, reflective thinking is a different intellectual domain. It is rather the development of epistemological assumptions. In other words, it concerns assumptions individuals hold about nature of knowledge. Individuals may assume if knowledge is an absolutely right or wrong existence, uncertain and contextual, or is constructed as an outcome of reasoning inquiry process. Second, for a reflective thinking to occur, people must “justify their beliefs when they are faced with ill-structured problems” – the problems that involve “real uncertainty” (King & Kitchener, 1994, p.41).

Figure 2.4. Summary of Reflective Judgment Stages

	View of Knowledge	Concept of Justification
Pre-Reflective Thinking		
Stage 1	Persons at this stage assume that knowledge exists as absolute, certain and concrete. Knowledge is obtained by direct observation because there is an absolute correspondence between what is seen or perceived and what is.	Persons at this stage do not perceive alternate beliefs nor the existence of controversies. Beliefs and truth are assumed to correspond to each other and they are both absolute. Thus, beliefs need no justification.
Stage 2	Persons at this stage assume knowledge to be absolutely certain or certain but not known to everyone. To obtain knowledge, one may directly use direct observation, or may obtain it from authority figures	Persons at this stage assume that most issues have a right answer. Disputed issues can be resolved by having the issues or the alternate beliefs justified by their correspondence with the beliefs of authority figures
Stage 3	Persons at this stage assume that knowledge is either absolutely certain or temporarily uncertain. In some areas, even authorities may not currently have the truth but the truth exists and will be discovered in the future.	In areas in which certain answers exist, persons at this stage make judgments on beliefs based on authority's views. In areas in which answers do not exist, they make judgments based on their personal opinion.
Quasi-Reflective Thinking		
Stage 4	Persons at this stage believe that one cannot know with certainty. Knowledge is also understood as an abstraction and not only limited to concreteness.	Persons at this stage understand justification as an abstraction and regard justification as a process that involves providing reasons and evidence as an essential part of an argument. Knowing is understood as idiosyncratic to the individual
Stage 5	Persons at this stage assume that knowledge depends on contexts and is limited because it is based on individuals' subjective interpretations of evidence. Only interpretations of evidence, events, or issues may be known.	Persons at this stage justify beliefs by considering contexts as frames of inquiry and interpretations of evidence.
Reflective Thinking		
Stage 6	Persons at this stage believe that knowing is a process that requires action on the part of the knower. For individuals to make conclusions about ill-structured problems, information from a variety of sources is needed. Knowledge is derived from a process of interpretations based on evaluations of evidence across contexts and on the evaluated opinions of reputable others.	Persons at this stage justify beliefs based on comparison of evidence and opinion from different perspectives on an issue or across different contexts and on construction of solutions. Such solutions need to be evaluated by certain criteria including the weight of evidence, practicality of the solution and pragmatic need for action.
Stage 7	Persons at this stage believe that knowledge is constructed from the process of reasonable inquiry that leads to solve ill-structured problems. The knowledge is validated on the basis of most reasonable or probable according to the current evidence, and can be later reevaluated when new evidence, perspectives, or tools of inquiry are available.	Persons at this stage justify beliefs employing the most probable variety of interpretive considerations. They defend conclusions as representing the most complete, plausible, or compelling understanding of an issue in relation to the current availability of evidence.

Research Validity and Reliability

To assess the reflective judgment level of an individual, the Reflective Judgment Interview or RJI is employed. Originally, King (1977) and Kitchener (1977-1978) developed four standard problems. Later on, King, Kitchener, and Wood (1990) developed the fifth problem for a ten-year longitudinal study. King and Kitchener (1994) explained that the questions focus on problems only in the intellectual domain. An illustration of the questions is “There have been frequent reports about the relationship between chemicals that are added to foods and the safety of these foods. Some studies indicate that such chemicals can cause cancer, making these foods unsafe to eat. Other studies, however, show that chemical additives are not harmful, and actually make the foods containing them more safe to eat.” The standard problems are supposed to be read aloud before trained interviewers asks the respondent a series of standard probe questions and ask follow-up questions as necessary to clarify or refocus responses. There are 7 probe questions such as “What do you think about these statements?, How did you come to hold that point of view?, On what do you base that point of view?, How is it possible that experts in the field disagree about this subject?”. The interviewers were trained to focus on the epistemological aspects of the given problems.

Since RJI is an interview, its reliability has been examined by means of various procedures including interrater reliability and agreement, test-retest reliability, and internal consistency. The interrater reliability, which is the most common index of rater consistency, is calculated by using a Pearson’s product-moment correlation. The interrater reliability fluctuates according to the heterogeneity of scores within samples

because the range of scores being compared affects the size of a correlation. Interrater agreement is another index used to indicate consistency across raters and to reflect actual agreements among raters in their assignment of the same score to a given protocol. Among 32 studies that employed RJI and are reported in King and Kitchener (1994), the interrater reliability ranges from .51 to .97 with two studies indicating .34 and .46 interrater reliability. From the same source of information, the interrater agreement ranges from 52%- 100%. In terms of the test-retest reliability, which indicates the degree to which respondents' scores remain stable over a relatively short period of time, King and Kitchener (1994) reported two studies that examined the test-retest reliability. One is a study by Sakalys (1984). From a three-month research with a sample of 25 senior nursing students, the test-retest reliability was found to be .71. Another study by Kitchener, Lynch, Fischer, and Wood (1993), with two of the standard RJI problems administered two weeks apart to 52 high school, college, and graduate students aged between 14 to 28 found the test-retest reliability to be .87.

Research Studies Based on Reflective Judgment Model

Studies that are based on King and Kitchener's Reflective Judgment Model examine different aspects of the theories. Some studies explore the application of theories with population of different backgrounds, some focus on factors that influence reflective judgment development, some explore the relationship between reflective judgment and other abilities, disposition, and personality, and others examine the measurement instruments to assess reflective judgment.

Dale (1995) explored reflective judgment in undergraduate college for clergy in conservative Christian college while Samson (2001) explored the reflective judgment in Hispanic college students. Dale (1995) examined the reflective judgment as described by King and Kitchener (1994) of students in a conservative Christian college. Participants were nontraditional freshmen, seniors, and faculty from an undergraduate college for clergy education. The measurement instruments employed in the study were Kitchener and King's Reflective Judgment Interview (RJI), two of the WAIS-R subtests to control for verbal ability, and an Impact of Faith questionnaire. Data revealed that freshmen's mean RJI score across three dilemmas indicated stage 3 on the Reflective Judgment Model and the mean for seniors indicated stage 4. However, controlling for verbal ability between freshmen and seniors, no a statistically significant difference in reflective judgment was found between freshmen and seniors. In addition, faculties' mean score across the three dilemmas indicated stage 5 and when controlling for verbal ability, a statistically significant difference in mean RJI scores was found between faculty and students. Considering content of the dilemmas, a statistically significant difference in participants' mean RJI scores was found between a religiously oriented dilemma and the dilemmas with secular content. In addition, there was a statistically significant negative correlation between mean RJI scores and responses to the Impact of Faith questionnaire for the creation dilemma. Samson (2001) pioneered a study of reflective judgment levels in Hispanic freshmen and senior college students. The association among their reflective judgment levels, their acculturation, and other background information was explored. King's and Kitchener's Reflective Judgment Interview and the Reasoning About Current Issues Test or RCI were employed to measure students'

reflective judgment levels. Data showed no statistically significant correlation between the RCI Overall and the RJI, or the RCI Sum High and the RJI. Statistical significant differences of scores between freshmen and senior students were found on the RJI Total mean score, the RJI Origin mean score, and the RCI Sum High score. However, no significant difference between freshmen and senior men was found on RCI Overall, or RJI Chemical mean scores. The overall regression, however, was only significant for the RCI Sum High and RCI Overall variable.

Following are studies that focus on factors that influence reflective judgment development. Cicala (1997) examined factors that influence college students' development of reflective judgment as described by Kitchener and King (1981) and King and Kitchener (1994). Based on a cross-sectional design, the study proposed and tested a model that involves students' overall involvement in their education as a mean to develop students' reflective judgment. Specifically, the study aimed to test the hypothesis that students' overall levels of involvement would be significantly positively correlated with their development of reflective judgment, and that such involvement would be more influential to such development than other variables. Participants were about 100 students in a single program of study in a single community college. They completed the College Activities Scales of the Community College Student Experiences Questionnaire by Friedlander, Pace, and Lehman (1990) and were interviewed with a modified version of the Reflective Judgment Interview. Data was analyzed with a focus on the relationships between students' ages, selected pre-college characteristics, educational progress, levels of academic and social involvement, and estimated levels of reflective

judgment. Findings suggested no statistically significant relationships between reflective judgment and any of the other variables in the proposed model, nor between overall student involvement and other outcome measures, such as grade point average or numbers of credits earned. However, statistically significant positive correlations were found between reflective judgment and the frequency of students' participation in class discussions and asking questions relating to class discussions or readings. Statistically positive, significant correlation was also found between the two activities and other major course activities including those that involved faculty contact and discourse. Cicala concluded that, rather than overall involvement, the specific, academically-oriented forms of involvement are more closely related to college students' levels of reflective judgment.

Pirttilae, Anna, and Kajanne (2001), in their attempt to test the hypothesis that specific adult experiences were stimulating development and that exploratory orientation was related to development, examined the development of implicit epistemologies of 59 Finnish adults during the period of 10 years. The Reflective Judgment Interview was administered at an initial pretest (1986-1988) and follow up posttest (1993-1994) periods. Results clearly indicated an increase in Reflective Judgment mean stage scores during the two study periods. Findings also suggested that education, especially the one beyond a person's primary profession, strongly predicts development. Encountering diversity and exploratory orientation was also found to relate to development. The associations were, however, more complicated. While no gender differences were found, the positive changes in thinking and reasoning was found to take place during adulthood.

Some studies explored the relationship between reflective judgment and other abilities, disposition, and personality. Jensen (1998) investigated the role of a dispositional/motivational variable, such as need for cognition, in relation to reflective judgment gains in college when age, verbal ability, and social desirability were controlled. Regarding reflective judgment as dependent variable, main effects for education level were found with effect sizes from .687 to .410, depending upon the number of covariates. As for gender, no gender main effects were found for reflective judgment. In addition, gender was not found to have significant effects on need for cognition or social desirability scores. Regarding need for cognition, it was found to predict reflective judgment at the freshman and the advanced graduate levels, but not at the senior and beginning graduate levels. It was also found to correlate positively and significantly with reflective judgment as were age and education level. In terms of social desirability, it was found not to have an effect on reflective judgment scores. Significant negative correlation was found between social desirability and reflective judgment. Undergraduate students scored significantly higher on the measure of social desirability than did the graduate students.

Guthrie (1997) explored the relationship between college students' levels of reflective thinking and their tolerance of diversity. The students' reflective thinking is defined based on King and Kitchener's Reflective Judgment Model and their tolerance was defined as low levels of prejudice towards African Americans and homosexuals. Guthrie employed three measures of intellectual development; two measures of prejudice (New Racism Scale, Jacobson, 1985) and Heterosexuals' Attitudes Toward Lesbians and

Gay Men Scale (Herek, 1988); a basic demographic questionnaire created for this study; and an interview about tolerance-related issues. The measures of intellectual development included two versions of the Reflective Thinking Appraisal (RTA) and a Reflective Judgment Interview (RJI). The relationship between tolerance and intellectual development was assessed using a correlational research design. The study aimed to further investigate the RTA and test the utility of creating a content-specific dilemma. Participants were 48 undergraduate and graduate students enrolled at a public university in the midwest. Sixteen of them were at each reflective thinking level. Findings indicated moderate positive correlations between the three measures of intellectual development and the combined tolerance score created from the two measures of prejudice. Guthrie concluded that there was a relationship between intellectual development and levels of tolerance, and that there is a critical level of reflective thinking ability for truly tolerant responses to different others: scoring above the mean on tolerance calls for reasoning at or above the quasi-reflective level, specifically Stage 4 reasoning.

Friedman (1996) studied the relationship between intellectual disposition and reflective judgment in college women. Participants were 43 college women who enrolled in education programs. They included 14 seniors, 15 master's candidates, and 14 doctoral candidates. Considering their majors as undergraduates, 14 were humanities majors and 27 were education majors. The participants were matched for educational level and intellectual disposition as measured by the Omnibus Personality Inventory (OPI). Their reflective judgment stages were obtained using the Reflective Judgment Interview (RJI).

Results from an analysis of variance indicated a significant relationship between educational level and reflective judgment. Individuals in a higher educational level were found to score higher in RJI. When intellectual disposition was treated as a discrete variable, a significant relationship between intellectual disposition and reflective judgment score was found only for the religion dilemma of the RJI. When intellectual disposition was treated as a continuous variable, Intellectual Disposition Category correlated significantly with Reflective Judgment. Six personality variables of the OPI correlated significantly with Reflective Judgment: Thinking Introversion, Response Bias, Altruism, Autonomy, Complexity, and Theoretical Orientation. Taking into account the academic majors, a statistically significant relationship between undergraduate major and reflective judgment was found. Those who majored in humanities scored significantly higher in reflective judgment than those who majored in education.

Kitchener and King (1981) examined the relationship between reflective judgment and age and education among 60 participants including high school, college, and graduate school students. The researchers hypothesized that verbal ability, Piagetian formal operation, socioeconomic status, and verbal fluency might affect reflective judgment scores. To obtain the data, in addition to RJI, they employed four other instruments. The Concept Mastery Test was used to assess students' verbal ability. Chemicals and pendulum task was employed to assess students' Piagetian formal operations. Two Factor Index of Social Position was used to obtain socioeconomic status, and number of words spoken during the interview was taken into account to represent verbal fluency. Highly significant differences were found on reflective

judgment level among the three-age/educational groups that were not statistically accounted for by scores earned on measures of the other 4 factors. Kitchener and King noted that although verbal ability was found to be closely related to reflective judgment level, this relationship could not account as a single factor for the differences between groups on reflective judgment.

There are also studies that examine measurement instruments to assess reflective judgment. Brabeck (1983) examined the relationship between 2 constructs that attempt to describe what constitutes good thinking: critical thinking as defined by the Watson-Glaser Critical Thinking Appraisal (CTA) and reflective judgment as described by Reflective Judgment Model. Participants were 119 females including high school seniors, college sophomores and seniors, and graduate students. They were matched on high and low extremes of CTA scores and were compared on the basis of their RJI scores. Results suggested the relationship between critical thinking ability and reflective judgment. Students at higher educational levels were found to achieve higher scores on the RJI. High critical-thinking students not out-performed low critical-thinking students on the RJI, but they also had significantly greater variability of RJI scores. Low critical-thinking students, however, were homogeneously low in reflective judgment.

Mines, King, Hood, and Wood (1990) studied the relationship between reflective judgment and standardized critical thinking tests. They also examined whether component critical thinking skills (CTSs) were present at some reflective judgment stages and not at others. Participants, including were 20 freshmen, 40 seniors, and 40 graduate students, completed 3 instruments. These instruments include the Reflective Judgment

Interview and the Cornell Critical Thinking Test. The graduate students had higher overall scores for each measure than the seniors, and the seniors had higher overall scores than the freshmen. Regarding the relationship between reflective judgment and critical thinking, students who reasoned using the assumptions of higher stages of reflective judgment are found to demonstrate better critical thinking skills than students who used lower stage assumptions. This finding suggested an association between developmental stage and acquisition of critical thinking skills.

The claim that Graduate Management Admission Test's (GMAT) Analytical Writing Ability (AWA) was a test of critical thinking had never been empirically validated. Boyd (2001), in an effort to validate such claim, examined if the differences in reflective judgment explain a significant portion of the variance in the Graduate Management Admission Test's (GMAT) Analytical Writing Ability (AWA) after differences in vocabulary and verbal ability were partialled out. The 44 state university students took GMAT, the vocabulary portion of the Nelson-Denny Reading Test and Reflective Judgment Interview. The analysis of statistical data revealed that after differences in vocabulary and verbal ability had been partialled out, RJI scores explained a significant portion of the variance in AWA essay scores. Boyd concluded that the GMAT essay was a valid measure of reflective judgment and the RJI explained other educationally relevant behavior.

Kozak (1996) explored the relationship between epistemological development and adaptation in terms of tolerating stressful affect and making causal attributions of interpersonal motivations. An experiment was conducted with the participation of both

male and female 100 undergraduates, half of which were randomly assigned to a stressful affect condition and half of which were assigned to a neutral control condition. Analysis of statistical data from Reflective Thinking Appraisal (RTA), Impact of Event Scale, and Thematic Apperception Test (TAT), provided support the connection between the adaptive capacities of affect-tolerance and attributions of pragmatic structure under conditions of stressful affect.

Hansen (1998), on researching among nursing staff aged between 18-35, employed the Reflective Thinking Appraisal to measure the relationship between smoking status and four cognitive variables including reflective judgment. The study revealed that there was no differences between the general reflective judgment levels of smokers, ex-smokers and non-smokers.

Dings (1998) investigated Reflective Thinking Appraisal or RTA, which is a paper-and-pencil version of Reflective Judgment Interview, in order to find evidence on the validity of the measure. Four studies were conducted to examine different features of the measure. The first study explored the problematic features of the instrument with the use of think aloud study of the RTA. The second study looked into the possibility if the RTA was limited in terms of its scope of generalization of ill-structured problems included in it by constructing a topic-independent, forced-choice measure. The third study sought to find out the concurrent validity of RTA by using the measure constructed in second study, the RTA, the RJI, and other plausible rival hypothesis measures. The last study experimented on the susceptibility of scores on the new measure and the RTA to faking. Results of these studies show that the RTA scores do not validly infer about

students' levels of reflective judgment because they are sensible to students' attempts to fake the scores. In addition, the scores neither correlated with each other nor as highly correlated with those in RJI as they should. Furthermore, the scores from RTA were found to correlate too highly with other measures. Moreover, the test items may be interpreted differently by test-takers than the intended meaning.

Like Perry's Scheme of Intellectual and Ethical Development, King and Kitchener's Reflective Judgment Model has been applied across different fields of studies, age groups, and educational and cultural background, for example. The following section describes Harvey, Hunt, and Schroder's Conceptual Development Theory (1961).

Harvey, Hunt and Schroder's Conceptual Development Theory

Harvey, Hunt, and Schroder (1961) defined the conceptual system as "a schema that provides the basis by which the individual relates to the environmental events he experiences. It describes, in part how he will perceive and experience these events" (p.244-245). They explained that the characteristics of conceptual systems include the interactive effects of both the individual and the environment the development and functioning of concepts depend on both external or stimulus factors and internal or motivational factors. They noted, "If we wish to predict what response a person will make, or to specify the conditions necessary for him to make a particular response, we approach the questions by considering both the person and the conditions in system-relevant terms" (p.245). In addition, conceptual systems are assumed to vary both in

terms of cognitive variables or information processing (degree of differentiation, integrative complexity), and in terms of motivation variables or interpersonal orientation (independence-dependence, empathic concern). To elaborate this concept he gives an illustration of persons at stage II compared to those at stage I. Persons at Stages II are both more highly differentiated and more independent than persons at Stage I.

Conceptual development is a continuous process that proceeds from concrete to the most abstract stages in order for persons to adapt to the environment and to relate self to others. On the basis of conceptual work that characterizes each stage of development, the developmental sequence is organized as follows:

Stage I. At this stage, the major developmental work is to define the external boundaries, and learn the generalized standards, which apply to both self and others. This is the basic assimilation of cultural norms and expectations. Persons at this stage are concerned with rules and appliance. They interpret events as either “good” or “bad”.

Stage II. Through a process of breaking away from the standard developed in Stage I, persons develop their self-delineation, which provides the basis for beginning to accept individual responsibility for outcomes. Even though persons at this stage tend to exaggerate their expression of independence, they first become aware of their own feelings as cues for differential action.

Stage III. Once persons have developed their self-delineation, they come to understand the feelings and experiences of others as being similar to, or different from, their own feelings and experiences. Even though persons at this stage may still

discriminate between others on the basis of acceptable standard at the concrete stage, their awareness of others make them generate a more highly differentiated interpersonal orientation.

Stage IV. Persons at this stage apprehend that both self and others occupy different positions on the same transcendent dimension under the same standards applicable to both self and others.

Under optimal conditions, the ultimate desired state of the development is expected to be abstract conceptual structure. Hunt (1961) noted “We believe that abstract conceptual structure and its associated characteristics of creativity, flexibility, stress tolerance, and broad-spectrum coping power is a desirable, adaptive stage. In short, it is a good thing!” (Quoted in Hunt, 1966, p.279). He also noted that conceptual evolvment is particularly concerned with an increase in effectiveness of adaptability to change rather than constant environment (Hunt, 1966).

As for the generality of developmental stages, Hunt (1961) explained, “A person need not reach the same level of abstractness of subject-object ties in all areas of development” (p.111). While some reach the fourth stage in many areas of development, others may function at the fourth stage in some areas, but at the second stage in other areas. Hunt (1966) referred to Carr’s (1963) investigation on the pattern of self-other relatedness at different stages using the Interpersonal Discrimination Test. The study indicated that the pattern of self-other orientation at Stage I, II, and III was exactly as

theoretically predicted. In other words, persons at different stages have exactly different characteristics in self-other orientation as described by the model.

Handcock (1994), refers to Hunt, Butler, Noy, and Rosser's explanation of conceptual level or CL as "a learner characteristic indexes on one's (a) conceptual complexity, as indicated by the ability to discriminate, differentiate, and integrate information and (b) interpersonal maturity, as indicated by self-definition and self-other relations. Low-conceptual-level (LCL) persons have relatively few cognitive structures and tend to minimize or avoid ambiguity; high-conceptual-level (HCL) persons are structurally more complex and use alternative thinking processes" (p.2).

Understanding that each individual has different level of flexibility and sophistication of the processing process, teachers and educators become aware that students' needs to learn have to be attended differently. Hunt (1966) also suggested a matching model of instructional environments to the level of the individual's conceptual functioning in order to maximize opportunities for learners to reach the desired state of abstractness.

Research Validity and Reliability

The Paragraph Completion Method (PCM) is a semi projective instrument, which has been used in a variety of contexts and it successfully defines and measures conceptual complexity. Mean interrater reliability has been reported as .86 with 1 year test-retest reliability intervals of .45 to .56 (Hunt, Butler, Noy, & Rosser, 1977).

The PCM test measures conceptual levels ranging from 0 to 3.0. A score of 0 indicates a low Conceptual Level. A score of 1 indicates moderately low Conceptual Level. A score of 2 indicates a moderately high Conceptual level, and a score of 3 indicates a high Conceptual Level. The test is composed six sentence items which are (1) "What I think about rules. . . "(Rules), (2) "When I am criticized . . . "(Criticism), (3) "What I think about parents. . . "(Parents), (4) "When someone disagrees with me. . . "(Disagreement), (5) "When I am not sure. . . "(Doubt), and (6) "When I am told what to do . . . "(Being Told). Test takers have three minutes to complete those sentences and write at least three additional sentences in response to each item.

To obtain test takers' score, a trained rater needs to assign a score of 0 to 3 to each of the six items and to aggregate these separate scores into a total CL score. The assignment of the 0 to 3 score depends on the extent to which an individual reveals conceptual complexity as indicated by the ability to discriminate, differentiate, and integrate information and interpersonal maturity as indicated by self-definition and self-other relations. The values 0-3 indicate the following (adapted from Hunt, Butler, Noy, and Rosser, 1977):

Score 0: A person is totally self-centered and resists being controlled by others. The person may either react impulsively to situations in a negative unsocialized manner, by losing her or his temper or becoming aggressive; or react defensively by withdrawing, ignoring the situation or blaming others. (These reactions are not associated with concern for right or wrong, desire for independence, or awareness of alternatives.)

Score 1: A person concerns with behaving in a socially acceptable way, and polarized or dichotomous thinking or behavior.

Score 2: A person is open to other people's ideas and evaluates alternatives. Even though the person has an increase in tolerance of uncertainly, ambiguity and difference of opinion, he or she is very much concerned with his or her own thoughts and feelings. Still, the person makes no attempt to integrate the evaluation of alternatives with the solution or decision.

Score 3: A person considers and weighs alternatives before making decision on the best possible solution to a particular problem. The person shows concern for his or her own and other's ideas and feelings, and about the possible consequences of his or her decision. He or she will accept full responsibility for the consequences of the decision.

A person's CL is calculated by averaging the highest three responses from a composite of six scores. Total CL scores of 0 to 3 indicate generic referents on the CL dimension as follows: (a) a score of 0, low CL; (b) a score of 1, moderately low CL; (c) a score of 2, moderately high CL; and (d) a score of 3, high CL.

Studies Based on Conceptual Development Theory

Morgan, Morgan, Foster, and Kolbert (2000) conducted a quasi-experimental study to investigate the effect of a Deliberate Psychological Education intervention (DPE) on the moral reasoning and conceptual development of law enforcement officer trainees and college students studying criminal justice; and the effects of new role-taking

on moral and conceptual development within a Deliberate Psychological Education intervention. The study was participated by 33 law enforcement officer trainees(31 males and 2 females) at a criminal justice academy in southeast Virginia (ages between 22 to 54 with an average of 30 years); and 31 students (15 males and 16 females) enrolled in a criminal justice course at a junior college in southeast Virginia (ages between 17 to 25 with the an average of 20 years). Overall the average age of the participants was 25. Participants were intact samples of police trainees and students so they were not randomly selected or assigned to the intervention an comparison groups. The experimental group was composed of 16 police trainees and 16 college students who enrolled in criminal justice course during the autumn semester. The course was taught using the Deliberate Psychological Education Model. Hunt's conceptual level matching model (CLMM) was also incorporated into this DPE programme to facilitate conceptual development or critical thinking. As for the control group, 17 police trainees and 15 college students enrolled in the criminal justice course during the spring semester following the intervention. The course was taught using the lecture method with minimal student discussion. Each course lasted for 10 weeks. Both courses were taught by the criminal justice professor except that for the former course the professor was supported by an assistant responsible for implementing the Deliberate Psychological Education Model to lead discussion and respond to writing assignments. Pre and post tests, using the Defining Issues Test (DIT) for measuring level of moral reasoning and the Paragraph Completion Method (PCM) for measuring level of conceptual complexity, were administered to both the experimental and control groups. Results of the experiment show that the intervention group scored significantly higher than the comparison group

on the DIT post-test. As for the PCM, no significant difference between the two groups was found even though the intervention group scored higher than the comparison group and the scores for the intervention group moved in a positive direction. In terms of the role-taking component of the DPE, since only the police trainees had opportunity to take a significant role in a helping profession, the results of the two tests between the police trainees and the college students were also compared. The results revealed the significant pre- to post-test change in moral development for both students and law enforcement officer trainees. The change in conceptual development, however, although not significant, was positive.

Konke (1984) investigate the teachers' conceptual level in relation to their preferences for support in the supervisory tasks of staff development, curriculum development, and instructional improvement. The 249 teachers from eleven elementary schools voluntarily participated in the study. The majority or 93% of the participants was females, 75% of them aged between 22-41, 58% of them had less than ten years teaching experience, 41% taught at the K-3 level, and 54% held teaching certificates at the T-5 (Master's) level. Teachers' conceptual levels were measured using the Paragraph Completion Method (PCM), and their preferences were measured by the Preferences of Supervisory Support Questionnaire. Findings reveal that only 14 teachers participating in the study had low conceptual level (but not below 1.0), 82 had moderate conceptual level, and 153 had high conceptual level. Although, overall, teachers indicated a strong preference for involvement with others in the three tasks of supervision and wanted flexibility and fluency in their teaching. No statistically significant correlation was found

between teachers' conceptual levels and the degree of their preferences for support in the supervisory tasks of staff development, curriculum development, and instructional improvement.

Coren and Suedfeld (1995) measure conceptual complexity of 277 university students using the Paragraph Completion Test (PCT). Results show that complexity was reliably related to a number of personality characteristics. Students with high conceptual level were found to be more extroverted, gregarious, and socially adapt, and showed more warmth and nurturance in social interactions. They also had higher sensation seeking tendencies and preference for nonconformity. The researchers mentioned that these findings are consistent with earlier studies regarding conceptual complexity as an independent characteristic combining aspects of intellectual and personality factors.

Harris (1981) examined conceptual complexity as measured by Paragraph Completion Test as a mediator of response to environmental stress. He studied 120 female undergraduates and found that those with low CL were more vigilant (behaviorally and cognitively) than those with high CL when anticipating a temporally unpredictable shock. Those with low CL also reported more subjective anxiety than those with complex CL. In addition, results suggest that those with complex CL exhibited higher levels of skin conductance overall than those with simple CL.

Bruch, Heisler and Conroy (1981) conducted 2 studies with 146 undergraduates with different conceptual complexity to compare their performance on various measures relevant to competent assertive behavior. The first study shows that students at high CL

demonstrated greater content knowledge, direct delivery skill, and fewer negative self-statements that inhibit assertiveness. The second study, which focused on the role of conceptual complexity in assertive encounters, reveals that high CL females were more assertive in difficult situations but did not differ in simple situations. In addition, the high conceptual complexity students were found to be more assertive in extended interaction tests, express consideration of the needs of others, and were more flexible in sex-role orientation.

This section has identified scope of the review, described related theories including Perry's Scheme of Intellectual and Ethical Development, Baxter-Magolda's Epistemological Reflection Model, King and Kitchener's Reflective Judgment Model, and Harvey, Hunt, and Schroder's Conceptual Development Theory. It has also provided information on assumptions, and validity and reliability issues of the theories. In addition, it has reviewed literature of studies related to these theories. The following section reviews related measurement instruments.

Review of Related Measurement Instruments

Since this dissertation compares studies that employed different measurement instruments, it is worth discussing the comparability of the scores obtained from different instruments. Regarding measurements for Perry's Scheme of Intellectual and Ethical Development, there have been studies on measurement instruments to depict intellectual differences or to assess epistemological development among college students described by Perry. The original unstructured open-ended interview employed by Perry and his associates is considered to provide "the highest degree of accuracy" (Baxter Magolda &

Porterfield, 1985, p.434) and achieved maximum respondents' freedom to project their own frame of reference, which was regarded as essential to accurate measurement for measuring cognitive development (Baxter Magolda & Porterfield, 1985). It also "establishes rapport and allows presuppositions and frames of reference of the interviewee to emerge" (Belenky, Clinchy, Goldberger, & Tarule, 1986, p.10). Furthermore, the interview data is considered to be "the richest source of information about the complex forms of personal epistemology and cognition" (Moore, 1987, p.39). However, interview requires extensive training of the interviewer and raters. This affects research reliability. In addition, the interviewing process is time-consuming, thus inefficiently cost-effective. (Baxter Magolda & Porterfield, 1985; Moore, 1987). The drawbacks of the interview approach brought about needs to establish measurement tools employing different techniques. The instruments later constructed include production, comprehension, and recognition tasks (Moore, 1987). Moore (1983b) noted that Knefelkamp and Widick created the Measure of Educational, Personal, and Vocational Concerns or —shortened to the Knewi in honor of its two co-authors. Benack (1984) and Slepitz (1983) constructed structured interviews. Taylor (1983) developed Measure of Epistemological Reflection (MER). Moore (1987) developed Learning Environment Preferences (LEP). Due to the scope of this dissertation, this section only discuss the MID and MER as measurement instruments to assess students' intellectual development as described by Perry's scheme.

Measure of Intellectual Development (MID)

Knefelkamp (1974) and Widick (1975) developed the first major alternative to the interview format, a combination of sentence stems and semi-structured essay tasks initially call the Measure of Educational, Personal, and Vocational Concerns—shortened to the KneWi in honor of its two co-authors. As noted by Moore (1982), the Maryland Rating Group developed a comprehensive set of rating cues to score responses for the two KneWi essays. Knefelkamp adapted the essays to be used as a separate instrument, the Measure of Intellectual Development (MID).

As described by the Office of Institutional Research and Planning, the University of Texas at Arlington, the MID is designed to measure patterns of longitudinal intellectual development of students or for pre- and post- evaluations of specific courses or groups of courses. Regarding the reliability and validity issue, an inter-rater reliability of 51.2% was based on results from more than 1200 essay interviews. In addition, individual studies have consistently shown inter-rater reliability ratings to be over .90. One test-retest correlation over a period of a month was reported to be .87. Moore (1982) reported interrater reliabilities for the MID of .58 for absolute agreement among expert raters and .83 for dominant position agreement. Rogers, Mentkowski, Hart, and Minik (2001) reported the inter-rater reliability in the Alverno Longitudinal Study to be .82 for essay A, .78 for essay B, and .75 for essay C. The MID was validated by its correlation with related measures including the Conceptual Level instrument. The MID has been shown to detect developmental change over time. It is also noted that concurrent validity studies did not reflect differences according to gender.

The MID is limited to measuring positions two through five on the Perry scheme. Respondents write three short essays describing (a) a best class, (b) a recent decision, and (c) their career choice (Knefelkamp, 1974; Knefelkamp & Slepitz, 1976; Widick, 1975). Two trained raters score the responses individually before discussing the rating and reconciling the differences of the two ratings. Moore (1986a) explained that a 3-digit number is used to represent individual rating. The number reflects the dominant and subdominant position rated in the essay. This system ‘stretches’ the Perry scheme continuum from 4 step –2, 3, 4, and 5—to 10. The scores assigned are 222, 223, 233, 333, 334, 344, 444, 445, 455, and 555. While the solid ratings 222, 333, 444, and 555 reflect a ‘stable’ position, the two steps between each stable position reflect transitions between positions. Moore elaborated that while 223 represents dominant position 2 opening to position 3, 233 indicates dominant 3 with trailing position 2. Moore (1986a) noted that the ratings of the essay indicate the respondents’ cognitive complexity with respect to classroom learning along a linear, simple stage model continuum. In addition to the 3-digit numbers, the ratings may also be treated as an interval scale through the conversion of rating to numerical scores. An example includes the conversion of the score 222 to become 2.0, 223 to become 2.67 (Moore, 1986b).

Although scholars agree that the MID has been the most valid and reliable among measures designed to assess students’ development as described in Perry’s scheme (King, 1978; Mines, 1982; and Moore, 1983; and Perry, 1981), Mines (1982) critiqued that scoring procedures could possibly lead to underestimation of students’ level of understanding complex reasoning. In addition, since MID requires respondents to write

answers, there are possibilities that although respondents could recognize higher levels of complex reasoning, their writing ability might limit them to accurately demonstrate their levels of complex reasoning.

Measure of Epistemological Reflection (MER)

Taylor (1983) developed the Measure of Epistemological Reflection or MER including two features relating to Perry's scheme and the interview instrument. The two features are domains of development and stimuli to elicit students' thinking. Taylor selected six domains of development that occurred in the context of the formal learning process. These domains are (a) decision making; (b) role of the learner; (c) role of the instructor; (d) role of peers; (e) evaluation; and (f) view of knowledge, truth, or reality. The MER is composed of series of questions for each domain. The first question of each series focuses on the respondent's thinking on each particular domain. This question is followed by four probes that focus on elicitation of respondent's reasoning. An illustration of a series of questions on role of the instructor includes: "During the course of your studies, you have probably had instructors with different teaching methods. As you think back to instructors you have had, describe the method of instruction that had the most beneficial effect on you." Follow-up probes are, for example, "What made that teaching method beneficial? Please be specific and use examples." "What are the most important things you learned from the instructors' method of teaching?" (Baxter Magolda, 1992b, p.423). Along with the construction of MER, Taylor also constructed a scoring manual composing of six sections for the six domains. Respondents' protocols were separated by domain and within domain by assigned position and were reviewed to

identify categories of responses based on reasoning structure. The reasoning structure was defined as the basic justification for the respondent's thinking and regarded as the unit of analysis in the scoring process. To obtain the overall scores or the Total protocol ratings (TRPs) for each respondents, raters need to establish the modal reasoning structure across the six domains and account for reasoning structures that occurred in two or more domains. Baxter Magolda and Porterfield (1985) noted that the MER scoring manual was designed to address Positions 1 through 5 of the Perry scheme. This is due to the fact that Position 5 was considered a logical boundary for an initial scoring process because of the relativistic cognitive structure that served as a basis for the latter positions on the scheme and because of the lack of clarity of the cognitive-structural aspects of the latter positions.

During the MER construction process, its reliability and validity were evaluated. Regarding the reliability, Cronbach's alpha coefficient and interrater agreement were employed to assess the MER internal consistency. As for the validity, in an original study the MER scores were compared to MID scores for concurrent validity, and scores were analyzed by level of education for construct validity. Pearson product-moment correlations calculated for the MER and the MID were found to be extremely low: .11 and .04 for the derivation and cross validation samples. Baxter Magolda and Porterfield (1985) noted that correlation between the MID and the MER was low because of the lack of similar variability in the MID scores. While the MID means revealed very little variance across levels of education and did not increase with level of education, the MER differentiated between levels of education. They also noted that the MER is an accurate

measure of intellectual development on the Perry Scheme and its format enhance the quality of the response and still maintain the production of response. Baxter Magolda (1987b) reported a correlation of .93 between the MER interview and essay formats. As well, the standardization of the MER was reported including the correlation of .80 ($p < .0001$) for interrater reliability achieved by two expert raters on a sample of 752 and by interrater agreement ranging from 70 to 80 percent (Baxter Magolda & Porterfield, 1985).

Reflective Judgment Interview (RJI)

King and Kitchener (1994) explained that the RJI questions focus on problems only in the intellectual domain. Originally, King (1977) and Kitchener (1977-1978) developed four standard problems. Later on, King, Kitchener, and Wood (1990, 1991) developed the fifth problem for a ten-year longitudinal study. In addition, discipline-based problems were also developed (2 psychology problems by Phillip Wood and Kurt DeBore, 2 business problems by Karen Kitchener and Cindy Lynch, and 2 chemistry problems by Patricia King and David Finster). King and Kitchener explained that the questions focus on problems only in the intellectual domain. The interviewers were trained to focus on the epistemological aspects of the given problems. The five problems are concerned with (a) the construction of pyramids, (b) the objectivity of news reports, (c) religious beliefs and scientific explanations about the origin of the world and human beings, (d) the safety of chemical food additive, (e) safety of nuclear energy. After the interviewer reads each problem aloud to the respondents, they respondents are asked to state and justify their opinion about the issue and to respond to follow-up questions. These standard problems

are supposed to be read aloud before trained interviewers ask the respondent a series of standard probe questions and ask follow-up questions as necessary to clarify or refocus responses. The probe questions include “What do you think about these statements? On what do you base that point of view? How is it possible that people have such different points of view about this subject? How is it possible that experts in the field disagree about this subject?” The standard probe questions aim to elicit respondents (a) reaction to the problem, (b) derivation to a particular viewpoint, (c) basis of their viewpoint, (d) assumption on certainty of knowledge, (e) basis of their justification of alternative interpretations, (f) perception of their understanding of differences in perspectives and opinions, and (g) perception of their understanding about the use of experts’ or authorities’ viewpoint in their decision making. Responses are scored based on stages of development with focuses on respondents’ perception of nature of knowledge and nature of justification.

Since RJI is an interview, its reliability has been examined by means of various procedures including interrater reliability and agreement, test-retest reliability, and internal consistency. The interrater reliability, which is the most common index of rater consistency, is calculated by using a Pearson’s product-moment correlation. The interrater reliability fluctuates according to the heterogeneity of scores within samples because the range of scores being compared affects the size of a correlation. Interrater agreement is another index used to indicate consistency across raters and to reflect actual agreements among raters in their assignment of the same score to a given protocol. Among 32 studies that employed RJI and were reported in King and Kitchener (1994),

the interrater reliability ranges from .51 to .97 with two studies indicating .34 and .46 interrater reliability. From the same source of information, the interrater agreement ranges from 52%- 100%. In terms of the test-retest reliability, which indicates the degree to which respondents' scores remain stable over a relatively short period of time, King and Kitchener (1994) reported two studies that examined the test-retest reliability. One is a study by Sakalys (1984). From a three-month research with a sample of 25 senior nursing students, the test-retest reliability was found to be .71. Another study by Kitchener, Lynch, Fischer, and Wood (1993), with two of the standard RJI problems administered two weeks apart to 52 high school, college, and graduate students aged between 14 to 28 found the test-retest reliability to be .87.

Paragraph Completion Method (PCM)

Hunt, Butler, Noy, and Rosser (1977) explained that the PCM is a semi-projective method to assess Conceptual Level (CL). Responses are considered as thought samples: how the person thinks. The scoring process requires rater's clinical judgment so raters need training to score the responses. The PCM require respondents to write ideas about six topics. Prior to taking the PCM test, respondents are informed that there is no right or wrong answer and they are supposed to indicate their actual feeling about the topic as oppose to the way others feel they should feel about those topics. They are also informed that they should write at least three sentences on each topic and they have about three minutes to write responses on each topic. The six topics include: (a) What I think about rules..., (b) When I am criticized..., (c) What I think about parents..., (d) When someone does not agree with me..., (e) When I am not sure..., (f) When I am told what to do... .

These topics are included to elicit different samples of thinking such as how the respondents handle conflict or uncertainty (topics b, d, and e), and how they think about rules and authority relations (topics a, c, and f).

A person's CL score is obtained through 2 steps: 1) assigning a score from 0 to 3 to each of the six responses, and 2) aggregating three separate scores into a total CL score. The focus of the scoring is on the way persons think in relation to an increase in conceptual complexity as indicated by discrimination, differentiation, and integration, and an increase in interpersonal maturity as indicated by self-definition and self-other relations. In addition, both thought structure and content of the responses are considered for scoring. Following are general characteristics of thinking at different levels of conceptual development indicated by scores 0-3 as explained by Hunt, Butler, Noy, and Ross (1977).

Respondents with a score 0 may react either impulsively to situations in a negative unsocialized manner or react defensively by withdrawing, ignoring the situation or blaming others. With the impulsive reaction, persons at this CL might lose their temper or become aggressive. They are totally self-centered, focusing on what they want, like, feel and believe. They do not consider other people's thoughts and/or feelings. In addition, they resist being ruled or controlled by other people. With defensive reaction, persons may express their hostility or anger differently, which may indicate characteristics of different CL. Respondents with a score of 1.00 try to behave in a socially acceptable way concerning with rightness and wrongness or good and bad behavior. They are sensitive to what authority figures such as teachers or parents would

perceive about their behaviors. They try to conform their behaviors to those that are considered correct, acceptable, and expected by authority figures. Respondents with a score of 2.00 open to other people's ideas and evaluate alternatives. This implies that they have an increased tolerance of uncertainty, ambiguity and difference of opinion. To solve problems or make decisions, however, persons at this CL are still so concerned about their own thoughts and feelings that they do not try to integrate the evaluation of alternatives in their solution or decision. They are striving for independence. Respondents with a score of 3.00 make decisions based on the best possible solution to particular problems. They take into consideration alternatives for their decisions, their own and others' ideas and feelings, and possible consequences of their decisions. As opposed to compromising their values, principles or beliefs to please others or to conform, they make decisions based on all possible alternatives to benefit the majority regardless of how others think about them. In addition, they will accept full responsibility for the consequences of their decisions. Persons assessed at CL 3.00 are secure in their independence, awareness of themselves, perception by others, and relationship with others.

Hunt, Butler, Noy and Rosser (1977) reported that the PCM has been administered to several thousand persons. Between 1964-1976, across 26 research studies including about 1051 respondents, the median inter-rater coefficient was found to be .86. Based on one-year intervals of data from 503 students in grades 6-13, the test-retest reliability coefficients for CL was found to range from .45 to .56. Based on shorter

intervals (from three months) of data from 36 college students, Gardiner and Schroder (1972) reported the test-retest reliability coefficient for CL of .67.

The description of the four instruments MID, MER, RJI, and PCM suggests that the resulting data from the instruments are compatible. Although the questions are different, all these four instruments aim to elicit some common information including: (a) respondents' personal reaction and opinion about the prompts or problems raised (This reflects their current position, stage, and level of thinking.); (b) their perception about the problems (This implies their assumptions about knowledge, its certainty and sources e.g. authorities.); (c) the rationale underlying their opinions or how they reach their opinions (This implies how they justify their beliefs). In addition to the elicitation of common information, these four instruments are based on the connected theories of which characteristics of position, stage, and level of development are comparable.

Regarding the description of the scores in relation to the theories, despite of the differences in numbers assigned to responses, the description of the development is comparable. Figure 2.5 provides description of characteristics of position, stage, and level of development as indicated by scores. It is noticeable that there is an overlap of characteristics of persons as described by different score scale in relation to how they view knowledge, authority figures, sources of knowledge, perspectives of others, and process of arriving at conclusions about issues or problems. The horizontal line in the figure marks where each position, stage, or level overlap as indicated by different scores from different instruments.

Figure 2.5 Characteristic of position, stage, and level of development as indicated by scores of different instruments.

Position in Perry's Scheme as assessed by MID*, MER	Stage in Reflective Judgment Model as assessed by RJI	Conceptual Level in Conceptual Complexity Development Theory as assessed by PCM
		Score 0 Self-centeredness. Persons only think that what they want, like, feel and belief is important than what others think or feel. Rules and controlled are resisted.
Score 1 Knowledge is absolute, certain, concrete, objective, right or wrong. Authority knows the answer for everything. Rightness is an outcome of hard work and obedience.	Score 1 Knowledge is absolute and concrete and not understood as an abstraction. Direct observation brings about knowledge. Beliefs need no justification and alternate beliefs are not perceived.	Score 1 Knowledge is polarized or dichotomous --Rightness and wrongness or good and bad. The rightness and wrongness is judged by authority figures and social norm.
Score 2 Uncertainty of knowledge and diversity of opinion is confusing and interpreted as lack of qualification by Authority.	Score 2 Knowledge is absolutely certain but not immediately available. Direct observation and authority figures are sources of knowledge. Most issues have a right answer and beliefs can be justified based on the beliefs of an authority figure.	
Score 3 Diversity and uncertainty are accepted as legitimate but still temporary until Authority discovers the answer.	Score 3 Knowledge can be either absolutely certain or temporarily uncertain. Beliefs are justified based on reference to authorities' views in areas of absolute certainty and based on personal opinion in areas of uncertainty.	Score 2 Uncertainty, ambiguity and difference of opinion exist. Alternate perspectives should be taken into consideration when solving problems or making decision. However, persons care more about their own thoughts and feelings so they do not try to integrate the evaluation of alternatives in their solution or decision. They are striving for independence.
Score 4 Uncertainty is legitimate and depends on personal opinion. Still rightness and wrongness exist and can be justified by qualitative contextual relativistic reasoning of Authority.	Score 4 Knowledge is uncertain and idiosyncratic to individual. Reasons and evidence are used to justified beliefs. The choices of reasons and evidence are idiosyncratic.	
Score 5 All knowledge and values are perceived as contextual and relativistic. The dualistic belief is a special case in context.	Score 5 Knowledge is contextual and subjective. Knowledge claims are the outcomes of interpretations of evidence, events or issues based on individuals' perceptions and criteria for judgment.	

*MID is limited to measuring positions two through five on Perry Scheme.

Figure 2.5 Characteristic of position, stage, and level of development as indicated by scores of different instruments. (Cont.)

Position in Perry's Scheme as assessed by MID*, MER	Stage in Reflective Judgment Model as assessed by RJI	Conceptual Level in Conceptual Complexity Development Theory as assessed by PCM
	<p style="text-align: center;">Score 6</p> <p>Knowledge is constructed through the process of interpretations of information from variety of sources and evaluations of evidence across contexts, and opinions of experts or reputable others. To justify beliefs, evidence and opinion from different perspectives or across different contexts are compared. Then, a set of criteria is used for evaluating the constructed beliefs. Such criteria might take into account the weight of the evidence, the utility of the solution, or the pragmatic need for action.</p>	<p style="text-align: center;">Score 3</p> <p>Persons make decision based on the best possible solution to a particular problems. They take into consideration alternatives for their decision, their own and others' ideas and feelings, and possible consequences of their decision. As opposed to compromising their values, principles or beliefs to please others or to conform, they make decision based on all possible alternatives to benefit the majority regardless of how others think about them. In addition they will accept full responsibility for the consequences of their decision.</p>
	<p style="text-align: center;">Score 7</p> <p>Knowledge is constructed through the process of reasonable inquiry that solves ill-structured problems. Beliefs are justified probabilistically as representation of the most complete, plausible, or compelling understanding of an issue based on current available evidence.</p>	<p>Persons with CL 3 are secure in their independence, awareness of themselves, perception by others, and relationship with others.</p>

So far, this chapter has addressed the importance of intellectual development in higher and teacher education and provided historical contexts for related research. It has described the four selected theories, validity and reliability issues, and related literature. It has also discussed how MID, MER, RJI, and PCM as measurement instruments used for evaluation of developmental growth are comparable. The following section proposes connections among the four theories.

A Proposed Connection among the Selected Conceptual and Epistemological Development Theories

It is worth noting again that Hofer and Pintrich (1997) categorized Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reflection Model (1992), and King and Kitchener's Reflective Judgment Model (1994) as epistemological development theories. They, however, did not include Harvey, Hunt, and Schroder's Conceptual Level Theory (1961). Pascarella and Terenzini (1991) categorized Perry's Scheme of Intellectual and Ethical Development (1970), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt, and Schroder's Conceptual Development Theory (1961) as cognitive structural theories. Pascarella and Terenzini explained that common features among these cognitive theories include hierarchical series of stages of development. Each stage of development is a prerequisite to movement on to the next. Such progression is believed to be irreversible, universal and transcultural. The major emphasis of these theories is on "how meaning is structured, not on what is known or believed" (p.28).

In fact, when King and Kitchener (1994) created their Reflective Judgment Model, they also used the work of Harvey, Hunt and Harvey, Hunt, and Schroder (1961). Although King and Kitchener did not explain how they use the work, they noted, "while the Reflective Judgment Model was built on the work of both Perry and Broughton (as well as the work of Harvey, Hunt, and Schroder, 1961, and Loevinger, 1976), it differs from them both" (p. 38). In the critique of their scheme, Perry and associates also related their work to the work of Harvey, Hunt, and Schroder (1961). One of the remarks they

made was that their work was parallel to that of Hunt's matching model, in which students were grouped according to "stages" for instructional purposes. (Perry, 1999, p.231).

In fact, all these four theories assume stage-like development. King and Kitchener (1994) remarked that general findings from research on Reflective Judgment Model suggest that development follows a stage-related pattern described by the Reflective Judgment Model and relate to age, and amount of academic involvement. In addition, findings from research that compared student and nonstudent adults, those without degrees, with lower degrees and with advanced degrees suggested that educational settings facilitate development, and educational activities tend to help individuals improve their reasoning about ill-structured problems.

Hunt Conceptual Level (CL) theory describes persons on a developmental hierarchy of increasing conceptual complexity, self-responsibility, and independence. Hunt (1971) explains that the development is a continuous process that can be developed to the highest conceptual level under optimal conditions evolves in a given order.

Perry and associates also related their work to Piaget's concept of formal operations. They note that their scheme traced the movement of the assimilation and accommodation process with the emphasis on "the structural changes in a person's assumptions about the origins of knowledge and of value" (Perry, 1970/1999, 229). Widick (1977) noted that the Perry scheme is appropriately classified as a cognitive developmental model since it describes development with a focus on the students'

internal cognitive structure. In addition, Perry's scheme describes intellectual development as it occurs in a generally irreversible sequence of stages. Each stage represents a qualitatively different structure for perceiving the meaning of knowledge.

Baxter Magolda (2002) noted that she is a social constructivist and her work is based on, among others, Piaget's (1950) notion of cognitive structures of equilibration process, Kegan's (1982) notion of meaning making through experience, and King and Kitchener's (1994) notion of epistemic assumptions.

Reiman (1999a) notes that the assumptions of cognitive structural theories are drawn from the cognitive sciences, cognitive-developmental psychology, and the biological sciences. The assumptions can be summarized as follows.

1. All humans process experience through cognitive structures, and all persons have an intrinsic motivation to be competent.
2. Such cognitive structures are organized in hierarchical stages or plateaus from less complex to more complex.
3. Growth occurs first within a particular stage and then only to the next leap to a significantly more complex system of processing experience.
4. Growth is not automatic not unilateral but occurs only with appropriate interaction between the person and the environment.
5. Growth occurs across a series of partially independent domains.
6. Behavior can be determined and predicted by an individual's level of development. Predictions are not, however, exact.
7. At the micro-developmental level, persons must acquire new skills within a level of stage in order to progress within and between stages. The difference between the optimal level or upper limit of a person's capacity and their functional level represents their developmental range or zone of proximal development.

(Reiman, 1999a, p.96-97)

Perry's Scheme of Intellectual and Ethical Development (1970), and Baxter-Magolda's Epistemological Reflection Model (1992b), King and Kitchener's Reflective Judgment Model (1994), and Harvey, Hunt, and Schroder's Conceptual Development Theory (1961) have features that are consistent with the cognitive-development assumptions. As well, they have common assumptions about information, knowledge, and problems. Assumptions about intellectual development include the following:

1. The hierarchical stages or plateaus of development move from less complex to more complex levels;
2. Development from one stage or position to the next stage takes place first within the current stage of development of each individual prior to the developmental transition to the next stage, which marks significant qualitative transformation change;
3. Development occurs as a result of interaction between individuals and their surroundings, thus, there is a contextual dynamic in stage growth;
4. There is a connection between stages of development and individuals' actions or behaviors, however, the relationship is not one-to-one;
5. Individuals need to assimilate and accommodate their cognitive structures in order to be able to develop from one stage to others; and
6. All persons have an intrinsic need to be competent and self-determining, and such competence grows when there is positive interaction in a supportive and optimally challenging environment.

Regarding assumptions about information, knowledge, and problems, these theories assume the following:

1. Information/knowledge/problems are derived not only from one, single source but multiple sources;
2. The structure of information/knowledge/problems can be either simple and well-structured or complex and ill-structured;
3. For development to take place, individuals need to be able to differentiate the structures of information/knowledge/problems; and
4. Intellectual development is best assessed when persons engage in complex new experiences with ill-structured problems.

Common features among these theories about knowledge can be summarized in Figure 2.6. The features are organized on the basis of plateau of development, contexts of development, position of self in relation to environment, perception about knowledge and its complexity, and perception about sources of knowledge. Since development move from less complex to more complex levels, only the polarity of aspects of development is demonstrated. The arrow indicates the movement from the starting point of development to the highest point of development as described by each theory.

Figure 2.6. Common features among the four selected theories.

	Scheme of Intellectual and Ethical Development (Perry, 1970)	Epistemological Reflection Model (Baxter Magolda, 1992b)	Reflective Judgment Model (King and Kitchener, 1994)	Conceptual Development Theory (Harvey, Hunt, Schroder, 1961)
Plateau of Development	Dualistic ↓ Relativist	Absolute knowing ↓ Contextual knowing	Pre-reflective ↓ Reflective	Concreteness ↓ Abstractness
Context of Development	Interaction with environmental demand such as academic requirements, campus social life	Interaction with environmental demand such as academic requirements, campus social life	Interaction with Environmental demand such as ill-structured problems requiring justification	Interaction with new concepts that lead to refutation and confirmation of current knowledge about concepts
Position of Oneself and Others	Initially focus on self, self-responsibility, obedience ↓ Personal commitment	Obtain knowledge ↓ Integration, exchange, sharing knowledge	Obtain knowledge ↓ Perspective taking	Self-center ↓ Subject-object relatedness, Self-other relatedness
Perception about Knowledge and its Complexity	Dualistic, absolute certain ↓ Validation of self identity	Certain ↓ Contextual and justifiable	Certain and concrete, non existence of abstract ↓ Temporarily, contextually validated	Concrete (externally controlled) ↓ Abstract (internally controlled)
Perception about Sources of Knowledge	Existence of absolute knower as a source of knowledge ↓ Relativistic (Personal commitment)	Existence of absolute knower responsible for transferring knowledge ↓ Multiplicity of perspectives from different sources	Existence of absolute knower ↓ Knowledge is constructed based on reasonable inquiry and can be reevaluated to verify or falsify	Absolutistic, concrete, ready-made conceptual criteria ↓ Abstract, alternative criteria

The theories reviewed are connected as summarized in Figure 2.6. Research based on these four theories reveals the stage-related patterns of progress from viewing knowledge as absolute, concrete and right or wrong to relative, abstract and contextual. While persons at the earlier stages of development were found to depend on authority as knowledge generator, those at later stages depend more on themselves. Persons at higher stages also viewed knowledge as more complex and ambiguous. Movement from one stage of development to others was found to be in sequence without skipping any stages. As well, the development is caused by stimuli that demand a different cognitive structure than the existing current ones to solve problems. Furthermore, persons at higher stages of development take others' perspectives. They also become more independent thinkers and make judgments that are more appropriate.

It has been stated earlier that chapter two is organized in accordance with the order of research questions posed in this dissertation. So far, this chapter has addressed the importance of intellectual development in higher education and teacher education. It has provided historical contexts for research on intellectual development in higher education. It has also reviewed the literature related to common assumptions and connections among four selected theories. The next section presents a review of literature related to research question four: what are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

The two convenient samples of four-year longitudinal studies examined the effectiveness of a Deliberative Psychological and Professional Education (DPPE)

program for teacher education students. Why did the two samples employ the DPPE program as an intervention to foster developmental growth in teacher education students? The detailed description of the DPPE is presented in chapter four. The following section reviews literature on the DPPE program.

Review of Research on a Deliberative Psychological and Professional Education Program (DPPE)

As described by Sprinthall, Reiman, and Thies-Sprinthall (1996), deliberative programs can be defined as psychological development-oriented education programs that integrate conditions that promote cognitive developmental growth. The programs are based upon longitudinal and clinical research that indicates a relationship between cognitive developmental stage and personal decision-making and professional behaviors. In addition, these programs have applied design principles (e.g., new social role taking, guided inquiry, balance, support and challenge, and continuity) with the goal of promoting development. These programs are also based on the concept that the level of cognitive developmental stage maturity can be promoted under five certain conditions as first described by Mosher and Sprinthall (1971) and summarized in Sprinthall, Reiman, and Thies-Sprinthall (1996). These conditions include learners' new role-taking helping experiences, careful reflection of the experiences, balance of the real experience with reflective inquiry, and the sustained continuity of the three.

Thies-Sprinthall (1984) conducted two studies: One in 1979-1980 and the other in 1981-1982. The purpose of the research was to examine if a training program designed

based on cognitive developmental theory could positively affect classroom supervising teachers. Research participants included a total of 10 classroom supervising teachers in the first study, and 12 teachers in the second. The program employed as an intervention was a cognitive-developmental intervention. It provided conditions needed to promote psychological growth. These conditions are 1) new-role taking, 2) careful and continuous guided reflection, 3) balance between real experience and discussion/reflection, 4) instruction that provides for both personal support and challenge, and 5) continuity of the program. The training approach had four components: describing and understanding the mode, viewing the model, planning and peer teaching, and adapting the model. The study used two assessment instruments: the Hunt Conceptual System Test and the Rest Defining Issues Test. The two instruments were administered as pre and post tests. Results of the first study indicated some positive trends since research participants slightly gained their scores on both instruments tests after 6 months. However, no statistical significant gain was found at the .05 level. After the first study, the researcher considered that learning tasks had not been adequately differentiated according to level of CL of research participants. Consequently, the program was changed around more structure and concrete directions for low CL participants and less structure and added theoretical readings and research projects of the high CL participants. Results of the second study were more positive. To elaborate, the Hunt CST mean change from pretest to posttest was from 1.60 to 1.80 ($t=2.35 < .043$). On the DIT, the mean change was from 12.4 (41%) to 15.8 (53%). The researcher noted that findings of the studies indicate a possibility to create educationally meaningful programs to affect the level of psychological stage development.

Sprinthall and Scott (1989) conducted an experimental study with 30 secondary school 11th-grade girls randomly selected from a small rural public school. The experimental and control groups had the same number of subjects, 15 in each group. The participants in the experimental group volunteered for tutoring, and the comparison group volunteered as teacher aids. Prior to the 15 weeks of the experiment, the participants in the experimental group were trained to tutor 4th and 5th grade students. The training program was developed by school guidance counselor and a secondary school math teacher by following guideline for DPE. The program provided the participants in the experimental group with fundamentals of psychological role-taking designed to help them understand and meet the learning needs of elementary students. In addition to regular tutoring to 5th graders and preparation of instruction and materials, the participants meet weekly with math teacher and counselor in a seminar. The seminar provided the participants with opportunities to reflect on and to evaluate the previous tutoring sessions. They were also provided with support for the planning of following tutoring, In addition, they were provided with time to record their experiences in journals.

The tutees, 5th graders were also divided into experimental and control groups. The experimental groups attended the tutoring sessions by the trained 11th grade experimental participants while the control group did not.

The comparison group of teacher aids spent the same amount of time as those in the experimental group. They received a series of brief instructions from each teacher on skills such as correcting papers, collating materials, explaining homework assignments, and monitoring the class when the teacher was not present. In other words, the

participants in the control group were involved in a structured-service learning experience without the use of journals or reflective discussion, but positive reinforcement and attention from the teachers.

To assess the development in 11th graders, the Hunt conceptual level (CL) test and the Rest (1979) Defining Issues Test (DIT) were used. To assess achievement and attributions for success and failure of the 5th graders, California Achievement Test and Math Attribution Scale were used. Results indicated that the process of role taking under the condition of DPE supported significant gains in levels of psychological development. The gains were explained to be associated with greater independence in thinking and autonomy in value judgment. The development was supported by evidence from weekly journals. In addition, there was a development in conceptual levels in 11th grader experimental participants.

Regarding the 5th grade tutees, the experimental group improved by more than 11 points on the California Achievement Test compared with a gain of more than 4 points by the controls.

The most update research on the effectiveness of DPPE programs was conducted by Reiman (2001). His comparative analysis includes 12 studies that employed the programs. Findings reveal significant averaged positive effects for role-taking interventions ranging from +.48 (N=9, conceptual/epistemological thinking), to +.59 (N=12, moral reasoning), to +.59 (N=3, ego development). Some of the studies found correlation between the intervention and performance. Seven out of nine studies reported

the significant positive changes at the .01 level. Nine of the twelve studies reported an effect of the intervention on performance. The effectiveness of the program includes changes in techno stress, increases in ability to engage tutee, increases in responsiveness to students' ideas and feelings, increases in motivating interactions with learners, and increases in posed inquiry-based questions.

This chapter has addressed the importance of intellectual development in higher and teacher education. It has provided historical contexts for research on intellectual development in higher education. It has also reviewed the literature related to common assumptions and connections among four selected theories. This chapter finally reviews literature related to the DPPE program. Chapter three provides detail of research methodology for this comparative analysis study.

CHAPTER III

RESEARCH METHODOLOGY

The major aim of this dissertation is to explore selected theories of epistemological and conceptual complexity development in order to propose connections among them. In addition, it reports the results of two convenient samples of four-year longitudinal studies (1988-1992 and 1989-1993) of a Deliberative Psychological and Professional Education program (DPPE) for teacher education students at North Carolina State University. It compares the results of these two convenient samples to other selected longitudinal research and reports the effect size among these studies. Moreover, it investigates in detail the elements of programs that are intentionally designed to promote conceptual complexity development in teacher education and compares them to other programs within the reviewed studies.

Specifically, the four research questions to be examined are:

1. What are common assumptions and connections across selected theories of intellectual development?
2. What are findings from the convenient samples of the two four-year longitudinal studies (1988-1992 and 1989-1993) that incorporated the Deliberative Psychological and Professional Education program (DPPE) for teacher education students?

3. What are the relationships between the findings from the two convenient samples and other longitudinal samples based on the connected theories of intellectual development?
4. What are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

This chapter describes research methodology employed in this dissertation.

Research Design

The comparative analyses of longitudinal research studies concerning intellectual development are based on four theories: Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reasoning Model (1992b), King and Kitchener's Reflective Judgment Model (1994) and Harvey, Hunt and Schroder's Conceptual Development Theory (1961). As described in chapter two, despite the difference in the terms employed to explain each of these theories, they have common theoretical assumptions and assumptions about knowledge. Two longitudinal sample studies on conceptual complexity development based on Harvey, Hunt and Schroder's Conceptual Development Theory (1961) are reported. Findings from these samples are compared to those from eight selected one- to four-year longitudinal studies based on the four theories just mentioned. The comparative analysis aims to examine the magnitude and trends of college students' intellectual development to nourish the realm of knowledge in this field. The ten studies, depending on which of the four theories they are based on, employ different research instruments. It is, thus, fundamental to examine the

comparability of these 10 studies in terms of their instruments. This is included in chapter two in the section that reviews related measurement instruments. Following sections provide information on research methodology, participants, and variables.

Methodology

In chapter two, connections were made among the four theories: Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reasoning Model (1992b), King and Kitchener's Reflective Judgment Model (1994) and Harvey, Hunt and Schroder's Conceptual Development Theory (1961). In accordance with the common conceptual assumptions among the theories, studies are selected to be compared and analyzed based on proposed criteria (included as a measurement instrument in this chapter).

As previously mentioned, since the 1970 publication of Perry's Scheme of Intellectual and Ethical Development, many research studies have employed the Perry Scheme as their theoretical framework. While some researchers extend and even create their own new theories, others try to find different types of instruments for data collection and assessment. Some of those research studies provide only qualitative evidence (Belenky et al., 1997) while others provide both qualitative and quantitative evidence (Baxter Magolda, 1990, 1992b; Baxter Magolda, & Porterfield, 1985). In addition, regarding research methodology, some studies are cross-sectional while others are longitudinal in scope.

The methodological focus of this dissertation is to compare the effect size of selected longitudinal studies to ascertain major trends of findings from studies based on connected theories. In addition to quantitative analyses, this dissertation qualitatively analyzes and compares the ten selected studies.

Research Participants

The target population of this study is undergraduate students of diverse background who participated in one to four year longitudinal research studies focusing on development or change in intellectual, conceptual complexity or level, and reflective judgment. This comparative analysis study includes research studies that meet the established criteria. Included in the analysis are two convenient samples of four-year longitudinal studies (1988-1992, and 1989-1993) of teacher education students at North Carolina State University. Other studies, as well as the two convenient samples, meet the established criteria.

Research Variables

The dependent variable for this study is change in level of intellectual development, conceptual complexity, and reflective judgment as a result of college attendance. Such change or development is indicated by an increase or gain in scores obtained from a measurement instrument administered at the end of junior or senior year as compared to the scores from freshman year.

The independent variable is college education from freshman to junior years.

This variable includes college curriculum in general and deliberative programs intentionally designed to promote students' development in the convenient samples. These programs are described in detail and compared to other programs described in other research samples.

Measurement Instruments

Measurement instruments include criteria for research sample selection, scheme for data encoding, and scheme for analysis of program characteristics.

Criteria for Research Sample Selection

One of the instruments to be used in this study is the set of criteria that is set for research selection both in terms of the basis of theories and measurements employed in those research studies. The researcher constructs this measure under the approval of academic advisors. Applying criteria for meta-analysis study as suggested by Lipsey and Wilson (2001), it is important for a comparative analysis study to establish eligibility criteria for inclusion of studies. The criteria should include such general categories of information as the distinguishing features of a qualifying study, research respondents, key variables, and research design. Consequently, eligible research studies to be included in this comparative analysis need to meet the following criteria:

1. Distinguishing features. The study must examine student development in related dimensions of intellectual domain of development based on common, cognitive structural assumptions such as those reviewed in the literature. The development

is determined by measurement tools in association with the developmental constructs that are previously reviewed in the literature in chapter two. The development is considered as a result of college attendance either with or without any particular intervention programs.

2. Research respondents. The study must involve college students as research participants, regardless of the diversity across race, ethnicity, and gender, who choose to attend colleges that provide undergraduate programs including teacher education and/or any other professional disciplines.
3. Key variables. The study must include quantitative outcome of conceptual and other related dimensions of development previously reviewed. They must be 1-4 years longitudinal studies that examine the development of during any of the undergraduate years or between any of those undergraduate years.
4. Research design. The study must include compatible assessment instruments that generate interval data on pretest and posttest scores. These assessment instruments are: King and Kitchener's Reflective Judgment Interview (RJI) (1994), Hunt, Butler, Noy, and Rosser's Paragraph Completion Method (PCM) (1977), Measure of Epistemological Reasoning (MER) (Taylor, 1983; Baxter Magolda & Porterfield, 1985; Baxter Magolda, 1987), and Measure of Intellectual Development (MID) (Knefelkamp, 1974). The quantitative data must also be accessible.
5. Time frame. According to an initial survey by the author, there are a few longitudinal studies with a focus on the effect of education in undergraduate college on students' conceptual development because such research takes time to

be completed. In this dissertation, studies to be analyzed include those conducted in the 1970s, if they meet all the criteria, up to the present.

6. Publication type. One of the main objectives of this dissertation is to report the results of two convenient longitudinal studies at College of Education, North Carolina State University. As well, this study aims to compare these results with other related studies and meta-analyze them. Both published and unpublished studies are deemed eligible to be included in this dissertation.

Scheme for Data Encoding

Another instrument is a coding scheme to be used for encoding information about characteristics of research samples and about empirical findings of them. The coding protocol to be used for collecting data in this research will include the following major units. (Part of this protocol is taken or adapted from Lipsey and Wilson (2001).

Figure 3.1. Scheme for data encoding.

<ol style="list-style-type: none"> 1. Characteristics of studies or study descriptors <ol style="list-style-type: none"> 1.1 Study title 1.2 Constructs 1.3 Samples <ol style="list-style-type: none"> 1.3.1 Sample source 1.3.2 Sampling procedures 1.3.3 Demographic information 1.4 Methodology <ol style="list-style-type: none"> 1.4.1 Design 1.4.2 Measurement 1.4.3 Quality of measures 1.4.4 Variables 1.4.5 Threats to internal and external validity 1.4.6 Forms of data analysis 1.5 Findings 	<ol style="list-style-type: none"> 2. Effect size coding <ol style="list-style-type: none"> 2.1 Variables represented in the effect size 2.2 Point in time when variables measured 2.3 Subsample information 2.4 Sample sizes (effect size specific) 2.5 Means or proportions 2.6 Standard deviations or variances 2.7 Calculation procedure (effect size specific) including estimation methods 2.8 Missing data 2.9 Reliability of variables represented in effect size 2.10 Type of statistical test of effect used in study
--	---

Scheme for Analysis of Program Characteristics

The third tool used in data analysis in this dissertation is the scheme for analysis of the deliberative programs designed to foster intellectual growth. This scheme will be used to analyze detailed characteristics of the programs. The scheme is created by the present researcher based on ideas from Program Evaluation Handbook (Serow, 1998)

Figure 3.2. Scheme for analysis of Deliberative Psychological and Professional Development Education (DPPE) and other programs.

-
1. Composition of the programs
 - 1.1 Purposes of the programs
 - 1.2 Detail description of the program
 - 1.3 Persons involved and their roles
 - 1.4 Materials

 2. Program implementation process
 - 2.1 Intervention activities
 - 2.2 Any specific required activities
 - 2.3 Sites for intervention implementation
 - 2.4 Period and length of the program implementation
 - 2.5 Evaluation of students' achievement

 3. Program Evaluation
 - 3.1 Instrumentation
 - 3.2 Type of evaluation e.g. formative or summative
 - 3.3 Follow-up

 4. Research findings about the effectiveness of the programs
-

Data Analysis and Procedures

Procedures to be followed are described following each research question respectively.

Research Question One

What are common assumptions and connections across selected theories of conceptual development?

Chapter two answers this research question following qualitative analysis procedures. First, selected theories on epistemological and conceptual development are described in detail. Then, common features among the selected theories are analyzed. The analysis is based on an analysis of common assumptions among the selected theories. These common assumptions are established from comparative analysis of the detailed elements of the selected theories extensively described and presented in chapter two.

Research Question Two

What are findings from the convenient samples of the two four-year longitudinal studies (1988-1992 and 1989-1993) that incorporated the Deliberative Psychological and Professional Education program (DPPE) for teacher education students?

To answer this question, the two convenient samples of four-year longitudinal studies are reported. These studies employed Hunt, Butler, Noy and Rosser' Paragraph Completion Method (1977) as a tool to assess students' Conceptual Level (CL) at the

beginning of their freshman year and at the end of their senior year in teacher education program at College of Education, North Carolina State University. The pre- and post-test scores, as well as the gain scores of each study are reported separately by employing a statistical significance. In addition, the statistical results of these two studies will also be converted using effect size statistic (e.g. “standardized mean gain of pre-post contrasts”). These two studies are then compared with other studies that use different measurement instruments. The data analysis procedure is applied from procedures for meta-analysis as suggested by Lipsey and Wilson (2001). To compare two variables that differ only with regard to time of measurement, pre-post contrasts is used. Since the selected studies to be compared employed varied measurement instruments (e.g., Paragraph Completion Method or PCM, Reflective Judgment Interview or RJI, and Measure of Epistemology Reasoning or MER, Measure of Intellectual Development or MID), it is important that pre-post contrasts be standardized in such a way that the values can be meaningfully compared across samples and studies vis a vis effect size.

Research Question Three

What are the relationships between the findings from the two convenient samples and other longitudinal samples based on the connected theories of intellectual development?

The answer to this question includes a comparative analysis of the two convenient samples and eight selected studies, and a quantitative analysis of effect sizes of the ten studies. As noted by Lipsey and Wilson (2001), Becker (1988) developed an effect size

statistic for pre-post contrasts in the form of the standardized difference between the Time 1 and Time 2 means for meta-analysis applications. This statistic is defined as:

$$ES_{sg} = \frac{\bar{X}_{T2} - \bar{X}_{T1}}{s_p} = \frac{\bar{G}}{s_g / \sqrt{2(1-r)}}$$

$$SE_{sg} = \sqrt{\frac{2(1-r)}{n} + \frac{ES_{sg}^2}{2n}}$$

$$\omega_{sg} = \frac{1}{SE_{sg}^2} = \frac{2n}{4(1-r) + ES_{sg}^2}$$

\bar{X}_{T1} is the mean at time 1 and \bar{X}_{T2} is the mean at time 2. \bar{G} is the mean Time 2 minus Time 1 gain score, s_p is the pooled standard deviation of the Time 1 and Time 2 scores. To be specific $\sqrt{(s_{T1}^2 + s_{T2}^2)/2}$, s_g is the standard deviation of the gain scores, n is the common sample size at Time 1 and Time 2, and r is the correlation between the Time 1 and Time 2 scores. (Lipsey & Wildon, 2001, p. 44)

In this dissertation, effect sizes are calculated from the mean gain score (mean Time 2 – mean Time 1) divided by the pooled standard deviation of the Time 1 and Time 2. After the enumerative findings of each selected studies are converted into standardized mean gain of pre-post contrasts, they are compared and presented in a table.

Research Question Four

What are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

The two convenient samples of four-year longitudinal studies (1988-1992 and 1989-1993) incorporated the Deliberative Psychological and Professional Education program (DPPE) for teacher education students. One of the purposes of these studies is to examine the effectiveness of a role taking/reflection curriculum purposefully designed to foster conceptual growth. This type of program has been used elsewhere in research literature. As well, two of the selected studies provide information of their programs. This study will consequently examine these programs in terms of its common characteristics. It should be noted, however, that since not all research included in this study has this information available, only those that have information available will be included. From the detailed description of the analysis and the common characteristics of the programs, conclusion will be drawn about programs characteristics that foster students' development.

Threats of Validity

One of the threats to internal validity of this design might be due to the differences in measurement instruments that each research sample uses. However, this threat is minimized by first providing information on the common theoretical assumptions of different types of measurements from the available studies in published and referred literature. Another way to minimize such a threat is to standardize the numerical data using appropriate effect size statistics such as “standardized mean gain of pre-post contrasts” before effect size across studies are compared. The major threat of external validity will be due to the number of research studies to be compared. However, since there are not many relevant longitudinal studies available in literature, the threat,

however, is minimized by the inclusion of the samples in its entirety or as many samples available as possible.

Ethics and Human Relations

Since this study used findings from convenient samples and from research studies available, there will be no threats posed to research participants.

CHAPTER IV

RESULTS AND DISCUSSION

This comparative analysis research addresses the need for a better understanding of higher education effects on undergraduate students' intellectual development. One of the aims of this study is to clarify different terms and definitions, which have been employed regarding constructs and theories of college students' intellectual development. The rationale underlying the exploration of this issue is to identify common assumptions across four selected theories of intellectual development: Perry's Scheme of Intellectual and Ethical development (1970); Baxter Magolda's Epistemological Reflection Model (1992b); King and Kitchener's Reflective Judgment Model (1994); and Harvey, Hunt and Schroder's Conceptual Development Theory (1961). In addition, in light of the availability of research based on these four theories, the present study attempts to add more research evidence through analysis of two convenient samples of undergraduate teacher education students. Trends of college students' intellectual development are analyzed through the process of comparative analysis of two convenient samples and other selected longitudinal studies of intellectual, conceptual and reflective development as a part of college education.

The four examined theories are employed in a variety of ways in classroom contexts. Intervention programs are now being designed to promote intellectual development. Thus, this dissertation also explores and analyzes some intervention programs in order to summarize common elements of interventions that support

development. Such an exploration is expected to enhance the design of curriculum, which serves to foster undergraduate students' intellectual development. This research, consequently, attempts to answer the following four questions:

1. What are common assumptions and connections across four theories of intellectual development?
2. What are findings from the convenient samples of the two four-year longitudinal studies (1988-1992 and 1989-1993) that incorporated the Deliberative Psychological and Professional Education program (DPPE) for teacher education students?
3. What are the relationships between the findings from the two convenient samples and other longitudinal samples based on selected theories of intellectual development?
4. What are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

Since research question one is included in chapter two as a part of the literature review, this chapter provides answers for research questions two, three, and four. First, findings are reported for the two convenient four-year longitudinal samples that examine conceptual development of teacher education students at North Carolina State University. Then, findings from the two convenient longitudinal samples are compared with other eight selected studies that are based on the four selected theories (question three). Both qualitative and quantitative findings are reported. Analysis then turns to research

question four: distinctive features among three education programs designed to foster intellectual development. Chapter four follows this outline:

1. Two convenient longitudinal samples: participants and intervention, research methodology, findings, discussion, and summary of the studies
2. Relationships across ten longitudinal sample studies
 - a. Part one: quantitative findings
 - b. Part two: qualitative findings
3. Analysis of distinctive features of education programs designed to foster intellectual development
 - a. Part one: description of programs
 - b. Part two: analysis of program features

Findings related to the second question are now described. The findings address both quantitative and qualitative data from two convenient samples of teacher education students who participated in two four-year longitudinal studies.

Research Question Two

Two Convenient Longitudinal Samples

Participants and the Intervention

Two longitudinal studies were conducted at the College of Education, North Carolina State University, during 1988-1992 and 1989-1993 academic years. These studies aimed to describe and monitor the development of the students participating in

this program during their undergraduate years. The research participants were from the population of pre-service teacher education students who enrolled in a Teaching Fellow Program, chose to attend North Carolina State University, and completed all requirements and graduated in four years. The first group of participants was composed of 49 students including 35 females, 13 males, and one student without gender identification. The second group was composed of 44 students including 33 females and 11 males.

These teacher education students in both samples attended a teacher education program that was intentionally designed to foster learning and development across three related domains i.e. principled moral judgment, reflective/conceptual judgment, and relational judgment. The education program was designed based on an applied theory for cognitive structural development called the Learning/Teaching Framework (LTF). The LTF is composed of six developmental conditions and an instructional coaching model with four components. The six conditions are (a) contextual understanding and building trust, (b) complex new social role-taking, (c) guided inquiry for self-analysis of performance and reflection, (d) balance between action and reflection, (e) continuity in interaction with conditions over at least six months, and (f) support and challenge. The coaching model has four components: (a) model description, (b) model viewing/demonstration, (c) model practice with feedback, and (d) model adaptation. This theory-based and research-based program was implemented throughout the four-year period of each study.

Research Methodology

A program for new teacher education students was developed and tested that built on the LTF conditions just described. Given the appropriateness of intellectual development as a needed college and teacher education outcome variable, North Carolina State University decided to institute two longitudinal studies of the cognitive growth of teacher education students across their four-year college experience. There were 49 students in the first cohort and 44 students in the following year's cohort. All entering students were from North Carolina.

North Carolina State University is a Research I land-grant institution with 28,400 students. The longitudinal research was undertaken to better understand growth in a series of developmental domains including the dimension of conceptual complexity. An overall purpose was to provide faculty and administrators with an understanding of whether more deliberative educational programming that included a role-taking and guided inquiry framework would be more powerfully educative on a series of developmental dimensions including the moral judgment dimension. The Paragraph Completion Method, which measures conceptual complexity, was administered as one of three developmental measurement systems. The most complete analysis of the test was undertaken by Miller in 1981. The test emphasizes conceptual complexity and interpersonal complexity.

The teacher education students in the two samples were administered a battery of assessments including the Paragraph Completion Method (PCM) during their first week

at the University. At that time, students were informed about the purpose of the study, advised that participation was voluntary, and assured that all results would be confidential. Students completed a letter of informed consent to voluntarily participate. The battery of assessments took 60 to 90 minutes to complete. The PCM was the second instrument in the packet to complete. After testing was completed, participants placed all materials into their envelopes. Investigators collected all envelopes and then assigned a four digit number to each set of assessments. Only the principal investigators had access to the names and assigned numbers of the study participants. To protect participants' identities, numbers only were used in data analysis. Posttesting was completed during the month of April of the students' senior year.

Instrumentation

As previously described, the two studies employed a battery of tests. However, only the results of the relevant instruments to this dissertation are included. The Paragraph Completion Method or PCM was used to measure the students' conceptual level, or CL, or their cognitive complexity as well as their interpersonal maturity. Hunt, Butler, Noy, and Rosser (1977) created this semi-projective test. The test comprises six open-ended topic stems: three stems assess the way individual thinks and about conflict and certainty, and another three stems that assess the way individual thinks about rule, structure, and authority relations. According to Hunt, Butler, Noy, and Rosser (1977), completion responses are considered as "thought samples" which are scored according to how a person thinks" (p.1). Responses were scored under a two-step procedure. The first step is to assign a score from 0-3 to each of the six responses. The second step is to

aggregate the separate scores into a total conceptual level score. The focus of the scoring procedure is to look into conceptual level in terms of 1) an increase in conceptual complexity as indicated by discrimination, differentiation, and integration, and 2) an increase in interpersonal maturity as indicated by self-definition and self-other relations. The scores 0 to 3 represent some characteristics. A person's thinking is scored 0, for example, if his or her thinking reflects that he or she is totally self-centered and resists impulsively to situations of being controlled by others. A person's thinking is scored 1, if his or her thinking reflects that he or she thinks in such a way because of his or her concerns with behaving in a socially acceptable way, and polarized or dichotomous thinking or behavior. If a person is open to other people's ideas and evaluates alternatives but still concerned with his or her own thoughts and feelings, the person is scored 2. If, however, a person considers and weighs alternatives before making decision on the best possible solution to a particular problem without personal bias, the person is scored 3.

Findings: Conceptual Change in Teacher Education Students
in Two Longitudinal Samples

For both longitudinal studies, participants had higher gain scores on the PCM during posttesting. Findings from both studies suggest that over the period of four years, students developed conceptually. There were no significant differences according to gender.

Table 4.1 provides detailed results of the first study. The first cohort (1988-1992) was composed of 48 participants including 13 men and 35 women. The mean

scores obtained from Hunt’s Paragraph Completion Method or PCM refer to research participants’ conceptual level or CL. Pretest refers to the administration of the PCM at the beginning of the freshman year. Posttest refers to the administration of the PCM at the end of the senior year. The maximum numbers are the highest score level that research participants obtained and the minimum numbers are the lowest score level that research participants obtained. The effect size is the magnitude of change of CL during the beginning of the freshman year and the end of senior year.

Table 4.1

Conceptual Level of pre-service Teacher Education Students (Cohort 1988-1992) as Indicated by Hunt’s Paragraph Completion Method

	Variable	N	Mean	Std Dev	Max.	Min.	Effect Size
All participants	Pretest	48	1.65	0.28	2.50	1.00	1.07
	Posttest	48	2.01	0.39	2.80	0.80	
	Gain	48	0.36	0.40	1.10	-0.50	
Men	Pretest	13	1.70	0.34	2.50	1.20	
	Posttest	13	2.05	0.33	2.70	1.50	
	Gain	13	0.35	0.43	1.00	-0.50	
Women	Pretest	35	1.63	0.25	2.00	1.00	
	Posttest	35	1.99	0.41	2.80	0.80	
	Gain	35	0.37	0.39	1.10	-0.50	

Overall, the average PCM scores at the beginning of their freshman for all participants was 1.65. Among them, the highest score was 2.5 and the lowest score was 1.00. It is worth noting, however, that for the pretest, only 1 of 48 students scored 2.5. The majority of students scored between 1.00-2.0. By the end of their senior years, these students had their average PCM score of 2.01. The average gain score of all participants

was 0.36. The paired t-test shows that this gain score represents a statistical positive mean difference between the pretest and posttest with $t=6.32$, $p<.0001$. This indicates a significant increase in level of conceptual complexity development. Considering the gain scores for all participants, the maximum score gain was 1.10, which is more than one stage of conceptual development. The results imply that students in general developed conceptually during their college years.

Regarding gender, men had an average pretest or the freshman PCM score of 1.70 and the posttest or senior PCM score of 2.05. Their gain score, which was 0.35, also indicated a positive movement of conceptual level. As for women, their average PCM score from the pretest during freshman year was 1.63 and is similar to that of men. At the end of senior year, they also made similar progress as men did with a mean gain of 0.37. No statistical difference was found between the gain scores of men and women. Both men and women in this first study made similar progress in their conceptual complexity.

The second study was a replication of the first study. The second cohort included 44 students enrolling in the same program during 1989-1993. There were 33 women and 11 men in this longitudinal sample. The freshmen entering in 1989 had the mean CL score of 1.85, higher than that of the first cohorts (1.65). At the end of the senior year, their average score increased as shown by posttest mean CL score of 2.07, also higher than that of the first cohort (2.01). The mean PCM gain score of the second cohort was, however, lower than that of the first cohorts, 0.22 in the second study as opposed to 0.36 in the previous study. In fact, no statistically significant difference was found between the pretest and posttest for this second cohort. Table 4.2 summarizes findings.

Table 4.2

Conceptual Level of Pre-service Teacher Education Students (Cohort 1989-1993) asIndicated by Hunt's Paragraph Completion Method

	Variable	N	Mean	Std Dev	Max.	Min.	Effect Size
All participants	Pre	44	1.85	0.34	2.80	1.10	.59
	Post	44	2.07	0.40	3.00	1.20	
	Gain	44	0.22	0.47	1.00	-1.10	
Men	Pre	10	1.74	0.30	2.20	1.20	
	Post	10	2.12	0.43	2.70	1.20	
	Gain	10	0.38	0.57	1.00	-0.80	
Women	Pre	34	1.88	0.35	2.80	1.10	
	Post	34	2.06	0.39	3.00	1.30	
	Gain	34	0.18	0.43	1.00	-1.10	

As for the matter of gender, men in this study began the freshman year having lower CL (1.74) than women (1.88). However, they made more gain than women by the end of the senior year (0.38 compared to 0.18). In fact, this difference is statistically significant in favor of men making more gain than women ($t=6.9, p<.005$). It is also worth noting that for study one, men also had average CL scores higher than that of women both from pretest and posttest. Yet, their gain scores at the end of senior year were comparatively the same, 0.35 for men and 0.37 for women.

In general, at the end of senior year, students participating in the fellowship program between 1988-1992 and 1989-1993 developed conceptually as indicated by the PCM scores. Based on conceptual complexity theory, the students can be assumed to enter the program with polarized or dichotomous thinking. They viewed knowledge as certain and is either right or wrong. By their senior year, they recognized others' ideas

and alternatives, and understood the ambiguity of knowledge. Consequently, they can be assumed to have the potential to recognize their students' differences in public school classrooms, and to consider alternatives for teaching before making instructional decisions.

The last columns of Tables 4.1 and 4.2 contain effect size of the two studies. This statistic, as previously described in chapter three, is a quantitative way of describing the magnitude of the difference between the pretest and posttest scores of the PCM tests. The first study has the effect size of +1.07, while the second one has the effect size of +.59. In other words, the students in the first study developed more in terms of their conceptual complexity when compared to those in the second study. For further explanation, Tables 4.3 and 4.4 are included to show the percentage of program participants at different CL level at the beginning of the freshman years and at the end of their senior years.

Table 4.3 shows the percentage of students at different CL level at the beginning of freshman years and at the end of senior years for both cohorts. The conceptual levels are organized from higher complexity level to the lower complexity level. The pretest column shows how many students were at each level when they entered the freshman year and the posttest column shows how many of them were at each level at the end of their senior years.

Table 4.3
Proportion of Students at Different Conceptual Levels at Freshman and Senior Years

Conceptual Level	Cohort 1988-1992 (n = 48)		Cohort 1989-1993 (n = 44)	
	Pretest	Posttest	Pretest	Posttest
Higher than 2.0	2.00%	35.42%	22.73%	50.00%
2.0	12.5%	27.08%	22.73%	18.20%
Between 1.1-1.9	81.25%	34.52%	54.55%	31.80%
1.0	4.00%	0.00%	0.00%	0.00%
Lower than 1.0	0.00%	0.01%	0.00%	0.00%

The numbers in table 4.3 demonstrate dramatic change in students in their conceptual complexity level before and after attending the teaching fellow program at North Carolina University. Considering the first cohort, more than 90% of participants developed conceptually. By the end of the senior year, more than 35% of the students were at CL levels higher than 2.0, which is 15 times greater than at the time of pretesting (2%). Another more than one time, 27.08%, as much as that in the beginning, 12.5%, moved up to CL 2.0. In addition, more than 45% of students who were at CL between levels 1.1-1.9 in the beginning (81.25 % compared to 34.52%) moved to upper CL levels (2.0 and higher than 2.0). As for the second study, even though the percentages do not show as dramatic change as in the first study, the change is as well remarkable. More than 25% of the students who scored at lower CL levels in the pretest were at CL levels higher than 2.0 (22.73% compared to 50.00%).

Table 4.3 shows the percentage of students at different CL levels at the time before and after attending the teaching fellow program. It does not reveal, however, if any students stayed at the same level or scored lower in their posttest. Table 4.4 reports this data.

Table 4.4

Proportion of Students CL Development Based on Directional Trends

Trends	Cohorts 1988-1992 (n = 48)	Cohorts 1989-1993 (n = 44)
Upward	81.25%	61.36%
Stay the same	7.5%	13.64%
Downward	12.5%	25.00%

Despite the dramatic change in students' CL levels as shown in Table 4.3, Table 4.4 shows that not all students demonstrated upward trends. Not only did some students stay at the same CL levels during their college years (7.5% in study one and 13.64% in study two), but findings show a downward trend for some students. In other words, 12.5% of the students in the first study and 25% of the students in the second study had lower scores in their posttest. In general, developmental trends do not show regression. Three explanations might be applied to this situation. First, there is the possibility of a reactive effect during post testing. For whatever reasons, the students may not have invested their effort to complete the test. Another explanation might be that there is a possibility of a transitional arrestation or closedness in the conceptual system as explained by Harvey, Hunt, and Schroder (1961). In other words, there are points during the transitions from one conceptual level to others when individuals close their current system of thinking before assuming more complex thinking schemas. There is a possibility that some of the students at their transitional phase did not show changes as indicated by CL test scores. This, however, does not mean that they did not develop. The other explanation might be related to the concept of match and mismatch as explained by Hunt's (1971) matching model. Accordingly, students at different CL need different types of instructions and learning environments in accordance with their need

for structure. On one hand, if the structure of instruction, activities, and learning experiences match their current CL, the students will learn well. On the other hand, if teachers provide too much or too little structure in their teaching, the challenge can be too little or too great for the students and learning is arrested.

The information in Tables 4.3 and 4.4 seems to support the difference in the effect size of the two studies. As noted earlier, the magnitude of conceptual complexity development of the students in the first study was found to be greater than that of the students in the second study. After the four-year attendance at the College of Education, a higher percentage of students in study one developed in an upward direction (81.25% compared to 61.36%), and less percentage of them remained at the same CL (7.5% compared to 13.64%) or moved in a downward direction (12.5% compared to 25%).

What do all these numbers suggest about the students? What do CL scores mean or imply? How do they tell us how students think?

Discussion of Findings

At the time of pretesting in both longitudinal studies, the average CL scores of the students from cohort one and two were 1.65 and 1.85 respectively with most students, 81.5% in the first study and 54.55% in the second study, were at the CL between 1.1-1.9. Referring to the description of CL scores provided by Hunt, Butler, Noy, and Rosser (1977), students' thinking in this range can be characterized in a number of ways. Thinking includes caution about authority and authority figures. They are concerned about how to think or behave in such ways that teachers and parents would expect and

approve for them to do. They try to behave in a socially acceptable way. They are concerned about correctness of their thinking and behavior in accordance with social acceptability. They evaluate situations based on a dichotomous—absolutely right or absolutely wrong—way of thinking.

At the end of the senior year, the majority of the students, 35.42% in study one and 50% of the students in study two, moved to CL levels higher than 2.0. The majority of the seniors learned to evaluate alternatives rather than treating ideas as absolute—either as right or wrong. They tolerate some uncertainty, ambiguity and differences of opinions. They are striving for independence of their beliefs in authority figures by valuing their own perspectives. Despite the fact that they accept the variability of perspectives, when decisions need to be made, they are still concerned with their own thoughts and feelings rather than integrating others' perspectives. Between the two cohorts with the total number of 92 students, only one student was at conceptual level 3. Persons at CL level 3 tend to make decisions based on alternatives considered as best possible solutions rather than caring only about their own thoughts and feelings. They do not behave or make decisions only to please others. They are more independent and feel secure about being a part of their society. Other students in the two studies might begin to display these characteristics. Some might shift toward seriously taking into consideration other perspectives, weighing them, and objectively integrating those alternatives into their thinking in order to find best solutions before making decisions. As well, they became persons who accept full responsibility for the consequences of their decisions.

Considering the difference of conceptual development in relation to gender, the inconsistency of the findings from the two studies is inconclusive as to whether men or women developed more for their conceptual complexity over the four-year period.

Summary of Two Convenient Samples

Findings from the two studies suggest that the Deliberative Psychological and Professional Education Program (DPPE) has supported the development of college students' in teacher education. After participating in the program for four years, their conceptual level changed in positive directions. The amount of change in one study can be described as medium and the other as very large as indicated by effect size numbers. Details of the findings are convergent with previous studies in some aspects and divergent in other aspects. The convergence and divergence of the findings to other studies are further discussed in detail when these two studies are compared with eight other longitudinal studies collected for this dissertation. How are these two studies related to other studies that are based on the four selected theoretical frameworks? The following section provides some answers to this question.

Research Question Three:

Relationships Across Longitudinal Sample Studies

Eight studies are identified following the searching procedures and criteria of research selection described in chapter three. The studies have available quantitative data adequate for the calculation for their effect sizes. In addition, these studies are considered theoretically comparable although they employed different instruments. The instruments

are discussed in detail in chapter two. Moreover, these studies are assumed to violate a minimum number of threats to validity and reliability based on the discussion of these issues available in dissertations and published journal articles.

The researcher found an emergence of common relationships among the ten studies. These relationships are described and include commonalities in objectives and commonalities in underlying assumptions of the studies. The studies have two common objectives and three common underlying assumptions.

Research Objectives

The ten studies have two objectives in common. The first common objective is to investigate changes in how students think, know and learn as a result of college education. The second common objective is to provide evidence that has not been researched.

To Prove that College Attendance Affects Changes in Students' Thinking

All ten research studies compared and contrasted “levels”, “stages”, or “positions” in student thinking in undergraduate higher education. The ultimate aim for all these studies is to analyze students’ judgments when faced with ill-structured problems. This aim is based on the assumption that intellectual development progresses from less complex to more complex reasoning over time. It also implies that more complex judgments enable students to more effectively problem solve, construct knowledge, and make appropriate judgments based on the preponderance of evidence. The studies

explored possible factors that promote students' intellectual development. Specifically, the possible factors that the majority of these studies investigated include gender (Baxter Magolda, 1990; Conway, 1994; Khalili & Hood, 1983; King, Kitchener, Davison, Parker, & Wood, 1983; Reiman & Parramore, 1994, 1996; Schmidt, 1983, 1985; Welfel & Davison, 1986), and instruction and types of programs (Conway, 1994; Mentkowski et al., 2000; Reiman & Parramore, 1994, 1996; Schmidt, 1983, 1985; Tomlinson-Keasey & Eisert, 1978). Four studies examined the effects of students' major fields of study on their intellectual development (Khalili & Hood, 1983; Schmidt, 1983, 1985; Welfel & Davison, 1986). Three studies examined age in relation to intellectual development (King & Kitchener, 1994; Mentkowski et al., 2000; Schmidt, 1983, 1985). One study investigated students' background relating to previous academic achievement (Conway, 1994) and one study examined students' age and background related to life experience (Schmidt, 1983, 1985).

To Provide Evidence that Has Not Been Researched

The other general objective of these studies is to fill in gaps in theories or previous studies. Baxter Magolda (1990) searched to find out if there is any particular pattern of thinking unique to men or women based on the established Epistemological Reflection Model. The gender-related pattern had never been studied. Conway (1994) examined the longitudinal effect of a cognitive developmental based intervention program for at-risk college students. Such long-term effect among at-risk students had never been studied. Khalli and Hood (1983) were the first to explore conceptual change in college students during their 4-year college education. King and Kitchener (1994)

attempted to document changes in diverse populations based on Reflective Judgment Model. Taking into account all possible factors involved in educational process, Mentkowski et al. (2000) investigated the continuity of effects of education in students' life. Reiman and Parramore (1994, 1996) examined among teacher education students their conceptual change as an effect of a deliberative program designed to foster conceptual change. Schmidt (1983, 1985), explored age and life experience in relation to impact of college on students' intellectual development. Tomlinson-Keasey and Eisert (1978) analyzed influences of a designed intervention program that integrate students' real-life experience to academic experience in order to promote students' thinking. Welfel and Davison (1986) examined longitudinal changes in reflective judgment of students during their four years in college and to clarify such changes in relation to gender. Such relationships had been unclear in previous research.

It is worth noting that the majority of these research samples center on the students' development in relation to their undergraduate education. These studies do not take into account the effect of institutional administration and other institutional environmental effects. Only the study by Mentkowski et al. (2000) explicitly explains the involvement of institutional administration and other institutional environmental effects as important factors to help students develop intellectually.

Another commonality among these ten studies are their underlying assumptions.

Underlying Assumptions of the Studies

The other common theme among these ten studies is that they have similar underlying theoretical assumptions. These assumptions concern conditions for cognitive change and the assessment of cognitive change.

Assumptions About Conditions for Change

All ten studies assume that college life contributes to students intellectual change and that changes occur as a result of integration of college education and social life as college undergraduate students. Within higher education, however, there are kinds of education that most contribute to students' learning and cognitive development. Three traditions in higher education may be particularly important: the residential liberal arts college, the college with service and value traditions, and the large land-grant university system with their intensive professional schools curricula. How learners learn, develop, and become more skilled within these traditions is being explored. Mentkowski et al. (2000) suggest better understanding to seven related tasks of higher education: (1) understanding learning; (2) understanding the learner; (3) envisioning learners as educative, mature, and effective over time; (4) preparing students in college to enact their learning as graduates; (5) fostering learning outcomes through more deliberative educational programming; (6) organizing colleges for learning; and (7) redefining inquiry on teaching and learning.

Assumptions About Assessment of Change

All ten research studies assume that appropriate measurement tools can accurately assess developmental changes, and that the pretest and posttest procedures within a longitudinal research design can validly indicate such changes. In addition, all these studies appear to assume that studies conducted longitudinally offer richer evidence of developmental changes over time.

As stated at the outset of this dissertation, one of the major aims of this research is to build up to the body of existing research based on four connected theories of intellectual development. The purpose is to provide insights into the understanding of trends of students' intellectual development. To achieve this aim, the researcher compares the two convenient longitudinal studies and other eight selected longitudinal studies. Three major findings across the ten studies include trend of intellectual development, magnitude of the development, and factors that affect intellectual changes. These findings are reported in two parts. Part one is the quantitative findings based on statistical analysis of the ten studies. This section is concerned with the trend and magnitude of students' intellectual change. Part two is concerned with the qualitative findings from the comparative analysis of the studies.

Part One: Quantitative Findings from Comparative Analysis Across the Ten Studies

What do these ten studies suggest about trends and magnitude of students' intellectual development?

Trends of Development

Generally, the ten studies suggest upward or positive change in intellectual development in students. It is worthwhile, however, to investigate in detail across the studies to see if this trend of positive change in development is consistent among all the studies. To begin, Table 4.5 presents the students' cognitive stages, conceptual levels, intellectual positions, or reflective judgment level at the beginning of each study and at the end of each study. The information is organized on the basis of measurement tools in alphabetical order to facilitate the viewing of students' changes in intellectual development as described by different measurement by different terms. As previously mentioned, most of these studies employed a battery of tests. However, for the purpose of this dissertation, only the data from relevant measurement tools from each study are included. These measurement tools are Measure of Epistemological Reflection (MER), Measurement of Intellectual Development (MID), Paragraph Completion Method (PCM), and Reflective Judgment Interview (RJI). The MER and MID scores ranged between 1-5 positions and are interpreted based on Perry's Position of intellectual and ethical development. The PCM scores ranged from 1-3 conceptual levels and are interpreted based on Hunt, Butler, Noy, and Rosser's conceptual levels. The RJI scores ranged from 1-7 thinking stages and are interpreted based on King and Kitchener's Reflective Judgment Model. The pretest and posttest mean scores indicate where the students were at the beginning and at the end of the studies respectively. The last column of the table provides description of the meaning of the scores obtained by the participants in the ten

studies. The detailed description of positions, levels or stage in relation to scores devised by different measurement instruments is included in the review in chapter two.

Although the scores across ten studies are different depending on the measurement instruments employed, the description of the scores for each instrument demonstrates a common picture of college students at the entrance and subsequent years in colleges. It seems apparent that at the entrance to their colleges, the students in most studies generally tended to view knowledge as absolute and certain and can be regarded either as right or wrong. The source of uncertainty or ambiguity of knowledge came from the fact that experts or knowers had not yet discovered it. They tended to depend on teachers to justify what was right or wrong.

During the college years, their ways of thinking and knowing changed. The majority of them began to perceive that knowledge was not always certain and absolute. They also started to tolerate or legitimately accept the uncertainty of knowledge. Knowledge could be judged as relatively right or relatively wrong with references to evidence claimed by experts or knowers, or by individuals' personal opinions. Still the rightness and wrongness of knowledge was seen as whether it was compatible with what the experts or knowers had claimed. While some students still justified their beliefs based on reasons in reference of rightness of knowledge as claimed by experts, some students began to think that the experts' claims could also be relative to those experts' individual personal opinions.

Table 4.5

Students' Intellectual Stages at the Beginning and at the End of Each Study as Indicated
by Mean Scores from Different Instruments

Tools	Studies	Mean Pretest Score	Mean Posttest Score	Score Description
MER	Baxter Magolda (1990)	2.50	2.93*	The scores between 2.50 to 2.93 indicate that the students were between Perry's Position 2 and 3. Students at Position 2 view knowledge as certain and known to Authority Students at Position 3 accepted diversity and uncertainty of knowledge with the reason that Authority has not yet discovered such knowledge.
MID	Menkowski & Associates (2000)	3.30	3.70*	The scores between 3.30 and 3.70 indicate that the students were between Perry's Positions 3 and 4. Students at Position 3 accept diversity and uncertainty of knowledge with the reason that Authority has not yet discover such knowledge. Students at Position 4 accept diversity and uncertainty of knowledge in relation to the right of individuals to have diverse opinions. Still, such knowledge could be right or wrong and contextual relativistic reasoning depends on Authority.
PCM	Conway, Jr. (1994)	1.51	1.72**	The scores between 1.51 to 2.07 indicate that these students were between CL 1 and 2. The students at CL 1 think in a polarized or dichotomous way. Knowledge is certain and is either right or wrong. They behave in a "right" way only because such rightness is socially accepted. The students at CL 2 recognize others' ideas and alternatives. They have an increase in tolerance of uncertainty and ambiguity of knowledge. Although they are open to others' ideas and evaluated alternatives, they are too concerned with their own thoughts and feelings to integrate others' ideas in their decision.
	Khalili & Hood (1983)	1.52	2.05	
	Reiman & Parramore (1994)	1.65	2.01	
	Reiman & Parramore (1996)	1.85	2.07	
	Tomlinson- Keasey & Eisert (1978)	1.69	1.91***	
RJI	King et al. (1983)	3.75	4.18	The scores between 3.36 to 4.19 indicate that these students were between stage 3 (pre-reflective thinking) and stage 4 (quasi-reflective thinking). Students at pre-reflective thinking stage 3 view knowledge as absolutely certain. Knowledge can be uncertain only because absolute knowledge is not yet discovered. They justify their beliefs with reference to authorities' views or defend their beliefs as personal opinions. Students at quasi-reflective thinking stage 4 view that knowledge could be uncertain and ambiguous. They justify their beliefs with reference to reasons and evidence. Such reasons and evidence could also be idiosyncratic.
	Schmidt (1985)	3.36	3.56*	
	Welfel & Davison (1986)	3.62	4.19	

In most studies, pretests were administered at the freshman year except one by King et al (1983), in which pretest was administered at the junior year. Posttests were administered at the senior year except:

* Posttest at junior year

** Posttest at years 4, 6, or 7 after starting freshman year (This group was considered "at-risk" by academic achievement from high schools.)

*** Posttest at the end of freshman year

The overall picture of where the students were at the beginning and at the end of each study clearly illustrates that college education has impact on students' intellectual development. Did all students benefit from college attendance the same way? Some studies provide detailed information on the trend of development of the students who participated in their studies.

Table 4.6 shows directions of students' intellectual development in six studies. The other studies do not provide this information. The numbers in the "upward" trend column show the proportion of students who gained on their posttest scores when compared to pretest scores. The numbers in the "stay the same" column show the proportion of students who did not gain on their posttest scores when compared to pretest scores. The numbers in the "downward" trend column indicate the proportion of the students whose posttest scores dropped.

Table 4.6

Directions of Students' Intellectual Development

Studies	Trends of Development		
	Upward	Stay the Same	Downward
Khalili & Hood (1983)	73.75%	5.00%	21.25%
King et al. (1983)	64.30%	25.00%	10.00%*
Reiman and Parramore (1994)	81.25%	7.50%	12.50%
Reiman and Parramore (1996)	61.36%	13.64%	25.00%
Schmidt (1983, 1985)	58.00%	31.00%	11.00%
Welfel & Davison (1986)	92.00%	-	8.00%

*From Welfel & Davison (1986, p.213)

The data in Table 4.6 suggest that although not all students benefited from college attendance in terms of their intellectual development, the majority of them did. Why did some students not develop? Khalili and Hood (1983) explained that there was a lack of connection between CL development and college experience. By referring to Hunt's concept of match and mismatch of students' current level of conceptual complexity, Khalili and Hood noted that the college environment should be organized to comply with students' need in terms of the matching between students CL and provision of "structure". As an example, they elaborated that depending on students' current conceptual levels, some students learned better from highly structured instruction such as lectures with visual aids and structured reading, assignments, and coursework. Other students, however, who have higher CL better learned from doing independent projects, class discussions or debates. Some students might come to college with higher CL than others. When the instructions, for example, were not differentiated to meet the different learning needs of students at different CL, they did not develop consistently. Khalili and Hood recommended that in order to maximize students' CL development, college environments should be arranged to meet students' need of structure not only in classrooms but also other aspects of college life such as residence hall environments, student activities, and career planning programs.

Findings from the ten studies not only provide information on the trend of intellectual development in college students, but also the amount of change or magnitude of such development. The next section reports statistical findings that indicate magnitude of students' intellectual development.

Magnitude of Intellectual Changes

Table 4.7 presents detail of statistical findings among the ten studies. The first column provides information about the researchers and the dates of the studies reported. The studies are organized in clustered on the basis of measurement tools including Measure of Epistemological Reflection (MER), Measurement of Intellectual Development (MID), Paragraph Completion Method (PCM), and Reflective Judgment Interview (RJI). The third column describes the duration of each research study implemented. The pretest scores indicate the beginning intellectual status as indicated by research measurement tools. In some studies, follow-up tests were conducted during the process of research implementation. The follow-up test scores include the first follow-up and the second-follow up scores. The second follow-up test scores are regarded as the final posttest scores. Accordingly, the total gain scores are the difference between the second follow-up posttest and the pretest scores. The last column shows the effect size of each study. The effect size is calculated from the total gain scores divided by the pooled standard deviation of the pretest and that of the second follow-up tests.

The length of these ten studies ranges from 1 to 4 years during college education. The majority of these studies examined the students' intellectual development from freshman to senior years except those by Tomlinson-Keasey and Eisert (1978), Mentkowski et al. (2000), and King et al. (1983). Tomlinson-Keasey & Eisert (1978) examined their students' abstract formal thinking development during their freshman year as a result of a particular designed program for freshman students. They administered the pretest at the beginning of freshman year and the posttest at the end of the freshman year.

Table 4.7

Magnitude of Change in Students' Intellectual Development

Studies	Tools	Length of Study	Pretest Scores			Follow-up Test Scores						Total Gain Scores	Effect Size
			N	M	SD	First Follow-up			Second Follow-up				
			N	M	SD	N	M	SD	N	M	SD		
Baxter Magolda (1990)	MER	3 yrs.	101 (Fr.)	2.50	0.31	77 (So.)	2.83	0.34	64 (Jr.)	2.93	0.34	0.43	1.34
Menkowski et al. (2000)	MID	3 ½ yrs.	126*, 140, 141 (Fr.)	3.30	1.10	126-141 (So.)	3.50	1.10	126*, 140, 141 (Sr.)	3.70	1.20	0.40	0.35
Conway, Jr. (1994)	PCM	4-7 yrs.	52NT (Fr.)	1.51	0.30	52 (Fr.)	1.68	0.37	52***	1.72	-**	0.20	0.60
Khalili & Hood (1983)	PCM	4 yrs.	169 (Fr.)	1.52	0.37	38 (Fr.)	1.77	0.47	76 (Sr.)	2.05	0.37	0.53	1.43
Reiman & Parramore (1994)	PCM	4 yrs.	48 (Fr.)	1.65	0.28	-	-	-	48 (Sr.)	2.01	0.39	0.36	1.07
Reiman & Parramore (1996)	PCM	4 yrs.	44 (Fr.)	1.85	0.34	-	-	-	44 (Sr.)	2.07	0.40	0.22	0.59
Tomlinson-Keasey & Eisert (1978)	PCM	1 yr.	28 (Fr.)	1.69	0.45	-	-	-	28 (Fr.)	1.91	0.36	0.22	0.54
King et al. (1983)	RJI	2 yrs.	40 (Jr.)	3.75	0.72	-	-	-	27 (C)	4.18	0.80	0.43	0.57
Schmidt (1985)	RJI	3 yrs.	40 T 18NT (Fr.)	3.36 3.47	0.26 0.35	-	-	-	34T 11NT (Jr.)	3.56 3.52	0.29 0.34	0.20 0.05	0.73 0.15
Welfel & Davison (1986)	RJI	4 yrs.	32 (Fr.)	3.62	0.28	-	-	-	25 (Sr.)	4.19	0.52	0.57	1.43

* The average numbers of students could not be obtained because different numbers of students took different essays.

**SD is not available, the average of the two SDs available are rounded up for effect size calculation.

***This includes those who still enrolled, prepare to graduated, and graduated (pre-tested at freshman year from cohort 1987, 1988, 1990, second follow-up tested in 1993).

Fr., So., Jr. and Sr. stand for freshmen, sophomores, juniors and seniors respectively.

T stands for traditional, and NT stands for nontraditional.

C stands for college completion.

The study by Mentkowski et al. (2000) is a ten-year long study. Beginning from entering freshman year, Mentkowski et al. administered a series of tests to follow-up the development of their students one and a half year after entry, three and one-half years after entry, and five and one-half years after most participants had graduated. As well, King et al. (1983) explored college students' reflective judgment starting from their junior year and two years later. The posttest is assumed to be administered after the graduation even though the investigators did not state the fact in the study. In fact, some of these students were further followed-up 8 years later. However, since this dissertation focuses on students' development during the college years, only the relevant statistical data are extracted to be included. In addition, because the magnitude of the students' intellectual development from these studies is compared, it is important that the statistical data that indicate the development after college years should not be included. After college, individual research participants exposed themselves to different environmental situations than college. They also have varieties of experience such as that of personal, professional, and family. These experiences influence the students' thinking, attitudes, worldview, and so forth differently. Thus, even though additional statistical data are available from the studies by Mentkowski et al. (2000) and King et al. (1983), they are excluded.

Regarding research participants, almost all of these studies explored traditional college students except those by Conway (1994), Schmidt (1985), and Mentkowski et al. (2000). As for Conway (1994) his research participants were nontraditional in the sense that there were "considered at-risk for academic failure within the institutions they have

attended” (p.24). These students were described as having “deficits in what are considered basic skills areas (reading, writing, and computational skills), underdeveloped study habits and social habits inappropriate to the academic setting” (p.24). Some research participants in Schmidt’s (1985) study were described as nontraditional regarding their age when entering the college freshman year. Schmidt compared reflective judgment of freshman students with traditional average age of 18 to those with nontraditional age of 21, and compared these same nontraditional freshmen to traditional juniors with the average of 21 years. In the study by Mentkowski et al. (2000), participants included students who came to college directly from high school and those who delayed college.

Regarding the numbers of research participants, in most of the studies, different numbers of students took pretest and the follow-up tests. In addition, towards the end of the study, a smaller number of students took posttests. This is typical for longitudinal studies. However, the participants at the end of each study were from the same population as those who took pretests.

Considering the total gain scores, Table 4.7 obviously shows that the mean scores of all follow-up tests of all these studies are higher than those of all pretests. In other words, the first follow-up mean scores are higher than the pretest scores, and the second follow-up mean scores are higher than the first follow-up mean scores and the pretest scores. The maximum total gain score among these studies is .57, which can generally be referred to as a half of a stage change. The minimum total gain scores is .05. This, however, can be regarded as an outlier when compared to the average gain score of the other studies, which can be rounded up as .35 (total up of all other total gain scores

divided by 10). Regardless of the nontraditional group in the study by Schmidt (1985), the total gain scores ranged from .20 to .57. In fact, the significance of the difference of scores gained between the first and follow-up tests are reported by Baxter-Magolda ($F(2,112)=62.22, p<.0001$); Khalili and Hood ($t=8.69, p<0.0001$); King et al. ($p<.05$); Reiman and Parramore (1994) ($t=6.32, p<.0001$); Schmidt ($F(1,45)=12.14, p<.05$); Tomlinson-Keasey and Eisert; and Welfel and Davison ($F(24,24)=3.50, p<.01$).

As discussed earlier for Table 4.5, gain scores from each instrument indicates the quality of development of students from one position, or conceptual level, or stage of thinking to a higher position, or conceptual level, or stage of thinking. The magnitude of development is exemplified by the effect sizes. The effect sizes indicate a comparable magnitude of students' intellectual development among studies. What do these numbers mean? How does one interpret effect size value? This dissertation follows the guidelines provided by Bowen (1977). As proposed by Bowen, the effect sizes of 0.10-0.39 are considered small, those of 0.40-0.69 are moderate, and those of 0.70-0.99 are large, and those of 1.00 and above are very large (Bowen, 1977, p.103).

Table 4.7 obviously shows that the ten studies yield a magnitude of effect for development ranging from small to very large. It is worth noting that 40% of these studies have very large effect sizes (+1.07, +1.34, +1.43, and +1.43), and the other 40% have medium effect sizes (+.54, +.57, +.59, and +.60). The average effect sizes of all these studies by rounding up all the effect sizes is +.87 excluding the outlier (+.05), and +.80 including the outlier. This average effect size is considered large. The very large effect sizes are obtained from four studies employing MER, PCM and RJI. Thus, there

seems to be no connection between the magnitude of development and measurement instruments. As well, no apparent connection emerged as a relation between either the number of research participants, nor the length of the study and the effect sizes.

What do these effect size numbers tell us? Among these ten studies, one of the common themes they focus on is students' intellectual development over time as a result of college attendance. All these selected studies collectively support college students' positive development in intellectual development. The resulting list of effect sizes from the ten studies indicate that over time only one study yields small magnitude of development in students' intellect, and other nine studies positively affect a medium to a very large magnitude of change in students' intellectual development. None of these studies yielded negative results.

What might influence intellectual development? These ten studies explore details of factors that are hypothesized to be related to students' intellectual development. Findings on factors influencing students' intellectual development is reported in part two.

Part Two: Qualitative Findings from Comparative Analysis Across the Ten Studies

The majority of these studies investigated possible factors that might have influenced intellectual development. Specifically, they examined intellectual development while taking into consideration gender, types of programs, students' major field of studies, age and life experience, and previous academic achievement. What do findings of these studies inform about such factors?

Gender Differences

Among the ten studies, eight of them (Baxter Magolda, 1990; Conway, 1994; Khalili & Hood, 1983; King et al., 1983; 1994; Reiman and Parramore, 1994, 1996; Schmidt, 1983, 1985; Welfel & Davison, 1986), investigated gender differences in relation to intellectual development. The focus of the exploration on this issue is, however, different. This includes patterns of thinking and trend of development.

Gender differences and patterns of thinking.

Baxter Magolda (1990) examined the details of gender difference at the cognitive stage level as well as within cognitive stages. Employing Measure of Epistemological Reflection as a tool to assess her students' epistemological reflection after each consecutive year of college attendance, she found unclear gender differences at cognitive structure level but an existence of gender differences at levels within cognitive structures.

Regarding cognitive structure level, even though men had higher mean scores as group as indicated by MER mean scores year two and year three than did women, the major position of both genders did not evidently differ. The overall means at the end of the third year were found to place both genders in the same overall position. Still, the longitudinal data show that movement between stages by each gender was found to be different. While more men shifted from certainty of knowledge to limited uncertainty of knowledge during the freshman year, more women experienced this shift after the first 2 years of college. The difference in this shift between men and positions 2 to 3 disappeared by the third year. Referring to previous findings by Benack (1982), Baxter

Magolda (1990) explained that women were found to withhold their discovery of uncertainty while male-orientated reasoning patterns were found to be more conducive to cognitive structure change.

Considering the levels within cognitive structures, Baxter Magolda (1990) noted that meaningful differences appeared consistently in reasoning structure patterns within cognitive structures. While reasoning structure of men was found to be as initiator and objective, that of women was receiver as well as subjective. These differences in reasoning structure appeared consistently across domains and positions. In addition, this finding is consistent with previous research. Baxter Magolda also remarked that the consistency of the difference in reasoning structures between men and women over the 3 year supports the existence of gender-related patterns as identified theoretical expectations. These theoretical expectations are derived from the work of Perry (1970), Belenky et al. (1997), and Benack (1982).

Gender differences: Trend and magnitude of development.

Other selected studies focus on trend and magnitude of development when gender is taken into consideration.

Khalili and Hood (1983) are the first researchers to explore 4-year longitudinal change in conceptual complexity among college students. With the examination on gender difference in conceptual development, they focused on trend and magnitude of conceptual complexity development. Employing the Paragraph Completion Method, they found that the mean CL of female students at the beginning of the freshman year

was higher than those of male freshmen. By the end of their senior year, all these students CL scores were higher but no significant differences were found in terms of magnitude of change between men and women. When analyzing scores on particular items with different contents, they found scores on each stem increased in a similar fashion for each of the retestings—during sophomore and senior years—and for each gender. Findings from the studies by Schmidt (1985), Welfel and Davison (1986), and King et al. (1983) support this finding. Schmidt concluded that no significant difference was found between men and women in their intellectual development during the testing period of her study. Welfel and Davison noted that there was no significant difference in reflective judgment development between men and women ($F(1,23)=.578, p>.05$). King et al. (1983) also found no significant difference in reflective judgment scores between men and women when examining time X sex interaction. As for the trend of development, although posttest scores of some participants in their study were found to drop from those of pretest, King et al. concluded that the analysis of response patterns supported for upward and sequential movement on the reflective judgment model.

Conway (1994), however, reported a different finding. Although he found that overall research participants had significantly higher scores from two follow-up tests than from pretest, women scores dropped at the second follow-up test. Men were found to have higher scores from the first follow-up test than the pretest and from the second follow-up test than the first follow-up tests (from 1.54 to 1.66 to 1.87). Women, however, were found to have higher scores from the first follow-up test than the pretest, but lower scores from second follow-up test than the first one (from 1.49 to 1.69 to 1.61).

Conway did not anticipate the cause of the decline in mean PCM scores for women but noted that the finding was consistent with previous studies such as that by Mentkowski and Dohertz (1984) who found senior female students' decline in self-confidence.

Regarding the two convenient samples, in study one (Reiman & Parramore, 1994) no statistically significant difference was found between men and women in terms of magnitude of development. The other study (Reiman & Parramore, 1996), however, shows statistically significant difference in terms of magnitude of development. As for the trend of development, however, both studies support the upward trend development in both genders.

Regardless of statistical significance, each study seems to indicate evidence of gender differences related to intellectual development. There seems to be no consistency of findings from these research studies and no conclusion could be reached on how different or similar men and women grow developmentally.

Types of Programs

Among these ten studies, some programs include deliberative efforts to foster students' intellectual development through the emphasis of the design of programs. The two convenient samples focus on the intellectual development of students in teacher education attending the fellow programs at North Carolina State University. The two studies explored the effects of a deliberative program, DPPE, designed to foster students' intellectual development. The study by Conway (1994) followed the students' intellectual development as a result of a deliberate psychological education-based

freshman program, DPE, designed to help at-risk students. Tomlinson-Keasey and Eisert (1978) implemented an ADPAT program that was designed to develop students' thinking processes that integrated with students' experiences. Mentkowski et al. (2000) examined the effectiveness of their efforts in developing college-wide programs to develop students' intellectual development. To what extent do these alternatives to instruction and programing influence students' intellectual development?

The DPPE and the DPE programs were designed based on psychological development theories. Both programs include similar conditions and training elements that foster intellectual growth. Regarding the DPPE program, the conditions and training elements were part of the curriculum, which was implemented over the four years at college. Findings from one study by Reiman and Parramore (1994) indicate that the students showed statistically significant growth in conceptual complexity. Findings from the other study (Reiman & Parramore, 1996) showed strong trends but no statistically significant growth. However, posttest scores in both studies were higher than pretest scores. The findings thus support the effectiveness of the DPPE program to foster students' intellectual development during college years.

Conway (1994) found evidence supportive to the effectiveness of DPE program. In his study, the research participants are academically at-risk students. Conway collected and combined research participants from three cohorts of three previous research studies by McAdams (1988), Lindsay, (1989), and Parker (1991). The previous three studies examined the effects of a deliberate psychological education-based freshman program,

DPE, designed to help at-risk students develop their cognitive complexity. The studies by Lindsay and Parker are the follow-up studies of that initiated by McAdams.

McAdams (1988) implemented the deliberate psychological education-based freshman orientation program for one semester. Even though, findings were not significant, students showed developmental growth. Following this research, Lindsay (1989) extended the implementation of the DPE program for two semesters. The first semester focuses on academic skill and the second one focuses on critical thinking and decision-making development. Again, Lindsay found no significance difference on developmental growth. However, it was recommended that cognitive developmental theory be employed as a framework for designing student development programs especially for at-risk college students because findings showed the positive trends in students' development. Parker (1991) refined the DPE program by differentiating methodology to match the students' conceptual level based on pretest scores. Although no statistically significant psychological growth was found between the differentiated DPE and non-differentiated DPE groups, Parker found that major developmental gains in the differentiated group were consistent across conceptual levels. The non-differentiated group, however, experienced gains only among lower conceptual level participants. Like McAdams and Lindsay, Parker concluded that positive trends in students' growth indicate the potential of differentiated DPE programming, especially for at-risk college students.

Unlike the two convenient longitudinal samples by Reiman and Parramore (1994, 1996), these three studies have comparison groups and were experimental in nature.

Although none of them report statistically significant difference of growth between experimental and comparison groups, the studies reveal positive trends in students' growth as a result of the attendance of the DPE program. The follow-up study of the three cohorts by Conway (1994) also supports the effectiveness of the DPE program. At the time of 4, 5, and 7 years after attending the DPE program, the students' conceptual complexity was found to continue in an upward trend.

Regarding the ADAPT program, Tomlinson-Keasey and Eisert (1978) found the program to achieve its goals in that it helps their students develop significantly in the realm of logical and abstract thinking. Compared to the students of the matched control groups, the students who had attended the ADAPT program seemed to have learned to conceptualize issues in a more differentiated way. Even though the students who attended the ADAPT program were from the bottom half of their graduating class and they had lower mean PCM scores on their pretest, they gained higher mean scores from the PCM posttest than the control group.

Another effort to design a program that effectively enables students to become intellectually mature was conducted at Alverno College. Mentkowski et al. (2000) not only conducted research longitudinally to cover the period of students' matriculation through college, but also followed students after their graduation. The program at Alverno College was found to significantly affect students' intellectual growth. The students were found to gradually develop from their original tendencies to approach learning and decision making with dualistic black-and-white points of view and absolute truths, toward being more open to alternatives and multiple perspectives.

Findings from these research studies demonstrate that if deliberately designed, programs can effectively foster students' intellectual development.

Students' Major Fields of Study

Some of the studies explored the relationship between students' major fields of study and intellectual development. Khalili and Hood (1983) examined the relationship between the students' majors and their CL development over the four years of college life. Their research participants were grouped according to the undergraduate majors including humanities, social sciences, natural sciences, business, and other. They found that the students in business had the smallest amount of change in CL scores. The few students in humanities were found to have the largest change. However, no significant difference in amount of change in conceptual level scores was found between the different undergraduate majors. This finding is consistent with that by Schmidt (1985). She compared the reflective judgment of students from a College of Liberal Arts (CLA) and those from an Institute of Technology (IT) during the freshman year and junior years. Although the CLA students RJ mean score was higher than that of the IT students, the difference did not reach a statistical significance. Other evidence supports the findings of these two studies. Welfel and Davison (1986) compared the engineering and humanities-social sciences students in terms of their reflective judgment. They found that the humanities-social sciences students had higher mean scores than did the engineering students. Still, the amount of change over the four years was not statistically significant. Referring to a previous study by Welfel (1982), they remarked that Welfel found no

significant relation between reflective judgment and academic major, but a combination of age and education was important and significantly related to reflective judgment.

In short, findings from the three studies suggest that the relationship between academic majors and conceptual complexity and reflective judgment are not significant. Yet, it is worth noting that of the three studies, students majoring in liberal arts tend to gain higher PCM and RJ scores when compared to other majors within each study. This could be an area for further inquiry.

Age and Life Experience

Age and life experience were examined as to whether they influence developmental growth during college years. As for the matter of age, Schmidt (1983) conducted her research to particularly answer the question of age in relation to reflective judgment development. One of the research questions in her study was whether the extent of students' intellectual development during college years is a function of age, maturation, or college education. One of the major findings was that the combination of age and education was associated with the highest level of RJ for the sample as a whole. Compared to traditional freshmen entering the college the same year as the nontraditional who averagely 3 years older, the nontraditional freshmen had higher RJ pretest scores. However, these nontraditional freshmen had lower average pretest RJ scores than the juniors, who had the same average age but more education. Taking gender into consideration, age seems to be a more significant factor for women than education on RJ level.

The study by Mentkowski et al. (2000) also suggests the influence of the combination of age and education on students' developmental growth. The students who came to study at Alverno College ranged in age from 17 to 55. Half of the research participants came directly from high school, and half were returning adults. Eighty percent were first generation college students. Mentkowski et al. found the results of MID essay indicated that students who came to college directly from high school generally began at a less complex levels of commitment to growth. Their older peers who delayed college were found to demonstrate an intermediate level of commitment to growth and learning. However, college experience was found to benefit both age groups in a similar fashion. All population groups were found to show comparable growth in terms of their position, attitude, and understanding toward knowledge, learning and self.

In their ten years of research, King and Kitchener (1994) pooled the scores of all individuals who had been tested one or more times in their study and categorized them by their age at the time of testing. They found a strong linear relationship between age and stage in the Reflective Judgment Model. They noted, however, that their research participants had been actively engaged in a variety of educational pursuits and did not represent adults in general. They anticipated the possibility of the predominant role of age or education or the combination of the two.

In terms of life experience, in their attempts to figure out what might be related to conceptual complexity development, Khalili and Hood (1983) examined variables including students' particular activities or experiences, extracurricular participation, and students' residence. They administered a questionnaire to the students in their senior

years to obtain the students' experience and demographic information and tried to analyze to their relationship with the students' CL growth. They found few significant relationships between changes in CL scores and the experiential and demographic variables. Using analysis of correlation between CL growth and campus activities, cultural events, recreational activities, and summer or part-time work experience, Khalili and Hood found no statistically significant correlation among them. However, they did find a statistical significance between CL and students' political commitment. One fourth of the seniors who stated that they had made a definite political commitment obtained significantly higher conceptual level scores as seniors than did the rest of their classmates.

Generally, it seems that when individuals get older they gain more experience. Findings from Schmidt's study suggested that life experience affected intellectual development especially for women. As well, Khalili and Hood found connections between some experience and intellectual development. Findings from these studies suggest that there is a lot of questions to be further explored regarding the influence of age, life experience and education on intellectual development.

Academic Achievement

Students' previous academic achievement has been related to intellectual development in some of the research. In his follow-up study, Conway (1994) described the students as at-risk students based on students' achievement prior to entering college and as indicated by the result of the Admission Index (AI), which was employed in the

admissions screening process. This index takes into account traditional academic variables such as high school GPA, class rank and SAT scores. Based on standard admissions criteria, none of these students would be admitted to the university. These students were, however, accepted to comply with a national policy to develop the nation's human resources. Conway categorized these students on the basis of their AI scores. Category 1 students had AI less than or equal to 1.60. Category 2 students had AI from 1.61 to 1.80. Category 3 students had AI equal to or greater than 1.81. Regarded as an independent variable, the three AI categories were analyzed to find the relationship with other dependent variables including first year grade point average, CL pretest, posttest, and gain scores, and pretest, posttest, and gain scores from Learning Context Questionnaire (an instrument used for projecting position on Perry's scale of intellectual development). Results indicated that the AI categories only related to first year grade point average. Conway further inquired into the significance of inclusion of other screening materials in addition to the Admission Index since the AI were related neither to other traditional academic or cognitive development variables. Conway also attempted to find out the relationship between CL scores and students long-term persistence. In doing so, he analyzed the high and low CL pretest scores of the graduates and dropouts. No statistical significance was found to confirm whether the students with high CL scores were more likely to persist. Another analysis of the relation between CL and persistence examined the students who gained and did not gain scores from CL posttest and their persistence. Again, no statistically significant difference was found among the gainers and non-gainers in relation to their persistence. When taking into account gender

difference, no statistically significant differences were found between men and women regarding the status of persistence among overall participants in his follow-up study.

For their four-year research, Khalili and Hood (1983) also questioned the relationship between students' CL and their persistence. By the end of their research, they found that only 101 students from the original 169 freshmen they studied were still on campus. To answer the question, they compared the CL pretest scores of those who were still on campus and the scores of the dropouts. The two sets of scores were found to be comparatively the same (1.52 and 1.53). They noted that the conceptual level scores elicited from freshmen seemed to be unrelated to persistence in college.

In real life, it is generally accepted that there are a lot of factors leading individuals to decide to withdraw from education. Level of academic achievement is only one factor. How academic achievement affects persistence in education is an area open for further research.

Summary of Findings from Research Analysis

Findings from the two convenient samples were analyzed and reported. As well, the two samples and other eight selected sample studies were compared. Quantitative findings about trends and magnitude of students' intellectual development from the ten studies were analyzed. A conclusion was reached to support the position that college attendance serves to elevate students' intellectual development, specifically their epistemological reflection (Baxter-Magolda, 1992b), intellectual and ethical position on Perry's scheme (1970), conceptual complexity (Harvey, Hunt, & Schroder, 1961), and

reflective judgment (King & Kitchener, 1994). The students tend to develop intellectually in an upward positive trend and the magnitude of this change tends to be large. Qualitative analysis among the ten studies examined issues or factors that had been anticipated to connect with intellectual development. Factors include gender, types of programs, students' major fields of study, age and life experience, and previous academic achievement. The comparative analysis shows both conclusive and inconclusive findings across the studies regarding these factors. Further research is needed. However, in this dissertation, findings suggest that in addition to academic programs, students' development might be affected by other factors such as gender, students' major fields of study, age and life experience, and academic achievement.

How can students' intellectual developmental change be maximized? Educators and researchers have been trying to answer this question. As well, this dissertation has an aim to contribute some information regarding this issue. The analysis of some features of deliberate educational programs with role taking/reflection components for teacher education students, as well as features of other programs in the collected studies provides a part of the answer to final research question of this dissertation.

Research Question Four: Distinctive Features of Programs Designed to Foster Students' Intellectual Development

This section describes characteristics or features of deliberative undergraduate programs that are intentionally designed to foster developmental changes. Among the ten studies, Tomlinson-Keasey and Eisert (1978) described the ADAPT program at the

University of Nebraska, Mentkowski and et al. (2000) described an ability-based learning program at Alverno College, Wisconsin, and Reiman and Parramore (1994, 1996) described the DPPE program at North Carolina State University. These three programs are now described in detail. In addition, some explanations are provided so as to give insights into how program features contribute to students' intellectual development.

The following section is organized in two parts. Part one describes the programs and part two analyzes and explains program features.

Part One: Description of Programs in Samples

The description of each program follows the same outline: 1) purposes, 2) program composition, 3) program implementation, 4) persons involved and their roles, 5) program evaluation, and 6) follow-up.

1. Deliberative Psychological and Professional Education Program (DPPE)

The DPPE can be described as a psychological development-oriented education program that integrate conditions that promote cognitive developmental growth (Sprinthall, Reiman, and Thies-Sprinthall, 1996). The program is based upon longitudinal and clinical research that indicates a relationship between cognitive developmental stage and personal decision-making and professional behaviors. In addition, it has applied design principles (e.g., new social role taking, guided inquiry, balance, support and challenge, and continuity) with the goal of promoting development.

Purposes.

The program is intentionally designed to foster intellectual development of college students in teacher education.

Program composition.

According to Reiman (2002), the development and evaluation of the program is guided by an applied theory called the Learning/Teaching Framework –LTF. The program is composed of two specific sets of design principles: six conditions for cognitive growth and a four-component coaching model. The six conditions for cognitive growth were developed based on cognitive developmental theories. Mead (1934) suggested that social experience and social perspective taking is an important condition for adult development. Recent educators such as Goodman (1985), Ross (1987), Schon (1987), and Zeichner and Liston (1987) asserted that self-reflection supports the development of student teachers' technical proficiency. Berliner (1985) suggested a combination of structured self-analysis and instructor feedback in field and laboratory experiences for students in teacher education. Educators Sprinthall and Thies-Sprinthall (1983), Sprinthall, Reiman, and Thies-Sprinthall (1996), and Reiman and Peace (2002) indicate the strong influences of psychological maturity or cognitive structures on quality of student teachers' performance. These six conditions are:

1. Contextual understanding and building trust. The condition requires that trust be initiated between educators and learners as well as to maintain

rapport between them. Educators should also acknowledge students' prior knowledge and experience.

2. Complex new social role-taking. The condition of learners in a complex new role such as student teachers, collaborative researcher, or mentees, precedes and shapes their intellectual consciousness.
3. Guided inquiry. Self-analysis of performance and ongoing discussions and journaling (reflection) are necessary conditions to promote growth.
4. Balance. Balance between action such as new role-taking and guided inquiry needs to be maintained.
5. Continuity. Both learning and development could be effectively fostered under a condition of continuous interplay between action and reflection. At least four to six months of the process is needed for significant developmental changes.
6. Support and challenge. Support (encouragement) and optimal challenge (prompting the learner to accommodate to new learning) are necessary for integrated learning.

In addition to these six conditions, the program is also composed of an instructional model of coaching that encourages acquisition of new performances when properly followed. The four components of the training model as suggested by Joyce and Shower's (1982) are:

1. Description of the model;
2. Viewing/demonstration of the model;
3. Practicing the model while receiving feedback;

4. Adapting the model.

The first component is concerned with the understanding of rationale underlying a particular teaching model or method. The instructor should introduce relevant theories and organize activities to allow student teachers to explore theories for their understanding about those theories. Next, the instructor should provide opportunities for learners to view a demonstration of the model. After having opportunities to view the model, learners should practice the model under simulated conditions and get feedback from peers and the instructor. Finally, student teachers apply the model in real contexts with support from concerned individuals such as peers, mentoring teachers, other teachers, and collaboration from school personnel.

Program implementation.

The programs were implemented over the period of four years—from freshman to senior years. For the first cohort, the program was implemented between 1988-1992 and between 1989-1993 for the second cohort. Students were required to take two one credit hour seminar courses: once during the fall semester of freshman year and another one during the last semester of the program. In addition, they were required to attend quarterly meetings among teaching fellows. Furthermore, during summer, they were required to participate in a program with teaching fellows from other campuses across the state. As well, students participated in a significant new role during each year of their program.

Persons involved and their roles.

For the program to be successfully implemented, educators have important roles to perform the following tasks:

1. To contextualize learning and instruction in accordance with learners' experience diversity by acknowledging their schema and developing rapport with them;
2. To support learners' new role and to guide inquiry so learners have opportunities to analyze and reflect on their experience while being in a new role;
3. To provide opportunities for learners to have a balance of opportunities to analyze and reflect on their new role;
4. To provide continuity of opportunity for learners to practice and simultaneously follow up by providing time to reflect and self assess the experience;
5. To support and challenge learners to step forward in their development by encouraging them with new learning challenges that are compatible with their existing intellectual level as well as providing appropriate challenge that is a little beyond their level but adequate for their accommodation of new learning.

Evaluation.

The program is evaluated using a battery of tests including the following:

1. Traditional measures of progression a university program, such as grades, and course completion;
2. Defining Issues Test (DIT) (Rest, 1986) to assess students' stage of moral reasoning (Kohlberg, 1969);

3. Paragraph Completion Method (PCM) to assess students' conceptual level (CL);
4. Constantinople Test (1967, 1969) to assess students' self-concept.

Follow-up.

A plan was made to survey those who began teaching in the fall to learn about their work during the first months of their teaching profession and to follow them through the two-year beginning teacher program.

2. Accent on Development Abstract Process of Thought (ADAPT)

Purposes.

This program was implemented at the University of Nebraska (Tomlinson-Keasey & Eisert, 1978). It is designed to encourage students to experience multiple academic disciplines in order to move them up to more formal thought patterns. The design is based on Piagetian's theory (Inhelder & Piaget 1972; Neimark 1975), and Brown's concept of student personal development within the academic area (1972).

Program composition.

The ADAPT program has the following elements:

1. A multidisciplinary curriculum (Ginsburg & Opper, 1969; Inhelder & Piaget, 1958) containing a combination of six disciplines: anthropology, economics, English, history, mathematics, and physics;

2. A learning process that moves students from concrete to formal thought, and that is composed of three phases: an active exploration phase, an invention phase, and an application phase for all class activities;
3. Activities that allow flexibility and wide applicability of the thinking skills.

Program implementation.

The program was implemented in freshman year during fall and spring semesters. The students took courses each of which comprised concrete exploration activities that introduced concepts important to the field to allow them to form hypotheses for further generalization to other topics. To generalize the concepts they had learned, students were required to participate in follow-up activities introducing formulations that are more abstract. The core underlying principle of this learning process is to allow opportunities for an integration of experience and ideas. Each learning cycle makes students' experiences the basis for subsequent learning. Students attended classes in group and participated in many group learning activities to allow them to learn from one another.

Person involves and their roles.

Multidisciplinary staff met weekly to integrate their teaching materials and relate the materials to real life experience. For example, to enable students to understand the functional relationships in economic, students were allowed to receive yearly incomes in play dollars. Then, they were required to allocate them to various expenses and to explore relationship between how the allocation of money by individuals, families and others in society is related to the abstract concepts of price indexes and inflation.

Evaluation.

The program was evaluated on the basis of pre- and post- tests. The pretest was administered in fall semester and the post test was administer in spring semester. A battery of tests was employed, including:

1. A series of paper-and-pencil tests adopted from Inhelder and Piaget's (1958) for evaluation of changes in formal operational thought processes;
2. The conceptual complexity test, Paragraph Completion Method (PCM) developed by Harvey, Hunt, and Schroder (1961) for evaluation of personality changes and development in social areas;
3. College Student Questionnaire (Part II) for evaluation of student's attitudes toward various facet of university life.

3. Ability-Based Learning Program

This ability-based learning program was implemented at Alverno College, Wisconsin (Mentkowski et al., 2000). The program has been developed since the early 1970s and used during freshman to senior years.

Purposes.

Alverno College is a private Womens College. Its aim is to “assist women of all ages and a variety of backgrounds to learn and develop personally and professionally so that they will be productive in a continually changing work environment and be prepared to assume responsibilities supportive of a more beneficent society” (cited in Mentkowski

et al., 2000 p. 62 from Alverno College Office of Academic Affairs, 1997, p.2.).

Specifically, the ultimate goal of the program is:

the development of each student as an educated, mature adult with such personal characteristics as:

1. A sense of responsibility for her own learning and the ability and desire to continue learning independently
2. Self-knowledge and the ability to assess her own performance critically and accurately
3. An understanding of how to apply her knowledge and abilities in many different contexts

(Alverno College Institute, Milwaukee, Wisconsin. In Appendix A in Learning That Lasts by Mentkowski et al. p. 416)

Program composition.

The ability-based approach is described as having the explicitly acknowledgement of the expectation that “students should be able to do something with what they know” (Alverno College Institute, Milwaukee, Wisconsin. From Appendix A in Learning That Lasts by Mentkowski et al., p. 416). In other words, students are expected to be able to apply their subject-matter knowledge through the performance of the expected abilities.

Eight specific expected abilities outcomes is the major component of the program. These abilities, which are identified by faculty, form a structural framework that helps students integrate their learning and faculty integrate their teaching. The development of each of these abilities is organized as a series of six developmental levels in accordance with student progress across college career. While taking general education courses, including humanities, fine arts, natural and behavioral sciences, and the introductory

courses in major and supporting areas of students, the students are expected to perform all eight abilities from levels 1 through 4. While taking specialized courses in majors and supporting areas of study, the students are expected to perform all eight abilities in all levels. The abilities are described as follows.

1. Communication
2. Analysis
3. Problem solving
4. Valuing in decision-making
5. Social Interaction
6. Global perspectives
7. Effective citizenship
8. Aesthetic responsiveness

The Alverno College Institute expects that the curriculum be regarded as a mean not an end. In other words, the Institute expects that the curriculum serve as models of “lifelong learners” for the students.

Program implementation.

Along with subject-matter contexts, the expected abilities are emphasized throughout the educational process. Criteria for the ability being performed were devised for each level of ability. These criteria are applicable and used throughout college because they are regarded as generic in the sense that they are not tied to specific courses. Guided by the criteria in accordance with developmental levels, teachers are able to

devise specifications of abilities and integrate them into contexts of specific subject areas. With the specifications of the abilities, teachers are able to transform the abilities into the forms that they are teachable. As well, they use the criteria for judging and certifying students' demonstration of the abilities.

As for the students, guided by the explicit articulation of expected abilities as learning outcomes and specific criteria, they have specific aims in mind to achieve for particular courses. Students also recognize that the same basic abilities are relevant in multiple course contexts and that they are refining those abilities in multiple applications.

To elaborate the implementation process, examples cited in appendix A from Mentkowski and associates (2000) from program description by Alverno College Institute are included. An expected outcome of students with an English major is to “communicate an understanding of literary criticism, question its assumptions, and use its frameworks to analyze and evaluate works” (p.419). This outcome connects abilities to communicate, analyze, value and involvement of arts at advanced levels. Another example is expected abilities for students with major in chemistry. The students must “use different models of chemistry to analyze and synthesize chemical data and to critique the data, strategies, and models of chemistry” (p.419). Although the primary focus of these outcomes is concerned with ability to independently apply theory, the students are also assumed to be able to draw upon their valuing ability to critique the underlying assumptions of the theoretical models, as well as her ability to analyze, criticize, and communicate her ideas effectively.

Persons involved and their roles.

The implementation of educational process at Alverno College was run under the cooperation among many parties using a collaborative inquiry process.

1. Teaching faculty who are developing and implementing curriculum, as well as academic and educational researchers
2. Administrators who consider common tasks important to the institution to provide resources and participate in consortia
3. Students who participate so as to make meaning of their own learning and growth
4. Departments and committees who critiqued and debated questions, methods, findings, interpretations, and related policy-making
5. A standing committee of faculty, administrators, and educational researchers who conceptualized systems of inquiry and their support, evaluation, and use
6. Colleagues from across higher education who critique and contributed to each other

Evaluation.

The evaluation process at Alverno College is not a traditional approach. To assess students' achievement that conforms with the expected learning outcomes, faculty designed an assessment approach that evaluated not only what students know but also how well they are able to conjure and contextualize that knowledge in accordance with situations. Regarding the indication of achievement, students received evaluative statements rather than traditional letter grades. Faculty and students are involved in the evaluation of strengths and weaknesses through observation, judgment, and feedback

guided by explicit criteria. In addition, the assessment process of students' achievement takes place periodically throughout the program. Alverno faculty are trained professionals from Milwaukee community who assess students in several key abilities or provide co-assessment in their major area of study. Referring to Locker and Mentkowski (1993), Mentkowski and associates (2000), noted:

This approach to assessment became key to continually transforming the way faculty defines student learning, created effective teaching approaches, and designed program and institutional assessment. ... It also became key to studying learning.

(Mentkowski et al., 2000, p. 65)

Part Two: Analysis of Three Intervention Programs

Evidence from the studies by Reiman and Parramore (1994 and 1996), Tomlinson-Keasey and Eisert, (1978), and Mentkowski and associates (2000) indicated the success of the programs deliberately designed to support college students' intellectual growth. What makes these programs achieve their goals? This section discusses features of the three programs in relation to how these features could possibly support college students' development.

The expected intellectual growth of students such as those addressed by the DPPE, ADAPT, and ability-based learning Alverno programs seem to effectively develop under specific conditions. Program design includes conditions supportive to growth. Program designers need to take into consideration three important components of the teaching and learning process. These three components include how the students learn,

what they should learn, and how to teach them to learn. The analysis of the programs features and its implementation process is presented following this outline:

1. How do the students learn?
2. What should the students learn?
3. How should teachers teach the students?

1. How Do the Students learn?

Learning from experiencing in real world context.

It seems that all these three programs assume a constructivist stance because they are designed based on developmental theories and they all include individuals' active meaning making as a core element of the program. Rogers, Mentkowski, Hart, and Minik (2001) view theories of human development of adult capacities that are rooted in developmental psychology to include a constructivist stance. These theories, including theories of cognitive development, epistemological development of self, moral development and ego development, "explain and observe structural development in terms of the individual's active meaning-making" (p.1). As proposed by Piaget (1975) in his theory of cognitive development, humans cannot understand or learn through the information being given to them, but they must construct their own knowledge. The construction of their knowledge can take place through experience. The experience further enables them to create schema, which can be changed through the processes of assimilation and accommodation.

The Deliberative Psychological and Professional Education Program (DPPE) includes conditions that require students to make meaning from their own experience. Taking a new role is a condition that requires students to engage themselves through experiential process. As opposed to role playing or simulation in classrooms, Reiman (1999b) explained that roletaking is “a complex new helping experience in a real world context” (p.603). For students in teacher education, an example of roletaking includes actual teaching of students. The ADAPT program focuses on experiential learning through the process of active exploration of concepts before students are required to form some hypotheses from the concept and finally generalize and apply the knowledge across fields. Life experience is also integrated into curriculum. The program provides opportunities for students to apply their life experience to solve problems in simulated activities. The ability-based program requires that students actively perform eight specific abilities in their subject areas throughout their educational process. These abilities include communication, analysis, problem solving, valuing in decision-making, social interaction, global perspectives, effective citizenship, and aesthetic responsiveness. The program requires students to learn from experience through regularly and actively performing these skills.

Learning through inquiry (self-assessment and reflection).

In addition to learning by experiencing, Reiman and Thies-Sprinthall (1998) noted that research has shown that complex new experiences without inquiry make no impact on moral and conceptual development. Students learn through active self-assessment and ongoing reflection. Thus, opportunities for inquiry are included as

another condition to support students' developmental growth. For inquiry on experiences to prompt learning and cognitive development to occur, Reiman (1999b) proposed that it should describe "a process of problem solving, reconstruction of meaning, and subsequent reflective judgments while persons are engaged in significant new activity" (p.598).

Two related conditions for learning are the balance and continuity between learning experience and inquiry. With the new role taking in progress, students will confront difficulties and problems. The inquiry process allows students to examine problems they have, to analyze and reconsider them in order to understand them, and to reach a conclusion or decision. In addition, the process of developmental growth or change in cognitive structures takes time with a consistent balance between experience and inquiry. Reiman and Thies-Sprinthall (1998) referred to the national survey by Howey, Yarger, and Joyce (1978) that demonstrated the ineffectiveness of short-term and discontinuous workshops for teachers as opposed to long-term continuously balance experience.

Learning with supports and challenge.

All the three programs seem to include support and challenge as conditions for learning. Learning takes place only when current knowledge of students is related or can be employed as background to understand new knowledge. If the new knowledge is too discrepant and beyond the situation where current knowledge can be used to understand it, students will not learn it. Like the DPPE program, the activities in ADAPT program are organized in phases that move the students through cycles of learning experience.

Each phase in the cycle is regarded as a stepping-stone to another phase, where previous knowledge is built up to cope with the new learning. As for the ability-based program, as the students move along the years they are gradually required to perform different and more complex levels of tasks within the specific abilities.

A lot of theories explain how challenge and supports are necessary for students to learn new knowledge.

Piaget's concept of equilibration explains the process of constructing new knowledge (Piaget, 1975). To construct new knowledge, which is built upon the current knowledge, a person assimilates and accommodates their cognitive system to cope with the demand of the task or the new knowledge to be learned. In other words, the person applies a kind of consistent system of problem solving to formulate knowledge. When the persons' current system is adequate for constructing knowledge, the person assimilates the knowledge. When the person's current system is inadequate, or ineffective, the person accommodates and adapts his or her the problem-solving system. That is to say, the person tries to balance the cognitive system to cope with the level of difficulties or demands of the new knowledge. Generally, when the system has been changed or adapted and becomes more complex, the new problem-solving system will replace the current one. This theory seems to be connected with conditions of teaching and learning process as suggested by John Dewey, a noted educational theorist and philosopher.

According to Reiman and Thies-Sprinthall (1998), Dewey advocated active learning which takes into account learner's interests and capacities. Dewey also described educative versus miseducative experience to explain the idea of matching and mismatching of educative experiences to the learners' curiosity. In addition, Dewey believed educators' responsibilities included providing conditions that allow opportunities for learners to have actual experience and to construct their knowledge. Furthermore, he recognized that learning leads to growth. A lot of pedagogy in practices applies Dewey's ideas of experiential learning with the management of environment or conditions that require equilibration. A related approach is constructivism originated by Lev Semenovich Vygotsky (1962). As explained by Davydov (1995), Vygotsky's work proposed that "psychological development occurs through teaching/learning and upbringing, which appears as its necessary and general forms..." and "in terms of its content, psychological development is an independent process, but it proceeds through interaction, through assimilation, and through teaching and upbringing" (p.18). Vygotsky regards teachers and the milieu, which is defined as "dynamic social surroundings that connect the teacher and child, the other adults and children with whom a given child actually lives and interacts" (p.17) as core factors for the creation of psychological development.

In addition to organized conditions for experiential learning to take place, Vygotsky also proposed the notion that social/cultural/historical forces affect development since an individual person cannot grow up in separation or isolation from where the person is born. Consequently, in order to educate a person to reach the optimum of their

development, the person needs support and challenge from external stimuli—such as parents and teachers. Vygotsky termed the arena of the difference of level of learning where development can be made with support and challenge from external stimuli as opposed to without such support and challenge as “zone of proximal development”.

So far, this section has discussed features of the programs that relate to students’ learning and development process. Those features include a focus on students’ experiential learning, the importance of role-taking and guided inquiry, and the value of challenge and support. Following this section is another look at the program in terms of the content of the programs.

What Should Students Learn?

Based on theories of human development and previous research studies of developmental growth, the DPPE or Deliberative Psychological and Professional Education Program was designed and implemented with a major aim to foster and monitor developmental growth of students in teacher education. Specifically, the program was designed for students in teacher education. These pre-service teachers were expected to develop both an awareness of their own development and their conceptualization of the teaching and learning process. The program was designed with an ultimate ideal aim to develop students’ psychological maturity in addition to and along with learning teaching skills. The program was designed with the expectation that the specific elements enable students to become competent. More ideal expectation includes the development of students’ ability to consider principles of decision-making; to be

empathic; to symbolize experience; and to be autonomous, reflective and self-directive. With these aims, the content of the program includes activities such as roletaking that require students to socially interact in real world contexts. It includes activities that require students to reflect on their experience in order to analyze their problems, and come up with alternative ideas before solving problems. Also, the inquiry activities require students to become aware of their own thought and thinking process. The program include the requirement of students' attending seminars to expose themselves to diversity, alternatives and perspectives, not only among their peers but with teaching fellows from other campuses across the state.

Like the DPPE, the ADAPT directs its aim toward moving students to become abstract thinkers. The aim of the program is established based on the position suggested by Piaget (1975) that persons at the lower level of thinking limit their thought only to concrete realities of the world and are unable to consider abstract concepts. Their performance is consequently restricted. They have difficulties dealing with such thinking as relating to reasoning about hypotheses and about situations where no absolute facts apply. Individuals attaining formal thought patterns, however, are able to understand abstract concepts and delineate alternative possibilities. This allows greater flexibility in thought patterns. To help students become abstract thinkers, the content of the ADAPT program includes activities that help students first understand concrete concepts of lessons followed by activities that require students to form hypotheses from that concrete concept before ultimately coming up with abstract concepts. In addition, it includes

activities that require students to apply the process of abstract concept formation across subject areas. This provides continuity and reinforcement of the learning cycle.

The ability-based learning program at Alverno College targets to graduates who develop personally so as to become educated, mature adults with “a sense of responsibility for her own learning and the ability and desire to continue learning independently; self-knowledge and the ability to assess her own performance critically and with balance; and an understanding of how to apply her knowledge and abilities in many different contexts” (Mentkowski et al., 2000, p. 65). Personal development is also believed to support cooperation, collaboration, and civic commitment in community activity and to facilitate understanding of oneself in relation to the world of work and civic duty, and in a multicultural context. The content of the program is linked to its aims in that the abilities included in the curriculum are those necessary in real life. Also, throughout the implementation process, regardless of academic area, these abilities are emphasized.

The three programs highly value students’ personal development or intellectual development, a general term used in this dissertation. With such an aim, they include content that prepares college students not only for students’ future profession, but also as committed and responsible citizens. Alfred North Whitehead (1861-1947), a contemporary philosopher who has some respects of idea similar to John Dewey, supported the importance of personal development as a part of education.

Whitehead stated:

A merely well-informed man is the most useless bore on God's earth. What we should aim at producing is men who possess both culture and expert knowledge in some special direction. Their expert knowledge will give them the ground to start from, and their culture will lead them as deep as philosophy and as high as art. We have to remember that the valuable intellectual development is self-development...

(Whitehead, 1929/1997, p. 262)

It is important that for a program to successfully reach its goals, there is a bridge between program elements and its implementation process. Teachers seem to play a vital role in the implementation process.

How Should Teachers Teach College Students?

Based on Vygotsky's idea, the practice of constructivism focuses on the provision of learning environment that facilitate students' learning process. Constructivist teachers provide opportunities for learners to apply principles and knowledge through practical experience. Authenticity and real-life context are regarded to help learners build up knowledge not only by individual learner alone but with peers and teacher through their interactions among one another. Learners play more roles in class. They do not only passively perceive the knowledge transferred only by teachers. Constructivists do not regard teachers as a body of knowledge but as facilitators.

Kearsley (1994-2000) noted that Piagetian theory has been generalized in educational contexts, especially about the teachers' role. Constructivist teachers are expected to provide a rich environment that encourages, motivates, and challenges

learners to explore or to become an active constructor of knowledge through their own experiences. Two key principles for teaching and learning generalized from Piagetian theory are that learning is an active process and learning should be authentic. With an introduction of knowledge from teacher that students can use as an aid to solve problems, students learn to solve problems.

Teachers or instructors play a vital role in organizing materials and activities suitable for students' experience, or the learning process could not successfully proceed. As described by Reiman and Parramore (1994), conditions for growth include not only student teachers in new roles to experience their learning, there are other conditions. The other five conditions include contextual understanding and building trust, guided inquiry, balance of action and reflection, continuity of practice and reflection, support and challenge. For these conditions to proceed in an effective way to impact students' success, teachers or instructors are very important. Teachers are the ones who take responsibilities to build relationship with students. They select materials to match students' prior knowledge. They guide students through the inquiry process with appropriate questions to stimulate students to think for themselves. They organize activities so as to support balance and continuity between students' opportunities to practice or experience and guided inquiry for students' reflection. The goal, of course, is to challenge and support the students to make them grow developmentally.

The training components suggested by Joyce and Showers (1982) that Reiman and Parramore (1994, 1996) incorporate into their teacher education program, plays a

vital part in introducing teaching models to students. The model includes demonstration and feedback for students' practice.

Regarding the ADAPT program, as explained by Tomlinson-Keasey and Eisert (1978), meetings were held weekly among multidisciplinary staff to integrate their teaching materials and relate the materials to real life experience. In addition, teachers or instructors played a core role throughout the program. For each cycle of experience as a basis for following cycles of experience, teachers or instructors had responsibility for introducing the meaning of each cycle of experience through connections and transitions among the cycles. Teachers are responsible for organizing and providing activities that allow flexibility as well as providing opportunities for students to apply varieties of thinking skills. Most of all, they need to appropriately challenge students in such a way that the challenges match the students' current level of ability so that they move from concrete thinkers to become abstract thinkers.

The operation of the ability-based learning program continues under a sustained activity and communication among faculty with structured criteria related to development of some specified expected abilities as learning outcomes. Learning principles are articulated in ways that are understandable and transportable among all concerned personnel.

In addition to teachers, the ability-based program stresses the cooperation among all concerned members such as academic and educational researchers to implement research for program improvement; administrators to consider common tasks important to

the institution, to provide resources, and to participate in consortia, and departments; and committees to critiqued and debate about questions, methods, findings, interpretations, and related policy-making, and so forth.

Chapter Summary

This chapter reported findings for research questions two, three and four of this dissertation. First, it reported findings from the two convenient samples. Then, it examined the relationship between the findings of the two convenient samples to other eight longitudinal studies selected as samples for this dissertation. The ten studies were compared both quantitatively and qualitatively. Finally, the chapter analyzed deliberative programs designed to foster students' intellectual development. This analysis included a description of the programs designed to foster students' intellectual development and the analysis of features that support students' development. The analysis also contained a theoretical explanation of how the programs features are supportive to the students' development.

Findings from the two convenient samples suggest that the Deliberative Psychological and Professional Education Program (DPPE) supports students' development positively. The majority of the two cohorts from 1988 and 1989 began college with CL average scores between 1.65-1.85 while four years later the majority of them moved to average CL scores 2.01 and 2.07. This change in conceptual complexity level reflects the students' ways of knowing about absolute rightness/wrongness, and authority figures, consideration of alternatives, and tolerance for uncertainty and

ambiguity of knowledge. Gender difference in CL development was not found in the two studies.

Quantitative findings across the ten studies show that college attendance supports the students' intellectual development positively in an upward trend. As well, the magnitude of development as indicated by effect sizes in the majority of the studies was found to be medium (+.54, +.57, +.59, and +.60), large (+.73), and very large (+1.07, +1.34, and +1.43).

Qualitative analysis across the ten studies suggests that these studies explored students' development taking into account gender, types of programs, students' major fields of study, age and life experience, and previous academic achievement. Regarding gender, one study that investigated thinking supports the existence of gender-related pattern consistent with theoretical expectation. Six studies that explored gender issue focused on trend and magnitude of development and found evidence of gender differences but were not statistically significant. In terms of types of program, findings among the studies suggest that programs can foster students' intellectual development when deliberately designed and implemented. As for the matter of major academic field of study, no statistical significance was found in findings from three studies in which the students' major was examined for relationships to intellectual development. Among four studies that investigated age and life experience in relation to intellectual development, there seem to be no conclusive findings. Findings suggest an interpretation that students' intellectual growth occurs most consistently as a response to education. Two studies examined the relation between students CL in relation to their persistence

and dropout and found no relation between conceptual level and rate of graduation or dropout.

The last part of this chapter provides description of three programs including the DPPE or Deliberative Psychological and Professional Development, the ADAPT or Accent on Development Abstract Process of Thought, and the ability-based program. Since the programs are designed to foster students' intellectual development, the program features are analyzed in connection with psycho developmental theories to provide better understanding on how the programs support students' development.

CHAPTER V

SUMMARY, CONCLUSION AND IMPLICATIONS

Summary

This dissertation examined longitudinal studies of college students' intellectual development including evidence from two convenient sample studies conducted at the College of Education, North Carolina State University. The rationale for this dissertation was to situate longitudinal studies of teacher education students' intellectual development within a larger longitudinal research context. Higher education is playing a larger role in assessing intellectual development. Reasons include an enlarged mission to prepare students for complex and ill-structured real world problems. In order to prepare the students, higher education must integrate curriculum which addresses students' personal development, namely intellectual development, in addition to subject-matter knowledge. Current research suggests that intellectual development will enable students to more successfully confront and resolve problematic situations in work and civic life based on reflective judgment—forming decisions based on a preponderance of evidence. Subject-matter knowledge will only be marginally helpful in such contexts. Intellectual development correlates with more flexible and adaptive problem solving. In step with the current trend of higher education, this dissertation aimed to provide insights of how higher education promotes students' intellectual development as interpreted through a collection of longitudinal studies of undergraduate students.

With this aim, this dissertation posed four research questions:

1. What are common assumptions and connections across selected theories of intellectual development?
2. What are findings from the convenient samples of the two four-year longitudinal studies (1988-1992 and 1989-1993) that incorporated the Deliberative Psychological and Professional Education program (DPPE) for teacher education students?
3. What are the relationships between the findings from the two convenient samples and other longitudinal samples based on the connected theories of intellectual development?
4. What are distinctive features of education programs designed to foster students' intellectual development in undergraduate programs?

In response to the four questions, the dissertation first explored selected theories that are related to intellectual development. Many theories describe the assumptions of “intellectual” development. Among the theories, some constructs are connected to the extent that they share common assumptions. Consequently, this dissertation examined four theories with related assumptions: Perry’s Scheme of Intellectual and Ethical Development (1970), Baxter Magolda’s Epistemological Reasoning Model (1992b), King and Kitchener’s Reflective Judgment Model (1994) and Harvey, Hunt and Schroder’s Conceptual Development Theory (1961). Theoretical relationships were drawn among these theories and were presented in chapter two as a part of literature review, and summarized in this chapter as a part of the conclusions.

Second, the dissertation reported findings from two four-year longitudinal samples based on Harvey, Hunt and Schroder's Conceptual Development Theory (1961). The two convenient samples participated in a Deliberative Psychological and Professional Development program designed to foster intellectual development of college students in teacher education. Findings from these samples were compared to those from eight selected longitudinal studies based on the four theories just mentioned.

Finally, the dissertation examined trends in deliberative programs within four selected longitudinal samples. Details of the programs are analyzed. In the analysis, theories are referred to as explanations for why some features of the programs might serve to develop students' intellectual growth.

This chapter presents conclusions, discusses implications and proposes recommendations for future research. The presentation of the conclusion of findings is organized in accordance with the order of the research questions.

Conclusions

Common Assumptions and Connections

Across Four Theories of Intellectual Development

What conclusions can be drawn about common assumptions and connections across the selected theories of intellectual development? The detailed description and analysis of the four theories as presented in chapter two as a part of the literature review suggests that Perry's Scheme of Intellectual and Ethical Development (1970), Baxter

Magolda's Epistemological Reasoning Model (1992b), King and Kitchener's Reflective Judgment Model (1994) and Harvey, Hunt and Schroder's Conceptual Development Theory (1961) have common general theoretical assumptions that are consistent with cognitive-developmental theories, as well as assumptions about information, knowledge, and problems. Assumptions about intellectual development include the following:

1. The hierarchical stages or plateaus of development move from less complex to more complex levels;
2. Development from one stage or position to the next stage takes place first within the current stage of development of each individual prior to the developmental transition to the next stage, which marks significant qualitative transformation change;
3. Development occurs as a result of interaction between individuals and their surroundings, thus, there is a contextual dynamic in stage growth;
4. There is a connection between stages of development and individuals' actions or behaviors, however, the relationship is not one-to-one;
5. Individuals need to assimilate and accommodate their cognitive structures in order to be able to develop from one stage to others; and
6. All persons have an intrinsic need to be competent and self-determining, and such competence grows when there is positive interaction in a supportive and optimally challenging environment.

Regarding assumptions about information, knowledge, and problems, these theories assume the following:

1. Information/knowledge/problems are derived not only from one, single source but multiple sources;
2. The structure of information/knowledge/problems can be either simple and well-structured or complex and ill-structured;
3. For development to take place, individuals need to be able to differentiate the structures of information/knowledge/problems; and
4. Intellectual development is best assessed when persons engage in complex new experiences with ill-structured problems.

With these connections among the four theories of intellectual development, it can be concluded that for intellectual development to take place, individuals must be intrinsically motivated to engage complex and ill-structured information/knowledge/problems in the environment. Without such motivation, individuals might not perceive the importance of change or development. As well, the theorists suggest that sustained interaction within the environment is needed. The interaction needs to involve individuals' deep engagement with the ill-structured problems and complex "human-helping" experiences. Such engagement includes the following: ongoing collaborative inquiry (analysis and reflection); complex human-helping experiences such as internships, service learning, tutoring, mentoring; support and optimal challenge; and continuity (at least three- to six- months of sustained interaction in the experience). In addition, individuals need to have adequate schema or background knowledge or competency in order that their experiential interaction effectively leads to the assimilation and accommodation of their cognitive structures. This competency includes the

acknowledgment of the multiplicity of information/ knowledge/problems and the differentiation of structures of information/knowledge/ problems. It can be concluded that the theories of intellectual development involve not only the inner process of individuals' cognitive development but also the motivation to stimulate the cognitive process to develop or change through an interaction with the world outside themselves such as environment.

The following section provides conclusions for research questions two and three: What are findings from the two convenient samples of four-year longitudinal (1988-1992 and 1989-1993) of Deliberative Psychological and Professional Education program (DPPE) for teacher education students? and What are the relationships between findings from the convenient samples when compared to other longitudinal samples based on selected theories of intellectual development?

Findings from the Two Convenient Samples and Their Relation to Findings from Other Studies

The two studies show positive changes in intellectual development. The Deliberative Psychological and Professional Education program (DPPE) may be contributing to students' conceptual change. However, experimental or quasi-experimental studies are needed to clarify this relationship. Reviews of the relationships have been summarized by Reiman (1999b) and Sprinthall (1994) and confirmed trends found in the longitudinal studies. Posttest CL scores of the two studies indicate change of conceptual complexity level that reflects the way students think from believing in

absolute rightness/wrongness and being cautious about authority and authority figures to taking into consideration alternatives and tolerating uncertainty and ambiguity of knowledge. No statistical trends of development were found regarding gender.

The quantitative analysis across the ten studies suggests that participation in college supports students' intellectual development positively in an upward trend. As well, the outcome effect sizes indicate small to very large magnitude effects, +0.15 to +1.43. Even though different measurement instruments were employed, consistency was found across the ten studies in terms of students' characteristics of the position, stages or levels of development at the entrance and later in college years. With respect to the four-year studies in particular, it seems apparent that at the entrance to higher education, the students generally tended to view knowledge as absolute and certain. Any uncertainty or ambiguity of knowledge came from the fact that experts or knowers had not yet discovered it. First-year students tended to depend on teachers to justify what was right or wrong. During the college years, their ways of thinking changed. The majority of students began to perceive that knowledge was not always certain and absolute. They also became more tolerant of the uncertainty of knowledge. Knowledge was judged as relatively right or relatively wrong with references to evidence claimed by experts or knowers, or by individuals' personal opinions. Still the rightness and wrongness of knowledge was interpreted as to whether it was compatible with what the experts or knowers claimed. While some students still justified their beliefs based on reasons of rightness of knowledge as claimed by experts, other students began to recognize that

experts' claims could also be relative to those experts' individual personal opinions and that knowledge is constructed by the knowers.

The ten studies examined students' development longitudinally. As well, the studies accounted for gender, types of programs, students' major fields of study, age and life experience, and previous academic achievement. Regarding gender, one study by Baxter Magolda (1990) found the existence of gender-related patterns of epistemological development: gender differences exist within stages in epistemological reasoning structures, however, there are no gender differences as males and females progress from one cognitive structure stages to next stages. Taking into account trend and magnitude of development, six studies examined gender difference and found evidence of differences in development between men and women but none of the differences were statistically significant. Based on this evidence, it seems sensible to conclude that there is no statistically significant difference in trend and magnitude of intellectual development between men and women.

Four research studies incorporated programs that were designed to deliberately foster students' intellectual development. Regarding students' developmental change, one study showed a small effect size (+.35), two studies showed medium effect sizes (+.54 and +.59), and one study showed a very large effect size (+1.07).

As for the matter of academic field of study, three studies explored this issue and found no statistically significant relation between intellectual development and academic majors. Yet, the evidence from all these studies suggest a tendency of students majoring

in liberal arts to show greater gains on PCM and RJI posttest scores when compared to other majors within each study.

Considering age and life experience, no conclusive findings were found among the four studies that investigated this issue. Findings suggest the interpretation that students' intellectual growth is most affected by a combination of education and the aging process.

Taking into account the relation between conceptual level and the rate of persistence and dropouts, two studies examined this relationship and found no statistically significant relationship. It was concluded that conceptual level is not associated with student persistence in college.

The next section describes conclusions related to research question four: What are features of a psychological and professional development education program and those of other two programs within the reviewed studies that support students' intellectual development?

Features of Educational Programs Supportive of Students' Intellectual Development

Three programs were described and analyzed in this dissertation including the Deliberative Psychological and Professional Education Program (DPPE), the Accent on Development Abstract Process of Thought (ADAPT), and the Ability-Based Learning Program. The educational programs fostered students' intellectual development by

attending to three basic elements of educational process: how students learn, what to teach, and how to teach.

Program features that are related to students' learning process are: a) students' experience in real-world contexts; b) students' reflection on and analysis of their experience; c) continuity of experience and reflection; and d) optimal challenge and support for students to equilibrate their cognitive growth.

It is more difficult, of course, to identify the most proximate causes of intellectual growth. Other studies have examined this question more carefully using experimental designs and casual models. Studies by Mentkowski and Strait (1983), McNeel (1994), Reiman (1999b), Sprinthall (1994), for example, provide strong evidence for curricular influence on both intellectual and moral development. Based on longitudinal and casual data, it was concluded that progress in a developmentally sequenced curriculum that includes progressively more complex activities/experience/new roles that are designed with ongoing self-assessment and reflection does lead to growth in intellectual reasoning—the capacity to think through and embrace abstract and reflective concepts.

What conclusions can be drawn from the analysis of features of the three programs? For a program to foster students' developmental growth, the student should be considered not only as learner but also as developing person. Most importantly, they need to be educated as effective and responsible graduates and professionals.

Considering the student as learner, the cultural contexts and formal curricula of the three educational programs were designed to shape the way students think, feel, and

relate to others. Each of the programs features constructivist approaches that involve intensive student interaction with other persons. As well, all programs encourage progressively more complex engagement in learning experiences. Among the features described which support the student as learner include student assessment (with its focus on performance, self-assessment, and feedback), experiential learning, written reflective inquiry that is dialogic, classroom discussion, work in groups and skill in communication.

Regarding the student as developing person, an enduring assumption of each of the “deliberative” developmental curricula is that education should develop the students’ progressively more complex ways of knowing and meaning making. The programs view the individual learner as actively constructing meaning across a broad range of domains. Such meaning making becomes developmental when the student can construct and activate increasing complexity in real world contexts. Program designers argue that such development should be the unifying purpose of higher education. The programs encouraged curriculum designers and faculty to understand general patterns of student development.

Although it has become popular to dismiss stage theories as overly simplistic, findings from the reviewed longitudinal studies support their ability to describe the student as developing person in the intellectual and personal domains. As well, the reviewed deliberative programs link student involvement in a breadth of activities such as reading, complex human-helping activities, ongoing reflection and assessment, and civic leadership.

Above all, the designs of the programs are based on views of students as effective and responsible graduates and professionals. Each of the deliberative programs aimed to transfer students' deepest learning to postcollege contexts. Thus, the programs make the case that performance is essential to the professions and to civic engagement. Performance can be coached, and it invites an emphasis on effective and responsible action. This definition implies standards of excellence in situational and discretionary contexts.

Performance, as defined by these programs, cannot be understood simply in behavioral terms. Effective and responsible performance illuminates the character of professional judgment. It also joins a dynamic interplay of a learner's intentions, thoughts, and feelings. As such, these programs suggest that developmental domains or dispositions transfer across a wide range of personal and professional settings.

Implications of Findings

Theoretical Implications

Implications of the findings relating to theories are discussed in terms of grounded theory and theory application.

Grounded Theory

Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reasoning Model (1992b), and King and Kitchener's Reflective Judgment Model (1994) were constructed from information elicited from

students, while Harvey, Hunt, and Schrodgers' Developmental Theory (1961) was generated from theory. In addition, the former three theories were constructed in educational contexts while the conceptual system was constructed in psychological contexts and later applied to educational contexts. What are the implications of the employment of students' thinking as a basis for theory construction? What are the implications of different context of theory construction?

Students' thinking as a basis for theory.

Regarding the employment of students' voice as a basis for college student theory construction, Perry (1970/1999) was the first to lead the way. Perry highly valued students' standpoint, positionality or perspectives acquired from one's lived life. Working with students, Perry (1981/1997) found that when the same students returned for his follow-up interviews, students startled him with their reinterpretation of their lives. He emphasized that students not only have an ability to construct meanings but also to reconstruct them. As students engage complex experiences, they shift or change their views as they accommodate complexity in knowledge and learning. Following the creation of Perry's scheme of intellectual and ethical development, Kitchener and King (1994) and Baxter Magolda (1992b) grounded their theories with information from students' perspectives. An advantage of employing students' perspectives as a basis for grounding developmental theories for college students seems to be related to the fact that such perspectives provide insights into how students are thinking, knowing and learning.

However, for the students' perspectives to be reliable and useful for theory construction, they need structured prompts in terms of what information they should provide. In their first trial efforts, Perry (1970/1999) and his associates asked open-ended questions and found that students would rather be addressed with appropriate interview questions. Perry explained that students typically came with expectations about the interviewers' interest. Without appropriate, specific questions that are relevant to interviewers' interest, some students had difficulties answering the questions. Baxter Magolda (1992b) had a similar experience when she first interviewed her students and her questions had to be revised for follow-up interviews.

It is also worth noting that when the students' perspectives are included as a basis for theory construction, researchers need to be cautious about their assumptions and bias when they construct theories. It is apparent that these three theories are different in details because the theorists viewed and interpreted the information with different focuses based on their personal beliefs and assumptions.

The study by Perry and his associates (1970) originated from their impression of students' varieties of means of responding to relativism in pluralistic environments at Harvard College in 1953. From their experience working with the students, they argued that the sequential changes of students' worldview are a response to the relativism of knowledge and values in a pluralistic culture. Such change is a personal process of maturation. Consequently, their outcome model focuses on students' worldview. Following are the illustrations of how variations of connected theories could possibly be affected by the agenda of the theorists.

Baxter Magolda (1992b) interviewed her students to elicit information to be used as a basis for constructing the Epistemological Reasoning Model. Based on previous research, Baxter-Magolda observed that Perry's theory and Belenky and Associates' (1997) theory suggested differences in ways of thinking between males and females. The interview protocols from students were analyzed to focus on the difference in patterns of thinking relating to gender.

King and Kitchener (1994) took the stance of a concept of reflective thinking grounded in Dewey's (1933, 1938) definition relating to his observation that uncertainty is a characteristic of the search for knowledge and that true reflective thinking is initiated only after there is recognition of the existence of real problems. Individuals also form reflective judgment in order to bring closure to uncertain situations. Consequently, the Reflective Judgment Model focused on the description of a developmental progression relating the way individuals understand the process of knowing and how they justify their beliefs about ill-structured problems. The ultimate outcome of this progression is expected to be the ability to make reflective judgments.

Psychological and educational contexts of theory construction.

Another implication is related to the integration of psychological knowledge in educational contexts. Psychology and education seem to be inseparable if one aims to foster students' intellectual growth. The four theories assume developmental maturity as an outcome of college education. If there is a goal of helping students as persons,

researchers need to understand how students grow psychologically, not only in educational environments, but also in personal and professional surroundings.

Another way to examine the implications of the theories is to consider how they can be applied.

Theory Application

Since the creation and publication of Perry's Scheme of Intellectual and Ethical Development of college students, numerous theories have been developed to better understand college students. The four reviewed theories have been applied in many research studies with the focus on trend and magnitude of students' development. Such development complies with the expectations of these theories. Cumulative research evidence seems to confirm the reliability and validity of these theories. This evidence implies the practicality of the theories as constructs for understanding the college student as learner. Not only can these theories be applied in a variety of educational contexts, but the development of students can be monitored to some extent.

As discussed earlier, a connection among these theories is that they share common assumptions of cognitive-developmental theories and as assumptions about information, knowledge, and problems. These assumptions are applied universally in any situation where learning is to take place. When applied in college educational contexts, students bring with them their current perspectives and potential for development to interact with the multiplicity of curriculum and college life experiences. They then accommodate their cognitive structures and their views change. How students change

depends on what they will experience in college both academically and personally. The quality and quantity of student changes depends on how they are guided, supported, and monitored through the progressive challenges of higher education. Understanding the elements of the developmental process can assist educators as they guide students along their potential developmental trajectories.

The next section discusses implications of findings from research studies.

Research Implications

Findings suggested that participation in college supports students' intellectual development. In addition to academic curriculum, there are co-curricular experiences that influence students' development. These factors can be categorized as internal and external factors.

The internal factors include students' gender, age, personal experience, and competency (as indicated by score on academic success). Considering internal factors, students attending college firstly have a significant potential to develop because of their age –late adolescence—and secondly they have expectations to change. In other words, college students as young adults have distinctive psychological capacities and motivations which strongly shape personality (Sprinthall & Collins, 1995). In addition, as noted by Boyer (1987), students view college as a step toward life fulfillment. College students bring with them expectations that the college experience will develop their competency toward becoming more complete human beings.

External factors include curriculum, major fields of study, and social context in college. Varieties of factors imply that the nature of intellectual development is complex and involved many interconnected factors. Pascarella and Terenzini's (1991) synthesis of research on college students' development suggests such factors including family background, socio-economic status, race and ethnicity. It is important that students' potentials for development, including external factors that influence students' developmental growth, be taken into consideration for the design of curriculum and the management of college environment if colleges are to foster students' intellectual development. Although this area is worth researching, it is not the focus of this study.

Theoretical and research implications have been discussed. The following section provides recommendations for further studies.

Future Directions

Recommendations address theory development, research and educational programming applications. As well, recommendations are made for teacher education and higher education.

Theory

As discussed earlier, Perry's Scheme of Intellectual and Ethical Development (1970), Baxter Magolda's Epistemological Reasoning Model (1992b), and King and Kitchener's Reflective Judgment Model (1994) are grounded in large numbers of student interviews over many years whereas Conceptual Development Theory was generated

from theories. Generalization of these theories is limited to comparable populations. In the study by Perry and his associates, research participants were white males from a liberal arts college. In studies by King and Kitchener (1994), participants were diverse in age, educational background, gender and ethnicity. However, there are no trans-national studies. In the study by Baxter Magolda (1992b), the majority of the students were white and 18 years of age. In Hunt's research, participants were diverse by age, ethnicity, and gender. As well, there were numerous studies of teachers. Baxter Magolda and Terenzini (2000) suggest that the diversity of students today may be greater than at any time in the history of American higher education. It is important that more research be done to add evidence accounting for such salient characteristics of today's students as socioeconomic status, race and ethnicity, sexual orientation, and learning and physical ability.

Apart from adding more evidence through the inclusion of students' diversity, it is recommended that further research examine students' change process. How do students' develop their intellectual capacity? What do developmental transitions look like? The four theories tend to focus on how and how much change occurs in college students' life rather than on the conditions that foster students development. This question requires further research. Although description of trends in stage change is important, there is a need for better understanding of what the change looks like, and how such student change can be fostered. Such questions have been encouraged by Siegler and Crowley (1991) who call for more research on the microgenetic processes of cognitive development. Both King and Kitchener (1994), Reiman (1988) and Sprinthall, Reiman, and Thies-

Sprinthall (1996) have begun to explore these microgenetic processes as applied in teacher education and teacher professional development. As well, Astin (1984/1997) proposes a developmental theory for higher education namely “the theory of student involvement” that focuses on the quantity and quality of the physical and psychological energy that students invest in the college experience. His study shows that the concept of involvement is a critical element in the learning process.

Research

Flowers, Osterlind, Pascarella, and Pierson (2001) noted that there were limitations in previous research concerning what students learn in college. One of the limitations was the lack of control for conditions that did not impact students apart from net impacts of college. That is to say, the studies were conducted in college contexts without comparison groups of participants who did not attend college. Findings of this dissertation suggest that this limitation exists in all the study samples. In addition, research on the effect of participation in college on students’ intellectual development could be conducted to compare the trend and magnitude of students’ development across varied colleges and university programs during the same period. The results might give insights into curricula development that are optimal.

Moreover, most of the research on individuals’ development has taken place in academic contexts. It is recommended that more studies be conducted among individuals in nonacademic settings to examine their developmental trajectories and tendencies for

growth. As well, more research should be done on the application of intervention programs.

Above all, in accordance with the recommendation for further research on theories about the factors that influence students' development, more research is needed to explore such factors as internal and external motives. What conditions encourage students to want to change and develop?

Applications

One of the recommendations for applications of developmental theories is experimentation on program designs and implementation. This seems to be a vast area for further exploration taking into account diversity of students, subject areas, levels of education and other factors.

Regarding the excellence of institutions, Pascarella (2001) notes that it is reasonable to assume that an excellent undergraduate education is most likely to occur at those colleges and universities that maximize good practices and enhance students' academic and social engagement or effort. To make judgments about "effective practices," Pascarella suggested that rather than rational linkages, the empirical linkages to student growth and development should be measured. In other words, the effective educational practices approach needs to be guided by a body of evidence on college impact rather than factors that are hypothesized to be important. This view is consistent with earlier observations by Widick and Simpson (1978) and Hunt (1978).

Widick and Simpson (1978) noted that an essential requirement for successful instruction includes the incorporation of a developmental view to teaching, which requires a “parallel artistry: a counselor’s intuitive sense and capacity to listen, empathize, and respond at many levels”. Hunt (1978) remarked that the idea of “parallel artistry” is most important. He alluded that what worked best when working with teachers was to discover their knowing system. He shared that with his experienced working with junior and senior high school teachers, he found:

Contrary to what I expected, the theory is much less important than teachers’ capacity to adapt to students through what I called “reading” and “flexing”. A theory may serve as an exemplar or metaphor for such adaptation (if it coordinates student differences with different educational approaches), but it cannot be “applied” directly. As Widick and Simpson observed, even a step-by-step manual will not guarantee success (p.65).

It follows that education needs to better understand students. Such understanding indicates: (a) respect for students as learners; (b) acknowledgement of learners as developing persons; and (c) recognition that graduates must be effective and responsible professionals.

As noted earlier, each of the deliberative programs described in this dissertation aimed to understand the student as learner and developing person, and then to encourage new transformations in students’ epistemological and intellectual conceptions. As well, two of them aimed to transfer students’ deepest learning into postcollege contexts such as teaching in K-12 classrooms. Thus, these deliberative programs sought to transform students’ conceptions and performance. Performance is essential to professions such as teaching and teacher education, and based on these programs, can be guided or coached.

Such performance encourages an emphasis on effective and responsible action. This goal implies standards of excellence in discretionary contexts.

The longitudinal studies help educators understand how students learn and develop particularly in the intellectual domain. The deliberative programs are designed to foster both development and changes in college and postcollege performance. However, only two deliberative programs address the importance of responsible action.

What is now needed are careful studies that examine the nexus between intellectual development, effective performance, and responsible action. How is responsibility seen in professional actions? What are the relationships between reflective judgment and greater tolerance? How does development express itself in the actions of teachers in the process of communicating and caring for each individual student or family member? These questions lead researchers to better understand the quality of dispositional action – how reasoning, method, and manner (e.g. tolerance, caring, industriousness) interact. Such questions are critical in teacher education, and they also relate to higher education's aims of fostering effective professionals who are responsible and civically engaged.

REFERENCES

References marked with an asterisk indicate studies included in the comparative analysis.

- Adams, M., & Zhou-McGovern, Y. (1994, April). The sociomoral development of undergraduates in "Social Diversity" course: Developmental theory, research and instructional applications. Paper presented at the American Educational Research Association, New Orleans, LA.
- Alishio, K. C., & Schilling, K. M. (1983, May). Sex differences in intellectual and ego development. Paper presented at the 55th, Annual Meeting of the Midwestern Psychological Association, Chicago, IL.
- Anson, C. M. (1989). Response styles and ways of knowing. In C. M. Anson (Ed.), Writing and response: Theory, practice, and research (pp. 332-367). Urbana, Ill: National Council of Teachers of English.
- Astin, A. W. (1997). Student involvement: A developmental theory for higher education. In P. G. Altbach (Series Ed.) & K. Arnold & I. C. King (Vol. Eds.), Contemporary higher education: International issues for the twenty-first century: Vol. 4. College student development and academic life: Psychological, intellectual, social, and moral issues (pp. 251-262). New York: Garland Publishing, Inc. (Reprinted from Journal of College Student Personnel, 25, pp. 297-308, 1984)
- Astin, A. W. (1999). Rethinking academic 'excellence'. Liberal Education, 85 (2), 8-18.
- Astin, A. W., & Astin, S. H. (2000). Leadership reconsidered: Engaging higher education in social change. Battle Creek, MI: W. K. Kellogg Foundation.
- Atkin, E. (1996, May). Students' learning stages and academic assessment. Paper presented at the 36th Annual Forum of the Association for Institutional Research, Albuquerque, NM.
- Atkin, E. (1998, May). Identifying dominant stage of intellectual development: Knowing and learning. Paper presented at the 38th Annual Forum of the Association for Institutional Research, Minneapolis, MN.
- Austin, W. W. (2000). The research paper in cyberspace: Source-based writing in the composition classroom. Dissertation Abstracts International, 61 (08A), 3152.
- Baer, E. M. (1988, March). An evaluation of integrating literature in the composition class. Paper presented at the 39th Annual Meeting of the Conference on College Composition and Communication, St. Louis, MO.

- Baxter Magolda, M.B. (1987a). The affective dimension of learning: Faculty-student relationships that enhance intellectual development. College Student Journal, 21, 46-58.
- Baxter Magolda, M. B. (1987b). Comparing open-ended interviews and standardized measures of intellectual. Journal of College Student Personnel, 28 (5), 443-448.
- Baxter Magolda, M.B. (1988). Measuring gender differences in intellectual development: a comparison of assessment methods. Journal of College Student Development, 29 (6), 528-537.
- *Baxter Magolda, M. B. (1990). Gender differences in epistemological development. Journal of College Student Development, 31, 555-561.
- Baxter Magolda, M. B. (1992a). Cocurricular influences on college students' intellectual development. Journal of College Student Development, 33 (3), 203-213.
- Baxter Magolda, M.B. (1992b). Knowing and reasoning in college: Gender-related patterns in students' intellectual development. San Francisco: Jossey-Bass Publishers.
- Baxter Magolda, M. B. (1992c). Students' epistemologies and academic experiences: Implications for pedagogy. Review of Higher Education, 15 (3), 265-287.
- Baxter Magolda, M. B. (1993a, April). The convergence of relational and interpersonal knowing in young adult's epistemological development. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA.
- Baxter Magolda, M. B. (1993b, November). The development of contextual knowing in graduate and professional education settings. Paper presented at the 18th Annual Meeting of the Association for the Study of Higher Education, Pittsburgh, PA.
- Baxter Magolda, M. B. (1995). The integration of relational and impersonal knowing in young adults' epistemological development. Journal of College Student Development, 36 (3), 205-216.
- Baxter Malgoda, M. B. (1999). The evolution of epistemology: Refining contextual knowing at twenty something. Journal of College Student Development, 40 (4), 333-344.
- Baxter Magolda, M. B. (2002). Epistemological reflection: The evolution of epistemological assumptions from age 18 to 30. In B.K. Hofer & P.R. Pintrich (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp.89-102). Mahwah: NJ, US: Lawrence Erlbaum Associates, Inc.

- Baxter Magolda, M., & Porterfield, W. D. (1985). A new approach to assess intellectual development on the Perry scheme. Journal of College Student Personnel, 26, 343-351.
- Baxter Magolda, M., Terenzini, P. T., & Hutchings, P. (2000). Learning and teaching in the 21st century: Trends and implications for practice. American College Personnel Association. [on-line].
Available: <http://www.acpa.nche.edu/seniorscholars/trends/trends.htm>
- Becker, B. J. (1988). Synthesizing standardized mean-change measures: British Journal of Mathematical and Statistical Psychology, 41 (2), 257-278.
- Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). Women's ways of knowing. New York: Basic Books.
- Belenky, M. F., Clinchy, B.M., Goldberger, N.R., & Tarule, J.M. (1997) Women's ways of knowing: The development of self, voice, and mind (10th Anniversary Edition). New York: Basic Books.
- Benack, S. (1982). The coding of dimensions of epistemological thought in young men and women. Moral Education Forum, 7 (2), p. 3-22.
- Berliner, D. C. (1985). Laboratory setting and the study of teacher education. Journal of Teacher Education, 16 (6), 2-8.
- Berman, D. S., & Davis-Berman, J. (1995). Outdoor education and troubled youth (Report No. ED-RC-95-5). Charleston, WV. (ERIC Digest No. ED385425)
- Borko, H., & Michalec, P. (1997). Student teaching portfolios: A tool for promoting reflective practice. Journal of Teacher Education, 48 (5), 345-358.
- Bowen, H. W. (1977). Investment in learning: The individual and social value of American higher education. San Francisco: Jossey-Bass.
- Boyd, K. B. (2001). Critical thinking tests and higher education research. Dissertation Abstracts International Section A: Humanities and Social Sciences, 62 (3-A) 984.
- Boyer, E. L. (1987). College the undergraduate experience in America. New York: Harper & Row Publishers.
- Brabeck, M. M. (1983). Critical thinking skills and reflective judgment development: Redefining the aims of higher education. Journal of Applied Developmental Psychology, 4 (1): 23-34.

- Brabeck, M. M. (1984). Longitudinal studies of intellectual development during adulthood: Theoretical and research models. Journal of Research and Development in Education, 17 (3), 13-27.
- Brogan, D. W. (1960). America in the modern world. New Brunswick: Rutgers University Press.
- Brown, R. D. (1972). Student development in tomorrow's higher education: A return to the academy. Student Personnel Series, (16). Washington, DC: American College Personnel Association.
- Bruch, M.A., Heiseler, B.D., & Conroy, C.G. (1981). Effects of conceptual complexity on assertive behavior. Journal of Counseling Psychology, 28 (5), 377-358.
- Call, P. E. (2000). Reflective questioning: A strategy to review notes. Journal of Adolescent & Adult Literacy, 43 (5), 487-489.
- Carr, L. N. (1963). "Social-action" as related to verbalized attitudes, self-concept and background factors. Dissertation Abstract, 24 (5), 2121.
- Case, R. (1985). Intellectual development: Birth to adulthood. Orlando: Academic Press, Inc.
- Cicala, J. J. (1997). The relationship between student involvement and reflective judgment for students attending a community college. Dissertation Abstracts International Section A: Humanities and Social Sciences, 58 (6 A), 2098.
- *Conway, Jr., T. E. H. (1994). A longitudinal follow-up study of change in cognitive complexity among participants in a cognitive developmental based intervention program for at-risk college students. Doctoral dissertation, North Carolina State University, Raleigh.
- Coren, S., & Suedfeld, P. (1995). Personality-correlates of conceptual complexity. Journal of Social Behavior and Personality, 10 (1), 229-242.
- Dale, J. L. (1995). Reflective judgment in a conservative Christian college of clergy education. Dissertation Abstracts International Section A: Humanities and Social Sciences, 56 (4 A), 1289.
- Davydov, V. V. (1995). The influence of L.S. Vygotsky: Theory, research, and practice. Educational Research, 24, 12-21.
- Dewey, J. (1896). The university school. University [of Chicago] Record, 1, 417-419.

- Dewey, J. (1933). How we think: A restatement of the relation of reflective thinking to the educative process. Boston: D.C. Heath and Company.
- Dewey, J. (1938). Logic: The theory of inquiry. Troy, Mo.:Holt, Rinehart & Winston.
- Dieker, L. A., & Monda-Amaya, L. E. (1997). Using problem solving and effective teaching frameworks to promote reflective thinking in preservice special educators. Teacher Education and Special Education, 20 (1), 22-36.
- Dings, J. G. (1998). Validity issues in using forced-choice, self-report measures of reflective judgment. Dissertation Abstracts International: Section B: The Sciences and Engineering, 58 (8B), 4514.
- Erlich, T. (2000). Civic responsibility and higher education. San Francisco: Jossey-Bass.
- Erwin, T.D. (1983). The Scale of Intellectual Development: Measuring Perry's scheme. Journal of College Student Personnel, 24, 6-12.
- Flowers, L, Osterlind, S. J., Pascarella, E. T., & Pierson, C. T. (2001). How much do students learn in college? Journal of Higher Education, 72 (5), 565.
- Flowers, L., Pascarella, E.T., & Pierson, C.T. (2000). Information technology use and cognitive outcomes in the first year of college. Journal of Higher Education, 71, (6), 637.
- Franklin, M. (1995). The effects of differential college environments on academic learning and student perceptions of cognitive development. Research in Higher Education, 36 (2), 127-153.
- Freeman, B. (1994). A study of the efficacy of a developmental instruction design for non-traditional adult learners (nontraditional students). Dissertation Abstracts International, 55, (08A), 2248.
- Friedman, A. A. (1996). The relationship between intellectual disposition and reflective judgment in college women. Dissertation Abstracts International Section A: Humanities and Social Sciences, 56 (9 A), 3468.
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). Educational research: An introduction (3rd ed.). New York: Longman Publisher USA.
- Gilligan, C. (1977). In a different voice: Women's conceptions of self and of morality. Harvard Educational Review, 47, 481-517.
- Ginsburg, H., & Opper, S. (1969). Piaget's theory of intellectual development: An introduction. Englewood Cliffs, NJ: Prentice-Hall.

- Glaser, E.M. (1985). Critical thinking: Educating for responsible citizenship in a democracy. National Forum, 65 (1), 24-27.
- Goldfarb, M. E. (2000). The educational theory of Lev Semenovich Vygotsky (1896-1934) [On-line]. (E. G. Rozycki & M. F. Goldfarb). Available:<http://www.newfoundations.com/GALLERY/Vygotsky.html>
- Goodman, J. (1984). Reflection and teacher education: A case study and theoretical analysis. Interchange, 15 (3), 9-26.
- Goodman, J. (1985). What students learn from early field experiences? Journal of Teacher Education, 25 (6), 42-48.
- Griffiths, M. & Tann, S. (1992). Using reflective practice to link personal and public theories. Journal of Education for Teaching, 18 (1), 69-85.
- Guthrie, V. L. (1997). The relationship of levels of intellectual development and levels of tolerance for diversity among college students. Dissertation Abstracts International Section A: Humanities and Social Sciences, 57 (11 A), 4666.
- Hallinan, M. T., & Khmelkov, V. T. (2001) Recent developments in teacher education in the United States of America. Journal of Education for Teaching, 27 (2), 175-185.
- Hancock, D. R. (1994). Motivating adults to learn academic course content. Journal of Educational Research, 88, 102-109.
- Hansen, R. N. (1998). Relationship of specific cognitive factors to smoking status among nursing staff. Dissertation Abstracts International: Section B: The Sciences and Engineering, 58 (10B), 5644.
- Harris, R. M. (1981). Conceptual complexity and preferred coping strategies in anticipation of temporally predictable and unpredictable threat. Journal of Personality and Social Psychology, 41 (2), 380-390.
- Harvey, O. J., Hunt, D. E., & Schroder, H. M. (1961). Conceptual systems and personality organization. New York: John Wiley & Sons, Inc.
- Hatcher, J. A., & Bringle, R. G. (1997). Reflection. College Teaching, 45 (4), 153- 159.
- Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. Teaching and Teacher Education, 11 (1), 33-49.
- Herek, G. M. (1988). Heterosexuals' attitudes toward lesbians and gay men: Correlates and gender differences. Journal of Sex Research, 25 (4), 451-477.

- Hofer, B. (2002) Personal epistemology as a psychological and educational construct: An introduction. In B. K. Hofer and P. R. Pintrich (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (pp. 3-14). Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc.
- Hofer, B., & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. Review of Educational Research, 67 (1), 88-140.
- Howey, K., Yarger, S., & Joice, B. (1978). Improving teacher education. Washington, DC: Association of Teacher Educators.
- Hunsaker, L., & Johnston, M. (1992). Teacher under construction: A collaborative case study of teacher change. American Educational Research Journal, 29 (2), 350-372.
- Hunt, D. E. (1966). A conceptual systems change model and its application to education. In O.J. Harvey (Ed.), Experience Structure & Adaptability (pp.277-302). New York: Springer Publishing Company, Inc.
- Hunt, D. E. (1971). Matching models in education: The coordination of teaching methods with student characteristics. Ontario Institute for Studies in Education.
- Hunt, D. E. (1976). Teachers are psychologist, too: On the applications of psychology to education. ACT Research Report (73). The Research and Development Division.
- Hunt, D. E. (1977). Conceptual level theory and research as guides to educational practice. Interchange, 8 (4), 78-90.
- Hunt, D. E. (1978). A comment on developmental concepts in college instruction. In C.A. Parker (Ed.), Encouraging Development in College Students (pp.64-66). Minneapolis, MN: University of Minnesota Press.
- Hunt, D. E., Butler, L. F., Noy, J.E., & Rosser, M. E. (1977). Assessing conceptual level by the Paragraph Completion Method. Ontario Institute for Studies in Education.
- Hunt, D.E., & Harvey, O, J. (1961). A conceptual systems change model and its application to education. In O.J. Harvey (Ed.), Experience structure and adaptability (pp.).New York: Springer Publishing Company, Inc.
- Inhelder, B., & Piaget, J. (1958). The growth of logical thinking from childhood to adolescence. New York: Basic Books.

- Irwin, R. R. (1996). Intellectual development and computer-assisted instruction: The improvement of dialectical thinking and argumentative writing. Dissertation Abstracts International Section B: The Sciences and Engineering, 56 (10B): 5795.
- Jacobson, C. K. (1985). Resistance to affirmative action: Self-interest or racism? Journal of Conflict Resolution, 29 (2), 306-329.
- Jensen, L. L. (1998). The role of need for cognition in the development of reflective judgment. Dissertation Abstracts International Section B: The Sciences and Engineering, 59 (4 B), 1884.
- Johnson, J. B. (1999). A comparison of cognitive development between Whites and African Americans based on William Perry's scheme of intellectual and ethical development. Dissertation Abstracts International, 61 (02A), 522
- Johnson, J. B. (2000). A comparison of cognitive development between White and African Americans based on William Perry's scheme of intellectual and ethical development. Dissertation Abstracts International Section A: Humanities and Social Sciences, 61 (2 A), 522.
- Joyce, B., & Showers, B. (1982). The coaching of teaching. Educational Leadership, 40 (1), 4-9.
- Kearsley, G. (1994-2000). Overview of Cognitive Constructivism. In Explorations in learning & Instruction: The Theory Into Practice Database [On-line]. Available: <http://home.sprynet.com/~gkearsley>
- Kegan, R. (1982). The evolving self: Problem and process in human development. Cambridge, MA: Harvard University Press.
- Kegan, R. (1994). In over our heads: The mental demands of modern life. Cambridge, MA: Harvard University Press.
- *Khalili, H., & Hood, A. B. (1983) A longitudinal study of change in conceptual level in college. Journal of College Student Personnel, 24, 389-394.
- King, P. M. (1977). The development of Reflective Judgment and formal operational thinking in adolescents and young adults. Dissertation Abstracts International, 38, 7233A.
- King, P.M. (1978). William Perry's theory of intellectual and ethical development. In L. Knefelkamp, C. Widick, & C. Parker (Eds.), Applying new developmental findings: New directions for student services, no. 4 (pp. 35-51). San Francisco: Jossey-Bass.

- *King, P.M., & Kitchener, K.S. (1994). Developing Reflective Judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults. San Francisco: Jossey-Bass Publishers.
- King, P.M., & Kitchener, K. S. (2002). The Reflective Judgment Model: Twenty years of research on epistemic cognition. In B. K. Hofer and P. R. Pintrich (Eds.), Personal epistemology: The psychology of beliefs about knowledge and knowing (p.37-61). Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc.
- *King, P. M., Kitchener, K. S. , Davison, M. L., Parker, C. A., & Wood, P. K. (1983). The justification of beliefs in young adults: A longitudinal study. Human Development, 26, 106-116.
- King, P. M., Kitchener, K. S., & Wood, P.K. (1990, November). Changes in students' reasoning over time: A ten-year longitudinal study of the development of reflective judgment. Paper presented at the annual meeting of the American Association of the Study of Higher Education, Portland, OR.
- King, P. M., Kitchener, K. S., & Wood, P.K. (1991, November). Moral and intellectual development beyond the college years: A ten-year study. Paper presented at the annual meeting of the Association for Moral Education, Athens, GA.
- Kish, C. K., & Sheehan, J. K. (1997). Portfolios in the classroom: A vehicle for developing reflective thinking. High School Journal, 80 (4), 254-261.
- Kitchener, K. S. (1977-78). Intellectual development in late adolescents and young adults: Reflective judgment and verbal reasoning. Dissertation Abstracts International, 39, 956B.
- Kitchener, K. S. (1983). Cognition, metacognition and epistemic cognition: A three-level model of cognitive processing. Human Development 4, 222-232.
- Kitchener, K. S., & King, P.M. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. Journal of Applied Developmental Psychology, 2, 89-116.
- Kitchener, K. S., King, P. M., Wood, P. K., & Davison, M. L. (1989). Sequentiality and consistency in the development of reflective judgment: A six-year longitudinal study. Journal of Applied Developmental Psychology, 10, 73-95.
- Kitchener, K. S., Lynch, C. L., Fischer, K. W., & Wood, P. K. (1993). Developmental range of reflective judgment: The effect of contextual support and practice on developmental stage. Developmental Psychology, 29 (5), 893-906.

- Kliebard, H.M. (1995). The struggle for the American curriculum 1983-1958 (2nd ed.). New York: Routledge.
- Knefelkamp, L.L. (1974). Developmental instruction: Fostering intellectual and personal growth of college students. (Doctoral dissertation, University of Minnesota, 1974). Dissertation Abstracts International, 38, 7233 A. (University Microfilms No. 78-09, 683).
- Knefelkamp, L. L., & Slepitz, R. A. (1976). A cognitive developmental model of career development: An adaptation of the Perry Scheme. Counseling Psychologist, 6 (3), 53-58.
- Knight, B. (1996). Reflecting on 'reflective practice.' Studies in the Education of Adults, 28 (2), 162-185.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D.A. Goslin (Ed.), Handbook of socialization theory and research (pp.347-480). Chicago: Rand McNally.
- Kolb, D. A. (1981). Experiential learning: Experience as the source of learning and development. Englewood Cliffs, NJ: Prentice Hall.
- Konke, K. (1984, April) A study of relationship of teacher conceptual level with perceptions of teachers in regard to staff development, curriculum development, and instructional improvement. Paper presented at the 68th Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Kozak, A. I. (1996). Epistemological development and adaptation: Reflective judgment and stressful affect and their relationship to affect-tolerance and attributions of pragmatic structure. Dissertation Abstracts International: Section B: The Sciences and Engineering, 57 (1B), 0699.
- Kroll, B. M. (1992). Reflective inquiry in college English class. Liberal Education, 78 (1), 10-14.
- Kubinski, D. J. (1999). Cognitive development as an outcome measure of a remediation program provided to high-risk senior baccalaureate nursing students. Dissertation Abstracts International, 61 (03B), 1324.
- Kuhn, D. (1991). The skills of argument. Cambridge, England: Cambridge University Press.
- Lazere, D. (1987). Critical thinking in college English studies. ERIC Digest. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills. ED284275

- Lerner, R. (1986). Concepts and theories of human development (2nd ed.). New York: Random House.
- Lincoln, Y. S., & Guba, E. (1985). Naturalistic inquiry. Newbury Park, DA: Sage Publications.
- Lindsay, G. A. (1989). Promoting the development of high-risk college students through a deliberate psychological education-based freshman orientation program—a follow-up study. Dissertation Abstracts International, 50, 3156A.
- Lipman, M. (1988). Critical thinking: What can it be? Educational Leadership, 46 (1), 38-43.
- Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. Applied Social Research Methods Series (Vol. 49). Thousand Oaks: Sage Publications.
- Loacker, G., & Mentkowski, M. (1993). Creating a culture where assessment improves learning. In T. W. Barita and Associates (Eds.), Making a difference: Outcomes of a decade of assessment in higher education (pp. 5-24). San Francisco: Jossey-Bass.
- Loevinger, J. (1976). Ego development: Conceptions and theories. San Francisco: Jossey-Bass.
- Loevinger, J., & Wessler, R. (1970). Measuring ego development (vol.1). San Francisco: Jossey-Bass.
- Martin, J. E., Silva, D. G., Newman, J. H., & Thayer, J. F. (1994). An investigation into the structure of epistemological style. Personal Individual Difference, 16 (4), 617-629.
- Matthew, B. & Jessel, J. (1998). Reflective and reflexive practice in initial teacher education: A crucial case study. Teaching in Higher Education, 3 (3), 231-244.
- McAdams, C. R. (1988). Promoting the development of high-risk college students through a deliberate psychological education-based freshman orientation program. Dissertation Abstracts International, 49, 1375A.
- McMinn, M. P. (2001). Preparing the way for student cognitive development. Multicultural Education, 9 (1), 13-15.
- McNeel, S. P. (1994). College teaching and student moral development. In J. Rest & D. Narvaez (Eds.), Moral development in the professions: Psychology and applied ethics (pp.27-49). Hillsdale, NJ: Erlbaum.

- Mead, G. H. (1934). Mind, self, and society. Chicago: University of Chicago Press.
- Mentkowski, M., & Doherty, Z. (1984). Ability that last a lifetime: Outcomes of the Alverno experience. AAHE, 36, 3-14.
- *Mentkowski, M., Roger, G. Doherty, A. Loacker, G., Hart, J. R., Rickards, W., O'Brien, K., Riordan, T., Sharkey, S. Cromwell, L., Diez, M. Bartles, J., & Roth, J. (2000). Learning that lasts: Integrating learning, development, and performance in college and beyond. San Francisco: Jossey-Bass.
- Mentkowski, M. & Strait, M. (1983). A longitudinal study of student change in cognitive development, learning styles, and generic abilities in and outcome-centered liberal arts curriculum (Final Report to the National Institute of Education: Research Report No. 6). Milwaukee, WI: Alverno.
- Mewborn, D. S. (1999). Reflective thinking among preservice elementary mathematics teachers. Journal for Research in Mathematics Education, 30 (3), 316-342.
- Middleton, J. A. (1999). Curricular influences on the motivational beliefs and practice to two middle school mathematics teachers: A follow-up study. Journal of Research in Mathematics Education, 30 (3), 349-358.
- Miller, A. (1981). Conceptual matching models and interactional research in education. Review of Educational Research, 51 (1), 33-84.
- Mills, M., & Satterthwait, D. (2000). The disciplining of pre-service teachers: Reflections on the teaching of reflective teaching. Asia—Pacific Journal of Teacher Education, 28 (1), 29-39.
- Mines, R. A. (1982). Student development assessment techniques. In G. R. Hanson (Ed.), Measuring student development: New direction for student services, no. 20 (pp. 65-91). San Francisco: Jossey-Bass.
- Mines, R. A., King, P. M., Hood, A. B., & Wood, P. K. (1990). Stages of intellectual development and associated critical thinking skills in college students. Journal of College Student Development, 31 (6), 538-547.
- Morgan, B., Morgan, F., Foster, V., & Kolbert, J. (2000). Promoting the moral and conceptual development of law enforcement trainees: A deliberate psychological educational approach. Journal of Moral Education, 29 (2), 203-219.
- Mosher, R. & Sprinthall, N.A. (1971). Deliberate psychological education. Counseling Psychologist, 2, 3-82.

- Moore, W. S. (1983a). The Maryland career course. Stage/style interactions: The Perry Scheme and the Myers-Briggs type indicator (Brief Report. No. 3). VA.
- Moore, W. S. (1983b). The Measure of Intellectual Development: A brief review. Farmville, VA: Center for Applications of Developmental Instruction.
- Moore, W. S. (1986a). Center for Applications of Developmental Instruction. Farmville, VA: Center for Applications of Developmental Instruction.
- Moore, W. S. (1986b). Interpreting rating on the Measure of Intellectual Development. Farmville, VA: Center for Applications of Developmental Instruction.
- Moore, W. S. (1987). The Learning Environment Preferences: Establishing preliminary reliability and validity for an objective measure of Perry scheme of intellectual development. Unpublished doctoral dissertation. University of Maryland.
- Moore, W.S. (1989). The Learning Environment Preferences: Exploring the construct validity of an objective measure of the Perry scheme of intellectual development. Journal of College Student Development, 30, 504-514.
- Neimark, E. (1975). Intellectual development during adolescence. In F.D. Horowitz (Ed.), Review of Child Development Research. Chicago, Ill: University of Chicago Press.
- Norton, J. L. (1997). Locus of control and reflective thinking in preservice teachers. Education, 117 (3), 401-411.
- Oakes, J. (1985). Keeping track: How schools structure inequality. New Haven, conn: Yale University Press.
- Ortiz, A. M. (1995). Enhancing student development in community colleges. Community College Review, 22 (4), 63-70.
- Pancer, S. M., & Hunsberger, B. (2000). Cognitive complexity of expectations and adjustment to university in the first year. Journal of Adolescent Research, 15 (1), 38-58.
- Parker, L. (1991). Improving the success of academically high-risk first-year college students through a differentiated deliberate psychological education. Dissertation Abstracts International, 50, 3156A.
- Pascarella, E.T. (2001). Identifying excellence in undergraduate education: Are we even close? Change. [online]. Available: <http://www.findarticles.com>

- Pascarella, E.T. and Terenzini, P.T. (1991). How college affects students. San Francisco: Jossey-Bass Publishers.
- Pascarella, E.T., Truckenmiller, R., Nora, A., Terenzini, P.T., Edison, M., & Hagedorn, L.S. (1999). Cognitive impacts of intercollegiate athletic participation. Journal of Higher Education, 70 (1), 1-2.
- Perry, W. G. Jr. (1970). Forms of intellectual and ethical development in college years: A scheme. New York: Holt, Rinehart & Winston.
- Perry, W. G., Jr. (1981). Growth in the making of meaning: Youth into adulthood. In A. W. Chickering (Ed.), The Modern American College (pp. 76-116). San Francisco: Jossey-Bass.
- Perry, W. G., Jr. (1997). Cognitive and Ethical Growth: The making of meaning. In P. G. Altbach (Series Ed.) & K. Arnold & I. C. King (Vol. Eds.), Contemporary higher education: International issues for the twenty-first century: Vol. 4. College student development and academic life: Psychological, intellectual, social, and moral issues (pp. 48-88). New York: Garland Publishing, Inc. (Reprinted from The modern American college, pp. 76-116, by A. W. Chickering and Associates, Eds., 1981, San Francisco: Jossey-Bass)
- Perry, W. G., Jr. (1999). Forms of intellectual and ethical development in college years: A scheme. San Francisco: Jossey-Bass Publishers. (Original work published 1970)
- Piaget, J. (1950). The psychology of intelligence (M. Piercy and D. Berlyne, Trans.). London: Routledge and Kegan Paul.
- Piaget, J. (1970). Science of Education and Psychology of the Child. New York: Orion Press (Translated from the French by Derek Coltman).
- Piaget, J. (1975). The development of thought: Equilibration of cognitive structure. New York: The Viking Press.
- Pirttilae, B., Anna M., & Kajanne, A. (2001). The development of implicit epistemologies during early and middle adulthood. Journal of Adult Development, 8 (2), 81-97.
- Placek, J. H., & Smyth, D.M. (1995). Teaching preservice physical education teachers to reflect. Physical Educator, 52 (2), 106-113.
- Pultorak, E. G. (1993). Facilitating reflective thought in novice teachers. Journal of Teacher Education, 44 (4), 288-296.

- Raines, P., & Shadiow, L. (1995). Reflection and teaching: The challenge of thinking beyond the doing. Clearing House, 68 (5), 271-275.
- Reed, J. H., & Kromrey, J. D. (2001). Teaching critical thinking in a community college history course: Empirical evidence from infusing Paul's model. College Student Journal, 35 (2): 201-215.
- Reiman, A. J. (1999a). Educating for teacher growth: The cognitive developmental perspective revisited. North Carolina Association for Supervision and Curriculum Development Journal, Spring.
- Reiman, A. J. (1999b). The evolution of the social roletaking and guided reflection framework in teacher education: Recent theory and quantitative synthesis of research. Teaching and Teacher Education, 15, 597-612.
- Reiman, A. J. (2001, October). Promoting moral, epistemological, and ego growth in teacher professional development programs. Paper presented at Association of Moral Education. Vancouver, Canada.
- Reiman, A.J. (2002, April). A comparison of four-year longitudinal studies of postconventional moral judgment reasoning in teacher education and other selected undergraduate samples. Paper presented at the American Educational Research Association, New Orleans, LA.
- *Reiman, A. J., & Parramore, B. (1994). Evaluation Project Teaching Fellows Program. Unpublished interim report, College of Education and Psychology, North Carolina State University.
- *Reiman, A. J., & Parramore, B. (1996). Evaluation Project Teaching Fellows Program. Unpublished interim report, College of Education and Psychology, North Carolina State University.
- Reiman, A. J., & Peace, S. D. (2002). Promoting teachers' moral reasoning and collaborative inquiry performance: A developmental role-taking and guided inquiry study. Journal of Moral Education, 31 (1), 51-66.
- Reiman, A. J., & Thies-Sprinthall, L. (1993). Promoting the development of mentor teachers: Theory and research programs using guided reflection. Journal of Research and Development in Education, 26 (3), 179-185.
- Reiman, A. J., & Thies-Sprinthall, L. (1998). Mentoring and supervision for teacher development. New York: Longman.
- Rillero, P. (1998). Process skills and content knowledge. Science Activities, 35 (3), 3-5.

- Rest, J. R. (1979). Development in Judging Moral Issues. Menneapolis: University of Minnesota Press.
- Robin, R.W., Fraley, R.C., Roberts, B.W, Trzesniewski, K.H. (2001). A longitudinal study of personality change in young adulthood. Journal of Personality, 69 (4).
- Rogers, G., Mentkowski, M., Haart, J. R., & Minik, K. S. (2001, April). Disentangling related domains of moral, cognitive, and ego development. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Ross, D. D. (1987). Reflective teaching: Meaning and implications for preservice teacher educators. Paper presented at the reflective inquiry conference, Houston, Texas.
- Royce A., Jr. S. (1998). Connecting the academic and social lives of students: The Holy Cross first year program. (Freshman program of the College of Holy Cross in Massachusetts. Change. <http://www.findarticles.com>..
- Sakalys, J. (1984). Effects of an undergraduate research course on cognitive development. Nursing Research, 33 (5), 290-295.
- Sanders, S. E., & Morris, H. (2000). Exposing student teachers' content knowledge: empowerment or debilitation?. Educational Studies (Carfax Publishing), 26 (4), 397- 409.
- Samson, A. W. (2001). Latino college students and reflective judgment. Dissertation Abstracts International: Section A: Humanities and Social Sciences and Engineering, 61 (9A), 3491.
- Shapiro, N. S. (1985, March). Rhetorical maturity and Perry's scheme of cognitive complexity. Paper presented at the 36th Annual Meeting of the Conference on College Composition and Communication, Minneapolis, M.
- Shaw, J. D. (2001). An application of Baxter Magolda's Epistemological Reflection Model to Black and Latino students. Dissertation Abstracts International, 62-01A, 100.
- *Schmidt, J. A. (1983). The intellectual development of traditionally and nontraditionally aged college students: A cross sectional study with longitudinal follow-up. Doctoral dissertation, University of Minnesota.
- *Schmidt, J. A. (1985). Older and wiser? A longitudinal study of the impact of college of intellectual development. Journal of College Student Personnel, 26, 388-394.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. Journal of Educational Psychology, 82, 498-504.

- Schommer, M. (2002). An evolving theoretical framework for an epistemological belief system. In B. K. Hofer and P.R. Pintrich (Eds), Personal epistemology: The psychology of beliefs about knowledge and knowing (103-118). Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc.
- Schon, D. A. (1987). Educating the reflective practitioner. San Francisco: Jossey-Bass.
- Serow, R. C. (1998). Program evaluation handbook. National Society for Experiential Education. MA: Simon & Schuster Custom Publishing.
- Siegler, R. S. & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. American Psychologist, 46 (6), 606-620.
- Slepitz, R. A. (1983). Commitment within the Perry scheme: A question of structural change. Unpublished doctoral dissertation. University of Maryland, College Park.
- Smith, D. G. et al. (1997). Diversity works: The emerging picture of how students benefit. Washington, D.C.:AAC&U.
- Sprinthall, N. A. (1994). Counseling and social role taking: Promoting moral and ego development. In J. Rest & D. Narvaez (Eds.), Moral development in the profession (pp.85-100). Hillsdale, NJ: Lawrence Erlbaum Associate.
- Sprinthall, N. A., & Collins, W. A. (1995). Adolescent psychology: A developmental view. New York: McGraw-Hill.
- Sprinthall, N. A., Reiman, A. J., & Thies-Sprinthall, L. (1996) Teacher professional development. In J. Sikula (Ed.), Second handbook of research on teacher education. (pp. 666-703). New York: Macmillan.
- Sprinthall, N. A., & Scott, J.R. (1989). Promoting psychological development, math achievement, and success attribution of female students through deliberate psychological education. Journal of Counseling Psychology, 36 (4), 440-446.
- Sprinthall, R. C., Sprinthall, N. A., & Oja, S. N. (1998). Educational Psychology: A developmental approach (7th ed.). Boston, MA: McGraw-Hill.
- Sprinthall, N. A., & Thies-Sprinthall, L. (1983). The teacher as an adult learner: A cognitive-developmental view. In G. A. Griffin (Ed.), Staff development: Eighty second yearbook of the National Society for the Study of Education (pp. 13-35). Chicago: University of Chicago Press.

- Stanley, D. (1998). A framework for teacher reflectivity. TESOL Quarterly, 32 (3), 584-591.
- Starnes, B. A. (1999). The foxfire approach to teaching and learning: John Dewey, experiential learning, and the core practices (Report No.EDO-RC-98-6). Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. (ERIC Digest No. ED426826)
- Stewart, D. K. (1994). Reflective teaching in preservice teacher education: An annotated bibliography from the ERIC database. Journal of Teacher Education, 45 (4), 298-303.
- Stonewater, B. B., Stonewater, J. K., & Hadley, T. D. (1986). Intellectual development using the Perry scheme: An exploratory comparison of two assessment instruments. Journal of College Student Personnel, 27 (6), 542-547.
- Stuessy, C. L., & Naizer, G. L. (1996). Reflection and problem solving: Integrating methods of teaching mathematics and science. School Science and Mathematics, 96 (4), 170-178.
- Sweeney, N. M. (1996). The epistemological perspectives and cognitive development of college nurse faculty. Dissertation Abstracts International Section B: The Sciences and Engineering, 57 (3B), 1716.
- Swick, H. M., & Others. (1991, April). Fostering the professional development of medical students. Paper presented at the American Educational Research Association Annual Meeting, Chicago, IL. (ED330283)
- Tatum, B. D., Calhoun, W. R., Brown, S.C., & Ayvazian, A. (2000). Implementation strategies for creating an environment of achievement. Liberal Education, 86 (2), 18-26.
- Taylor, M.B. (1983). The development of the Measure of Epistemological Reflection. Dissertation Abstracts International, 44, 1065A. (University Microfilms No.DA83-18, 441).
- Tharp, M. L., & Lovell, C. (1995, October). Achieving cognitive equity in the mathematics classroom. Paper presented at the 17th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Columbus, OH.
- Thies-Sprinthall, L. (1984). Promoting the developmental growth of supervising teachers: Theory, research programs, and implications. Journal of Teacher Education, 35 (3), 53-60.

- Thorndike, R. M., et al. (1989). Impact of life experiences on cognitive development (Report No. TM014826). U.S.; Washington. Clearinghouse: Tests, Measurement and Evaluation. (ERIC No. ED318768)
- Tillema, H.H. (1997). Promoting conceptual change in learning to teach. Asia-Pacific Journal of Teacher Education, 25 (1), 7-17.
- *Tomlinson-Keasey, C., & Eisert, D. C. (1978). Can doing promote thinking in the college classroom? Journal of College Student Personnel, 19, 99-105.
- Torosyan, R. (2000). Encourage consciousness development in the college classroom through student-centered transformative teaching and learning. Dissertation Abstracts International, 60 (12A): 4347.
- Vygotsky, L. S. (1962). Thought and language. Cambridge, MA: The M.I.T. Press.
- Welfel, E. R. (1982). How students make judgments: Do educational level and academic major make a difference? Journal of College Student Personnel, 23, 490-497.
- Welfel, E. R., & Davison, M. L. (1986, May). The development of reflective judgment during the college years: A 4-year longitudinal study. Journal of College Student Personnel, 209-216.
- White, B.C. (2000). Pre-service teachers' epistemology viewed through perspectives on problematic classroom situations. Journal of Education for Teaching, 26 (3), 279-306.
- Whitehead, A. N. (1997). The aims of education and other essays. In S. M. Cahn (Ed.), Classic and contemporary reading in the philosophy of education (pp. 262-273). New York: McGraw-Hill. (Reprinted from The aims of education and other essays by A. N. Whitehead, 1929. Macmillian Publishing Company)
- Widick, C. (1975). An evaluation of developmental instruction in the university setting. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.
- Widick, C. (1977). The Perry scheme: A foundation for developmental practice. The Counseling Psychologist, 6 (4), 35-38.
- Widick, C. & Simpson, D. (1978). Developmental concepts in college instruction. In C. A. Parker (Ed.), Encouraging development in college students (pp.27-59). Minneapolis: University of Minnesota Press.

- Woodhouse, J. L., & Knapp, C. E. (2000). Place-based curriculum and instruction: Outdoor and environmental education approaches. (Report No. EDO-RC-00-6). Charleton, WV: ERIC Clearinghouse on Rural Education and Small Schools. (ERIC Digest No. ED 448012)
- Woodward, H. (1998). Reflective journals and portfolios: Learning through assessment. Assessment & Evaluation in Higher Education, 23 (4), 415-424.
- Wubbels, Th., & Korthagen, F.A.J. (1990). The effects of a pre-service teacher education program for the preparation of reflective teachers. Journal of Education for Teaching, 16 (1), 29-44.
- Zeichner, K.M. (1983). Alternative paradigms of teacher education. Journal of Teacher Education, 34(3), 3-9.
- Zeichner, K. M., & Liston, D. P. (1987). Teaching student teachers to reflect. Harvard Educational Review, 57 (1), 23-48.
- Zhang, L. F. (1999). A comparison of U.S. and Chinese university students' cognitive development: The cross-cultural applicability of Perry's theory. Journal of Psychology, 133 (4), 425-439.