INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Educating Preservice Teachers to Teach for an Evaluative View of Knowledge and Critical Thinking in Elementary Social Studies

by

Carole Lorraine Ford
B.Ed., University of British Columbia, 1979
M.A., University of British Columbia, 1988

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Social and Natural Sciences

We accept this dissertation as conforming to the required standard

Dr. L. D'Yore, Supervisor (Department of Social and Natural Sciences)

Dr. R. Fowler, Departmental Member (Department of Social and Natural Sciences)

Dr. T. Riecken, Departmental Member (Department of Social and Natural Sciences)

Dr. A. Preece, Outside Member (Department of Communications and Social Foundations)

Dr. I. MacPherson, Outside Member (Department of History)

Dr. I. Wright, External Examiner (Department of Curriculum Studies, University of British Columbia)

© Carole Lorraine Ford, 1998
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopying or other means, without the permission of the author.
Supervisor: Dr. Larry D. Yore

ABSTRACT

This one-semester pretest-posttest case study of 3rd and 5th year female preservice teachers at the University of Victoria in British Columbia (N=8) involved teaching for critical thinking and an evaluative view of knowledge in elementary social studies. Philosophical and psychological perspectives of critical thinking provided an evaluative view of knowledge, intellectual resources, and cognitive tasks for responses to critical challenges about belief and action (Bailin, Case, Coombs, & Daniels, 1993; Facione, 1991; Siegel, 1992). This integrated conception of critical thinking guided instruction, instrument selection, and interpretation of qualitative evidence. Instruction utilized an interactive constructive approach that involved social and pedagogical challenges appropriate to adults students but related to the elementary social studies curriculum.

Multiple measures of critical thinking dispositions, view of knowledge, and argument proficiency revealed moderate dispositional strength toward critical thinking, mixed views of knowledge, mixed argument proficiencies, and small positive gains over the duration of instruction. Pretest-posttest measures included Facione and Facione’s (1992) California Critical Thinking Disposition Inventory (CCTDI), Kuhn’s (1991) interview protocol for view of knowledge and argument proficiency, and a written argument paralleled the posttest interview evidence.
Most changes aligned with the nature of instruction and instrumentation. CCTDI entry-level results were aligned with results reported for samples from other college and university studies whereas exit-level results displayed more positive change than reported in other studies. Posttest views of knowledge were mixed (evaluativists=4, multiplists=1, absolutists=3), largely consistent with the pretest, and exhibited more evaluativism than reported in other studies. Proficiency in argument was somewhat less than found in other studies, but increased slightly, particularly the generation of alternative theories over the duration of the study. Compared with interviews, written arguments revealed stronger rebuttals and somewhat weaker use of evidence. Inconsistencies across qualitative results and formal results aligned with the nature of the instruction, assessment tasks, evaluation criteria, and some problematic aspects of instrumentation. Participants stated that interactive constructivism; justification of ideas against explicit criteria; an early emphasis on developing a rationale for teaching to an evaluative view of knowledge; the use of examples, non-examples, and borderline examples to generate criteria for key ideas; and instructor-student interaction to monitor and adjust instruction to maximize clarity were positive features of instruction. An excessive concept load and inadequate compatible prior learning experiences were identified as impediments to clarity.
Examiners

Dr. L. D. Yore, Supervisor (Department of Social and Natural Sciences)

Dr. R. Fowler, Departmental Member (Department of Social and Natural Sciences)

Dr. T. Riecken, Departmental Member (Department of Social and Natural Sciences)

Dr. A. Preece, Outside Member (Department of Communications and Social Foundations)

Dr. I. MacPherson, Outside Member (Department of History)

Dr. I. Wright, External Examiner (Department of Curriculum Studies, University of British Columbia)
TABLE OF CONTENTS

ABSTRACT

ACKNOWLEDGEMENTS

CHAPTER ONE: PROBLEM FOCUS AND SIGNIFICANCE

Introduction

Problem Focus

Research Question

Significance of the Study

Limitations of the Study

CHAPTER TWO: PERSPECTIVES OF CRITICAL THINKING AND VIEWS OF KNOWLEDGE: TOWARD AN INTEGRATED FRAMEWORK

The Integrated Critical Thinking Framework—An Overview

Views of Knowledge

Epistemology—What is It?

Epistemology in the Social Sciences

View of Knowledge, Society, and Education: Three Profiles

Epistemological Journeys: Selected Research
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measures</td>
<td>103</td>
</tr>
<tr>
<td>The California Critical Thinking Disposition Inventory</td>
<td>104</td>
</tr>
<tr>
<td>Kuhn’s Interview</td>
<td>110</td>
</tr>
<tr>
<td>Other Measures</td>
<td>113</td>
</tr>
<tr>
<td>Procedures</td>
<td>117</td>
</tr>
<tr>
<td>Intervention</td>
<td>118</td>
</tr>
<tr>
<td>Analysis of Data</td>
<td>130</td>
</tr>
<tr>
<td>Dispositions for Critical Thinking</td>
<td>131</td>
</tr>
<tr>
<td>View of Knowledge</td>
<td>132</td>
</tr>
<tr>
<td>Proficiency in Argument</td>
<td>134</td>
</tr>
<tr>
<td>Perspectives of Participants</td>
<td>142</td>
</tr>
<tr>
<td>CHAPTER FOUR: ANALYSIS OF DATA AND RESULTS</td>
<td>144</td>
</tr>
<tr>
<td>The Research Focus</td>
<td>144</td>
</tr>
<tr>
<td>Research Results</td>
<td>145</td>
</tr>
<tr>
<td>Entry-Level Description: Dispositions for Critical Thinking,</td>
<td>146</td>
</tr>
<tr>
<td>View of Knowledge, and Proficiency in Argument</td>
<td>146</td>
</tr>
<tr>
<td>Dispositions for Critical Thinking</td>
<td>146</td>
</tr>
<tr>
<td>View of Knowledge</td>
<td>148</td>
</tr>
<tr>
<td>Proficiency in Argument</td>
<td>150</td>
</tr>
<tr>
<td>Exit-Level Description: Dispositions for Critical Thinking,</td>
<td>159</td>
</tr>
<tr>
<td>View of Knowledge, and Proficiency in Argument</td>
<td>159</td>
</tr>
<tr>
<td>Dispositions for Critical Thinking</td>
<td>159</td>
</tr>
<tr>
<td>View of Knowledge</td>
<td>161</td>
</tr>
<tr>
<td>Proficiency in Argument</td>
<td>164</td>
</tr>
<tr>
<td>Pretest-Posttest Differences</td>
<td>172</td>
</tr>
<tr>
<td>Dispositions for Critical Thinking</td>
<td>172</td>
</tr>
<tr>
<td>View of Knowledge</td>
<td>176</td>
</tr>
<tr>
<td>Proficiency in Argument</td>
<td>177</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Written Argument Evidence: Consistency With Interview</td>
<td>181</td>
</tr>
<tr>
<td>Written Argument: Exit-Level Description of Proficiency in Argument and View of Knowledge</td>
<td>183</td>
</tr>
<tr>
<td>Comparison of Written Argument and Exit-Level Interview Data</td>
<td>184</td>
</tr>
<tr>
<td>Qualitative Evidence: Perspectives of the Effectiveness of Intervention</td>
<td>189</td>
</tr>
<tr>
<td>The Preparatory Phase</td>
<td>191</td>
</tr>
<tr>
<td>The Application Phase</td>
<td>204</td>
</tr>
<tr>
<td>The Consolidation Phase</td>
<td>220</td>
</tr>
<tr>
<td>CHAPTER FIVE: SUMMARY AND CONCLUSIONS</td>
<td>257</td>
</tr>
<tr>
<td>Methodology</td>
<td>257</td>
</tr>
<tr>
<td>Results</td>
<td>260</td>
</tr>
<tr>
<td>What Were Entry-Level and Exit-Level Dispositions for Critical Thinking, View of Knowledge, and Proficiency in Argument?</td>
<td>260</td>
</tr>
<tr>
<td>What Pretest-Posttest Change was Evident in Students’ Dispositions for Critical Thinking, View of Knowledge, and Proficiency in Argument?</td>
<td>262</td>
</tr>
<tr>
<td>Was the Written Argument Evidence Consistent with the Interview Evidence?</td>
<td>264</td>
</tr>
<tr>
<td>What did Participants’ Informal Responses Suggest About the Effectiveness of Intervention?</td>
<td>264</td>
</tr>
<tr>
<td>Discussion</td>
<td>265</td>
</tr>
<tr>
<td>Results Relative to Ideals</td>
<td>266</td>
</tr>
<tr>
<td>Resistance to Change</td>
<td>267</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1: University of Victoria Bachelor of Education Program (Elementary Curriculum)-------------------------------------------------------------- 102

Table 2: Sequence of Instructional Events--------------------------------------------------------------- 121

Table 3: CCTDI—Entry-Level Dispositions for Critical Thinking: Cross-Student and Cross-Subscale Scores----------------------------------------- 147

Table 4: Entry-Level Cross-Topic, Cross-Student Views of Knowledge ----- 149

Table 5: Entry-Level Argumentation: Cross-Dimension, Cross-Topic, Cross-Student Level of Success----------------------------------------------------- 151

Table 6: Entry-Level Evaluation of Evidence: Cross-Student, Cross-Category, and Cross-Topic Level of Success------------------------------ 155

Table 7: Entry-Level Epistemological Reasoning: Cross-Student and Cross-Topic Level of Success-------------------------------------------------- 158

Table 8: CCTDI—Exit-Level Dispositions for Critical Thinking: Cross-Student and Cross-Subscale Scores----------------------------------------- 160

Table 9: Exit-Level Cross-Student, Cross-Topic Views of Knowledge------ 162

Table 10: Exit-Level Argumentation: Cross-Student, Cross-Category, Cross-Topic Level of Success-------------------------------------------------------- 165

Table 11: Exit-Level Evaluation of Evidence: Cross-Student, Cross-Category, and Cross-Topic Level of Success----------------------------- 169

Table 12: Exit-Level Epistemological Reasoning: Cross-Student and Cross-Topic Level of Success-------------------------------------------------- 171

Table 13: CCTDI—Pretest-Posttest Difference---------------------------------------------------------- 173
Table 14: Cross-Sample Pretest-Posttest Comparison of CCTDI Subscale Mean Scores
Page 175

Table 15: Cross-Student, Cross-Topic Pretest-Posttest Difference in View of Knowledge Scores
Page 176

Table 16: Cross-Student, Cross-Dimension, Cross-Topic Pretest-Posttest Difference for Argumentation
Page 178

Table 17: Cross-Student, Cross-Dimension, Cross-Topic Pretest-Posttest Difference for Evaluation of Evidence
Page 179

Table 18: Pretest-Posttest Difference in Epistemological Reasoning
Page 180

Table 19: Exit-Level Cross-Student, Cross-Category, Cross-Dimension Written Argument Data
Page 184

Table 20: University of Victoria Students: Comparison of Exit-Level Interview and Written Argument Data for View of Knowledge and Proficiency in Argument
Page 185

Table 21: The Difference of Written Argument Scores Relative to Interview Scores
Page 186

Table 22: Cross-Student, Cross-Evidence Comparison of View of Knowledge
Page 247

Table 23: Cross-Evidence Comparison for Proficiency in Argument
Page 249
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Integrated critical thinking framework</td>
<td>18</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Critically thoughtful response in an integrative framework</td>
<td>74</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Critical thinking—The path makes a difference</td>
<td>93</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Instructional framework for critical thinking</td>
<td>114</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

Special thanks to my committee members:

Larry Yore, my supervisor, for his enduring mentorship and scholarship;

Alison Preece, for her affirming support and guidance in illuminating the story behind the numbers;

Bob Fowler, for his encouragement and ready responses to my queries;

Ted Riecken, for his warm integrity, keen insights, and encouragement to follow my bliss;

Ian MacPherson, for opening windows to the past that helped me better understand the present;

My external examiner, Ian Wright, for sharing his epistemological journey and his zest for critical thinking;

The eight preservice teachers whose critically thoughtful perspectives about teaching for critical thinking in elementary social studies made this study and what follows possible;

The community of critical thinking educators who, through the Internet, willingly shared their ideas toward the improvement of education for critical thinking;

And in particular, my husband, for his continuing love, support, and inspiration throughout this odyssey that we shared.
CHAPTER ONE

PROBLEM FOCUS AND SIGNIFICANCE

Introduction

The educated global citizen needs to be knowledgeable about the grounds for belief and to be caring enough to strive for adequate grounds for beliefs when the outcomes of those beliefs matter. Education for this ideal requires commitment to teaching for critical thinking. This vision guided the present study that was designed to help preservice elementary teachers acquire the knowledge, proficiencies, and habits of mind to teach for critical thinking and an evaluative view of knowledge in elementary social studies. Implementation was evaluated by utilizing pretest-posttest evidence of (1) dispositions for critical thinking using the California Critical Thinking Disposition Inventory (CCTDI; Facione & Facione, 1992) and (2) views of knowledge and argument proficiencies using Kuhn’s (1991) interview protocol. Qualitative artifacts provided additional evidence of participants’ perspectives and proficiencies. These artifacts included a written argument that paralleled the interview protocol, instructional plan assignments, journals, in-class responses, two course evaluations, and the instructors’ journal and instructional plans.

The significance of this study is found in today’s global community. Decisions once limited in impact now have potential global implications.
Technological and demographic changes are blurring boundaries that have separated societies. Social, economic, and environmental judgments that might affect others are reasons for shared concern and responsibility. The vision of a global community is becoming a reality that many citizens who share responsibility for decisions are not adequately prepared.

The global community is an interconnected ecosystem involving diverse groups of people. Diversity is reflected in contexts, experiences, needs, and beliefs—differences not well-served by any single absolute world view or the belief that multiple views are equally acceptable. Risks are too great to leave judgments to either absolutists who are close-minded to different views of reality or to multiplists whose acceptance of all positions is not helpful when a judgment with potential for profound social impact is required. Judgments require an evaluativist view of knowledge, open to multiple perspectives and yet cognizant that some positions are better supported than others. Judgments must consider the common good. Ill-informed decisions may have implications for the survival of humanity and for the planet Earth. Isolationism is no longer a viable option. Social responsibility now extends to family, community, country, and the world (Parker & Jarolimek, 1984).

Citizens need to be well-informed, socially responsible, open to the views of others, and committed to sustainable judgments that can endure evaluation
against competing arguments. Sustainable judgments go beyond self-serving sophist thinking that strives to win; rather, they strive for alignment with the best reasons or evidence available, remaining open to change if more compelling evidence or reason becomes available. Yet research indicates that citizens may not be adequately prepared for this responsibility. Commitment to and proficiency for sustainable judgments do not evolve naturally (Brabeck, 1980; Kuhn, 1991; Perry, 1968; Woods, 1997). Level of education appears to be an important but insufficient influence. Undergraduates, preservice teachers, practicing teachers, graduate students, and postsecondary faculty have displayed inadequate commitment to and resources for sustainable judgment (Brabeck, 1980; Court & Francis, 1993; Facione, Facione, & Giancarlo, 1992; Goodlad, 1990; King & Kitchener, 1994; Kuhn, 1991; Paul, Elder, & Bartell, 1997; Perry, 1968; Unks, 1985; Wright, 1995b).

The assumption that preservice teachers acquire resources to teach for critical thinking through their teacher education programs (Sears & Parsons, 1991) does not correspond with evidence of incompatible instruction in their programs and faculty who do not display such capacity (Goodlad, Soder, & Sirotnik, 1990; Paul et al., 1997). Yet the only practical arena—the crucial context to facilitate the qualities necessary for global citizenship—is the public education institutions. By definition, public schools do not have a mandate to promote the beliefs of one
group, even though a contextual veneer of a “common countenance” (Tompkins, 1984, p. 2) has supported such curricula in the past (Department of Education, British Columbia, 1934; 1936). However, public schools do have a moral responsibility to facilitate rational or critical thinking. Rationality, arguably the most justifiable purpose of a liberal education in a public school system, addresses citizens’ moral rights to education in a critical manner and provides a means by which citizens’ moral rights may be preserved (Siegel, 1980; Strike, 1982).

The nature of thinking evident in British Columbia curriculum documents for over 60 years has not reflected rationalism. For example, in 1925 “straight” thinking was desirable of a compliant citizenry (Putnam & Weir, 1925, p. 45) and in 1939 “careful thinking” was important to national interests in the face of the perceived communist threat (King, 1939). In 1941, “critical, reflective thinking” (Department of Education, British Columbia, 1941, p. 28) adhered to one objective interpretation of reality in which history was revealed (Department of Education, British Columbia, 1934; 1936). This absolutist perspective was closed to other potential perspectives of reality and incompatible with the ideal of rationality. The intent of critical thinking was not explicated in 1941 or clarified in the 1983 elementary social studies curriculum (British Columbia Ministry of Education, 1983), even though critical thinking and problem solving were the sole vehicles for realizing the four main curriculum goals (Ford, 1988). It is not clear
whether the status of the more recent curriculum reflects a hidden curriculum intended to promote a fixed view of knowledge with one correct answer (Ginsburg & Clift, 1990) or whether it reflects decision makers who lacked relevant educational experiences.

Public school curriculum mission statements have increased in alignment with rationalism in their depiction of the ideal educated citizen—a knowledgeable, adaptable, critically thoughtful, socially responsible decision maker who participates through decisions and actions to further and improve upon the ideals of a democratic society (British Columbia Ministry of Education, 1990; National Council for the Social Studies, 1994). This image of the educated citizen provides a goal to strive for, but alone, is insufficient to help teachers who wish to educate for this ideal. Teachers need educational experiences and support materials to help them understand the centrality of critical thinking to this image, what critical thinking involves, and how to teach for critical thinking. British Columbia assessment evidence suggests that pupils’ critical thinking has been inadequate (Bognar, Cassidy, & Clarke, 1996; Cassidy & Bognar, 1991). These results are better understood in light of the nature of the 1983 British Columbia elementary social studies curriculum and the support materials that were developed. Neither curriculum nor support materials explicated the intent of critical thinking and problem solving (Ford, 1988). Curriculum developers had ventured into "virgin
territory" (Ford, 1988, p. 98) and did not recognize the limitations of their representation of critical thinking. As one curriculum worker recalled, "Never did it become significant enough for us to have requested expert advice....[We] were happy enough with the way it was" (Ford, 1988, p. 94). The curriculum developers did not have the responsibility to operationalize critical thinking. An obligation that did not encompass reflection about the implementation and evaluation of the unfamiliar and complex ideas associated with critical thinking likely contributed to limited evidence of critical thinking in curriculum, support materials, and in assessment results.

Change that would facilitate the ideals of critical thinking has been elusive. Sheelian (1986) claimed that "social, political, economic, technological, and educational conditions need to merge in order for a fundamental change to occur" (p. 52). Stimulus for change also exists when the consequences of not changing exceed the costs that change would incur. But by then, it could be too late to make a difference. Kurfiss (1988) claimed that absolutism and multiplication are key misconceptions and that educators have a responsibility to move students towards evaluativism. Some educators such as those involved with The International Conference on Critical Thinking in Sonoma, CA and The Critical Thinking Cooperative, Vancouver, B. C. are dedicated to this ideal. But to suggest that prospects are optimistic is likely unwarranted, given the dismal
history of curriculum change generally and more specifically in social studies (Cremin, 1961; Lybarger, 1991; Sheehan, Wilson, & Jones, 1986). “Unwarranted optimism,” Fullan’s (1998, p. 10) definition for hope, seems more appropriate. Inspiration is provided by Havel, the Czech Republic president, who stated, “It is hope, above all, that gives us the strength to live and to continually try new things, even in conditions that seem hopeless” (cited in Fullan, 1998, p. 10).

To abandon hope is morally untenable—an acceptance of a citizenry largely unprepared for judgments with serious potential social impact. This moral impetus guided selection of the context for this study, an elementary social studies course in a preservice teacher education program. Teachers completing this program have potential to make a difference to their students. Teachers also contribute to society personally as citizens, curriculum workers, parents, and teachers of teachers, influencing social, political, and institutional contexts directly and indirectly through their decisions and actions. Social studies teachers, in particular, are responsible for citizenship education. Moreover, the social in social studies represents ideas, events, and issues of social consequence. Decisions that have potential for social impact require the moral accountability of an evaluative perspective—open to multiple views of reality and to the relative merit of both beliefs and actions. Social studies provides a crucial context for
education toward critical thinking and an evaluative view of knowledge by nature of its social fabric and its citizenship mandate.

**Problem Focus**

The global community has need for critically thoughtful citizens. Yet research indicates that teachers' prior education experiences have not adequately prepared them to facilitate this desired goal. Significant attributes of critical thinking, as they relate to the nature of social studies, need to become part of the teacher education program in a manner that is meaningful, practical, and provocative for preservice teachers. These goals guided the present study, the purpose of which was to build a pedagogical plan that was responsive to a class of preservice elementary social studies teachers and that would facilitate commitment to and understanding of teaching for critical thinking and an evaluative view of knowledge in elementary social studies.

This research problem involved (1) the articulation of a critical thinking model, (2) the development of an instructional framework for critical thinking as it applies to teaching elementary social studies, and (3) the implementation and evaluation of instructional strategies to promote teaching for critical thinking and an evaluative view of knowledge in the context of social studies education. A synthesis of the literature to build a model, analysis of curriculum and reform
documents to establish a framework, a series of pilot studies, and analysis of related literature contributed to decisions about these sub-problems.

Cognitive tasks of self-regulation, deliberation, judgment, and justification were infused with philosophical dimensions of critical thinking as they relate to elementary social studies. The assumption was that the current education focus on cognitive processes (Ford, 1988; Ford, Yore, & Anthony, 1997) is inadequate for educators who need to understand what it is that determines the quality of thinking. Without knowledge, proficiency, and commitment necessary to sustainable judgments educators arguably lack sufficient resources to plan instruction for critical thinking and to evaluate its effectiveness.

**Research Question**

Insights about effective instruction for critical thinking and an evaluative view of knowledge in elementary social studies were facilitated by evaluating the pedagogical framework, instruction strategies, and classroom activities in light of the nature of individual students and their responses. The instructional effectiveness was further evaluated in terms of students’ critical thinking dispositions, how they viewed knowledge, their argument proficiencies, and their responses to the instruction. The researcher-instructor also provided a detailed account of instructional decisions and accompanying rationales. In addition to providing a case study for external evaluation and a comparative base for future
curricular decisions, it was hoped that this exploratory research would point to some promising hypotheses for further research.

The following research question focused the inquiry: What were particular student teachers' critical thinking dispositions, epistemic views, and argument proficiencies, and what were their responses to the instructor's pedagogical decisions made to enhance accessibility, practicality, and commitment to teaching for critical thinking and an evaluative view of knowledge in elementary social studies? More specifically:

1. What were the entry-level critical thinking dispositions, views of knowledge, and argument proficiencies of the preservice elementary social studies teachers sampled?

2. What were the exit-level critical thinking dispositions, views of knowledge, and argument proficiencies of the preservice elementary social studies teachers and were there changes in these attributes over the duration of the 13-week advanced social studies curriculum and instruction course?

3. Did the written argument confirm the interview protocol evidence?

4. What did qualitative evidence of participants' perspectives (instructor's journal and instructional plans, students' journals, instructional plan assignments, 1 in-class survey, in-class activities, and 2 course evaluations) suggest about the effectiveness of instruction?
It was anticipated that results might suggest associations among some of the following variables: preservice elementary teachers' education background, view of knowledge, critical thinking dispositions, argument proficiencies, and responses to instruction for critical thinking and an evaluative view of knowledge in elementary social studies.

Significance of the Study

The value of this study is in its potential to generate hypotheses for further investigation. The evidence may offer insights into teaching for critical thinking and an evaluative view of knowledge in elementary social studies. Some people unjustifiably assume that the poor state of critical thinking in public schools is related to the inadequate application of teachers' critical thinking. However, in this study it is assumed that the inadequacies in critical thinking reflect the nature of teachers' prior educational experiences. Until elementary social studies teachers are adequately prepared with the resources necessary to fulfill curriculum mission statements, attributing responsibility to them for less than satisfactory implementation is unwarranted. Education researchers have a responsibility to seek answers to enhance critical thinking in teachers, teachers of teachers, curriculum workers, and researchers if public schools' moral mandate to facilitate the goal of the educated global citizen is to be realized.
A detailed description of critical thinking and an instructional framework that attempts to relate recent curricular foci on cognitive tasks (Ford, 1988) with philosophical dimensions of critical thinking provided a basis for external comparison and evaluation. A blending of qualitative and quantitative methods was employed to strengthen conclusions and provide information not currently apparent in the literature (Armento, 1991; Fraenkel & Wallen, 1991). A variety of formal and informal measures taken over time, a description of participants’ ongoing evaluative responses to instruction, and an explication of the pedagogical decisions related to classroom instruction were offered to facilitate future decisions about the education of preservice teachers for critical thinking and an evaluative view of knowledge in elementary social studies. The potential moral benefits to helping students develop the resources for well-grounded judgments justify the cognitive demands imposed by the unfamiliar educational experiences (Hammack, 1997; Siegel, 1980). Additionally, the natural classroom context and educational activities utilized did not ethically compromise normal educational practice that is responsive to students’ needs and institutional expectations (Hammack, 1997), limiting conflict from the dual instructor-researcher role.

The instruments used to determine participants’ critical thinking dispositions, view of knowledge, and argument proficiencies have been verified individually, but do not appear to have been triangulated with each other and with
other informal measures as they were in this study. Furthermore, Kuhn’s interview protocol and criteria for view of knowledge and argument proficiencies have not apparently been used in a pretest-posttest design. The potential limitations of a single response modality, interviewer influence, and response time were better apprehended by use of more than one measure, as was topic-specific test effect. The research design made it possible to examine evidence of preservice teachers’ views of knowledge and proficiency in a written argument completed near the end of the course without the response demands imposed by an interview. The written argument focused on a topic distinct from the interview protocol, addressing potential for topic-specific test effect. A pretest-posttest design provided difference scores for participants’ critical thinking dispositions, views of knowledge, and argument proficiencies over the 13-week course duration.

**Limitations of the Study**

Prospects for generalizable conclusions were seriously limited by a pretest-posttest case study of a small intact convenience sample (N=8). Subjects were volunteers from a class of 10 female preservice elementary social studies teachers, unrepresentative of general education students by selecting to take the advanced curriculum and instruction course in elementary social studies and by gender. The sample was further distinguished by four students who had already taken a basic
social studies methods course and 1 other student who had taken basic teaching methodology courses 15 years earlier but reported remembering little about social studies. The remaining 3 participants had not taken the basic requisite social studies methods course.

Potential confounds for internal validity were found in the history of the study where other extraneous factors might contribute to change over time; maturation of subjects unrelated to the intervention; pretest influence on subjects’ posttest responses; and instrument decay if the researcher’s administration of instruments, particularly the interviews, varied from pretest to posttest (Campbell & Stanley, 1963). Relevant to this latter confound was the potential for bias and influence through the investigator’s dual role as instructor and researcher (Borg & Gall, 1983).

The study focused on the participants’ responses to the instruction and related pedagogical decisions within the context of a one semester course of 39 hours over 13 weeks. Some research has found that intervention of such brief duration has been inadequate to effect much change (Collins, 1992; Hatcher, 1995; Hatcher & Price, 1998; King & Kitchener, 1994; Perry, 1968; Woods, 1997). This study did not explore the implementation of critical thinking in public school classrooms.
The degree of implementation expected was also limited by the complexity of critical thinking that requires understanding the many associated concepts as they apply across varied contexts (Hare, 1993; Paul, 1996; Scriven 1988). This study of critical thinking was confined to exploration of critical thinking dispositions, views of knowledge, argument proficiencies, and their correspondence with instruction.
CHAPTER TWO

PERSPECTIVES OF CRITICAL THINKING AND VIEWS OF KNOWLEDGE: TOWARD AN INTEGRATED FRAMEWORK

The literature is replete with divergent accounts of critical thinking—conceptual and operational. Some disparities are so significant that a single synthesis is problematic. Facione (1991) attempted to facilitate a consensus about critical thinking with a group of forty-six scholars of critical thinking. Although there was some agreement, consensus among the participants was lacking. Some did not subscribe to the final image (R. H. Ennis, personal communication, August 4, 1995; S. P. Norris, personal communication, August 12, 1995) and one participant noted that the consensus that might have existed in 1991 did not necessarily reflect a current consensus (M. Weinstein, personal communication, January 5, 1996).

Nonetheless, it is necessary to illustrate the dynamic diversity of views about critical thinking to understand the complexity of this study, its evidence, its warrants, and its claims. To this end, a purposeful selection was made of critical thinking perspectives. The critical thinking framework foundational to this study addresses several prevailing issues within the critical thinking movement: views of knowledge, integration of psychological and philosophical dimensions of critical thinking, generalizability, gender bias, and instruction and assessment.
An advanced organizer for this critical thinking framework (derived largely from Bailin, Case, Coombs, & Daniels, 1993; Facione, 1991; Siegel, 1992) that depicts relationships among dimensions is provided in Figure 1 to facilitate comprehension of issues and judgments that pertain to this framework. Figure 1 illustrates how critical thinking—thinking directed toward sustainable judgments about what to believe or about what to do—is embedded in responses to critical challenges that stimulate integration of cognitive tasks and intellectual resources, as orchestrated by metacognitive self-regulation. Each of these constituent parts will be elaborated later.

**The Integrated Critical Thinking Framework—An Overview**

The critical thinking framework (Figure 1) attempts to integrate perspectives from cognitive psychology and philosophy. The reasons for critical thinking are found in a critical challenge, a problematic situation that invites a decision about what to believe or what to do, a situation in which the answer is not readily available and in which the overall judgment may be of importance to someone. Thinking through a critical judgment may involve several recursive cognitive tasks for purposes of self-regulation, deliberation, judgment, and justification. The quality of these responses will depend, in part, upon metacognition, represented here as (1) the knowledge necessary to distinguish the particulars of critical thinking as they apply to the challenge and (2) self-regulatory competency in planning, on-
going evaluation, and adjusting responses to improve the sustainability of judgments. Consistent critical thinking responses would reflect critical thinking habits of mind that characterize the ideal critical thinker, habits nurtured by an
emotional disposition toward critical thinking and commitment toward sustainable judgments. Integral to such commitment is knowledge of the epistemological position underlying critical thinking—understanding what a conception of critical thinking implies for how knowledge is realized and what is evidence of knowing. Intellectual resources important to critical thinking responses include background knowledge that is conceptually related to the problematic situation and relevant context, as well as knowledge of critical thinking and its many associated concepts, criteria and standards, habits of mind, and thinking strategies appropriate to the judgments involved. An adequate justification or an external evaluative judgment would appeal to relevant intellectual resources to explain how well the available evidence or reasons sustain the judgment.

A description of some key distinctions among views of knowledge and critical thinking perspectives will provide a context for considering related research, curricula, and pedagogy for teaching elementary social studies. This review provides background information important to understanding the issues, the complexities, and the critical thinking instructional framework that are pertinent to this study.

**Views of Knowledge**

Epistemological positions or views of knowledge have substantial implications for curriculum, instruction, and society. Teachers need to
comprehend deeply the consequences that the view of knowledge prevalent in curricula and pedagogy have for both their students and for society. Educators have a responsibility to empower students with the intellectual resources necessary for sustainable judgments about what to believe or do, decisions that may have potential to impact the survival of our world (e.g., uninformed decisions about environmental pollutants could contribute to worldwide contamination of food and air). Toward this goal, the nature of epistemology and its relationship to disciplines relevant to social studies, key research, and relevant connections with curricula are represented. Furthermore, provocative implications of view of knowledge for society and education are discussed.

**Epistemology—What is It?**

Epistemology has been described as the study of the nature of knowledge—what it is, why it merits pursuit, how it is constructed, and by whom (Tiles & Tiles, 1993). How one views knowledge is integral to how one views critical thinking. Critical thinking involves, in part, bringing one’s intellectual resources to bear on decisions about what counts as knowledge. Perceptions of what counts as knowledge will influence views of critical thinking. Thus epistemological consensus is difficult to achieve. Coffey (1917) proposed that epistemology involves “the philosophical investigation of human knowledge itself, from the standpoint of the certitude, validity, or truth-value of this knowledge” (p. 1).
Knowledge and epistemic justification—grounds for belief, remain central epistemological concerns (Chisholm, 1977; Pollock, 1986). But recognition of the fallibility of human inquiry corresponds with epistemological perspectives that are moving away from notions of certitude (confidence in the possibility of a one-to-one correspondence between reality and objective representation) toward acceptance of the fallibility of claims of knowledge (Rescher, 1980). In other words, judgments currently sustained by evidence remain open to new evidence that could show the belief to be unwarranted. This does not rule out the potential for true belief; rather, it recognizes that we might never know how near we are to the truth (Hamm, 1989).

**Epistemology in the Social Sciences**

Epistemological dissension flourishes in the halls of academia. The objectivism of the late 19th and early 20th centuries was seriously challenged by scholars who addressed the limitations of the scientific method and the need for historians to examine, clarify, and broaden their frames of reference (Beard, 1934; Becker, 1932). More recently, Novick (1988) suggested that “work in the history, philosophy, and sociology of science has made...[historians] increasingly aware of the influence of external and social factors in theory choice, in deciding what is ‘a fact,’ and even in defining ‘rationality’” (p. 10). New perspectives have corresponded with history’s differentiation into the social sciences and growing
specialized areas of study that reflect increasing variation of purposes, questions, assumptions, methods, conceptual foci, and theories (Furet, 1983). For example, variation is apparent in the questions asked by and goals of social scientists: anthropologists may strive for description and understanding of cultures, historians may seek a best explanation to account for the past, economists may pursue evidence to model events and support prediction, and philosophers may analyze concepts including epistemological views of objectivity and moral concerns about justice. With specialization has come fragmentation and increasing difficulty communicating within and among sub-disciplines, complicating accessibility of ideas and limiting peer evaluation to smaller communities (Monkkonen, 1986). These communication constraints pose impediments to fruitful epistemological conversation. Yet, specialization has benefits in new questions, new evidence, and new methods that have led to information that challenge incomplete or misleading evidence and claims (Davis, 1975, 1987; Keeley, 1984; Sutherland, 1976; Valverde, 1991). The new found empowerment has informed the quality of research and has stimulated more complete representations of the past and honest explicitness about their strengths and limitations (Davis, 1975, 1987).

Traditional absolutists believe in a one-to-one correspondence between historical representations of reality and reality. This assumption of a reality that
can be represented objectively is being challenged as historians face new evidence that contradicts past accounts of 'historical truth.' Post-modernists have rebelled against modernist's imposition of an ethnocentric absolute authority who discounts contrary views of reality as being incorrect, if indeed, it recognizes alternative perspectives at all. Postmodernists' responses to such exclusionary abuse range between the extremes of deconstructivism and constructivism. Deconstructivists have abandoned modernist's notions of certitude that could not be sustained in the presence of contrary evidence and they deny objective representation of reality (Tiles & Tiles, 1993). Radical deconstructivism supplants the extreme of absolutism with the excess of relativism. Relativism embraces a multiplist view of knowledge that recognizes the particularity of experience and perceptions. Reality is inaccessible and truth is contextual, a matter of perspective. If truth is subjective, it follows that one view is as acceptable as the other. As such, universal generalizations are deemed unwarranted. Relativists, by accepting the equal status of positions, endorse counter views, raising problems with the logic of their position. Relativism's denial of general claims is itself an exemplar of a general claim, putting the position on shaky epistemological grounds (Siegel, 1996). In spite of such incongruities, relativists' concerns about the inadequacies of modernity are important to the advancement of scholarship.
However, the abandonment of the pursuit of truth is an overreaction to an absolutist view of knowledge that has not served us well.

The logical implications of an extreme relativist position for shared scholarship and the blinders of absolutism suggest that a more fruitful path would be to retain or revise what has merit in these epistemological positions toward a more sustainable evaluativist theory of knowledge. An evaluativist position acknowledges reality and yet recognizes the limitations of observation and report of reality; discerns multiple ways of experiencing the world while accepting that some views are more supportable than others; and grasps the implications of judgments for which evidence or reasons cannot be sustained publicly over time. Source credibility is important to the evaluativist. Unlike the absolutist, however, the evaluativist goes beyond appeal to authority to evaluate the quality of argument—seeking the ‘best fit’ conclusion based on the most sustainable evidence and reasons available (King & Kitchener, 1994). This evaluative view of knowledge is a defensible position for a society characterized by multiple perspectives and the requirement that its citizens have the capacity and commitment for sound judgments. While scholars continue to debate these epistemological issues, social studies curricula portray mixed messages, offering rhetoric but inadequate evidence of knowledge or commitment to help teachers
and students toward an evaluative view of knowledge and the critical thinking resources to attain such knowledge (Ford, 1988; Ford et al., 1997).

**View of Knowledge, Society, and Education: Three Profiles**

It could be argued that the predominant view of knowledge in a society corresponds with who has choice, the vision of the ideal educated citizen, and pedagogy toward this educated citizen. The following contextualized profiles of absolutism, multiplicism, and evaluativism are this researchers' attempt to illustrate this congruency and provide a background context to aid comprehension of associated research. These contextualized profiles have been inferred from epistemological research (King & Kitchener, 1994; Kuhn, 1991; Perry, 1968; Woods, 1997), critical thinking literature (Bailin, Case, Coombs, & Daniels, 1993; Siegel, 1992, 1980; Paul, 1992) and examination of curriculum-related documents (e.g., Department of Education, 1934, 1936, 1941; Ford, 1988; Ford et al., 1997; Putnam & Weir, 1925). It is also intended that the implications of alternative epistemological profiles provide a compelling case for an evaluative view of knowledge.

An absolutist does not perceive opposing perspectives of reality and thus has little need to justify conclusions. Historical truth is revealed to historians; it is not constructed. Beliefs are grounded in their congruence with experts’ perspectives or based on personal experience. The quality of thinking important
to the ideal citizen in an absolutist context is “straight” thinking aligned with experts (Putnam & Weir, 1925, p. 45). This view is consistent with a society characterized by shared beliefs or with a benevolent or totalitarian government where individual choice is not an option. Absolutism is not compatible with a society comprised of citizens with multiple perspectives and responsibilities for decisions that would impact others. Curricula represents the truth to be delivered didactically by the teacher or other accepted authorities, or found in the textbook. Little attention is given to evidence and quality of argument; authority is sufficient. Students’ accurate regurgitation of experts’ conclusions is evidence of ‘knowledge.’ Successful students might attribute to an absolutist educational experience a facility in recounting facts and details, perseverance in memorization, and adherence to rules. Students would not acquire the resources for judging the sustainability of arguments from experiences that did not facilitate the commitment, understandings and proficiencies to evaluate the quality of available evidence relative to the conclusion.

A multiplist views knowledge as subjective constructions of reality, truth being unique to each individual according to experiences and contexts that filter perceptions of reality. Multiple reports of reality are equally acceptable and expertise is not given preference. The ideal citizen in a multiplist context does not criticize competing views as each person is entitled to his or her own conclusions.
Multiplism is compatible with a context in which the average citizen is not responsible for decisions that have potential for serious implications for others. Multiplism becomes problematic if decisions are expected and citizens lack the resources to judge what counts as a good argument. Multiplism can leave citizens vulnerable to manipulation and put at risk the ideals of openness and democracy. Curricula that represents a multiplist view of knowledge promotes different points of view but is not directed toward shared general understandings. As such, curriculum is likely to be fragmented. A compatible education experience involves a supportive context where multiple views of reality are shared. Students are expected to generate beliefs according to their unique resources, beliefs, values, and logic. Justification is not a priority. All evidence is equal, thus observations, measurements, opinions, beliefs, need not be compared. Students demonstrate personal knowledge by drawing their own conclusions from multiple perspectives. Success is a matter of internal coherence, not a matter of the adequacy of grounds for belief. From a multiplist educational experience, students might acquire empathy or tolerance for different perspectives and sensitivity to the feelings of others in cooperative situations, attributes necessary but not sufficient for the ideal citizen. Proficiency in evaluating the relative merit of alternatives and in justifying sustainable arguments would have to be acquired elsewhere.
An evaluativist views knowledge as human constructions of reality that are sustained by arguments that justify evidence quality and its adequacy to support the conclusion over alternatives. The evaluativist recognizes that variance in human experiences and contexts corresponds with different filters through which reality and reports of reality are perceived. Awareness of human fallibility motivates the evaluativist to seek out and strive to understand multiple perspectives and to remain both vigilant to disparities in the quality of arguments and open to new evidence. Contrary to the multiplist, the evaluativist respects experts' more informed knowledge base, but also recognizes that when the outcome matters, there is a need to evaluate the sustainability of experts' arguments. An evaluativist self-regulates thinking toward qualities such as clarity, precision, specificity, accuracy, plausibility, relevance, sufficiency, consistency, coherence, significance, fairness, and adequacy to purpose (Paul, 1992). An evaluativist is compatible with a society in which citizens contribute to judgments that have potential to impact others. The ideal citizen in an evaluativist context would have the knowledge, habits of mind, and proficiencies to care about the impact of decisions, to judge what counts as sustainable beliefs or actions, and to take action in a socially responsible way. Curricula to facilitate this evaluativist citizen makes salient the connections among intellectual resources, discipline-based knowledge, and compatible instruction. Instruction for an evaluative view
of knowledge involves teaching in a critical manner, sharing with students and expecting of them justification to sustain judgments when the outcome matters to someone, and facilitating for students the intellectual resources to do so (Siegel, 1980). Students demonstrate knowledge through justifications of the adequacy of grounds for their decisions. The criterion for success is the degree to which students' judgments are sustained by reasons and how adequately students' justifications reflect the quality of argument. An evaluativist educational experience provides students with the opportunities and instruction compatible with facilitating the intellectual resources necessary to seek, judge, and realize sustainable arguments in order to satisfy their social responsibilities as informed and involved decision makers.

Pedagogy for an evaluative view of knowledge infuses students' experience with a compatible experience in critical thinking (Siegel, 1992). It can be argued that an evaluativist educational experience is the only option with some prospect of helping citizens acquire the intellectual resources necessary to sustain and improve democracy (Kurfiss, 1988). It can also be argued that students have rights as persons to being educated in a critical manner (Siegel, 1980). Teaching in a critical manner honors students' rights to reasons and justifications for disciplinary knowledge, invites students to question and challenge ideas, expects students to provide reasons and justifications to sustain their beliefs, and facilitates
students' acquisition of the intellectual resources to do so. As persons, students are entitled to learn about the adequacy of the grounds for disciplinary knowledge. Evaluativism and the critical thinking conception implied are thus the only morally sustainable positions for education in a democracy.

**Epistemological Journeys: Selected Research**

Commitment to an evaluative view of knowledge alone is inadequate to recognizing and justifying the sustainability of knowledge. Congruency is needed among the vision of the ideal citizen, the view of knowledge, the critical thinking conception, and pedagogy. Research has identified a paucity of evaluativism; disparity in views of knowledge; a correspondence among view of knowledge, level of education, nature of education, critical thinking proficiency, intellectual and ethical development, and gender; and an emotional commitment to view of knowledge that is resistant to change even with intervention (Belenky, Clinchy, Goldberger, & Tarule, 1986; King & Kitchener, 1994; Kuhn, 1991; Mines, King, Hood, & Wood, 1990; Perry, 1968; Wood, 1997). If such research is representative, the potential for citizens to acquire the commitment to and the resources for sustainable judgment is problematic as few people experience the level and nature of education important to attaining the intellectual resources necessary to value, recognize, build, and justify sustainable judgment.
Key epistemological theories foundational to this research have built upon Perry’s (1968) two consecutive 4-year longitudinal studies of 140 undergraduate college students (mostly males). Perry identified from his unstructured interview evidence 9 major intellectual and ethical positions of how “students construed their worlds” and 3 transitional responses that together represented students’ “epistemological Pilgrims Progress” (p. 44, italics in original). The earliest position was Basic Duality, an absolutist view in which answers are right or wrong with no room for neutrality or values of relative strength between such dipoles. An authority’s ethnocentric representation of reality is believed to correspond exactly with reality. Counter views are considered incorrect. Positions 2 and 3 illustrate a gradual movement toward recognition and acceptance of the legitimacy of other points of view with forms of multiplicity vying with dualism for dominance. Position 4 is characterized by either a Multiplicity Correlate or a Relativism Subordinate stance. The Multiplicity Correlate view accommodates both absolute dualism and multiplicity, retaining dualism for when authority is certain and multiplicity for when uncertainty prevails and one opinion is as good as another. Students who adhere to the alternative route of the Relativism Subordinate position conform to what authority wants by entertaining multiple views and using critical procedures that will eventually lead to the right answer. This consideration of multiple views enhances the potential to discover a
foundation for the Relativism of position 5. Position 5 is pivotal to Perry’s framework, representing a major epistemological reorganization during which relativism gradually gains dominance over multiplicity and dualism is relegated to subordinate or special case status. Perry’s depiction of relativism shares multiplism’s recognition of multiple perspectives of reality. But unlike multiplism, the relativism of position 5 accepts potential for an objectivity that is “qualified by the nature of the contexts in which one stands back to observe” (p. 126), thus limiting generalizations to similar contexts. Furthermore, a person at position 5 comprehends the need for critical thinking skills for “comparative, and hence, relativistic contextual thought” (p. 96), necessary to weighing the relative merit of alternatives. Position 6 through 9 involve movement toward ethical ideals of personal commitment to beliefs and acceptance of responsibility for choices and consequences, characterized by the development of critically thoughtful habits of mind.

Perry also observed transitional behaviors of delay, detachment, or retreat as students found ways to cope with the demands of epistemological challenges and responsibilities. Individuals may delay or rest in any position over a full year, perhaps considering the implications of that position or gathering strength to move on. Detachment, apparently not possible until position 5, provides escape from responsibility and commitment to belief and action. Retreat to earlier positions
may reflect hostility, perhaps evolving from uncertainty and anxiety when faced with rejecting the security of dualism.

Kitchener and King's (1981) Reflective Judgment model built on Perry’s work and others, providing fuller elaboration of evaluative attributes. Data from 60 students (matched for verbal ability and balanced for gender and education level—high school seniors, college juniors, and doctoral students) represented structured interview responses that accounted for 7 developmental stages of reality, knowledge, and justification. Stage 1 assumes objective representation of reality or absolute knowledge, a lack of perceived difference among beliefs, and no need for justification. Stage 2 differs in that some claims are believed to be false. Justification to distinguish absolute knowledge from falsity is based on direct observation or appeal to authority. Stage 3 recognizes objective reality, false claims, and includes uncertainty. Beliefs depend on evidence and authority, or where certainty is lacking, temporary beliefs are a matter of opinion. Stage 4 accepts objective reality but not absolute knowledge, given that one cannot be certain how close one is to truth. Beliefs are justified according to idiosyncratic evaluations or not at all. Knowledge is subjective for stage 5 people in which neither objective reality nor objective knowledge exists. Justification differs significantly from stage 4 as judgments are evaluated for relative strength or how well they adhere to simple context-specific rules for inquiry and evaluation.
Judgment is withheld if evidence is inadequate. Stage 6 people accept an objective reality but deny an objective understanding of reality. Judgments are subjective but some are more defensible than others. Alternative positions (including experts' views) are evaluated for fit with generalized rules of inquiry and evidence appropriate to the context. Justifications are based on personal judgment of adherence to appropriate rules and evaluated authoritative views. The seventh stage of reflective judgment assumes an objective reality and fallible constructions of knowledge. Certainty exists only in the relative merit of claims. Critical inquiry and evaluation of evidence and expert claims lead to judgments about which conclusions best approximate reality. Judgments must remain open to new evidence. Stage 7 is also distinguished by a capacity for cognitive self-evaluation as well as sufficient cross-disciplinary knowledge to enhance synthesis of evidence, methods, and criteria from these domains to build more adequate grounds for justifying multi-disciplinary judgments.

Belenky et al. (1986) found Perry's model inadequate to account for data from 135 females who varied in age, ethnicity, life circumstances, and educational background. Belenky et al. also expressed dissatisfaction with developmental categorizations which they suggested imply inadequacy and alienation. Ruddick (1996), with whom Belenky et al. agreed, suggested that impersonal dominant authorities have treated people with different epistemological views with
condescension, even though "their standpoint, epistemologically speaking, may be less partial, more 'objective,' than that of those who enact dominant epistemologies" (p. 252). Belenky et al.'s interview, more structured than Perry's, revealed evidence of 5 positions. Silence is the response to authority by one who perceives little room for the development of personal mind or voice. Received knowledge involves the reproduction of ideas from authority but does not generate original ideas. Subjective knowledge involves denial of authority in pursuit of a personal or subjective truth and knowledge. Procedural knowledge involves "separate and connected knowledge" (p. 100). Separate knowledge is realized through the impersonal rules of logic for decisions about what to believe and the communication and justification of these ideas. Connected knowledge is oriented toward understanding points of view, underlying reasons, and the social implications of decisions, not toward the critical evaluation and justification of ideas. Constructed knowledge represents the more mature epistemological perspective of one who values intuition, expertise, reason, and who is passionate about approximating reality in a quest for truth, recognizing the limitations of human inference. Knowledge is contextual, constructed by individuals who strive to balance separate and connected knowing and a sense of moral responsibility to make a difference.
Kuhn's (1991) study of experts (N=15) and non-experts (N=140) used structured interviews that distinguished absolutist, multiplist, and evaluativist epistemological positions. The absolutist perceives an objective truth and thus potential for certainty. Justification appeals to authority or personal observation. Truth is subjective to the multiplist and thus authoritative sources are no more credible than anyone else, potential for expert certainty is denied, and justification of relative merit irrelevant. The evaluativist also respects different viewpoints, recognizes that different experiences and prior knowledge can influence interpretations, but holds that some beliefs are more defensible than others. The evaluativist denies potential for certainty but as a non-expert, defers to the more informed position of the expert.

Research related to these epistemological models (Belenky et al., 1986; Kitchener & King, 1981; Kuhn, 1991; Mines, King, Hood, & Wood, 1990; Perry, 1968; Wood, 1997) revealed scarce evidence of more mature levels of evaluativism. Higher scores corresponded with education level. Perry reported few if any students exhibiting position 9 attributes. Most students displayed positions 3, 4, or 5 at the end of the first year and senior students displayed attributes representative of positions 6, 7, and 8.

Reflective judgment research found an association between level of education and reflective judgment level with scant evidence that graduate students
realized the higher levels of the Reflective Judgment model. Kitchener and King (1981) found that students' reflective judgment stages varied across 3 age-education groups but were not gender-related. The 16 to 17 year-old college students exhibited more multiplist views; while graduate students (aged 24 to 34) exemplified the more advanced reflective judgment levels. Even so, only 3 of 60 students evidenced stage 7 attributes. Mines, King, Hood, and Wood (1990) found that without exception, students (N=100) across 3 university levels (freshmen, seniors, graduate students) and balanced by gender and area of study (mathematics and social sciences) achieved increasingly high scores for each of 3 measures (Reflective Judgment Interview, Watson-Glaser Critical Thinking Appraisal, Cornell Critical Thinking Test) as their education level increased. These researchers found that the critical thinking dimensions that distinguished reflective judgment stages included interpretation and weighing evidence and identifying generalizations that are warranted beyond a reasonable doubt, detecting arguments that violated logic rules, reasoning deductively from premises to conclusions, and analyzing inferential strength. They concluded that these proficiencies are theoretically consistent with and perhaps necessary for realizing the upper reflective judgment stages. Mines et al. noted that in spite of strong relationship between educational level and reflective judgment stage, the 40 graduate students' reflective judgment mean score was only 4.76 and only 1 of 99
subjects achieved the highest reflective judgment level of stage 7. Unlike Mines et al., Brabeck (1980) reported that while critical thinking proficiency was associated with reflective judgment scores, proficiency in critical thinking was insufficient to realize the highest reflective judgment levels (high proficiency in critical thinking corresponded with variable reflective judgment scores; scores for reflective judgment increased with educational level but not with critical thinking proficiency). Brabeck concluded that critical thinking and reflective judgment were different constructs. Brabeck's female students (4 education levels: high school seniors, college sophomores, college seniors, and graduate students) who exhibited higher reflective judgment levels (scores ranged from 2.5-5.3) had also attained a higher level of education, congruent with other research findings. It is possible that the expectations and context of the Watson-Glaser test depict an evaluative competence that differs from that demanded by the more generative expectations and context of the Reflective Judgment Interview. Moreover, Mines et al.'s use of both the Watson-Glaser and the Cornell tests, the latter which addresses credibility of sources and observations and semantic aspects of critical thinking, provided a more comprehensive measure of critical thinking, perhaps accounting for conclusions that differed from Brabeck's findings. Even so, neither test includes critical thinking dispositions. Critical thinking test results also must be interpreted in light of Norris and Ennis' (1989) observations that existing
multiple choice tests are problematic in that good critical thinkers may choose ‘wrong’ answers, influenced by their prior knowledge, proficiencies, and dispositions to respond in unanticipated ways.

Kuhn examined a wider range of subjects for argument skill and view of knowledge. The main sample of 160 subjects was balanced across 4 age intervals (teens, 20s, 40s, and 60s); gender (males and females), and education level (noncollege and college). Argument skill was determined by the nature of subjects’ beliefs or causal theories about 3 separate complex social issues, the structural complexity of theories, and the quality of supporting evidence, alternative theories, counterarguments, and rebuttals that justify the relative merit of positions. Only 16% of subjects provided genuine evidence for all 3 issues, 33% generated alternative theories for all topics, while 21% to 32% could justify the relative merit of theories across all 3 issues. Epistemological evidence indicated that only 9% to 22% of subjects across topics understood the integral relationship among deliberation, evaluation, and knowledge. Kuhn found that gender and age differences were not significant but that higher education corresponded with more evaluativist responses. Only 2 subjects were consistently evaluativist and only 11 subjects were predominantly evaluativist.

Kuhn found that topic expertise made little difference to reasoning quality for her small sample of experts’ responses (N=15: 5 teachers, 5 parole officers,
and 5 philosophy doctoral candidates). She observed that the 5 philosopher candidates who lacked specific topic expertise all exhibited reasoning expertise across the 3 social issues, perhaps a reflection of the nature of reasoning that characterized their doctoral studies. Kuhn's assumptions of expertise might be unfounded, if research is representative that suggests that undergraduate education provides limited opportunities for evaluating the relative merit of theoretical ideas and related research (Goodlad, 1990; Paul, 1996). The 5 advanced philosophy students' argumentation competency and consistent evaluative view of knowledge may represent the benefits of advanced graduate studies or perhaps more specifically, the nature of a philosophy education.

Developmental research (Belenky et al., 1986; Kitchener & King, 1981; Kuhn, 1991; Mines et al., 1990; Perry, 1968) has found that epistemological change is slow and that there is an association between epistemological development and educational level, suggesting that evolution is not simply a matter of natural development. Movement from one position to another can take more than two years (Perry, 1968; Woods, 1997). Intervention can meet with overt resistance when an instructor's epistemological views are not congruent with students' prior educational experience and views of knowledge (Kurfiss, 1988; Perry, 1968). Woods (1997) reported in his secondary analysis of reflective judgment research that pretest-posttest intervention studies of short duration made
little difference, limiting the influence of pretest effect. Kuhn (personal communication, December 15, 1997) questioned the value of pretest-posttest studies because of the relative stability of views of knowledge.

Inconsistencies across research instrumentation and the nature of students’ prior education could account for discrepancies between students’ critical thinking proficiency and reflective judgment maturity as well as for their limited epistemological maturity. Instrumentation inconsistencies include critical thinking tests that do not include generative responses of interviews which include the justificatory dimensions that are better captured by the Reflective Judgment Interview. In addition, multiple choice critical thinking tests do not offer a comprehensive range of critical thinking dimensions and may incorrectly represent the intent of respondents. Paul et al. (1997) found that almost half of the postsecondary faculty surveyed did not adhere to an evaluative view of knowledge; nor did the majority of this group of faculty members demonstrate a clear understanding of important critical thinking concepts. Sirotnik (1990) reported that the majority of over 500 preservice teachers whose interview responses paralleled survey data for 3000 students apparently lacked opportunity for and the language required for critical discussion of epistemological concerns about educational purpose and means. This finding paralleled Norris and Phillip’s (1994) research that revealed that top grade 12 science students (N=91) exhibited
inadequate understanding of epistemological language—conclusion, evidence, and justification—likely contributing to their inadequate preparation to judge the adequacy of support for belief about science magazine reports. Olson and Babu (1992) also reported that relevant concept knowledge is important to critical thinking as it might be the "language of instruction; [but] it is not the object of instruction" (p. 193). One possible explanation Olson and Babu (1992) posited for apparent reasoning inadequacies was a "failure to recognize the genre" (p. 195)—in this case the distinction between logical and empirical claims. Subjects did not accept the premises as hypothetically true in order to focus on logical relationships among claims when they had empirical evidence that contradicted the assumption that the premise was true (see Scribner, cited in Norris, 1995, for similar findings). Bailin (1995) theorized that to adequately understand particular epistemological concepts, one must comprehend the integral associations among the many concepts that comprise the whole of the critical thinking conceptual network. One must also "see the point of the practice..., a practice essentially epistemological in nature" (p. 5) and fundamental to inquiry and knowledge.

If relatively few people attend graduate school where evidence of an evaluative view of knowledge is most apparent and if the nature of graduate studies further influences peoples’ view of knowledge, there is reason for concern. The majority of citizens responsible for decisions that have potential to seriously
impact all of humanity do not attend graduate school. If research is representative, it follows that most citizens lack an evaluative view of knowledge. The integral relationship among deliberation, evaluation, and knowledge represent epistemological understandings that provide reason for and justify an adequate conception of critical thinking (Siegel, 1997). These understandings are fundamental to seeing the point of critical thinking, necessary to knowledge that reflects an evaluativist epistemological position. It is essential that critical thinking theorists, curriculum workers, teacher educators, and teachers understand the integral relationship between epistemology and critical thinking as they relate to each area of study if they are to impact education so that citizens are adequate to the challenges of a global village.

**Critical Thinking**

Critical thinking is espoused throughout curriculum reform documents as an important attribute of the critically literature educated citizen. Yet, while there is increasing evidence that what is required to think critically is given prominence in some reform documents, seldom is what counts as critical thinking elaborated. Nor are attributes essential to critical thinking well integrated throughout and across most documents (Ford et al. 1997). In British Columbia, the cognitive task dimension of critical thinking, rooted in Bloom, Engelhart, Furst, Hill, and Krathwohl’s (1956) taxonomy of cognitive objectives and salient in the 1983
elementary social studies curriculum, has been a mandatory component of expected learning outcomes in current curricula (British Columbia, Ministry of Education, Curriculum Development Branch, 1983; British Columbia Ministry of Education, Skills, & Training, 1996; Ford, 1988; V. Grout, personal communication, 1996) and central to related assessments (Bognar, Cassidy, & Clarke, 1996). A plethora of materials were developed to facilitate implementation of the 1983 elementary social studies curriculum, a document in which the 4 main curriculum goals were to be realized through "critical thinking and problem solving skills" (British Columbia, Ministry of Education, Curriculum Development Branch, 1983, pp. 5-6). In spite of the centrality of critical thinking to these curriculum goals, and the materials developed to facilitate implementation of the curriculum, provincial assessments have found that evidence of successful critical thinking was inadequately demonstrated across elementary and secondary students (Bognar, Cassidy, & Clarke, 1996; Cassidy & Bognar, 1991). It could be that critical thinking as represented in curricula does not help teachers understand what counts as critical thinking to help guide instruction and assessment (Cassidy & Bognar, 1991; Court & Francis, 1993; Ford, 1988), or if research is representative, teachers are not adequately prepared to teach for critical thinking (Goodlad et al., 1990; Paul, 1996; Unks, 1985; Wright, 1995b). It is also possible that multiple factors operate together to challenge prospects for critical thinking.
This study is guided by the judgment that cognitive tasks, as represented in recent social studies curricula, have not been adequately elaborated to help teachers comprehend what is necessary to integrate critical thinking (Cassidy & Bognar, 1991; Ford, 1988; Ford et al., 1997). Cognitive tasks represent purposes for intellectual engagement while thinking through a problematic situation toward decisions about what to believe and what to do. Engagement in cognitive tasks might or might not adhere to the intellectual standards of critical thought.

To address this perceived inadequacy, and to try to illustrate connections between current curricula emphases and the intellectual resources that are important to critical thinking, this selected literature review focused on accounts of critical thinking from cognitive psychology and philosophy to identify issues, emphases, definitions, and dimensions of critical thinking. This selected literature review of critical thinking included material accessible to curriculum workers and teachers (e.g., popular journals, teaching method texts, general publications, videos, the 1983 British Columbia elementary social studies curriculum guide, and recent reform documents in Ford et al., 1997) as well as sources relevant to this interpretation of perspectives of critical thinking. Accounts were examined for explicit and implicit definitions, dimensions, relationships, and assumptions. The goal was to interpret, select, and integrate dimensions of critical thinking from cognitive psychology and philosophy that had promise to move education for
elementary social studies preservice teachers closer to evaluativist ideals. The integrated critical thinking model represents an evaluativist view of knowledge, the only position that this researcher-instructor accepts as morally sustainable for public school education.

Interest in critical thinking, both in cognitive psychology and philosophy, is directed toward the improvement of thought. Issues, emphases, definitions, and dimensions, and relationships relevant to critical thinking are examined here to provide some background for this study.

**Issues**

Issues about critical thinking that have relevance for social studies education include distinctions between critical thinking and creative thinking, the notion of generalizability, the process-product debate, and a concern that critical thinking might bias gender and cultures. Misconceptions could persuade educators that critical thinking is pursued to the neglect of creative thinking (de Bono, 1984). Some claim that critical thinking and creative thinking are distinct (Marzano, Brandt, Hughes, Jones, Presseisen, Rankin & Suhor, 1988) and that creative thinking is non-judgmental (Torrance, cited in Nickerson et al., 1985). This researcher agrees with arguments for an integral relationship between creative thinking and critical thinking (Bailin, 1992; Lipman, 1991; Nickerson et al., 1985; Scriven, 1976; Sternberg, 1995). Bailin's (1992) argued that critical
thinking "...involves the intimate interplay of generation and evaluation" (p. 94). Even though novel ideas may be discovered by accident, Bailin explained, there is good reason to believe that the production of ideas worthy of attention are "constrained in some way by the same criteria and principles that constrain evaluation....Created products have their roots in the problems and paradigms of existing traditions, and conform to a considerable degree to their principles, procedures, and rules" (Bailin, 1992, pp. 89-90). It follows that critical and creative thinking are not mutually exclusive kinds of thinking. Critical thinking entails generation of alternatives when considering what to believe or what to do; creative products require evaluation for decisions about relative merit.

The issue of critical thinking generalizability has implications for educational goals, instruction, and evaluation (Ennis, 1992; McPeck, 1990; Norris, 1992; Paul, 1992; Siegel, 1992; Smith, 1990; Toulmin, 1956). In order to understand related research about what aspects of critical thinking, if any, are generalizable, and under what situations generalization occurs, it is necessary to distinguish both the intent of critical thinking and of generalizability. What counts as critical thinking in some research is unclear or misrepresented (Bers, McGowan, & Rubin, 1996; Collins, 1992; McKee, 1988; Onosko, 1989; Orton & Lawrence, 1990). For example, McKee (1988) defined critical thinking as "a dynamic process of questioning and reasoning; active inquiry as opposed to
passive accumulation of knowledge. To think critically is to question definitions, actions, and beliefs” (p. 444). It is not clear in this definition that instruction involved critical thinking. One could engage in such questioning without knowing what counts as sustainable judgment. Bers et al. (1996) reported that teachers participated in critical thinking workshops but no specifics about the nature of instruction were available. Assumptions of critical thinking education would be unwarranted. Research that purports to focus on critical thinking might not involve critical thinking at all. It must be clear that critical thinking is implemented successfully in one context before one is justified in judging whether related understandings, dispositions, and proficiencies have been applied in new contexts. Research that has not clearly represented critical thinking and research that has reported inadequate understanding of critical thinking (Court & Francis, 1993; Goodlad et al., 1990; Paul et al., 1997; Unks, 1985; Wright, 1995b) help explain limited evidence of critical thinking in public schools (Bognar, Cassidy, & Clarke, 1996; Cassidy & Bognar, 1991). One should be even less surprised, then, by limited generalization evidence.

This is not to discount an apparent growing body of empirical evidence and anecdotal reports relevant to successful critical thinking implementation (Hatcher, 1995; Leshowitz & Yoshikawa, 1996, Nelson, 1996) and possible transfer of critical thinking (Kuhn, 1991; Phillips, 1992). For example, Hatcher and Price
(1998) reported seven years of data for which a 2-semester college course that integrated critical thinking and written composition (argument genre) produced "significantly better outcomes than stand-alone traditional courses in critical thinking or composition and in the area of critical thinking, the degree of difference is great" (p. 26). The Ennis-Weir Critical Thinking Essay Test results for 1105 first-year Baker University students revealed a pretest-posttest mean score difference of +4.93 (p-value <.001). Comparisons were made with similar student populations in traditional programs (separate course in logic or critical thinking) in 3 different contexts (a state university, a state college, and a liberal arts college). Students who took a standard logic course (N=44) had a pretest-posttest mean score difference of -1.7; students taking a standard critical thinking course (N=23) attained a pretest-posttest mean score difference of +1.6; and another group of students taking a standard critical thinking course (N=67) had a pretest-posttest mean score difference of -0.10. Baker students who were tested again in their senior year (N=303) also attained statistically significant (p-value <.001) pretest-posttest difference mean score of +6.59. Hatcher and Price (1998) reported that while one semester made little positive difference, two semesters that allowed multiple applications of understandings to written arguments made a significant positive difference to students' critical thinking essay test scores. This
success is encouraging. It does not, however, address whether critical thinking proficiencies transfer beyond the context of written argument tests.

As for generalizability, Johnson’s (1992) distinction between notions of generalization and transference align with this researcher’s position. View of knowledge and disposition toward critical thinking might well be applicable and transferable to other contexts. However, they may not generalize or automatically appear across all domains once they have been demonstrated in the intervention context. Internalization of these epistemological understandings and critical spirit might increase the potential for their transfer to new contexts. Blair (1992) hypothesized that transference of critical thinking about observational reports will depend on whether principles for judgment were taught for transfer and the degree that motivation and experience applying critical thinking in similar situations coincide. Or it may be that one is more disposed toward critical thinking in some contexts (e.g., social studies) than in others (e.g., language arts, physical education). Reflective judgment research (King & Kitchener, 1994) suggested that stage 7 is exemplified by insights into cross-disciplinary epistemological and critical thinking connections that could reinforce stability of perspective in each area. A lack of clarity about what researchers and theorists intend by generalizability claims makes representation and discussion of ideas problematic.
Kuhn (1991) suggested that the 5 advanced philosophy doctoral students who exemplified consistent evaluative views of knowledge and exceptional proficiency in argument for each of 3 distinct social issue topics had transferred their learning from their philosophical studies. The relative simplicity of the content for the social issue scenarios suggests that the basic proficiency in argument that philosophy students would likely have by nature of their education specialization, would be sufficient to recognize the inadequacy of evidence represented. Whether these philosophy students would have also excelled in other forms of reasoning was not clear.

Some of the more salient questions about transference relate to domain-specificity, what aspects of critical thinking are transferable, and the stability of processes across contexts and individuals. While some have argued that critical thinking is subject specific, constrained by the nature of judgments and knowledge specific to a domain (McPeck, 1990; Toulmin, 1956), others have argued that many aspects of critical thinking are generalizable (Ennis, 1992; Norris, 1992; Phillips, 1992; Siegel, 1992). Some identify many inter-field commonalities that challenge a theory of domain specificity (Ennis, 1992; Resnick, 1987; Siegel, 1992). For example, Ennis (1996) and Pinto, Blair, and Parr (1992) delineated criteria for credible sources that would be common to other contexts where source credibility is a concern. Phillips (1992) reported that some cognitive tasks of self-
regulation were transferred across narrative reading and mathematics problem solving and across narrative reading and expository reading in science. Siegel (1992) argued that what counts are not discipline-specific constraints on judgments and criteria but rather, what types of claims are involved and what counts as evidence, elements that may be relevant to more than one domain.

Fully generalizable, Siegel (1992) claimed, are the epistemological understandings of the "goodness of reason, and of related issues concerning truth, fallibilism, rationality, and the like [as well as]...the critical spirit component of critical thinking" (p. 107). One with a critical spirit willingly seeks out sustainable judgments about what to believe or what to do and acts accordingly. Whether one actually transfers this epistemological perspective and critical spirit across all contexts remains a question for empirical research.

The notion of stable processes and skills that are generalizable across contexts has been widely criticized in both cognitive psychology and philosophy (Coombs & Daniels, 1991; McPeck, 1990; Resnick & Klopfer, 1989; Smith, 1990). For example, proficiency in analyzing chemical compounds does not necessarily transfer to analyzing other kinds of information (e.g., concepts associated with critical thinking, demographic data, qualitative research evidence). One needs knowledge specific to what is being analyzed to be successful. A process-oriented conception has also been challenged for its neglect of the essence
of critical thinking which is not fruitful if the goal is to enhance education for critical thinking (Bailin et al., 1993; Paul, 1992). Linear lock-step recipes or cognitive processes involved in inquiry, problem solving, decision making, or issue analysis may present occasions for critical thinking but do not alone help educators comprehend what counts as quality judgment. That is, cognitive processes that are not clearly directed toward an evaluative view of knowledge and integrated with the intellectual resources for judging quality are insufficient for critical thinking (Bailin et al., 1993).

In spite of such recognition of the inadequacies of thinking skills-processes conceptions of critical thinking, Perkins, Jay, and Tishman (1993) observed that

liberally interpreted, the term thinking skills conveys well what the majority of efforts to cultivate thinking have been. They treat good thinking as composed of a variety of general cognitive processes, such as generating ideas, exploring consequences, reviewing options, monitoring progress, and so on. They adopt the practical approach of offering learners stepwise strategies for such activities and providing plenty of practice, sometimes accompanied by metacognitive reflection. (p. 2)

It is argued that a conception of critical thinking that does not capture the essence of criticality—judgment that can be sustained against relevant criteria—fails as an adequate conception (Coombs & Daniels, 1991). It is also argued that a comprehensive conception that remains inaccessible to those for whom it is intended fails as an adequate conception. Philosophy’s contribution of conceptual analyses and principles related to good judgment is necessary for thinking
critically about what counts as good judgment. However, the ideas are complex and not always readily interpreted by the uninitiated (e.g., Ennis, 1996; Goodlad, 1990; Siegel, 1988). This could account, in part, for the paucity of critical thinking in curricula and in classrooms. It may be that understanding the more technical aspects of critical thinking is not sufficient for commitment to critical thinking. As Bailin (1995) suggested, students may need to grasp the point of critical thinking—it's epistemological rationale. This understanding could be linked to a strong psychological dimension that can dispose one to seek out and utilize knowledge of critical thinking for constructing knowledge (Facione, Sanchez, Facione, & Gainen, 1995). Furthermore, research into impediments to and conditions for critical thinking may also be important to improving conditions conducive to critical thinking education (Sternberg, 1995). Conceptual and empirical research from both cognitive psychology and philosophy may contribute to better prospects for critical thinking education.

Thoughtful discourse is emerging from rather heated debates that critical thinking is biased, particularly against females and some cultural groups (Alston, 1995; Bailin, 1995b; Belenky et al., 1986; Gilbert, 1994; Miller, 1995; Norris, 1995; Orr, 1989; Phelan & Garrison, 1994; Ruddick, 1996; Wheary & Ennis, 1995). This contention is of particular importance to the degree that critical thinking is a guiding principle for education (Norris, 1995). As Norris pointed
out, if public education promotes a way of thinking that excludes groups, there is
good reason for concern. Many critics have interpreted critical thinking more
narrowly than represented in current more comprehensive critical thinking
conceptions (e.g., Bailin et al., 1993; Ennis, 1996; Paul, 1992). A traditional logic
conception has been criticized for its concern for consistency among claims to the
exclusion of the adequacy of premises and their underlying reasons; a traditional
western rationalist search for universal truths through detached objectivism that
divorces the self from the object to be known and adheres strictly to rules and
principles is apparent in other accounts (Belenky et al., 1986; Gilbert, 1994; Orr,
1989; Phelan & Garrison, 1994; Ruddick, 1996).

Key concerns about critical thinking's gender bias include inadequate roles
in critical thinking for emotion, caring, connection between the knower and the
object to be known, context, and personal voice. These female-oriented thinking
attributes have not been legitimized by dominant male-oriented forms of reason
(Belenky et al., 1996; Gilbert, 1994; Phelan & Garrison, 1994; Orr, 1989;
Ruddick, 1996). Critical thinking proponents have responded that critical thinking
is more inclusive than are the traditional logic-oriented argumentation conceptions
that critics have denigrated. For example, Wheary and Ennis (1995) argued that
emotion provides the positive motivation to use critical thinking toward problem
resolution and that the emotions that underlie positions need to be considered in
critical thinking judgments. Wheary and Ennis also cautioned that emotions can impede critical thinking and blind one to reality. Caring disposes one to recognize the worth of all people and to fair-minded consideration of all relevant perspectives and their supporting arguments (Norris, 1995; Wheary & Ennis, 1995). Ideally, the fair-minded critical thinker strives to suspend personal beliefs in order to consider the best arguments in support of contrary beliefs as if they were one’s own. Such fair-mindedness aligns with Phelan and Garrison's (1994) discussion about the intimate link between self and object that characterizes a feminine way of knowing. Phelan and Garrison captured the ideal of connected knowing and its movement toward constructed knowing in Keller’s account of a “‘dynamic objectivity... a pursuit of knowledge that makes use of subjective experience...; an exquisite balancing act,’ [that] recognizes the role of the self in the act of knowing although it refuses to allow itself to be reduced to the romance of mere subjectivism” (p. 259). Dynamic objectivity “embrace[s] contraries” (Phelan & Garrison, 1994, p. 256) or alternative positions and increases the range of ideas important to the deep understanding necessary to judge the relative merit of positions. The range of ideas necessary to sustainable judgments and a constructivist view of knowledge are thus facilitated. Traditional absolutists who sought universal generalizations denied the impact of context on language, values, and perceptions of reality. Feminist concern for the particular align with the
centrality of context to critical judgments about belief and action (Bailin et al., 1993; Lipman, 1991; Paul, 1992). Nor is the personal voice precluded in a holistic critical thinking conception. Wheary and Ennis (1995) suggested that one must find a balance between personal and external knowledge sources and judge which ones are most helpful to realizing the best fit between judgments and reality. Bailin (1995b) summarized some critics’ examples of male bias as non-examples of critical thinking:

Thinkers...[who] failed to take the whole situation into account, failed to be sensitive to the feelings of others, failed to understand other perspectives, and failed to demonstrate a critical spirit [do not represent examples of critical thinkers]. (p. 193)

Some critics’ have focused on communication style rather than critical thinking itself, which can be engaged in either collaboratively or assertively (Gilbert, 1994; Miller, 1995). Neither communication style necessarily involves or precludes critical thinking (Bailin, 1995b; Miller, 1995). As Bailin observed (1995b):

What has not been recognized is the fact that critical thinking can be performed in the kind of collaborative style that is more usual for females, and further, that attention to the interpersonal context for critical thinking may better serve our purposes as critical thinkers. (p. 194)

The confrontational style of argumentation that has characteristized male communication in accounts of gender bias can intimidate and limit the open honest sharing of ideas and the potential for understanding disparate voices
(Gilbert, 1994; Miller, 1995; Norris, 1995; Phelan & Garrison, 1994; Ruddick, 1996). Instead, Norris (1995) proposed that a narrative style might be more suitable for encounters toward understanding and resolving bias concerns. Phelan and Garrison's (1994) argument for embracing contraries may move the narrative even further toward understanding. This is not to discount the need for other critical thinking resources for the construction and justification of sustainable arguments. To deny the value of critical thinking is to rob critics of the very resources necessary to building a sustainable argument for their position (Bailin, 1995b).

As for empirical evidence of critical thinking bias, Norris (1995) identified Kuhn's (1991) research evidence that did not find significant gender differences in epistemological positions and argument proficiencies. Research that has found limited evidence of critical thinking and an evaluative view of knowledge needs to be examined in light of gender difference claims (Goodlad et al., 1990; Kitchener & King, 1981; Mines et al., 1994; Paul et al., 1997; Woods, 1997). It is not clear that either males or females find critical thinking—narrowly or broadly conceived—a natural quality of thinking (Paul, 1992). Belenky et al.'s (1986) study of female subjects suggested that environment and life experiences can help or promote intellectual development (Orr, 1989). It is quite likely that males and
females can be negatively impacted by similar experiences. These findings may have important implications for critical thinking education.

Norris (1995) also shared cases of cultural groups for whom critical thinking is negatively biased. Some cultures value different behavioral principles and the imposition of critical thinking could be antithetical to their beliefs. Some groups accept claims as true, to be applied and explained but not criticized or questioned.

As with feminists' concerns for particularity, so too evidence of critical thinking bias must be judged in context. The principles that have served cultural groups well in their past contexts, may not suit them in their current context (Ross, 1992). Citizens responsible for choices that impact others must make judgments about reality to better insure informed decisions about belief and action. Critical thinking might be biased against specific group. If this is so, investigators need to consider the relevant context, the responsibilities that citizens within such marginalized groups share, and the implications that follow for those with disparate ways of knowing.

**Emphases**

Cognitive psychology emphasizes critical thinking as process—cognitive operations that enhance travel toward the ideal of good thinking. Philosophy focuses on the destination—judgment that can be sustained by criteria. These
emphases have corresponded with different contributions to education. Sternberg (1995) credited Neisser for defining cognitive psychology as “the study of how people learn, structure, store, and use knowledge” (p. 57). Education can benefit from these theories through research about the effectiveness of conditions, strategies, and aids for good thinking (e.g., Beck, Perfetti, & McKeown, 1982; Caine & Caine, 1991; Graesser, Magliano, & Tidwell, 1992; Jensen, 1998; Rumelhart, 1980; Sylwester, 1995). Success is judged by the effectiveness of moves taken toward solving problems. Philosopher focus more on what good judgment entails in order to generate criteria of good judgment, important to curriculum and instruction decisions (Bailin et al., 1993; Ennis, 1996; Lipman, 1991; Paul, 1992, 1996; Pinto et al., 1992; Scriven, 1976). The quality of thinking depends on to what degree judgments reflect the criteria of good thinking (Case & Wright, 1997; Lipman, 1988; Paul, 1992). Both emphases are important to education. Information about conditions that can enhance or impede students’ thinking can inform decisions about conditions for success. Information about what counts as quality thinking has utility for curriculum, instruction, and assessment.

**Definitions of Critical Thinking—What it Is and What it is Not**

Sternberg (1995) defined critical thinking as a response to a problem through the conscious control of mental processes when working recursively
toward a thoughtful solution. This definition reflects cognitive psychology's procedural emphasis, a purposeful response to a problematic situation through engagement in cognitive tasks of self-regulation, deliberation, drawing conclusions, and ideally, justification (Facione, 1991; Marzano et al., 1988; Nickerson, Perkins, & Smith, 1985; Sternberg, 1995). Critical thinking is not represented by discrete thinking skill definitions that depict a linear, lock-step, mechanistic, non-critical, response to a problem (Sternberg, 1995). Thinking or cognitive skills and mental processes, higher order thinking skills, intellectual skills, and thinking abilities have been used interchangeably (Baron & Sternberg, 1987; Costa, 1985; Perkins et al., 1993). However, inferences that have equated these terms may not reflect their original intent (Costa, 1985; Lipman, 1991). That is, some accounts of thinking skills never indicated an intent to represent critical thinking.

"Cognitive accountability" (Lipman, 1991, p. 118) is synonymous with philosophical conceptions of critical thinking that align with the etymology of critical thinking: the Greek krites—judge, kritikos—able to make judgments, and kriterion—standards for making a judgment (Ayto, 1990, p. 146). Judgments are purposeful determinations about problematic situations about what to believe or what to do. Critically thoughtful judgments are based on criteria, reasons against which we can justify and ideally, sustain the relative quality of decisions (Bailin et
al., 1993; Ennis, 1987, 1996; Lipman, 1991; Paul, 1992; Pinto et al., 1992; Siegel, 1992). Cognitive accountability also implies self-correction, holding judgment open to change in the face of more compelling evidence or reasons (Bailin et al., 1993; Ennis, 1996; Lipman, 1991; Paul, 1992; Siegel, 1992). Critical thinking contrasts to non-critical thinking that does not require such deliberation, judgments, and rationales. For example, Bailin et al. (1993) noted that daydreaming (a free-floating kind of thinking) and musing or pondering without any intent to judge among alternatives are not in themselves critical thinking. But daydreaming or musing may lead to critical thinking if cognition begins to focus on sustainable belief or action.

**Dimensions of Critical Thinking**

Dimensions of critical thinking include a reason for critical thinking that issues from a critical challenge or problematic situation; cognitive tasks that are directed toward a decision about belief or action; and intellectual resources by which the quality of judgments are evaluated and justified. Resources from both cognitive psychology and philosophy were used to interpret, adapt, elaborate, and translate critical thinking dimensions and view of knowledge for educating preservice teachers to teach for critical thinking and an evaluative view of knowledge in elementary social studies.
Definitions and elaborations of the dimensions of the integrated critical thinking framework for this study follow. Siegel (1992), prominent for his writing about epistemology and his justification for rationalism (1988; 1997), provided a philosophical perspective of the evaluative epistemological position that was foundational to the critical thinking conception in this study. Cognitive psychology provided a resource for thinking strategies and later, ideas for instruction. Habits of mind were common to both domains. Facione’s (1991) work, a multi-discipline collaboration, provided one of the more elaborated accounts of cognitive tasks and was adapted for this study. Cognitive tasks provided some key dimensions from current curricula which served as a connection for intellectual resources. Implementation literature has advocated relating existing beliefs and practices to new innovations (Fullan & Steigelbauer, 1991). It was hypothesized that explicit connections between cognitive tasks and intellectual resources would make the necessary role of intellectual resources more compelling. Bailin et al.’s (1993) conception of critical thinking was judged by the researcher-instructor to be one of the more useful critical thinking conceptions for distinguishing kinds of judgments and intellectual resources for educational purposes. It was also the conception commissioned by the Ministry of Education in British Columbia and was thereby relevant to the preservice teachers in the course. Lipman’s (1988) definition was valued for it’s simplicity in dimensions of
judgment, reasons, context, criteria, and self-correction and for elaborations that were useful for clarifying such concepts. Lipman's account was not as useful for demonstrating relationships among judgments and resources to suit the needs of the instructor. Similarly, Paul's (1992) conception and related representations facilitated interpretations and elaborations as well as understandings of instruction for critical thinking as it applied to social studies. Paul's conceptual organization was less useful to the instructor for clarifying associations among judgments, evidence, and criteria. Ennis' (1987) taxonomy of critical thinking dispositions and abilities that is foundational to many critical thinking conceptions was somewhat difficult to translate into practice. Pinto et al. (1992) provided useful information to aid interpretation and elaboration of judgments and criteria, as did Wright (1995a) whose textbook helped clarify concepts and translate some ideas for practice.

**An Integrative Framework for Critical Thinking**

Integration of the Bailin et al. (1993) and Facione (1991) conceptions of critical thinking provided the definition that *critical thinking is the self-regulated thinking through of problematic situations involving evaluative judgments about what to believe or what to do about meaning, relational, empirical, or value claims that clearly illustrates criteria appropriate to quality deliberations, judgments, and justifications.* Implicit in this definition are epistemological
assumptions of the nature of knowledge and a range of critical judgments, cognitive tasks, and intellectual tools.

**An Evaluative View of Knowledge**

The evaluative view of knowledge underlying this conception of critical thinking supports striving for a justified representation of reality as a goal of the critical thinker. Evaluativism acknowledges the fallibility of human inference and the need to remain open to the possibility of new evidence that challenges beliefs. Furthermore, this epistemological position recognizes that critical thinkers may interpret evidence, reasons, and experiences differently but they will strive to find out which judgments are most sustainable. Critically thoughtful judgments would be accompanied by a readiness and proficiency to share evidence or reasons for belief, as well as justification of why this support is relevant and sufficient. Knowledge claims are expressions of confidence that are backed by a proficiency to justify the grounds for beliefs.

**Critical Challenges and Embedded Judgments**

Social studies provides a rich venue for thinking critically about challenges in which are embedded one or more judgments about belief or action. Critical challenges that provide reason for critical thinking may involve one or more kinds of judgments.
Critical Challenges

Critical challenges ideally (1) address a problematic situation and are (2) meaningful, (3) provocative, (4) embedded in the core of the curriculum, (5) doable, and (6) acceptable (adapted from Case and Daniels, 1995). First, a problematic situation concerns a judgment about belief or action that matters to someone and for which a readily available answer is not apparent (Pinto et al., 1992). Nosich (1993a) captured this criterion of significance in the questions, “So what? Who cares?” Judgments that pose no serious benefit or loss to anyone provide little reason for critical thinking. A problematic situation must invite more than rote unconscious response, literal recall of discrete bits of information, mere guessing, or mindless application of an algorithm, formula or procedure. It requires purposeful cognitive engagement in thinking through a challenge to a judgment about which conclusion is most sustainable.

Second, critical challenges should be meaningful to students. That is, they should connect in some way to students’ prior knowledge and experiences. Meaningfulness might be realized through a situation of personal relevance for students or by building bridges—mediating for students the relationship between what they know and new information. The challenge might involve situated learning within a realistic context where judgments have potential for real-life impact and understandings develop in relevant contexts (Resnick, 1989).
Attention to meaningful connections is necessary if instruction is to benefit students (Ausubel, 1963; Resnick, 1989).

Third, provocativeness might be realized through challenges that engage students or have the potential to do so, as in situations that excite their curiosity or imagination, instill dissonance by challenging existing beliefs, arouse emotions, or invite decisions and actions by students with prospects for making a positive difference. Egan (1992) advocated that students’ interest and satisfaction might be enhanced through depth of inquiry and the expanding knowledge base that can lead to a sense of accomplishment.

Fourth, a critical challenge should be embedded in core aspects of the curriculum. The challenge should support the big ideas or important goals of the curriculum that connect beliefs and actions in past, present, and future.

Fifth, a critical challenge should be doable. The challenge should be within the capacity of students, given the time and resources available, and the challenge should be focused to limit the intellectual resources required. Instruction should be planned to facilitate the adequacy of resources—intellectual, material, and time.

Sixth, acceptability to relevant stakeholders has been appended to Case and Daniel’s (1995) criteria for critical challenges. Educators within a public school system have a responsibility to address this criterion. Strike (1982) distinguished
private and public concerns, suggesting that schools do not have a mandate to address private matters, such as religious values, unless they have potential for unacceptable social impact. As for social issues, many potential critical challenges would benefit from sensitivity to stakeholders who are important to the success of critical thinking education. It makes good sense to anticipate parental or administrative concerns to a sensitive topic and to plan ways to assuage their fears of an unbalanced, uncritical approach to the challenge. Plans and rationales could be shared with stakeholders who might be invited to contribute their ideas and resources. Building bridges may reap a more positive educational context for critical thinking.

**Judgments**

Critical challenges may encompass one or more associated judgments about what to believe or do about: (1) meaning, (2) logical relations, (3) empirical or descriptive claims, and (4) values one ascribes to an object, idea, or event. For example, at least 3 kinds of judgments are associated with the challenge, 'What should be done about racial discrimination?' The meaning of racial discrimination would have to be clearly understood to recognize the empirical evidence of racial discrimination. If empirical evidence was confirmed, judgment of the value of racial discrimination would be appropriate. Then inquiry could be conducted into what should be done about racial discrimination.
**Judgments about meaning.** Judgments about meaning may be associated with a critical challenge and may involve matters of intent, matters of definition, or matters of relationships. Judgments about meaning imply responsibility for the claimant and for the interpreter—the claimant to be understood and the interpreter to understand. One may have to select among a range of definitions the definition most applicable to a situation. Relational meaning claims found in conceptions or conceptual schemas are understood in terms of how parts relate to each other and how the meaning of one part can impact on the meaning of other parts (e.g., critical thinking). Judgments about a problematic situation are not warranted unless one is clear about the problem (Paul, 1993). Knowledge of context is important to understanding concepts across time and place as the words we use and the meanings we ascribe to them can change (Lipman, 1988; Resnick, 1989). Meaning is fundamental to all judgments for as Coombs (1987) argued, unless we are clear about the concepts we use, our reasoning may well be defective.

**Judgments about logical relations.** Pinto et al. (1992) suggest that judgments about logical relations are largely a matter of evaluating relationships among propositions, whether the truth of one has implications for another. Part of understanding these relationships is understanding the meaning that specific relational terms impart to the argument. For example, if one understands the
meaning implied by relational terms such as disjunctions (either, or), conjunctions (and, also, yet), and conditionals (if, then), one is better able to judge the intent of claims and what this implies about the conditional truth of propositions, including premises and conclusions. Relationships of equivalence, entailment, consistency, and incompatibility also have implications for deciding the strength of logical relationships and what they imply about the truth of the conclusion. For example, if the premises are known to be true, and if the conclusion is entailed in the premises (a valid relationship), then the conclusion would necessarily be true (a sound argument). Scriven (1976) suggested that such knowledge is “the main justification for improving one’s reasoning skills” (p. 58). It should be noted that the nature of logical relationships implies nothing about actual truth of premises or conclusions—only that if some claims are true or not true, this implies something about the truth or falsity of other propositions. Decisions about the accuracy or truth of premises requires judgments about empirical claims, necessary to any adequate evaluation of arguments.

Judgments about empirical or descriptive claims. Empirical or descriptive claims have to do with what was, is, or will be the case about reality or the state of the world, decisions that are theoretically verifiable through the senses. Judgments about an empirical claim concern whether evidence for a claim warrants belief. Inquiry into empirical claims might be initiated with the
questions, "Is it true or is it so?" and then "How do you know?" (Nosich, 1993a). The fallibility of human inference is evident in Pinto et al.'s (1992) question, "Is it plausible, given what is known?" Inquiry into a generalization would require evaluation of evidence for discrete cases, followed by evaluation of the collective strength of all the evidence as it relates to the general claim (a judgment of inferential strength). Even if evidence consistently supports a claim, it would not necessarily follow that contradictory evidence is lacking. It has been suggested that a more fruitful strategy is trying to disprove a generalization by seeking counter-evidence or counterfactuals (Scriven, 1976), or as in hypothetico-deduction, trying to falsify a tentative causal statement (Yore, 1992). This strategy must be viewed in the same light, however, as 'confirming' evidence. Just as we need to remain vigilant for evidence that counts against claims that have been accepted as true, so too we must remain open to the possibility that falsification evidence is not absolute (Stevenson, 1987). Conclusions should be qualified according to their perceived strength, recognizing that knowledge of absolute certainty may not be possible. This is not to suggest that any answer is equally acceptable. Empirical claims are attempts to objectively describe reality and are not a matter of opinion or value (Ennis, 1996). The consequences of accepting any point of view or using consensus as a criterion are unacceptable. Consider the consequences of leaving navigational judgments to a pilot's
preference or a committee of passengers—a potential air disaster! Clearly, any answer will not do.

**Judgments about value.** Judgments about value concern the worth, merit or desirability of objects, ideas, or actions. Critical judgment might focus on objects if the resulting decision is important to someone (e.g., structural strength of aircraft metal relative to airworthiness and the safety of travelers). Judgments of the value of ideas might include criteria—the ideal or goals we strive for (e.g., freedom, justice, critical thinking, democracy, success, credibility, precision, friendship, peace, wealth, clean air, fun). Plans about what to do to attain such ideals or goals might be judged for relative value, as might the action taken to implement the plans. Judgments about what to do might be planned in advance of an action, during an action, or following action. Judgment-in-action as during a performance or production might appear to lack conscious deliberation. The degree of critical thoughtfulness will depend on the prior knowledge and experience of the performer or the time available for monitoring application of appropriate intellectual resources. Plans for a range of potentialities might be considered in advance of a performance, facilitating availability of options and potential quality of judgments-in-action.

A critical challenge may require a range of judgments, depending upon what is necessary to thinking through the challenge to a sustainable conclusion.
Each embedded judgment involves engagement in cognitive tasks and utilizes intellectual resources. The degree of critical thoughtfulness depends on how well cognitive tasks and intellectual resources have been integrated (see Figure 2).

**Cognitive Tasks**

This critical thinking framework attempts to avoid the 'pitfalls' of many linear step-wise skills models of critical thinking while still recognizing that cognitive tasks are involved. The delineation of cognitive tasks in this framework is *not* intended as a general claim of what actually occurs in someone's mind. Rather, the tasks represent four cognitive purposes that are involved in responses to critical challenges toward judgments about what to believe or what to do: self-regulation, deliberation (interpretation, analysis, inference, evaluation), judgment, and justification. The response to a critical challenge will likely be recursive, with movement among the different tasks as needed.

**Self-regulation**

Responses to critical challenges may vary, depending in part on the control that one has over one's thinking—a metacognitive or self-regulatory enterprise that aims to improve one's thinking. Self-regulation is afforded extra prominence in this conceptual framework because of the executive control function it can serve and for the potential it might have for enhancing thinking (Facione, 1991; Jacobs & Paris, 1987; Wang, Haertel, & Walberg, 1993-1994).
Figure 2. Critically thoughtful response in an integrative framework.

Critically thoughtful response involves metacognition—knowledge and self-regulation of intellectual cognitive tasks, which include self-regulation, deliberation, judgment, and justification.

Metacognition—knowledge and self-regulation of cognitive tasks and goal, which include background knowledge about critical thinking concepts, criteria & standards, habits of mind, and thinking strategies.

- Background knowledge
  - about problematic situation
  - relevant context
- Critical thinking concepts
- Criteria & standards
- Habits of mind
- Thinking strategies
This framework extends metacognitive awareness to declarative, procedural, and conditional knowledge of distinctions among kinds of judgments, cognitive tasks, and intellectual resources. Self-regulation of one’s intellectual response to a critical challenge would involve planning, monitoring, and adjusting responses as necessary to ensure a good fit among kinds of judgments, cognitive tasks, and intellectual resources. The success of self-regulation may hinge on the adequacy of background knowledge of the situation and of critical thinking; the proficiency acquired from experiences, particularly in sharing and evaluating justifications of relative merit; the quality and timeliness of evaluative feedback received; and commitment toward an evaluative view of knowledge and the action it implies.

**Deliberation**

Deliberation includes thinking through a problematic situation toward a conclusion or judgment about what to believe or do. Deliberation may be directed toward interpretation—translating or communicating intent, analysis, inference, and evaluation. These are recursive and interdependent deliberation dimensions. Distinctions among these components are not always clear. They are differentiated here to foster a sense of different aspects of deliberation. Overlap between sub-tasks of deliberation and kinds of judgments has been difficult to avoid in this integrative endeavor. For example, interpretation has a counter-part
in judgments about meaning, analysis shares some qualities of judgments about logical relations, and evaluation has a partner in judgments of value.

**Interpretation and analysis.** An essential aspect of early deliberation is to understand the intent of a critical challenge. This may involve first recognizing that a problematic situation exists and then interpreting the intent of the problematic situation by examining explicit concepts, constituent concepts, and associated concepts in order to better understand how they relate to each other. Clarification continues throughout deliberation in a recursive manner. In this framework, the sub-task of interpretation ideally strives for an accurate representation of intended meaning, not speculation about what *might* be intended (an inferential matter), while recognizing the actual intent may be elusive. The adequacy of meaning interpretation depends much on selecting and applying appropriate intellectual resources to acquire adequate information to judge meaning and to withhold or qualify judgment until clarity is achieved.

Analysis of relationships is not distinct from interpretation. One may be better able to grasp intent by analyzing how ideas relate. Some kinds of analyses include identifying parts and how they relate to the whole and to each other, as in an argument or a conceptual network; considering patterns in data to decide if any theories or limited generalizations are implied; considering similarities and
differences between at least two alternatives, and inquiring into the fit between an interpretation of intent and actual intent or accepted usage.

**Inference.** An inference is made when prior knowledge, evidence, experience, and alternatives are considered and conclusions are drawn based on this available information. Deliberative inferences might involve inferring meaning where explicit information is inadequate; inferring what information might enhance judgments; inferring alternatives and implications; and inferring a plan to help realize and represent a sustainable judgment (Facione, 1991). For example, claims about unobserved events, predictions, or hypotheses generated from existing knowledge are inferential products.

**Evaluation.** Evaluation, central to the purpose of critical thinking, is ongoing throughout critical responses. Whenever one decides about the positive or negative value of support for some claim or action, an evaluation occurs. Defensible evaluation involves comparison of an idea, action, or event with an ideal (the desired criterion) to decide whether or to what degree the ideal is evidenced, a matter of standards. When one decides to investigate a claim or action, one has evaluated the effort as justified because the implications of what one believes or what one does are sufficiently important to someone.

If the impact of judgments matter, then all evaluations are not equal. Importance, intent, clarity, fit, consistency, relevance, sufficiency, accuracy,
plausibility, credibility, strength, validity, acceptability, effectiveness, efficiency, and sustainability are just some examples of criteria that might pertain to what is being evaluated. Appropriate selection of criteria is core to the quality of an evaluation (see Intellectual Resources for some related distinctions).

**Judgment**

A conclusion is inferred from relevant information when a judgment is made about what to believe or what to do about a critical challenge. This concluding judgment might represent the end product of many smaller decisions that occur throughout deliberation and which lead to a more global determination directly related to the judgment. Conclusions should be appropriately qualified to fit the degree of confidence that is justified by the support for the judgment.

**Justification**

Justification provides an argument that addresses the power of the grounds (reasons, evidence, data, and so forth) to sustain the conclusion (Facione, 1991; Siegel, 1992). If one shares a conclusion, one has a responsibility to be prepared to share the grounds for that judgment, or as Toulmin (1956) said, what one has to go on. Justification requires evidence or reason(s) to support the conclusion and reasoning that shows that the evidence or reason(s) sustain the conclusion. It might be necessary to qualify the conclusion to reflect the limitations suggested by discrepant evidence or by inadequate deliberation. Justification is fundamental to
an evaluative view of knowledge. A public justification or argument provides an appropriate vehicle for instruction and assessment of critical thinking.

**Intellectual Resources**

The integration of appropriate intellectual resources with cognitive tasks is important if responses to problematic situations are to involve critical thinking. These intellectual resources include commitment to sustainable judgment, evident in consistent striving for and using (1) background knowledge about the situation and the context to which it pertains, and (2) metacognitive knowledge about critical thinking—the distinctions and relationships among critical thinking concepts, criteria and standards, habits of mind, and thinking strategies as they relate to the different kinds of judgments and cognitive tasks.

**Background Knowledge**

Adequate background knowledge relevant to the problematic situation and the context to which the challenge applies contributes to a more comprehensive foundation for judgments. Relevant domain-specific knowledge and contextual considerations about culture, place, time, constraints, conditions, or extenuating circumstances might make a difference to interpretations, premises, points of view, underlying assumptions, options, implications, and values (Lipman, 1988).
Critical Thinking Concepts

Knowledge of the many concepts associated with critical thinking allows one to better generate, judge and justify the quality of a judgment (e.g., points of view such as self-serving, moral, legal, economic, and aesthetic perspectives; reason, evidence, conclusion, and justification; criteria and standards; observation and inference; valid and sound arguments; inductive and deductive arguments; necessary and sufficient conditions; cause and effect). Self-regulation of thinking to improve thinking requires understanding the critical attributes that distinguish these concepts. This knowledge is requisite to deciding how members of a conceptual cluster relate to each other. For example, one needs to distinguish what criteria is appropriate to specific kinds of judgments if one is to reflect on, evaluate, and justify thinking quality. Knowledge of what counts as evidence is requisite to judging evidence quality. The distinction between observation and inference is important to judging the strength of empirical evidence.

Criteria and Standards

Criteria provide the basis for evaluation—the attributes, dimensions, or desired ideals to be measured; standards provide a measurement of how much of a criterion is evident—the benchmarks, performance levels, or degree that a criterion has been met (Case, 1997). Criteria evolve in societies to serve functions, needs, and wants (Pinto et al., 1992). More specifically, criteria are
exemplified in particular laws, codes, customs, rules, guidelines, procedures, specifications, moral principles, attributes, and expectations that allow for comparison and accountability of related judgments (Lipman, 1988). Case (1997) stated that criteria may be descriptive—empirically observable expected elements (e.g., Provide 3 reasons for your decision.); or qualitative—the qualities that elements should exemplify (e.g., Reasons should be relevant and sufficient to support the decision.). Lipman (1988) delineated different senses of standards—informal comparison measures (e.g., The smog is worse than yesterday.) and more formal comparison measures. Formal measures include (1) instruments and standardized units set by accepted authorities (e.g., weights and measures); (2) an ideal comparison measure (e.g., perfect or error-free performance); (3) a minimalist comparison measure (e.g., the minimum grade average required for university admission); (4) a normative standard for comparison (e.g., one student’s average score is compared to the class mean score); or evaluative comparisons against established benchmarks depicting a range of levels of achievement for desired criteria. Case (1997) indicated that standards may be binary, with only two options (e.g., accurate or inaccurate; valid or invalid), or they may provide a range of measurements (e.g., unsatisfactory to outstanding; superficial to comprehensive; weak, moderate, or strong; or 1 to 10).
Criteria for meaning judgments hinge on what is being evaluated. Communications of intent are evaluated for clarity to recipients. Criteria for success would be the accuracy of recipients’ interpretative responses compared with the intended meaning. Evidence of success could be the claimant’s confirmation of such alignment or the interpretation’s congruency compared with other claimant’s claims. If the intent of specific words or concepts is unclear, one could check for evidence of accepted use of the word with competent language users who ideally would have knowledge of the relevant context (e.g., a comprehensive dictionary or relevant reference book). Criteria for successful definitions would be their accurate reflection of the situation and how the word is used. Representations of conceptions, such as critical thinking, can be judged for the strength of internal relationships—the relevancy, sufficiency, and coherence of the attributes—and for the adequacy of attributes themselves (e.g., Do the attributes accurately depict the real-life concept referent?). Interpretations of concept attributes and definitions can be tested for accuracy when compared with concept referents. Attributes that distinguish examples and non-examples could be identified, tested against further samples, and adjusted to better represent the concept referent.

Empirical judgments concern what was, is, or will be the case. We base our judgments on evidence and testimony, striving for truth, an accurate fit
between interpretations and representations of reality and reality itself. However, we might never know how accurately our interpretations and representations reflect reality. *Plausibility* might better reflect the possibility of incomplete evidence and flawed interpretations. Source *credibility*, evidenced by opportunity to know, capacity to know, and reputation for reliability, as well as alignment of source claims with other reputable sources, is theoretically empirically verifiable and contributes to the confidence one might have in a conclusion for which direct evidence is inadequate. Given human fallibility, source credibility does not ensure the accuracy or plausibility of empirical claims.

Judgments about logical relations concern the nature of relationships among propositions and what that relationship implies about the conditional truth of propositions. Criteria for judging logical relations include relevance, sufficiency, consistency, compatibility, and validity.

Judgments about value may apply to objects, ideas, or actions. Judgments about the value of objects can be personal or they may depend on what is valued in the domain by accepted authorities. Related criteria and standards may evolve over time, as different perspectives gain status, established criteria are challenged, boundaries of standards are extended, and understandings become more elaborated. For example, the insistence by art critics on one-to-one correspondence between a painting and physical reality was challenged by
impressionistic representations thought to better capture what the eye actually perceives (Gombrich, 1960). Over time, this challenge succeeded and new criteria accompanied new understandings.

Judgments about ideas might include evaluation of the criteria themselves. For example, truth and plausibility are criteria for judging empirical claims. Justification of their appropriateness as criteria might address the implications if they were not valued. The absence of honesty and scholarship and the removal of the need to warrant claims would leave one with little confidence in what to believe and thus vulnerable to manipulation. It would be difficult to build upon existing knowledge, plan for the future, maintain justice, and sustain society. The ideal of justice is another example, pertinent to judgments about action in varied contexts (law, ethics, family). Justice might be judged for its value by the satisfaction and perceived safety it bestows in contrast to other options, perhaps allowing for more predictability, less chaos, and wiser use of resources. This ideal might be sustained by considering the implications of life without justice—perhaps no life at all. A similar reason could be offered for valuing peace, relevant to local, national, and international contexts.

Judgments about action (e.g., what should have been done, what ought to be done) need to made with comprehensive consideration of who might be impacted by judgment outcomes (an empirical judgment). Each potential outcome
needs to be evaluated for acceptability to those impacted by the judgment. Personal benefit provides a criterion if judgment outcomes have only personal relevance. Moral principles provide criteria for guiding judgments that had, have or, or might have social impact (e.g., individual’s rights to freedom from harm).

Pinto et al., (1992) caution that a comprehensive range of alternatives should be considered and a comprehensive range of consequences should be realistically evaluated.

Point of view is relevant to testing consistency in applying criteria to decisions about what to believe or what to do: the role exchange test, the universal consequences test, and the new cases test (Wright and La Bar, 1987). For example, if there is doubt whether the criterion of justice is applied fairly or consistently, then the role exchange test could be applied. Those who agree that the judgment is just but who are not themselves negatively impacted by it might be asked, “Would you still agree with the judgment if you were negatively affected by it?” The universal consequence test asks whether the consequences would still be acceptable if everybody contributed to the action in question. That is, “what if everybody did it in the same situation for the same reasons?” The new cases test considers parallel cases to determine if application of the criterion was consistently applied. For example, Canadian legislation restricting Chinese immigration in 1900 appears inconsistent with the parallel case of Japanese
immigrants during the same time period who were not similarly restricted (Con, Con, Johnson, Wickberg, & Willmott, 1982).

Plans to realize ideals, values, or criteria such as truth, justice, and peace are judged for potential effectiveness. Efficiency in time, effort, and resources would also be a relevant criterion when judging the relative merit of plans (Pinto et al., 1992).

**Habits of Mind**

Habits of mind are dispositions that reflect consistent mindful tendencies toward thinking through problematic situations in a critically thoughtful manner. Habits of mind represent specific criteria embraced by the umbrella metacognitive self-regulation—monitoring thinking to improve thinking. Some cognitive research indicates the importance of feelings of self-efficacy and the benefits of self-regulation to cognitive responses—suggesting integration among cognition, motivation, and action (Rogoff, 1991; Wang et al., 1993-1994). Recognition of feelings and the reasons for them are important to monitoring thinking through problematic situations in a critically thoughtful way. Emotional dispositions for critical thinking may reflect one’s view of knowledge—whether claims are considered absolute and final, open to any interpretation, or in need of evaluation as some claims are more defensible than others. Dispositions that consistently move one to seek out relevant perspectives, to change one’s conclusions or to stick
to a position in face of opposition if the evidence warrants it; to seek appropriate reasons where possible before deciding what to believe or do; to persevere toward a sustainable conclusion, and to be prepared to justify how the evidence sustains the conclusion, characterize the ideal critical thinker. Critical thinking habits of mind compatible with an evaluative view of knowledge include (1) respect for reason and truth, (2) an inquiring attitude, (3) fair-mindedness, (4) open-mindedness, (5) independent-mindedness, (6) an intellectual work ethic, and (7) respect for others in interactive deliberations (Bailin et al., 1993; Ennis, 1987; 1996; Facione, 1991; Hare, 1993; Lipman, 1988; Marzano, Pickering, & McTighe, 1993; Paul, 1992).

**Respect for reason and truth.** Respect for reason and truth is exemplified in a commitment to consistency in seeking justified beliefs and actions. A person with this habit of mind would assume responsibility to consistently seek and offer evidence or reasons toward sustainable judgment. Reason is viewed as the most sustainable route toward personal and social benefit—restrained and thoughtful rather than compulsive. One would have patience with complexity, realizing that understandings evolve over time. Respect for reason and truth is further demonstrated when one withholds or qualifies judgment if evidence or reason is insufficient for confidence, and by taking a position when it is warranted. One who respects reason and truth would also be
inclined to use and cite credible sources, respecting the more informed position of legitimate experts who have gained their status in ways acceptable to the community to which they belong. This is not to say that experts are always correct. But when one lacks the knowledge to judge and has no reason to favor one argument over another, an expert’s judgment carries more weight than does the less-informed position of a non-expert.

**Inquiring attitude.** The ideal critical thinker recognizes opportunities to use critical thinking, seeks out problematic situations, and pursues his or her curiosity about a wide range of issues.

**Fair-mindedness.** The fair-minded critical thinker would hold an evaluative view of knowledge and would be committed to understanding and giving fair or equal consideration to alternative views, especially seeking out views in opposition to one’s own position. An effort would be made to consider the total situation where possible and relevant points of view belonging to others as if they were one’s own, against the same standards.

**Open-mindedness.** Self-correction is central to open-mindedness (Lipman, 1988). The open-minded person remains open to the potential for new evidence or reason that could count against current beliefs, recognizing the fallibility of human inference. The open-minded person is moved to change beliefs when confronted with more sustainable arguments.
Independent-mindedness. The independent-minded critical thinker has intellectual courage and honesty, a willingness to adopt or stand by a view in face of strong opposition. Humility, the disposition to acknowledge the limitations of one’s thinking, recognizing the tendency one has toward bias, also typifies this habit of mind.

An intellectual work ethic. The ideal critical thinker values quality products and performances and is committed to competency. An intellectual work ethic is evidenced in perseverance to do one’s best in spite of obstacles, striving to be well-informed, pushing one’s own limits of knowledge and abilities, seeking to resolve complex challenges, dealing with complexities in a systematic manner, and in reasoned application of appropriate criteria. Commitment to ongoing self-regulation of one’s thinking to improve one’s thinking exemplifies an intellectual work ethic.

Respect for others in group inquiry or group deliberation. The critical thinker demonstrates respect for others in collaborative inquiry or deliberation situations through sensitivity to those with different levels of knowledge. All participants would be afforded fair consideration of interests and feelings.
Thinking Strategies

Strategies, principles, procedures, and even skills have been erroneously equated. Furthermore, strategies and methods have been incorrectly viewed as sequential, lock-step procedures for solving a problems, paths that if selected, might enhance prospects for success in new but similar situations (Groner, Groner, & Bischof, 1983). Today the recursive nature of thinking through complex problems is recognized (Sternberg, 1995). Heuristics, “from the Greek heuriskin, meaning ‘serving to discover,’ [are currently recognized as action plans or devices that are more flexible than strategies or methods and that represent] only a good bet” (Nickerson et al., 1985, p. 74). Critically thoughtful paths of discovery are pursued with uncertainty (Groner et al., 1983) and represent more than mindless adherence to rules (Scriven, 1976).

Thinking strategies may be individualistic, benefiting some more than others, and they may be situation-specific. General thinking strategies have been criticized for their relative weakness but can be helpful for novice critical thinkers. The thinking strategies listed below and used in this study are examples of such strategies with broad application. More powerful strategies for critical thinking, exemplified in deductive reasoning rules in mathematics, have been described as domain-specific (Bailin et al., 1993; McPeck, 1990; Perkins & Salomon, 1989).
The range of thinking strategies is limited only by human judgment of what strategies work in enhancing the quality of thinking toward sustainable judgment and recognition of new situations that share similar characteristics to which the strategy might be applied. Some thinking strategies that were adapted for use in this study included: individual reflection followed by social interaction to extend ideas and understandings which were deepened by justifying decisions against shared criteria; writing about ideas to gain clarity and deepen understandings; posing counter-examples and alternatives to test the strength of conclusions; using visual organizers such as charts, frameworks, and concept maps to explore relationships; listing pluses and minuses on a t-chart for each option to enhance systematicity and fair-mindedness; using examples and non-examples to generate criteria, test the adequacy of interpretations, and refine criteria; and summarization to help focus on the main idea (Ausubel, 1963; Bailin et al., 1993; Brooks & Brooks, 1993; Caine & Caine, 1991; de Bono, 1985; Hyerle, 1996; McKeown, Beck, Omanson, & Pople, 1985; Paul, 1992; Rivard, 1994; Scriven, 1976).

**Argument**

Proficiency in argument is not separate from proficiency in critical thinking. An argument represents the culmination of deliberation about a problematic situation and includes conclusions, reasons and evidence involved in the decision. An argument may epitomize critical thoughtfulness or it may be
devoid of critical thinking. Its quality will depend on how well the evidence or reasons presented sustain the conclusion. Ideally, an argument includes a justification of the adequacy of conclusion support (Toulmin, 1956), providing explicit evaluation of the evidence or reasons and acknowledging the conclusion's strengths and weaknesses. In short, such an argument provides evidence of cognitive accountability (Lipman, 1988). Facione (1991) reported that arguments “give reasons for accepting some claims [and]...meet objections to the method, conceptualizations, evidence, criteria, or contextual appropriateness of inferential, analytical or evaluative judgments” (p. 18). A justificatory argument provides a framework for representing the results of critical thinking and was central to this study.

Critical Thinking—The Path Makes a Difference

This critical thinking framework proposes that critically thoughtful responses involve the appropriate integration of cognitive psychology’s cognitive tasks and strategies for effective thinking with philosophy’s intellectual resources. The path one takes in response to a critical challenge may lead to naïve contentment, uncertainty, or satisfaction, depending upon the adequacy of such integration (see Figure 3).
Critical Challenge: I don't know what to do about this situation. Which path will help me make the best decision?

I know my decision is the best one for now. It is most acceptable and least harmful to those affected.

START HERE

COGNITIVE TASKS
• Self-regulate
• Deliberate
• Judge or draw conclusion
• Justify

INTELLECTUAL RESOURCES
• Background knowledge
• Critical Thinking Vocabulary
• Criteria and Standards
• Habits of Mind
• Thinking Strategies

I sure worked hard! I hope my decision will be okay. But how can I be sure?

How's that for decisiveness?

IGNORANCE IS BLISS!

QUICK AND DIRTY PATH

UNCERTAINTY
CHAPTER THREE
RESEARCH METHOD

The central purpose of this study was to explore an instructional framework and the related teaching for critical thinking and an evaluative view of knowledge for preservice elementary social studies teachers. Responses from 3rd and 5th year education students were solicited about their disposition to think critically, their view of knowledge, their argument proficiency, and their reactions to the instruction. Descriptions of these variables were central to judging their pretest-posttest status and evidence of change. Data were analyzed in light of critical events and qualitative artifacts in order to better understand results (Preissle-Goetz & LeCompte, 1991).

A single group case study with a pretest-posttest design was used to document students' expressed dispositions toward critical thinking, their views of knowledge, and their argument proficiency; detect any changes in these dispositions, views, and proficiencies; and partially evaluate the effectiveness of intervention. Integration of qualitative data and quantitative data was beneficial to interpretation, allowing insights not apparent in quantitative data. Miles and Huberman (1994) noted that when good quantitative research is "combined with the up-close, deep, credible understanding of complex real-world contexts that characterize good qualitative studies, we have a very powerful mix" (p. 42). The
complexity and mutuality of the educational endeavor was made salient by this more holistic approach and multiple data sources, much of it gathered over the duration of a 39 hour/13-week course within the context of natural classroom activities (Bolster, 1983). Perceptions of participants were more accessible and a more responsive pedagogy was possible. The “complimentarity” of the qualitative and quantitative dimensions of this study allowed for stronger conclusions and ideally, will contribute to more fruitful decisions about future research (Roberts, 1982, p. 289).

Internal variables of history, maturation, testing, and instrumentation could confound this single group pretest-posttest design (Campbell & Stanley, 1963). History or other intervening events between pretest and posttest might allow factors other than the experiment intervention variable to influence critical thinking and view of knowledge. Students were not part of a cohesive group outside of class and major contextual events (e.g., critical thinking workshop or conference) were not apparent which makes this confound less plausible. However, some events had potential to limit prospects for positive change. For example, a reading break was scheduled for the latter half of the week of February 19th and influenced continuity for four students who chose not to attend the 3-hour Monday class in that week. Also, the clustering of major assignments toward the latter half of the course had the potential to raise anxiety, as did the administration
of the posttest interviews so near to final exams by increasing demands on students' time. On-campus job interviews were scheduled at this time, adding to the competition for the attention of some 5th year students. Students' personal situations (e.g., illness, family situations, moving) could also have influenced responses to intervention.

Students' natural maturation between pretest and posttest would also be a potential variable if development occurred systematically without external intervention. A person's view of knowledge, however, while apparently sequentially acquired, does not appear to occur simply as a naturally evolving phenomenon (Kuhn, 1991) and development of critical thinking understandings and proficiencies do not appear easily won (Collins, 1992; Hatcher, 1995; Paul, 1996). Potential for pretest-posttest influence has been investigated in several studies that used Kitchener and King's (1981) Reflective Judgment interview instrument, very similar to Kuhn's model. There was no significant change in the retest after two weeks or six months, even with specific interventions to alter people's views of knowledge (Wood, 1997). Wood noted that 8-10% of subjects may even exhibit regression. Hatcher (1995) reported that 2 semesters were better than 1 semester for realizing positive change in evidence of critical thinking in college students' written composition. In light of such stability in attitudes or resistance to desired change, a one-group design is justified (Borg and Gall, 1983).
A written argument toward the end of the course about a similar social issue allowed students a different modality for expression removed from the performance demands of oral language and from interviewer influence. Reliability of interpretations was enhanced by audio-taping interviews and brief field notes to better ensure reliable representation of students’ responses. Reliability of judgments about students’ view of knowledge and argument proficiency was facilitated by (1) sharing one complete interview analyses with Kuhn, the author of the interview protocol, and responding to her suggestions, and (2) through the instructor-researchers’ 4 complete analyses of interview evidence over 2 years with several partial recursive evaluations to realize consistent judgments and strengthen reliability (Goetz & LeCompte, 1984). Instructor-researcher potential for bias was compensated by consistent discounting the instructional impact throughout the data analysis and striving for accurate qualifications of conclusions (Bogdan & Biklen, 1992; Marshall & Rossman, 1995; Miles & Huberman, 1994). Internal validity was strengthened by using multiple measures to triangulate data, spaced over the 13 weeks of the course.

External validity was limited by a small intact non-random convenience sample from the target population (8 females between 18 and 25 years; 1 female between 35 and 45 years) who selected social studies as a specialization. It was possible that aptitude distinguished this group from others but the range in
perceived performance (93% to 50%) in the social studies assignments suggests a lack of aptitude bias. There are no claims intended for implementation beyond the context of this 13-week course and the 8 participating female students.

The pretest-posttest design includes 2 instruments that provided measures of three dependent variables: The California Critical Thinking Disposition Inventory (Facione & Facione, 1992) for dispositions toward critical thinking and a semi-structured interview (Kuhn, 1991) for epistemological theories and for basic proficiency in argument. The instructional framework provided a basis for interpreting artifacts specific to these students' prospective roles as elementary social studies teachers. For example, students' instructional plans and justifications were compared with the instructional framework to measure their success in integrating critical thinking and instruction in elementary social studies. Students' journals, in-class products, 1 in-class survey, and 2 course evaluations were gathered throughout the study as additional qualitative evidence of their dispositions to critical thinking, their view of knowledge, and their proficiency in argument. The instructor's journal, lesson plans, and instructional materials were collected as evidence of pedagogy and of the instructor's perspective and pedagogical decisions.
Subjects

Subjects in this convenience sample (N=8 females; 3rd year: N=3; 5th year: N=4; teacher recertifying: N=1) were 3rd or 5th year elementary teacher education students at the University of Victoria who elected to enroll in *Approaches to Teaching the Social Studies Curriculum (1-7)*, a 39 hour, 13-week advanced curriculum and instruction course (Education E 446) about current research trends, learning approaches, and instructional strategies as they apply to the elementary social studies curriculum. The students in this sample were from a target population of pre-service elementary school teachers with an expressed interest in social studies. Ten students were enrolled in the course, 8 of whom agreed to contribute to the improvement of a preservice teachers' course for teaching for critical thinking in elementary social studies. Of the 2 non-participating 3rd year students, 1 student volunteer missed the pretests resulting in an incomplete data set for analysis and the other student chose not to participate. Neither of the non-participating students appeared to differ in any obvious way from the study participants. Of the 8 participating females, four 5th year students had taken the basic curriculum and instruction course for teaching elementary social studies; 1 other student was renewing her teaching credentials after some 15 years and could not recall what social studies courses she had taken; 6 students had completed a course about interdisciplinary social science concepts, themes, and processes
related to the elementary social studies curriculum; 3 third year students had not yet taken their professional year and thus had not experienced the basic social studies methods course.

The researcher's background is relevant to comprehending the conceptual and pedagogical decisions in this study. Understandings of critical thinking as it applied to elementary social studies evolved sporadically over several years, initiated when intellectual resources, pedagogical understandings, and the social studies curricula seemed insufficient to facilitate adequate intellectual resources for her class of grade one students. Later, the quest was further provoked by a philosophy professor who challenged a 'bad theory' of critical thinking. A masters’ thesis about macro-implementation of critical thinking in the British Columbia elementary social studies curriculum (1983) revealed a need for more clarity, liaison, and capacity for curriculum workers desiring to promote critical thinking (Ford, 1988). These experiences inspired an earnest search to understand what was missing in a cognitive task approach to critical thinking. Undergraduate education experiences deprived of philosophical studies did not help bring clarity to a conception of critical thinking bereft of dimensions of judgment, criteria, and related concepts. Doctoral studies toward this goal included one directed studies critical thinking course with Jerrold Coombs, co-author of the Bailin et al. (1993) conception of critical thinking; and opportunities to explore understandings of
critical thinking and epistemology in other courses, including historiography, a directed studies course in the Canadian history of education, a directed research course that attempted to synthesize views of critical thinking and apply the resulting framework to a range of curricula; and involvement with two different curriculum and instruction courses for elementary social studies (SNSC 346; Ed E 446). Experiences as sessional instructor for three elementary social studies education courses (Ed. E 746 and Ed. E 446) provided the researcher opportunity to evolve a conceptual framework for teaching for critical thinking in elementary social studies and to explore related instruction. Outside events included participation in a pre-conference workshop and the international critical thinking conference at Sonoma State University; a three-day workshop conducted by Case and Daniels, co-authors of the Bailin et al. (1993) conception of critical thinking; informal involvement as assistant to Roland Case in a secondary social studies curriculum design course; e-mail correspondence with a critical thinking discussion list; and e-mail communications with other experts in the field (e.g., S. Bailin; R. Ennis; S. Norris; H. Siegel; M. Weinstein; P. Woods).

Context

The University of Victoria, situated in the capital city of British Columbia on Vancouver Island, has had degree granting status since 1961. The Bachelor of Education program for elementary education is a five year program (see Table 1).
## Table 1

**University of Victoria Bachelor of Education Program (Elementary Curriculum)**

<table>
<thead>
<tr>
<th>Social Studies-related courses</th>
<th>All students: 5 required Social Studies units</th>
<th>Social Studies concentration</th>
<th>Teaching area specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIST 130 (3 units) (History of Canada)</td>
<td>HIST 130</td>
<td>HIST 130</td>
<td>HIST 130</td>
</tr>
<tr>
<td>ED E 746 (2 units) Curriculm &amp; Instruction in Ele. S. S.</td>
<td>ED E 746</td>
<td>ED E 746</td>
<td>ED E 746</td>
</tr>
<tr>
<td>SNSC 346 (1.5 units) S. S. in the Ele. School</td>
<td>SNSC 346</td>
<td>SNSC 346</td>
<td></td>
</tr>
<tr>
<td>ED E 446 (1.5 units) Approaches in Teaching the S. S. Curriculum</td>
<td>ED E 446</td>
<td>ED E 446</td>
<td></td>
</tr>
<tr>
<td>ED E 438 A (1.5 units) Computer applications in the Instruction of Ele. School, Mathematics, Science, and Social Studies</td>
<td>ED E 438 A</td>
<td>ED E 438 A</td>
<td></td>
</tr>
</tbody>
</table>

**Social Science Electives**
- (e.g., History, Geography, Sociology, Anthropology, Political Science, Economics, Psychology, Philosophy)

**Electives 4.5 units**
- (3 x 1.5 unit course)

**Required:** Geography 101 A (1.5 units) and 101 B (1.5 units); Anthropology (3 units); Sociology 100 (1.5 units)

**Electives (10.5 units)**

---

**Note:** Abbreviations include SNSC (Social and Natural Sciences), S. S. (Social Studies), Ele. (Elementary), ED (Education), and E (SNSC courses).

Regular students enter after year 1 in the Faculty of Arts and Science and the University accepts transfer students from regional colleges all over the province.
for the second, third, or fifth year of students' programs. Admission and retention in the program requires a minimum C+ average. Year four focuses on required preparation courses for curriculum and instructional methods and practical teaching experience, qualifying students for a Standard Certificate upon successful completion. Year five emphasizes preferred teaching specializations, a foundation course, and a curriculum studies course leading to a Professional Certificate and a Bachelor of Education degree. Table 1 illustrates required and elective social studies courses (University of Victoria, 1994-1995).

Measures

Currently, comprehensive instruments for evaluating critical thinking have not yet been developed and thus the use of multiple and repeated measures over time is desirable (Borg & Gall, 1983; Norris & Ennis, 1989). No single objective measure of critical thinking has apparently been widely accepted for social studies (Newmann, 1992). Open-ended responses to objective test questions offer more authentic and valid measures of critical thinking, a prospect not viable with general multiple choice critical thinking tests (Norris & Ennis, 1989). Interviews have the potential to provide more valid information but they could be subject to interviewer influence (Borg & Gall, 1983; Norris & Ennis, 1989). Also, critical thinking measurements need to allow preservice elementary social studies teachers opportunity to apply the critical thinking instructional framework to new,
practically relevant, problematic situations; to exhibit appropriate dispositions, to justify conclusions, and to reveal possible misconceptions about critical thinking as it applies to elementary social studies instruction.

The California Critical Thinking Disposition Inventory (CCTDI) and Kuhn's interview and provided evidence of students' disposition toward critical thinking, their epistemological theories or views of knowledge, and their argument proficiency. The closed standardized feature of the CCTDI and the semistructured nature of the interview protocols allow for cross-study comparisons. Other evidence were produced during the course. Artifacts produced by students provided evidence of students' evolving critical thinking understandings, dispositions, and proficiencies relative to teaching elementary social studies. Instructor's course materials, a personal journal, students' journals, and student course evaluations offered evidence of the relationships among participants' perspectives and responses to each other and to instructional decisions.

The California Critical Thinking Disposition Inventory

The California Critical Thinking Disposition Inventory (Facione & Facione, 1992) was selected as the only apparent critical thinking disposition inventory. Apparent resistance from some students in the pilot classes preceding the present study, considered in light of research that identified similar resistance to disparate epistemological positions (Perry, 1968), suggested that an association
existed among dispositions for critical thinking, view of knowledge, and response to instruction for critical thinking. Triangulation of evidence would help illuminate such associations.

Each of the CCTDI scales, derived from the Delphi Consensus conception of critical thinking (Facione, 1991), were represented by 9 to 12 Likert-type items for a total of 75 items. Face validity of 150 items was deemed acceptable by college instructors. These items were piloted and 75 items were then selected for internal consistency and for discriminating among “a representative sample of undergraduates” (Facione, Facione, & Giancarlo, 1994, p. 17; N=267) tested at 3 comprehensive university settings (one Canadian and two American). Individual items for each scale were interspersed throughout the inventory and labels that might connect items to the intent of the inventory were omitted to reduce potential for socially desired responses.

Cronbach’s Alpha coefficients of internal consistency for individual scales ranged from .71 to .80 (overall value of .91) on the 1992 pilot sample (N=156 undergraduates, high school students, and post-baccalaureates) and .60 to .78 (overall value of .90) on a later 1992 and early 1993 pilots (N=1019 freshmen students, including a group of 587 new academically strong students from a selective, private, comprehensive university). The reliability values are viewed as moderately strong for attitude scales (Borg & Gall, 1983). It should be noted that
the truth-seeking dimension of the inventory had the lowest value (.60). Factor analysis was used to determine items retained in each scale. The authors reported that the alpha reliability and factor analyses of pilot responses “support the existence of seven common (but not necessarily discrete) factors” in the dispositional dimension of critical thinking (Facione et al., 1994, p. 27).

It was recognized that the Likert format of the CCTDI provided limited evidence of actual critical thinking dispositions over time and in various contexts and that some items might be subject to misinterpretation and provide inadequate opportunity to qualify responses (Norris & Ennis, 1989). The latter concern was addressed in the only departure from recommended protocol by having students circle items that they found unclear to help later interpretation of their responses. The CCTDI provided some indication of students’ thoughts about the target dispositions at the time of administration and when triangulated with other data, would facilitate interpretations and conclusions. The pretest-posttest design for the CCTDI offered some sense of change over time.

The CCTDI is separated into seven critical thinking disposition scales including truth-seeking, open-mindedness, inquisitiveness, systematicity, analyticity, critical thinking self-confidence, and maturity. The writings of Facione et al. (1995, pp. 4-8) and Facione et al. (1994, pp. 2-3) offered the following clarification.
Truth-seeking involves intellectual courage to strive for the best knowledge or defensible beliefs, even if it threatens existing beliefs and interests; honest and objective inquiry; and willingness to change views if evidence warrants it. Sample items with which a truth-seeker would strongly disagree include “If there are four reasons in favor and one against, I’ll go with the four” and “To get people to agree with me I’d give any reason that worked.”

Open-mindedness involves awareness of one’s own bias, striving to understand the views of others, and tolerance. Sample items include “It’s important to me to understand what other people think about things” and “It concerns me that I might have biases of which I am not aware.”

Analyticity involves alertness to problematic situations, valuing the use of reason and evidence to solve problems, anticipation of potential problems in meaning or consequences of ideas or actions, alertness to the need for intervention, and perseverance in spite of difficulties. Related items include “It bothers me when people rely on weak arguments to defend good ideas” and “People need reasons if they are going to disagree with another’s opinion.”

Systematicity is represented by one who values an organized, focused, diligent, and careful approach to solving problems. This disposition is assessed in the item “I always focus on the question before I attempt to answer it” but not in the item “My problem is I’m easily distracted.”
**Critical thinking self-confidence** relates to one's perception of and trust in one's capacity for critical thinking, which may or may not correspond with reality. This self-confidence may inspire one to lead others in thinking through problems to rational solutions. Sample items include "Tests that require thinking, not just memorization, are better for me" and "I take pride in my ability to understand the opinions of others." The authors suggest that fluctuations in critical thinking self-confidence might be induced by dissonance and reflect developmental progression, akin to Perry's (1968) transitory positions of intellectual development. Such oscillation might reflect coping responses to the apparent insecurity and frustration that may follow challenges to personal epistemological positions.

**Inquisitiveness** is a measure of intellectual curiosity, the desire to learn even when knowledge applications are not apparent. Such curiosity is evident in items "No matter what the topic, I am eager to know more about it" and "Learn everything you can, you never know when it could come in handy."

**Maturity** involves judiciousness—a balanced and discriminating approach to making, withholding, or revising judgment. The authors suggested that this scale reflects epistemological development or cognitive maturity, recognition that some problems may have more than one plausible alternative conclusion, that conditions might preclude certainty, and that there may be need for closure in the
absence of complete knowledge. The critical thinking mature person would strongly disagree with the items "The truth always depends on your point of view" and "Powerful people determine the right answers."

These subscales paralleled some aspects of the habit of mind dimension of the critical thinking instructional framework for this study. For example, the truth-seeking subscale items aligned with respect for truth and reason, independent-mindedness, and open-mindedness (2 items appeared irrelevant). The open-mindedness subscale items corresponded with some aspects of fair-mindedness (e.g., considering the views of others and racial tolerance) and intellectual honesty (some items represented multiplism). Some analyticity subscale items aligned with respect for truth and reason and an intellectual work ethic (the intent of some items is vague). Systematicity subscale items aligned with an intellectual work ethic. The CCTDI description of critical thinking self-confidence did not align with the habits of mind but some individual subscale items related to fair-mindedness, an inquiring spirit, and an intellectual work ethic. Inquisitiveness related to an inquiring spirit. The CCTDI description of the maturity subscale aligned somewhat with open-mindedness and an evaluative view of knowledge regarding potential for certainty. However, the items did not correspond well with the expressed intent. Given CCTDI subscales that were not discrete (Facione et al., 1994) and some problematic subscale items, and given
some correspondence between individual items and most habits of mind (respect for others in group deliberations was not included and habits of mind were not equally represented), it was decided to use the CCTDI as a general ordinal scale for critical thinking disposition.

**Kuhn’s Interview**

Kuhn’s (1991) semistructured interview used realistic problematic situations relevant to social studies to obtain data about epistemological perspectives and basic argument proficiency. The cause-effect focus represented only one form of empirical argument and the basic categories and dimensions of argument proficiencies were recognized as incomplete representations of argumentation (see analysis of data for related descriptions and scoring rubrics). Yet the issues were relevant to social studies, the two variables (view of knowledge and argument proficiency) were applicable to the goals of the study, and the interview had potential for more valid evidence of argumentation than did a multiple choice test (Borg & Gall, 1983; Norris & Ennis, 1989). The Reflective Judgment model (King & Kitchener, 1994) provided more elaborated criteria for students’ views of knowledge than did Kuhn’s interview. However, the Reflective Judgment interview topics were not as appropriate to social studies and only one variable was measured.
Kuhn reported interrater reliability percentages ranging from 83% to 97%. Queried about validity, Kuhn responded that “the instrument has only face validity as a measure of what it is trying to measure...The correspondences in performance across the three topics offer a form of validity...[in the identification of a ] broad construct assessing a kind of cognitive competence not wedded to a single topic” (D. Kuhn, personal communication, April 22, 1997).

The Kuhn interview was conducted over 2 sessions of about 1 hour each and spaced 2 to several days apart. The interviews were individually conducted during the first two weeks of the course (pre-interview), and again in the last week and a half of the semester (post-interview). Students were advised that their input would contribute to the improvement of instruction to teach for critical thinking in elementary social studies, that participation or lack thereof would not be reflected in course grades, and that responses would not be analyzed until after course grades had been submitted. Students were assured of anonymity and asked to sign a consent form authorizing audio-taping of the interviews and the use of interview and class responses as data (see Appendix A for sample consent form). Notes were recorded during the interview to facilitate preservation and interpretation of audio-tapes (Borg & Gall, 1983; Miles & Huberman, 1994). Participants were paid 10 dollars for each of the 4 interviews at the completion of the last interview in recognition of their substantial investment of time and travel.
The interviews focused on 3 different ill-structured problems about the cause of complex social issues, about which even experts cannot agree, and related research on two of the issues. Kuhn selected the topics of why children fail in school, what causes unemployment, and why prisoners return to crime as important and provocative social issues about which subjects would likely have considered and discussed with others. Topics were also thought to represent concerns that varied in relevance to subjects, school failure being closest to personal experience, unemployment perhaps less so, and prisoners return to crime likely the most remote.

The interview scenarios asked the student to determine sufficient cause, evidence to support causal claims, alternative arguments that contrast to original causal claims, counterarguments that would falsify their original causal claims, rebuttals that establish relative merit of their own argument and the counter-argument, and reasons for the degree of confidence they had in their conclusions. Two social problems (school failure and prisoners' return to crime) were placed in the first interview session. In both interviews, students were asked epistemological questions about how they would justify their causal positions, the degree of confidence in their beliefs relative to experts' positions, and perceived adequacy of their own knowledge of each social problem compared to the average person. They were also asked to rate the relative importance of the problem.
compared to a list of other social problems from their own perspective and for society as a whole. The second interview session concentrated on unemployment and then had students evaluate evidence for the 2 previous interview topics of school failure and prisoners' return to crime. For each topic, underdetermined or non-evidence was presented to students as a vignette of events in the life of a student who had failed school and a prisoner who returned to crime, in neither case establishing cause. In addition, overdetermined evidence was presented for each topic that was comprised of a cluster of 3 small studies, each study reporting evidence in support of a different cause, none of which established one cause over the other.

Other Measures

The critical thinking instructional framework provided a model against which most qualitative evidence were compared. Qualitative evidence were interpreted for their degree of alignment with framework conceptual distinctions and relationships (see Figures 1, 2, and 3 for the critical thinking framework and Figure 4 for the instructional framework for critical thinking).

Qualitative evidence included students' instructional artifacts, a written argument, two-way journals, course evaluations, biographical data, group products
Figure 4. Instructional framework for critical thinking

**Self-Regulate**
- plan
- monitor
- evaluate
- adjust

**Intellectual Resources**
- background knowledge
- critical thinking vocabulary
- criteria & standards
- habits of mind
- thinking strategies

for students responding to a critical challenge—a problematic situation requiring a decision about what to believe or what to do

**Critical Challenge:**

**Curriculum Link:**

---

**Dimensions of Critical Thinking Instructional Framework**

<table>
<thead>
<tr>
<th>Judgments</th>
<th>Cognitive Tasks</th>
<th>Intellectual Resources</th>
<th>Teaching Strategies</th>
<th>Student Activities</th>
<th>Assessment &amp; Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>What judgment is needed?</td>
<td>What cognitive tasks are involved?</td>
<td>Which intellectual resources do students need to judge what to believe or what to do?</td>
<td>What will the teacher do to help students acquire the intellectual resources they need to justify the relative merit of their judgments?</td>
<td>What kinds of meaningful activities will students engage in to help them acquire the intellectual resources to justify the relative merit of their judgments?</td>
<td>What format will assessment take that fits with how students learned?</td>
</tr>
<tr>
<td>What does it mean?</td>
<td>self-regulate?</td>
<td>deliberate?</td>
<td>conclude?</td>
<td>justify?</td>
<td></td>
</tr>
<tr>
<td>How do parts fit?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is it true?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How good is it?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What should be done?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>What criteria are students' responses to satisfy?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>What standards will show how well criteria are met?</td>
</tr>
</tbody>
</table>
and records of students' responses, an instructor's journal and instructional materials. Instructional plan assignments (lesson plans and a unit plan with accompanying justifications) for teaching elementary social studies provided evidence of students' distinctions and associations among dimensions of the critical thinking instructional framework (critical challenge, judgments, cognitive tasks, intellectual resources, and instructional ideas) as they apply to the elementary social studies curriculum. Plans and justifications provided relational arguments about the goodness of fit among parts, alignment with the ideal educated citizen, and consistency with course intent and learning experiences. Instructional plans and justifications were evaluated for alignment with the instructional framework and with the conception of critical thinking underlying this study in the context of elementary social studies.

A written argument about school violence produced toward the end of the course provided evidence to compare to the interview responses. This assignment allowed students time to respond after reflection without the influence of the interviewer. Students were to offer their views about cause and provide one or more alternative causes, evidence to support each position, and then evaluate the relative merit of each argument. Conclusions were to be justified for their alignment with evidence and with the quality of thinking that led to their
conclusion. Students were also asked to indicate how much confidence an external reader should have in their conclusions, relative to experts' conclusions.

Interactive journals between subjects and instructor focused on specific assigned topics that were central to instructional experiences (e.g., evaluative judgments, ideas, and concerns; criteria generation and application; responses to readings—the ideas, complexity, and suitability for the course; and reflections on strategies). The instructor asked the students to share their thoughts in their journals anytime throughout the course. They received verbal and written reminders that optional responses were welcome. Ten percent of the course grade was allocated to participation, attendance, and journals. The journal portion of the 10% was to be awarded for completion and some evidence of thoughtfulness, not for views expressed. These journals were intended to facilitate instructional planning; provide indications of understandings, misconceptions, and dispositions; and aid interpretation of other evidence.

Course evaluation forms (1 in-class survey, 1 formal evaluation completed for the university, and 1 instructor-constructed evaluation) provided indications of students' perceptions of the adequacy of instruction, the critical thinking instructional framework, and perspectives of teaching for critical thinking in elementary social studies. Subjects were assured that evaluations would not be read by the instructor until after course grades were submitted. These responses
were compared to the instructor’s personal journal reflections and instructional
decisions and contributed to later judgments about intervention adequacy.

Biographical information (age, education level, academic interests, and
background in the social sciences and critical thinking) that might relate in some
way to subjects’ responses was gathered by questionnaire and confirmed by
interview. Group-produced artifacts (e.g., charts, concept maps) and class
comments recorded by the instructor on chart paper provided some indication of
general class progress. An instructor’s journal, completed as soon as possible
after each class, provided a record of observations and inferences about the
classroom experience. Lesson plans, prepared course materials, and
accompanying notes on decisions and actions that digressed from the lesson plans
were retained to facilitate reflection, interpretations, and judgments.

Procedures

The instructor-researcher’s interpretations of conceptual and pedagogical
literature and of students’ responses before and during the course influenced the
questions guiding this study, pedagogical decisions, and selection of measures. An
overview of intervention and analysis procedures help contextualize the actual
data analysis and conclusions for this study.
Intervention

The original vision of critical thinking and the 13-week course relied on a well-documented framework of critical thinking, constructivist pedagogy, and traditional elementary social studies curriculum. The assumption was that the quality of educational research depends on the quality of its conceptual foundation (Scriven, 1988). Moreover, documentation of the conceptual model and the participants’ responses (instructor and students) to instruction provide evidence for external comparison.

Judgments about educating prospective teachers to teach for critical thinking in elementary social studies evolved concurrently with each new application by the instructor-researcher (e.g., analyzing curricula for evidence of critical thinking; analyzing relevant literature to help elaborate dimensions of critical thinking; discussing interpretations with educators on critical thinking bulletin boards and other internet communications; responding to literature and experts’ claims; and reflecting on reasons for students’ responses to instruction). The instructor’s comprehension of the whole of critical thinking was impacted each time new understandings were recognized. Analysis of preservice teachers’ misconceptions and questions contributed to such understandings. These data aligned with a view of knowledge that is constructed by an individual but enhanced through interactive experiences that facilitate access to other ideas and
alternative interpretations (Yore, Shymansky, Henriques, Hand, Dunkhase, & Lewis, 1998). Ongoing evaluation of the relative merit of alternative positions as they apply in similar and diverse contexts facilitates self-correction and strengthens judgments.

Such data were considered in light of preservice teachers' responsibilities to teach elementary social studies in a critical manner and to promote social studies understanding and critical thinking. These considerations contributed to instruction characterized by increasing simplicity, relevance, and practicality for students. Instruction included (1) making explicit the connections among students' classroom experiences, the curriculum, and the instructional framework; (2) in-depth concept attainment strategies for key concepts; (3) an emphasis on a rationale for an evaluative view of knowledge; (4) student-instructor liaison regarding experience of instruction to maximize effectiveness; and (5) expectations for justification to promote understanding and sustain arguments. An interactive constructivist paradigm for teaching and learning was central to instructional decisions (Yore et al., 1998). Students' current understandings formed the basis for instruction and were extended through small group interaction and whole class sharing during which ideas were generated, discussed, evaluated, and justified. Student's contributions to small group interactions were often utilized as the central focus of large-group discussion, illustration, and assigned
reading. While interactions contributed to the range of ideas available to each student, ultimately students decided for themselves through private reflection which of the range of possibilities was most sustainable.

Intervention involved three phases: preparation, application, and consolidation. Although not discrete, these phases represented different emphases. The preparation phase focused on students constructing a rationale for teaching for critical thinking and an evaluative view of knowledge in elementary social studies. Practical application of the critical thinking instructional framework as it related to teaching, learning and assessment considerations comprised the next phase, involving engagement with and articulation of related conceptual distinctions and associations. The final phase was directed to consolidating key conceptual distinctions and relationships (see Table 2).

Decisions important to all phases concerned rationale for course emphasis, course structure, general strategies, and contextual considerations. Course emphasis was influenced by the instructor’s teaching experiences and pilot studies that revealed evidence of student resistance to instruction for critical thinking in elementary social studies. Students’ comments from the pilot studies included “There is more to teaching social studies than critical thinking,” “[There was] too much educated citizen and justifications,” “too many big ideas,” an “overkill on educated citizen, challenge, [and] inquiry,” and advice to “get rid of actual
Table 2

Sequence of Instructional Events

<table>
<thead>
<tr>
<th>Date</th>
<th>Preparation phase—Lesson foci</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan. 8</td>
<td>Course overview, CCTDI administration</td>
<td>Read: Draft curriculum—The Educated Citizen</td>
</tr>
<tr>
<td>Jan. 10</td>
<td>Educated Citizen-evolve and justify attributes; summarize; evolve-justify criteria for a summary</td>
<td>Journal-generate &amp; justify own profile of Educated citizen; Read: Critical thinking &amp; S.S.</td>
</tr>
<tr>
<td>Jan. 15</td>
<td>Pedagogy for educated citizen</td>
<td>Journal: Educated citizen-implications for teaching ele. S.S.</td>
</tr>
<tr>
<td>Jan. 17</td>
<td>Interpreting meaning-Request strategy; evaluate strategy</td>
<td>Journal: own view-teaching for critical thinking in ele. S.S; Read Critical Thinking as an Ethic</td>
</tr>
<tr>
<td>Jan. 22</td>
<td>Interpreting meaning-critical reading; View of knowledge-implications for education, society, &amp; teaching elementary social studies (ele. S.S.)</td>
<td>Journal: view of knowledge readings; implications-pedagogy; Read fairly opposing world views</td>
</tr>
<tr>
<td>Jan. 29</td>
<td>Context for critical thinking; defensible view of knowledge &amp; critical thinking for public schools; Critical thinking-what it is, what it is not, the framework, &amp; critical challenge criteria</td>
<td>Study set of critical challenges; Read: Link history &amp; literature; epistemological concepts; complete White Jade Tiger</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>Application phase—Lesson foci</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb. 5</td>
<td>Epistemology concepts-implications for history &amp; teaching; critical challenges-sort &amp; justify, generate &amp; justify, evaluate criteria; Concepts-what they are &amp; how to teach-evaluate?</td>
<td>Journals: justify view of critical challenge criteria; justify critical challenge &amp; identify concept to teach; evaluate strategies</td>
</tr>
<tr>
<td>Feb. 12</td>
<td>Unit plans-building on a critical challenge; S.S. kinds of knowledge; pedagogy re-concept of racial discrimination; relate to framework</td>
<td>Ass. 1-prepare draft lesson plan-teaching a concept embedded in a critical challenge</td>
</tr>
<tr>
<td>Feb. 19</td>
<td>Lesson plans for teaching a concept—evaluate, justify, evolve criteria; justifications for lesson plans-evaluate, justify, evolve criteria; Peer evaluation of draft lesson plans</td>
<td>Ass. 1-complete lesson plan and justification</td>
</tr>
</tbody>
</table>

(table continues)
### Table 2 (continued)

<table>
<thead>
<tr>
<th>Date</th>
<th>Application phase—Lesson foci</th>
<th>Assignments</th>
</tr>
</thead>
</table>
| Feb. 26| Select and justify evidence of racial discrimination; judge empirical claims in novel; evolve criteria for credible sources; Jigsaw Strategy-experts inquire into empirical claims | Experts prepare to teach home group; Read opposing views about affirmative action  
Optional: archive visit Feb. 28 |
| Mar. 4 | Jigsaw-experts teach: Do sources support racial discrimination actions? Are sources credible? Distinguish meaning & empirical judgments; Ass. 2 criteria; apply lesson to framework | Ass. 2-lesson plan and justification re-teaching for judging empirical claims |
| Mar. 11| Look back-look ahead; Journals-concerns/ideas; Empirical claims-facts, opinions, judge general claims & primary sources, test cause & effect; evolve criteria for written argument | Ass. 3-begin to build written argument re-social issue |
| Mar. 18| Empirical claims-generalizations, curriculum, & pedagogy; argument criteria & standards; past & present; racial discrimination & moral principles; pedagogy re-judging what to do | Ass. 3-complete written argument; Ass. 4-lesson plan re-deciding what to do; Re-read: views about affirmative action |

<table>
<thead>
<tr>
<th>Date</th>
<th>Consolidation phase—Lesson foci</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar. 25</td>
<td>Sort-justify judgments &amp; criteria; Unit plans-predict &amp; confirm criteria; evaluate sample justifications; Conflict resolution strategy re-issu-affirmative action; evaluate strategies</td>
<td>Unit plans &amp; justifications</td>
</tr>
<tr>
<td>Apr. 1</td>
<td>Administer CCTDI; Critical thinking-generate, evaluate concept maps re-concept map criteria &amp; quality of dimensions; uses of concept maps; course summary; course evaluations</td>
<td>Unit plans and outstanding assignments due April 8</td>
</tr>
</tbody>
</table>

**Note:** Abbreviations include Jan. (January), Feb. (February), Mar. (March), Apr. (April), ele. (elementary), S. S. (social studies), and Ass. (Assignment).

*investigation—just have [the] teacher 'design one.'*

Such resistance to innovation was expected (Perry, 1968). Possible explanations include capacity of instructor for integrating the complexities of
critical thinking with instruction for teaching social studies, compatibility between the instructional approach and students’ prior education, and the characteristics of students attracted to teaching. Nelson (1996) reported of his efforts to teach for an evaluative view of knowledge in science that “students have become more enthusiastic as my goals became clearer and my teaching moves more effective” (p. 61), aligning with the experience of this instructor-researcher. A lack of congruity between instruction and students’ prior educational experience might also contribute to the resistance (Goodlad, 1984). Moreover, adherence to how one was taught is supported by the finding that “people attracted to teaching tend to favor the status quo” (Lortie, 1975, p. 230). Implementation literature also emphasizes building on current beliefs and practice (Fullan & Stiegelbauer, 1991).

Thus the decision to emphasize rationale and understandings central to an evaluative view of knowledge and the framework for critical thinking was influenced by the apparent inadequate description of critical thinking in the literature and the desired goal of education for critically thoughtful citizens. An instructional emphasis limited to conceptual and strategic aspects of critical thinking may be misdirected as one may have these resources but may not be disposed to use them. The nature of an evaluative epistemological position is well aligned with the constructivist perspective of teaching and learning. Thus the time afforded in the preparatory phase to building toward an understanding of
epistemological concepts, a rationale for an evaluative view of knowledge, and the conception of critical thinking implied appeared justified, in spite of the considerable limitations this would impose on the application phase of the study. There would be less time for exploring different kinds of judgments, distinguishing appropriate intellectual resources, and applying understandings to new situations in instructional plan assignments which would be necessarily placed later in the course.

The pilot students' responses to teaching for critical thinking in elementary social studies supported the crucial need to build conceptual understandings as sequentially as possible, and maintain practicality and relevance for teaching elementary social studies. Introductions and readings that began with a theoretical overview of a conception of critical thinking had appeared to overwhelm students. It was decided that embedded concepts needed to be developed if students were to comprehend more complex conceptual relationships central to critical thinking (Bos & Anders, 1990; Gagne, 1966). This decision presented a challenge, given the difficulty in isolating concepts while engaging in realistic critical challenges pertaining to the curriculum when so many concepts were integral to each judgment.

The response was to utilize both immersion and infusion approaches to instruction for critical thinking (Ennis, 1992). The preparation phase provided an
immersion experience of critical thinking where students were actively engaged in critical thinking challenges and judgments about meaning and values relevant to their prospective teaching roles, and in generating and justifying criteria without much direct explication of conceptual distinctions. Foundational concepts of the educated citizen, criteria, justification, and view of knowledge were constructed gradually.

The application phase blended immersion with an increasing emphasis on an infusion mode of instruction in which distinctions and associations among concepts in the critical thinking instructional framework were made explicit (Bangert-Drowns & Bankert, 1990). Students were provided brief summaries of dimensions of the critical thinking framework and an advanced organizer, an instructional framework that was elaborated concurrently with experiences in the application phase (Ausubel, 1963). In a further attempt to reduce the conceptual load, judgments associated with a critical challenge were approached in a logical order, even though this did not necessarily reflect the actual recursive nature of responding to challenges. Examples, non-examples, and sometimes border-line examples were employed for important concepts with expectations that students justify how the samples related to the essence or central tendencies of the concept (Coombs & Daniels, 1991). Concept maps were utilized in the consolidation
phase and students demonstrated and justified evolving understandings of the relationships among dimensions of critical thinking (Novak & Gowin, 1984).

Interactive constructivism provided the global instructional approach employed, sustained by the belief that knowledge is constructed gradually, enhanced through application, depth of inquiry, building connections between existing knowledge and new information, and through capacity, and indeed, obligation, to share and justify the relative merit of competing beliefs and actions (Johnson, 1992; McKeown, Beck, Omanson, & Pople, 1985; Resnick, 1989; Toulmin, 1956; Yore et al., 1998). These expectations were facilitated through instructor modeling, explicit attention to criteria development, application, and justification; as well as through alternating individual, small group, and whole group cognitive engagement. The instructor regularly challenged students responses by posing alternatives for them to consider, asking them to appraise other possible explanations and contradictory evidence, and having students test criteria to see if they continued to best represent students’ evolving ideals and understandings (Nosich, 1993b; Scriven, 1976). For example, criterial attributes of the educated citizen, summarization, and racial discrimination were delineated by students and later revisited to judge their adequacy in the light of new exemplars.
Other general strategies were utilized. Students’ responses were valued in different ways. The instructor used students’ journal entries for planning lessons and also integrated them anonymously into class discussions. Students were also surveyed verbally at times to gauge their responses to readings and strategies used. The instructor monitored these ideas and adjustments were made where appropriate.

Visual reinforcement of ideas was consistently modeled with the use of overhead transparencies, blackboard charts and diagrams, and prepared handouts. Blackboard and chart paper records of students’ responses facilitated extended related discussions. Small group products were recorded on charts or paper strips for sharing and critical response. Additional reinforcement was fostered through re-visiting important ideas, extending ideas to show connections, and sorting activities (e.g., students sorted examples, non-examples, and borderline examples of discrimination and then challenged and justified groupings, indicating why samples were or were not examples of discrimination, thus identifying criterial attributes of the concept).

The instructional approach emphasized the elementary social studies curriculum, embedding critical thinking activities in curriculum goals and core ideas during the application phase. Previous pilot students had expressed a desire for more attention to social studies curriculum content. The in-depth study of a
shared body of knowledge associated with one curriculum general claim (e.g., the
generalization that early settlers in Canada faced challenges provided the
curriculum link for inquiry and this claim was narrowed to a focus on the
experience of the Chinese in British Columbia in the 1880s) was intended to
promote a sense of mastery (Egan, 1992) and inspire a sense of intellectual
community appropriate for the introduction of complex ideas. A common body of
information was readily accessible through the historical novel, *White Jade Tiger*
(Lawson, 1994), providing a springboard for thinking critically about claims and
sources. Additional benefits included the novel’s relevance to British Columbia
elementary social studies curriculum and its British Columbia context. Integration
of the novel with archival evidential sources and relevant secondary references
provided a range of materials that could be evaluated for their relative merit. The
novel was well-suited for establishing meaningful and provocative connections for
students through the experience of the two young main characters and a time-
travel mystery.

Physical and emotional context were also central considerations. Students
were invited to decide on mutually agreeable class meeting times and office hours.
Considerable deliberation eventuated in one weekly 3-hour session rather than the
scheduled bi-weekly 1.5 hour classes (each 1 hour University of Victoria class
involves an actual 50 minutes of class time). To accommodate students’ wishes
for one weekly class, physical space changed from alternating between two classrooms for the first 4 classes to assignment to a large basement science laboratory with long fixed tables and a podium for demonstrations. The instructor tried to lessen the physical impediment of the podium by sitting on the same level as the students. The invitation for students’ ideas toward course improvement and responses to them were intended, in part, to enhance emotional climate. Students’ input was valued through the instructor’s written and verbal comments and actual adjustments to instruction and assignments where appropriate. Anticipation of student stress and inadequate class time for sufficient practice to promote secure understandings were addressed by inviting students to build on the instructor’s feedback and revise assignments for higher grades. Further to reducing stress and to foster more elaborated responses, individual deliberation and then collaborative experiences with partners/small groups preceded whole class sharing. Journal responses were encouraged throughout the course but specific assignments were most often optional later in the course to allow students to focus on assignments. In-class time was used to obtain students’ written evaluative feedback when level of concern appeared to accelerate. Individual lesson foci and related assignments are summarized in Table 2. An elaboration of the intervention is found in Appendix B.
Analysis of Data

Analysis of data reflected the research questions and the type of data utilized. The global form of analysis used to judge evidence most closely aligned with Goetz and LeCompte’s (1984) description of typological analysis that involved partitioning observations according to theoretical frameworks (Kuhn’s view of knowledge categories and argument proficiency dimensions; Facione and Facione’s (1992) critical thinking dispositions; and the critical thinking framework for this study) to facilitate consistency. Descriptive evidence were translated into ordinal numerical values and displayed in tables to aid analysis and comparison across students, variables, categories, dimensions, topics, and groups (Borg & Gall, 1983). The number of study participants was too small for meaningful statistical analyses or for comparative value judgments. Frequency counts facilitated more precise analyses and conclusions (Miles & Huberman, 1994). Numbers of central tendencies were used to describe whole group data and facilitate comparison across data (Borg & Gall, 1983).

Criteria for judging students’ responses to formal instruments aligned with established criteria (except for the addition of one criterion for epistemological reasoning in interview evidence—whether or not students offered evaluative comments about evidence quality). Qualitative evidence, where it varied from formal instruments, were interpreted in light of the critical thinking instructional
framework. Disparity in evidence or criteria limited some comparisons. Discrepancies in conclusions across different sources of evidence motivated closer examination for possible explanations.

**Dispositions for Critical Thinking**

The CCTDI provided evidence of *expressed* dispositions toward critical thinking. Pretest-posttest results provided evidence of consistency in expressions of dispositions. Procedures for scoring overall disposition aligned with CCTDI instructions, with the following exceptions. Students had been asked to circle items on the CCTDI that they were unclear about and some students qualified their responses. If a student’s response indicated lack of clarity or was unrelated to the intent of the item, the response was treated as an omitted inventory item. The items omitted were assigned the mean replacement score calculated for the other items within the disposition scale.

The norms for the CCTDI suggested that a total score of less than 280 was seriously weak in expressed disposition toward critical thinking and a total score of 350 or more was a strong expression of critical thinking disposition (Facione et al., 1994). Scores between these poles were interpreted as expression of moderate disposition to critical thinking. These standards were recommended by the authors, based on a 6-point ordinal Likert scale which required students to agree or to disagree, there being no mid-range choice. Strength toward a subscale item
was indicated by a response in the upper half of the scale (4, 5, or 6) whereas weakness was indicated by a response in the lower half of the scale (1, 2, or 3).

The instructional framework and the embedded model of critical thinking provided an informal instrument for measuring other forms of evidence for students' dispositions for critical thinking. Assertions about or manifestations of specific habits of mind in students' journals, the in-class survey, and 2 course evaluations were compared to the framework descriptions.

**View of Knowledge**

Interviews offered evidence of students' view of knowledge and argument proficiency that is fundamental to critical thinking. Kuhn's 3 categories for views of knowledge (absolutism, multiplist and evaluativism) were used in this study. An ordinal scale with numerical scores from 1 to 3 was used, with 3 denoting the more mature epistemological position. The absolutist (score=1) adheres to the potential for expert certainty. The multiplist (score=2) denies potential for expert certainty and does not favor expert judgment over personal certainty. The evaluativist (score=3) also denies potential for expert certainty but recognizes that expert's experience and knowledge may enhance their capacity for more informed judgments than non-experts.

Qualitative evidence were also examined for view of knowledge. Where evidence did not parallel Kuhn's criteria, as was sometimes the case in written
arguments and journal entries, recourse to more elaborated accounts in epistemological research literature allowed for adaptation of criteria to guide judgments (e.g., King & Kitchener, 1994). For example, absolutism (score=1) was inferred from the use of absolute adjectives (e.g., all) and associated universal claims and acceptance of experts’ claims without question (e.g., published sources). Multiplicity (score=2) was distinguished from absolutism by broad acceptance of multiple theories, denial of single causal theories that imply broad generalizations, qualified language that avoids absolute terms, and emphasis on individuals and contexts. Multiplists were distinguished from evaluativists by indications that personal confidence was equal to or greater than that of experts and inattention to relative merit. Evaluativists (score=3), like the multiplist, would deny single causal claims, embrace multiple views, and exhibit careful use of language that avoids absolute terms. The evaluativist was distinguished by qualifications to conclusions to reflect the quality of evidence, expressions of openness to new evidence in a quest for truth, evidence of respect for the informed position of experts, evaluation of the relative quality of arguments by both experts and non-experts, and justification of the relative merit of personal conclusions. Where qualitative evidence did not allow clear conclusions, judgments were qualified or withheld.
**Proficiency in Argument**

Interview and qualitative responses provided evidence of argument proficiency which is integral to critical thinking. Judgments were based on interpretations of Kuhn’s (1991) 3 categories for argument proficiency: (1) argumentation, (2) evaluation of evidence, and (3) epistemological reasoning. Kuhn’s descriptions of these dimensions roughly depicted standards for level of success and lent themselves to an ordinal numerical scale that allowed cross-student comparison. One exception was Kuhn’s epistemological reasoning dimension that did not provide descriptions of the different degrees of epistemological reasoning competence (success, partial success, and unsuccessful). Rating scores were assigned to specific responses that ranged from 1 through 3, 3 denoting success. Interpretations were shared with Kuhn who found them acceptable (personal communication, Sept. 22, 1997).

For interview evidence, mean scores were calculated across topics for each dimension of each category. For example, the category of argumentation included 4 dimensions that were scored: evidence, alternative theory, counterargument, and rebuttal. Students’ responses were evaluated for each dimension within a category across 3 topics (Return to Crime, School Failure, and Unemployment) and then were assigned a dimension mean score that was calculated by summing scores for each topic and dividing the total by the number of topics involved. The mean
score for the argumentation category was calculated by summing mean dimension scores and dividing the total by the number of dimensions (in this case 4). In a few situations where the interviewer misinterpreted protocol instructions and did not gather appropriate data, NA (not applicable) was assigned and that particular case was not included in calculating the mean scores for that dimension.

**Argumentation.** Argumentation included 5 dimensions: major causal theories, supporting evidence, alternative theories, counterarguments, and rebuttals. Causal theories that represented respondents’ views about the major cause of each social issue were not evaluated here as Kuhn expressed doubt about their compatibility with an ordinal scale (personal communication, Sept. 22, 1997). Four dimensions of argumentation were evaluated for this study: (1) evidence, (2) alternative theories, (3) counterarguments, and (4) rebuttals.

1. **Evidence.** Students’ responses were evaluated for the quality of evidence generated in support of their major causal theories. Genuine or successful evidence (covariation, external, and indirect evidence; score=3) impacts directly or indirectly on the correctness of a theory. Covariation evidence encompasses correspondence, covariation, and correlation evidence. Correspondence evidence, the weakest form of genuine evidence, merely mentions both the antecedent and effect without making the nature of their relationship explicit. Covariation evidence provides at least some explicit sense of comparison
and quantification. Cases of varying levels of the antecedent are linked with changes to incidence of effect. Correlation evidence explicitly addresses comparison and quantification, showing that change to the level of the antecedent co-occurs with a change to the outcome; that is, they both occur at the same time or in close succession.

External evidence (positive and negative) goes beyond the antecedent and effect to provide additional evidence that supports the relationship between the antecedent and the effect. For example, positive external evidence draws on other situations in which the antecedent is present and as its level changes, so too does the incidence of the effect, increasing the likelihood that a similar relationship between antecedent and effect applies to the case in question. Negative external evidence involves appeal to some factor external to the antecedent and outcome that is absent, increasing the probability that the antecedent is responsible for the outcome. For example, if a prisoner had been rehabilitated, assume that the prisoner would have a positive disposition toward lawful behavior. If, however, the prisoner does not have such a positive disposition, this would be negative external evidence that the prisoner was rehabilitated, increasing the potential that the prisoner returns to crime.

Indirect evidence (analogy, assumption, and full and partial discounting) bears indirectly on the causal claim. An analogy suggests that a cause-effect
relationship applying in one particular instance also applies to another case in a new context. An assumption invokes a generalization (e.g., it's human nature to respond that way) to account for a particular case. Full and partial discounting provide evidence that negate or reduce the power of a claim.

Partially successful evidence (score=2) might enhance the plausibility of a cause by depicting one or more causal sequences but it does not distinguish cause and effect or suggest a causal connection between cause and effect. Unsuccessful or non-evidence (score=1) includes irrelevant information, supports the outcome rather than the cause, or does not relate to the causal theory.

2. Alternative theories. Alternative theories represent causal claims that, if successful (score=3), contrast with students' major theories and students are clearly willing to argue against them. Partially successful alternative theories (score=2) are absorbed into major theories. Unsuccessful alternative theories (score=1) include non-attempts or responses that do not exhibit contrasting positions.

3. Counterarguments. Counterarguments, if successful (score=3), falsify or reduce the power of students' major theories, as in claims against causal sufficiency or causal necessity. They might also fully or partially discount a major causal theory, showing that the antecedent is not consistently causing the outcome, or diminish the adequacy of methodology and quality of evidence.
Partially successful counterarguments (score=2) include the generation of an alternative theory, which if accepted, would not necessarily impact on the correctness of the original theory. Unsuccessful counterarguments (score=1) which do not impact the students’ major causal theory might include nonattempts, rejection of the possibility of a counterargument, a focus on the effect rather than its causal antecedent, a focus on the antecedent rather than the antecedent’s causal relationship to the effect, and a focus on the remedy.

4. Rebuttals. Successful rebuttals (score=3) integrate both the major theory and opposing argument in an attempt to establish the relative merit of the major theory. An integrative rebuttal may argue why the major theory is preferable to an alternative theory or reconcile the two theories, showing how one contributes to the other. A more successful integrative rebuttal is directed toward a genuine counterargument, falsifying or reducing the power of the counterargument, thus restoring power to the major theory. Partially successful rebuttals (score=2) involve a simple non-integrative rebuttal that argues against an alternative theory without any mention of how it relates to the major theory. Unsuccessful rebuttals (score=1) include non-attempts, apparent lack of capacity, rejection of the possibility of rebuttals, responses that do not contradict the alternative theory, a contradiction of the major theory, and rebuttals by assertion only.
**Evaluation of evidence.** The evaluation of evidence category included 3 tasks: (1) interpretation of overdetermined evidence, (2) evaluation of overdetermined evidence, and (3) evaluation of underdetermined evidence (evidence terms are stipulated by Kuhn) across 2 topics (Return to Crime and School Failure). Overdetermined evidence portrays 3 separate investigations by 3 different authority figures into 3 distinct causes for 2 different groups of subjects (either 25 prisoners who returned to crime or 25 students who failed). In neither group did any of the 3 separate investigations establish a correlation between cause and effect. Underdetermined evidence was represented by two separate vignettes depicting a few events in the life of a prisoner or in the life of a student. Neither vignette provided grounds for certainty.

1. **Interpretation of overdetermined evidence.** Interpretations of overdetermined evidence are evaluated for evidence of selectivity and selective bias. Evidence of selectivity (neglects full range of causal theories implied by the evidence) and selective bias (theories selected from overdetermined evidence are biased toward personal initial causal theories) count against objective interpretation. Standards for selectivity include: successful responses (score=3) which exhibit low selectivity—either no reference is made to any of the 3 theories or some aspect of each theory is identified; partially successful responses (score-2) which display moderate selectivity—1 of the 3 theories is omitted; and
unsuccessful responses (score=1) that reveal high selectivity if 2 of the 3 theories suggested by the evidence are omitted. Standards for selective bias include: successful responses (score=3) that exhibit low selective bias if students' selection of theories do not favor personal initial theories; partially successful responses (score=2) that exhibit moderate selective bias if most theories selected align with initial theories; and unsuccessful responses (score=1) that exhibit high selective bias if the only theories selected align with initial theories.

2. Evaluation of overdetermined evidence. Overdetermined evidence is evaluated for its adequacy for establishing causal proof. Successful evaluations (score=3) are undeviating in conclusions that the evidence is inadequate for conclusive judgment. Partially successful evaluations (score=2) are inconsistent. Unsuccessful evaluations (score=1) accept the adequacy of overdetermined evidence to establish causal proof.

3. Evaluation of underdetermined evidence. Undetermined evidence is evaluated for the level of certainty it justifies. Successful evaluations (score=3) express low certainty, the only response justified by the inadequate nature of the information to sustain the conclusions. Partially successful evaluations (score=2) indicate moderate certainty. Unsuccessful evaluations (score=1) exhibit high certainty in the adequacy of underdetermined evidence to establish cause.
**Epistemological reasoning.** Epistemological reasoning involves the impact of the overdetermined evidence on students' initial causal theories. Students' responses are examined for modification of initial theories, rejection of evidence, assimilation of evidence as aligning with initial theories, acceptance of evidence as reinforcing or confirming initial theories, or an absence of epistemological reasoning. Responses are analyzed for their relationship to initial theories and the adequacy of evidence. For example, if students generate many initial theories that encompass the theories implied by the evidence, they would be justified in saying that the theories align with their beliefs. If, on the other hand, students posit only one single causal theory, assimilation would not be justified. The only departure from Kuhn's criteria was the inclusion of evaluative comments about specific factors that influence the quality of evidence. As epistemology concerns the nature of knowledge, the nature of the knower, and how the known is justified, it seemed appropriate that epistemological reasoning consider the adequacy of the support for knowledge claims. Successful epistemological responses (score=3) are justified in light of initial theories and they include some specific critical comment on the adequacy of evidence. Partially successful epistemological reasoning (score=2) involves mixed responses, some justified and some not, or scant evidence of epistemological reasoning. Unsuccessful responses (score=1) are unjustified or lacking in any epistemological reasoning.
Individual and group artifacts provided corroborating testimony, aiding interpretation and analysis. The written argument furnished evidence of proficiency in argumentation in a modality removed from direct instructor-researcher influence. Argumentation categories and analysis aligned with Kuhn's interview procedure. The instructional framework provided a means to measure other aspects of argument. Relational judgments embedded in instructional plan assignments and justifications were compared with the framework to see how well the judgment focus of the instructional plan aligned with the tasks, tools, activities, the ideal of the educated citizen, and with the elementary social studies curriculum.

Qualitative evidence provided documentation of dispositions, views of knowledge, and argument proficiency that extended pretest-posttest data. Assertions were anchored to the literature and with experts in the field.

**Perspectives of Participants**

Both students’ and the instructor's perspectives of instruction provided insights into this study’s results. Artifacts (journals, in-class survey, 2 course evaluations, and instructional plan assignments) were examined for expressions of confidence and concern in evolving understandings, responses to strategies about teaching for critical thinking and an evaluative view of knowledge in elementary social studies, and responses to the experience of intervention. Responses were
clustered according to common attributes and counted to determine how representative they were (Miles & Huberman, 1994). Participants' perspectives were analyzed for congruency with each other and with other measures to enrich conclusions about the effectiveness of conceptual and instructional decisions in educating preservice teachers to teach for critical thinking and an evaluative view of knowledge in elementary social studies.
CHAPTER 4
ANALYSIS OF DATA AND RESULTS

The Research Focus

This study explored the changes in elementary teacher education students’ dispositions toward critical thinking, their view of knowledge, and their argument proficiency as revealed by the participants’ responses to the CCTDI and Kuhn’s interview protocol before and after a 13-week elementary social studies course. A written argument completed toward the end of the study paralleled the Kuhn’s interview but provided data from another perspective. Data were analyzed to reflect the following specific research questions:

1. What were the entry-level dispositions for critical thinking, views of knowledge, and argument proficiencies of the preservice elementary social studies teachers sampled?

2. What were the exit-level dispositions for critical thinking, views of knowledge, and argument proficiencies of the preservice elementary social studies teachers and were there changes in these attributes over the duration of the 13-week advanced social studies curriculum and instruction course?

3. Was the written argument evidence consistent with the interview protocol evidence?
4. What did qualitative evidence (instructor’s journal, students’ journals, assignments, course evaluations) suggest about the effectiveness of the intervention from the perspectives of the instructor and students?

**Research Results**

The CCTDI and Kuhn’s interview protocol provided information about students’ entry-level critical thinking dispositions, view of knowledge, and argument proficiency relevant to the first research question. Likewise, the CCTDI and Kuhn’s interview protocol completed after the 13-week course provided information about the students’ exit-level critical thinking dispositions, view of knowledge and proficiency in argument. Comparison of exit-level and entry-level data provided indications of change to address the second research question. The written argument completed in week 11 or 12 allowed triangulation with Kuhn’s interview data for the third question. The instructor’s journal and students’ informal responses (students’ journals, group and individual assignments, course evaluations) provided evidence from the instructor’s and students’ perspectives regarding the effectiveness of intervention in facilitating the stated goals regarding dispositions for critical thinking, view of knowledge and proficiency in argument.

The sample size in this study was too small for meaningful statistical analysis, as was intended. Nor was the sample large enough to make any
comparative value judgments. Relevant research has been provided as a contextual backdrop for the reader to interpret this study’s results.

**Entry-Level Description: Dispositions for Critical Thinking.**

**View of Knowledge, and Proficiency in Argument**

Preservice elementary social studies teachers’ pretest mean scores revealed moderate positive dispositions toward critical thinking (322.5), varied views of knowledge (2.1), and partially successful proficiency in argument (2.1). These mean scores and individual scores for the 8 female students are described to provide baseline profiles.

**Dispositions for Critical Thinking**

The CCTDI pretest mean score of 322.5 for the 3rd and 5th year University of Victoria education students was within Facione et al.’s (1994) moderate positive range of scores (280-350). The University of Victoria students can be viewed in light of the mean scores reported for other university and college students. Facione et al. reported a CCTDI mean score of 304 for 267 American and Canadian undergraduate students in the norming study and a CCTDI mean score of 298.2 for a sample of 587 academically strong first-year university students. Bers et al. (1996) reported a pretest CCTDI mean score of 284.9 for 224 first-year college students.
Overall, data suggested that the University of Victoria education students’ held moderate strength dispositions toward critical thinking. The CCTDI mean score obscured the variation across individuals and subscales (see Table 3).

Individual students’ CCTDI scores ranged from 279 to 359. Student’s specific subscale scores ranged from a low of 32 to a high 59 (out of a possible 60). The mean entry level subscale scores ranged from a low of 40 for truth-seeking to a

Table 3

**CCTDI—Entry-Level Dispositions for Critical Thinking: Cross-Student and Cross-Subscale Scores**

<table>
<thead>
<tr>
<th>CCTDI subscales</th>
<th>Students</th>
<th>T</th>
<th>OM</th>
<th>A</th>
<th>S</th>
<th>SC</th>
<th>I</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>43</td>
<td>46</td>
<td>41</td>
<td>51</td>
<td>40</td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>42</td>
<td>50</td>
<td>42</td>
<td>48</td>
<td>47</td>
<td>46</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>32</td>
<td>44</td>
<td>37</td>
<td>34</td>
<td>38</td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>45</td>
<td>52</td>
<td>51</td>
<td>49</td>
<td>52</td>
<td>59</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>36</td>
<td>50</td>
<td>46</td>
<td>46</td>
<td>38</td>
<td>49</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>40</td>
<td>57</td>
<td>46</td>
<td>45</td>
<td>41</td>
<td>51</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>41</td>
<td>51</td>
<td>45</td>
<td>45</td>
<td>47</td>
<td>51</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>41</td>
<td>52</td>
<td>46</td>
<td>46</td>
<td>48</td>
<td>53</td>
<td>50</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>40.0</td>
<td>50.3</td>
<td>44.3</td>
<td>45.5</td>
<td>43.9</td>
<td>49.9</td>
<td>48.8</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Subscale abbreviations included truth (T), open-mindedness (OM), analyticity (A), systematicity (S), critical thinking self-confidence (SC), inquisitiveness (I), and maturity (M). Subscale scores ranged to 60: ≤ 40 denoted weakness; ≥50 denoted strength.
high of 50.3 for openness, extremes of the range of scores for moderate
dispositional strength.

**View of Knowledge**

Kuhn’s interview protocol provided entry-level evidence on subjects’ views
of knowledge. Absolutists (1.0) were distinguished from multiplists (2.0) and
evaluativists (3.0) by their adherence to expert certainty or potential for expert
certainty about causal theories. Multiplists and evaluativists both denied expert
certainty but they differed in their level of certainty relative to experts. Multiplists
were as certain or more so than experts whereas evaluativists expressed less
certainty than experts.

The mean score of 2.1 masked variability in views of knowledge across
topics and across students (see Table 4). The Return to Crime topic mean of 2.7
reflected 6 evaluativist responses and 1 absolutist response (data were not
obtained for one student) whereas the School Failure and the Unemployment
means of 1.9 were comprised of 3 evaluativist, 1 multiplist, and 4 absolutist
responses for each topic. Students’ predominant views of knowledge (consistent
across 2 or 3 topics) varied: 3 students (B, E, K) displayed only evaluativist views
of knowledge, 1 student (C) displayed mainly a multiplist perspective, 3 students
(A, G, J) displayed mainly absolutist positions, and 1 student (H) displayed only
an absolutist view of knowledge. Overall, the University of Victoria education
students exhibited mixed views of knowledge that were consistent across at least 2 of 3 topics (across 3 topics for 4 students and across 2 topics for 4 students).

Table 4

Entry-Level Cross-Topic, Cross-Student Views of Knowledge

<table>
<thead>
<tr>
<th>Topics</th>
<th>Students</th>
<th>RC</th>
<th>SF</th>
<th>U</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>2.7</td>
<td>1.9</td>
<td>1.9</td>
<td>2.1</td>
<td></td>
</tr>
</tbody>
</table>

Note. Abbreviations included RC (Return to Crime), SF (School Failure), and U (Unemployment). Scores denoted view of knowledge (evaluativism-3, multitivism-2, and absolutism-1). NA indicated that data were inadequate for that item.

The University of Victoria students' data can be viewed in light of Kuhn's (1991) research. While 37.5% of the 8 female students in the present study exhibited a predominantly evaluativist view of knowledge (across 2 or more topics), 8.1% of Kuhn’s main sample of 160 participants balanced for age, gender, and education level displayed a predominantly evaluativist view of knowledge. The percentages of University of Victoria students’ absolutist responses for each
topic ranged from 14% to 50% (Return to Crime, 14%; School Failure, 50%; Unemployment, 50%); their multiplist responses ranged from 0% to 13% (Return to Crime, 0%; School Failure, 13%; Unemployment, 13%); and their evaluativist responses ranged from 38% to 86% (Return to Crime, 86%; School Failure, 38%; Unemployment, 38%). Kuhn's college sample displayed a range of 38% to 56% absolutist responses for each topic (Return to Crime, 38%; School Failure, 39%; Unemployment, 56%); 30% to 41% multiplist responses (Return to Crime, 31%; School Failure, 41%; Unemployment, 30%); and 14% to 31% evaluativist responses (Return to Crime, 31%; School Failure, 20%; Unemployment, 14%). Kuhn's main sample displayed a range of 49% to 65% absolutist responses (Return to Crime, 49%; School Failure, 50%; Unemployment, 65%); 26% to 38% multiplist responses (Return to Crime, 29%; School Failure, 38%; Unemployment, 26%); and 9% to 22% evaluativist responses (Return to Crime, 22%; School Failure, 13%; Unemployment, 9%).

**Proficiency in Argument**

The entry-level mean score of 2.1 for proficiency in argument was a composite of 3 categories: (1) argumentation (2.1), (2) evaluation of evidence (2.4), and (3) epistemological reasoning (1.8). The mean scores obscured variation across category dimensions, topics, and students. Proficiency in argument results follow, as do Kuhn's (1991) data.
Argumentation. The entry-level argumentation mean score of 2.1 was comprised of average scores for (1) generating evidence to support theories (2.3), (2) alternative theories (2.2), (3) counterarguments (2.1), and (4) rebuttals (1.6).

Variability across students (see Table 5) was apparent in students' cross-topic average scores for evidence (1.7 to 3.0), alternative theories (1.7 to 3.0), counterarguments (1.0 to 2.7), and rebuttals (1.0 to 3.0).

Table 5

Entry-Level Argumentation: Cross-Dimension, Cross-Topic, Cross-Student

Level of Success

<table>
<thead>
<tr>
<th>Students</th>
<th>Evidence</th>
<th>Alternative theory</th>
<th>Counterargument</th>
<th>Rebuttal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>SF</td>
<td>U</td>
<td>RC</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>J</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>K</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Topic Average

2.1 2.4 2.4 2.2 2.0 2.5 2.4 2.1 1.8 1.7 1.4 1.6

Note. NA indicated that data were not available for that topic. Topic abbreviations included Return to Crime (RC), School Failure (SF), and Unemployment (U). Level of success scores ranged from 1 to 3, 3 denoting success.
Entry-level generation of evidence (mean score=2.3) was the most successful argumentation task. Genuine evidence was offered by 4 to 5 students for each topic, pseudo-evidence by 1 student for each topic, and non-evidence by 2 or 3 students for each topic. Students' success generating evidence for one topic did not predict success for the other 2 topics. The percentages of University of Victoria students who generated genuine evidence ranged from 50% to 62.5% for each topic (Return to Crime, 50%; School Failure, 62.5%; Unemployment, 62.5%). The percentages of Kuhn's (1991) college sample participants who generated genuine evidence ranged from 53% to 66% (Return to Crime, 61%; School Failure, 66%; Unemployment, 53%) and the percentages of the main sample participants who generated genuine evidence ranged from 39% to 48% (Return to Crime, 39%; School Failure, 48%; Unemployment, 40%).

The entry-level mean score for alternative theories of 2.2 represented 2 to 4 students who were successful for each topic, 1 to 4 students who were partially successful for each topic, and 2 students who were unsuccessful for 2 topics. Alternative theories were least successful for the School Failure topic. Success with a specific topic did not predict success across topics. The percentages of University of Victoria students who generated successful alternative theories ranged from 25% to 50% for each topic (Return to Crime, 37.5%; School Failure, 25%; Unemployment, 50%). The percentages of Kuhn's (1991) college sample
who generated successful alternative theories ranged from 71% to 81% (Return to Crime, 71%; School Failure, 81%; Unemployment, 78%). The percentages of Kuhn’s main sample who generated successful alternative theories ranged from 58% to 69% (Return to Crime, 58%; School Failure, 69%; Unemployment, 66%).

Students’ entry-level mean score for counterarguments of 2.1 included successful counterarguments for 1 to 5 students for each topic, partially successful responses for 1 to 4 students for each topic, and unsuccessful responses for 2 to 3 students for each topic. Each of the 8 students exhibited consistent performance for at least 2 of the 3 topics. The percentages of University of Victoria students who generated successful counterarguments ranged from 12.5% to 62.5% for each topic (Return to Crime, 62.5%; School Failure, 50%; Unemployment, 12.5%). The percentages of Kuhn’s (1991) college sample participants who generated successful counterarguments ranged from 50% to 56% (Return to Crime, 56%; School Failure, 54%; Unemployment, 50%). The percentages of Kuhn’s main sample participants who generated successful counterarguments ranged from 41% to 52% (Return to Crime, 48%; School Failure, 52%; Unemployment, 41%).

Entry-level rebuttals (mean score-1.6) were the least successful of the 4 argumentation tasks with only 2 students generating a successful rebuttal. Cross-topic consistency for individual students was relatively high (5 students exhibited consistent achievement across the 3 topics) but only 1 student was consistently
successful across topics. One to 2 students were successful for each topic, 1 to 3 students were partially successful for each topic, and 4 to 6 students were unsuccessful across topics. The percentages of University of Victoria students who generated successful rebuttals ranged from 12.5% to 29% for each topic (Return to Crime, 29%; School Failure, 12.5%; Unemployment, 12.5%). The percentages of Kuhn’s (1991) college sample participants who generated successful rebuttals ranged from 26% to 39% (Return to Crime, 26%; School Failure, 39%; Unemployment, 31%). The percentages of Kuhn’s main sample participants who generated successful rebuttals ranged from 21% to 32% (Return to Crime, 21%; School Failure, 32%; Unemployment, 22%).

**Evaluation of evidence.** The entry-level mean score for evaluation of evidence was 2.4. Data for three separate tasks, depicted in Table 6, contributed to this mean score: (1) interpretation of overdetermined evidence (2.8); (2) evaluation of overdetermined evidence (2.1); and (3) evaluation of undetermined evidence (2.2).

Interpretation of overdetermined evidence ideally involved objective representation that avoided imposition of personal beliefs and biases. The University of Victoria education students’ entry-level score for each topic for interpreting overdetermined evidence was 2.8, the highest evaluation of evidence response. All 8 students successfully avoided selectivity by representing the 3 theories implied by the overdetermined evidence for the Return to Crime topic and
### Table 6

**Entry-Level Evaluation of Evidence: Cross-Student, Cross-Category, and Cross-Topic Level of Success**

<table>
<thead>
<tr>
<th>Students</th>
<th>Interpretation of evidence</th>
<th>Evaluation of overdetermined evidence</th>
<th>Evaluation of underdetermined evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>SF</td>
<td>RC</td>
</tr>
<tr>
<td>A</td>
<td>2.5</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>B</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>C</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>E</td>
<td>2.5</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>G</td>
<td>3.0</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>H</td>
<td>3.0</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>J</td>
<td>2.5</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>K</td>
<td>3.0</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Topic Average</td>
<td>2.8</td>
<td>2.8</td>
<td>2.3</td>
</tr>
</tbody>
</table>

**Note.** Topic abbreviations included Return to Crime (RC), School Failure (SF), and Unemployment (U). Level of success scores ranged from 1 to 3, 3 denoting success.

5 students were successful for the School Failure topic. Partial success characterized by selectivity in the neglect of one major theory was displayed in the balance of responses. Five to 7 students successfully avoided selective bias (the unjustified weighting of theories that aligned with personal theories) for each topic. The remaining responses were only partially successful because they assigned priority to theories that aligned with personal theories. Eighty-seven and a half percent (87.5%) of the University of Victoria students successfully avoided
selectivity whereas 45.5% of Kuhn’s (1991) college sample and 34.5% of the main sample were successful in avoiding selectivity.

The entry-level task average score for evaluating the adequacy of overdetermined evidence as grounds for proof was 2.1. Variation in topic average scores (Return to Crime, 2.3; School Failure, 1.9) was of interest, given students’ educational background in elementary teacher education. For the Return to Crime topic, 4 students were successful, 2 students were partially successful, and 2 students were unsuccessful. For the School Failure topic, 2 students were successful, 3 students were partially successful, and 3 students were unsuccessful. Considerable variability in individual cross-topic average scores was apparent (1.0 to 3.0). The percentages of University of Victoria students who unsuccessfully expressed high certainty in the overdetermined evidence were 25% for the Return to Crime topic and 37.5% for the School Failure topic. The percentages of Kuhn’s (1991) college sample members who unsuccessfully expressed high certainty in the overdetermined evidence were 59% for the Return to Crime topic and 44% for the School Failure topic.

The entry-level mean score for evaluating the adequacy of underdetermined evidence as grounds for certainty (2.2) was comprised of topic average scores (Return to Crime, 2.3; School Failure, 2.1). For the Return to Crime topic, 4 students were successful evaluating the adequacy of underdetermined evidence, 2
students were partially successful, and 2 students were unsuccessful. For the School Failure topic, 3 students were successful, 3 students were partially successful, and 2 students were unsuccessful. The percentages of the University of Victoria students who unsuccessfully expressed high certainty about the underdetermined evidence were 25% for both topics. The percentages of Kuhn’s (1991) college sample students who unsuccessfully expressed high certainty about the underdetermined evidence were 33% for the Return to Crime topic and 18% for the School Failure topic.

**Epistemological reasoning.** Students’ entry-level epistemological reasoning mean score of 1.8 represented the impact that overdetermined evidence had on their thinking (assimilation, reinforcement, modification, rejection, critical evaluation, or absence of epistemological reasoning) and to what degree their response was justified by the evidence (see Table 7). For the Return to Crime topic, 1 student was successful, 5 students were partially successful, and 2 students were unsuccessful. For the School Failure topic, 6 students were partially successful and 2 students were unsuccessful. Four of the students’ responses were consistent across both topics. All students displayed some form of epistemological reasoning (limited to concerns about generalizability) but evaluation of evidence quality was seldom apparent.
Table 7

Entry-Level Epistemological Reasoning: Cross-Student and Cross-Topic Level of Success

<table>
<thead>
<tr>
<th>Students</th>
<th>Return to crime</th>
<th>School failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>B</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>C</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>E</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>G</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>H</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>J</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>K</td>
<td>2.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Topic Average 1.9 1.8

Note. Level of success scores ranged from 1 to 3, 3 denoting success.

In summary, the education students in the present study exhibited
moderately positive critical thinking dispositions, mixed views of knowledge, and
partially successful argument proficiencies at the beginning of the study.

Epistemological reasoning was only partially successful across students in the
present study with scant evidence of critical evaluation of evidence. The
University of Victoria students' lower view of knowledge scores for the School
Failure topic were of interest, considering that students' chosen professional
specialization was elementary teacher education. Entry-level evidence did not
support Kuhn's suggestion that successful responses for one topic predicted
success for other topics.
Exit-Level Description: Dispositions for Critical Thinking.

View of Knowledge, and Proficiency in Argument

The second research question focused on a description of the University of Victoria education students' exit-level dispositions for critical thinking, view of knowledge, and proficiency in argument as well as determination of change from the pretest profile. The exit-level CCTDI and Kuhn interview group mean scores aligned with modest positive critical thinking dispositions (329.8), mixed views of knowledge (2.3), and partially successful proficiency in argument (2.2). Changes in mean scores were small but positive. Closer examination of individual students' data revealed variability across dimensions, topics and subjects and more sizable changes than were reflected in the mean scores.

Dispositions for Critical Thinking

The CCTDI posttest mean score of 329.8 for the University of Victoria education students in the present study that remained within Facione et al.'s (1994) range of moderate positive scores masked variation across dispositions and across students (see Table 8). Strong dispositional mean scores (≥50) were achieved for open-mindedness, inquisitiveness, and maturity in decision making whereas truth-seeking reflected the weakest mean score (42.4). Individual students' exit-level scores ranged from 288 to 369.
Table 8

CCTDI—Exit-Level Dispositions for Critical Thinking: Cross-Student and Cross-Subscale Scores

<table>
<thead>
<tr>
<th>Students</th>
<th>T</th>
<th>OM</th>
<th>A</th>
<th>S</th>
<th>SC</th>
<th>I</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>45</td>
<td>40</td>
<td>38</td>
<td>47</td>
<td>38</td>
<td>45</td>
<td>53</td>
</tr>
<tr>
<td>B</td>
<td>35</td>
<td>52</td>
<td>51</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>46</td>
</tr>
<tr>
<td>C</td>
<td>35</td>
<td>47</td>
<td>39</td>
<td>32</td>
<td>38</td>
<td>48</td>
<td>49</td>
</tr>
<tr>
<td>E</td>
<td>45</td>
<td>55</td>
<td>53</td>
<td>54</td>
<td>51</td>
<td>57</td>
<td>54</td>
</tr>
<tr>
<td>G</td>
<td>42</td>
<td>46</td>
<td>56</td>
<td>45</td>
<td>54</td>
<td>55</td>
<td>46</td>
</tr>
<tr>
<td>H</td>
<td>44</td>
<td>52</td>
<td>44</td>
<td>40</td>
<td>41</td>
<td>50</td>
<td>49</td>
</tr>
<tr>
<td>J</td>
<td>46</td>
<td>53</td>
<td>44</td>
<td>38</td>
<td>41</td>
<td>51</td>
<td>52</td>
</tr>
<tr>
<td>K</td>
<td>47</td>
<td>55</td>
<td>50</td>
<td>45</td>
<td>48</td>
<td>53</td>
<td>51</td>
</tr>
<tr>
<td>Mean</td>
<td>42.4</td>
<td>50.0</td>
<td>46.9</td>
<td>43.9</td>
<td>45.3</td>
<td>51.4</td>
<td>50.0</td>
</tr>
</tbody>
</table>

Note. Subscale abbreviations included truth (T), open-mindedness (OM), analyticity (A), systematicity (S), critical thinking self-confidence (SC), inquisitiveness (I), and cognitive maturity (M). Sub-scale scores ranged to 60: ≤ 40 denoted weakness; ≥50 denoted strength.

The University of Victoria students’ posttest mean CCTDI score (329.8) can be viewed in light of Facione et al.’s (1994) first-year university students’ mean score (298.2) and the mean scores for two groups of Bers et al.’s (1996) first-year college students (284.9, 279.5). Posttest subscale scores can be similarly considered. For example, the 3rd and 5th year students in the present study attained the following subscale mean scores: truth-seeking, 42.4; open-mindedness, 50.0; analyticity, 46.9; systematicity, 43.9; self-confidence, 45.3;
inquisitiveness, 51.4; and maturity, 50.0. Bers et al.'s posttest subscale mean scores for first year students included: truth-seeking, 33.8; open-mindedness, 40.9; analyticity, 43.7; systematicity, 38.4; self-confidence, 43.8; inquisitiveness, 45.1; and maturity, 41.1. Facione et al. (1994) reported percentages for graduate students' (N=34) subscale scores that were below 40 (truth-seeking, 26%; open-mindedness, 9%; analyticity, 15%; systematicity, 26%; self-confidence, 6%; inquisitiveness, 1%, and maturity, 12%) and above 50 (truth-seeking, 12%; open-mindedness, 35%; analyticity, 29%; systematicity, 21%; self-confidence, 47%; inquisitiveness, 68%; and maturity, 38%). Inspection of Table 8 reveals the University of Victoria students' posttest subscale scores that were under 40 included truth-seeking (25%), analyticity (25%), systematicity (25%), and self-confidence (25%); their subscale scores that were over 50 included open-mindedness (50%), analyticity (37.5%), systematicity (12.5%), self-confidence (37.5%), inquisitiveness (62.5%), and maturity (50%).

**View of Knowledge**

Kuhn's interview protocol was administered toward the end of the course to document students' view of knowledge. The posttest mean score of 2.3 represented slight cross-topic variability, consistency for 2 to 3 topics for each student, and variance in predominant view of knowledge across students (see Table 9). Average scores for the Return to Crime topic (2.3) and the School
Failure topic (2.4) were somewhat higher than for the Unemployment topic (2.1).
Six students exhibited consistent views of knowledge across all topics and 2 students were consistent for 2 of 3 topics. Students' predominant views of knowledge (consistent across 2 or 3 topics) varied: 4 students (B, E, H, K) displayed entirely evaluativist views of knowledge, 1 student (C) displayed a predominantly multiplist position, and 3 students displayed either entirely absolutist views of knowledge (G, J) or a predominantly absolutist perspective (A). Overall, the University of Victoria education students' views of knowledge were mixed and largely consistent across topics.

Table 9

Exit-Level Cross-Student, Cross-Topic Views of Knowledge

<table>
<thead>
<tr>
<th>Students</th>
<th>Return to crime</th>
<th>School failure</th>
<th>Unemployment</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2.3</td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>H</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>K</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.0</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>2.3</strong></td>
<td><strong>2.4</strong></td>
<td><strong>2.1</strong></td>
<td><strong>2.3</strong></td>
</tr>
</tbody>
</table>

*Note.* Scores denoted view of knowledge (evaluativism-3, multiplist-2, and absolutism-1).
These posttest view of knowledge data can be viewed in light of Kuhn’s (1991) data. Fifty percent of the University of Victoria students displayed a predominantly evaluativist view of knowledge (across 2 or more topics) whereas 8.1% of Kuhn’s main sample did so. The percentages of University of Victoria students who displayed an evaluativist view of knowledge ranged from 50% to 62.5% (Return to Crime, 62.5%; School Failure, 62.5%; Unemployment, 50%). The percentages of Kuhn’s participants who displayed an evaluativist view of knowledge ranged from 14% to 31% for the college sample (Return to Crime, 31%; School Failure, 20%; Unemployment, 14%) and from 9% to 22% for the main sample (Return to Crime, 22%; School Failure, 13%; Unemployment, 9%). The percentages of the University of Victoria students who displayed a multiplist view of knowledge ranged from 0% to 37.5% (Return to Crime, 0%; School Failure, 12.5%; Unemployment, 37.5%). The percentages of multiplist responses by students in Kuhn’s samples ranged from 30% to 41% for the college sample (Return to Crime, 31%; School Failure, 41%; Unemployment, 30%) and from 26% to 38% for the main sample (Return to Crime, 29%; School Failure, 38%; Unemployment, 26%). The percentages of absolutist responses displayed for the University of Victoria students ranged from 25% to 37.5% (Return to Crime, 37.5%; School Failure, 25%; Unemployment, 37.5%). The percentages of absolutist responses for Kuhn’s participants ranged from 38% to 56% for the
college sample (Return to Crime, 38%; School Failure, 39%; Unemployment, 56%) and from 49% to 65% for the main sample (Return to Crime, 49%; School Failure, 50%; Unemployment, 65%).

**Proficiency in Argument**

The exit-level proficiency in argument mean score of 2.2 represents the composite of 3 categories: (1) argumentation (2.3), (2) evaluation of evidence (2.4), and (3) epistemological reasoning (1.9). These scores do not reflect the inconsistencies across category dimensions, topics, and students. Posttest proficiency in argument data follows, as do Kuhn’s (1991) data.

**Argumentation.** The exit-level argumentation mean score of 2.3 represented averages for 4 argumentation dimensions: evidence (2.4), alternative theories (2.6), counterarguments (2.2), and rebuttals (1.8). Table 10 depicts student variability in the range of topic average scores for evidence (1.3 to 3.0), alternative theories (1.7 to 3.0), counterarguments (1.0 to 3.0), and rebuttals (1.0 to 3.0).

Exit-level generation of evidence (mean score=2.4) represented success for 7 of 8 students for at least 1 topic. Genuine evidence was offered by 4 to 6 students for each topic, pseudo-evidence by 2 students for both the Return to Crime topic and the Unemployment topic, and non-evidence by 2 students for each topic. Success with one topic did not predict success across topics.
Table 10

Exit-Level Argumentation: Cross-Student, Cross-Category, Cross-Topic

Level of Success

<table>
<thead>
<tr>
<th>Students</th>
<th>Evidence</th>
<th>Alternative theory</th>
<th>Counterargument</th>
<th>Rebuttal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>SF</td>
<td>U</td>
<td>RC</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>J</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>K</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

| Topic Average | 2.3 | 2.5 | 2.4 | 2.6 | 2.6 | 2.6 | 2.1 | 2.1 | 2.3 | 1.9 | 1.9 | 1.5 |

Note. NA indicated that data for that topic was not available. Topic abbreviations included Return to Crime (RC), School Failure (SF), and Unemployment (U). Level of success scores ranged from 1 to 3, 3 denoting success.

While 37.5% of the University of Victoria students generated genuine evidence across all topics, 16.3% of the students in Kuhn’s main sample did so. One University of Victoria student (12.5%) was unsuccessful generating genuine evidence for any topic compared with 28.8% of Kuhn’s main sample participants.

Alternative theories (mean score=2.6) represented the most successful exit-level argumentation dimension. Five to 6 students were successful for each topic, 1 to 3 students were partially successful for each topic, and 1 student was
unsuccessful for 2 of the 3 topics. The percentages of University of Victoria students who were successful generating alternative theories ranged from 62.5% to 75% (Return to Crime, 75%; School Failure, 75%; Unemployment, 62.5%). The percentages of students in Kuhn’s (1991) samples who were successful generating alternative theories ranged from 71% to 81% for the college sample (Return to Crime, 71%; School Failure, 81%; Unemployment, 78%) and from 58% to 69% for the main sample (Return to Crime, 58%; School Failure, 69%; Unemployment, 66%). Fifty percent (50%) of the University of Victoria students were successful generating alternative theories consistently across all topics, compared with 37.9% of Kuhn’s main sample participants. None of the education students were entirely unsuccessful generating alternative theories across topics whereas 8.6% of Kuhn’s main sample participants were unsuccessful generating alternative theories across all topics.

The University of Victoria education students exit-level mean score for counterarguments of 2.2 included successful responses for 2 students for each topic, partially successful responses for 5 students for each topic, and unsuccessful counterarguments for 1 student for each topic. Cross-topic consistency was apparent for all students. Twenty-five percent (25%) of the University of Victoria students were successful generating counterarguments for each topic whereas Kuhn’s participants’ success ranged from 50% to 56% for the college sample
(Return to Crime, 56%; School Failure, 54%; Unemployment, 50%) and from 41% to 52% for the main sample (Return to Crime, 48%; School Failure, 52%; Unemployment, 41%). One University of Victoria student (12.5%) was unsuccessful generating counterarguments for each topic whereas the percentage of Kuhn’s (1991) main sample participants who were unsuccessful generating counterarguments for each topic ranged from 65% to 80% (Return to Crime, 68%; School Failure, 80%; Unemployment, 65%).

Exit-level rebuttals (mean score=1.8) were the least successful of the University of Victoria education students’ argumentation dimensions. One student (C) was successful across topics while 2 students (B, G) were unsuccessful across topics. Three students (E, H, K) were successful for just one topic. These students generated 1 to 3 successful rebuttals for each topic, superseded by the 3 to 5 unsuccessful rebuttals for each topic. The percentages of the University of Victoria students who generated successful rebuttals ranged from 12.5% to 37.5% for each topic (Return to Crime, 37.5%; School Failure, 25%; Unemployment, 12.5%). The percentages of successful rebuttals for Kuhn’s (1991) samples ranged from 26% to 39% for the college sample (Return to Crime, 26%; School Failure, 39%; Unemployment, 31%) and from 21% to 32% for the main sample (Return to Crime, 21%; School Failure, 32%; Unemployment, 22%). The percentages of the University of Victoria students who displayed unsuccessful
rebuttals for each topic ranged from 37.5% to 62.5% (Return to Crime, 50%; School Failure, 37.5%; Unemployment, 62.5%). The percentages of unsuccessful rebuttals for each topic by Kuhn's main sample participants ranged from 65% to 80% (Return to Crime, 68%; School Failure, 80%; Unemployment, 65%).

**Evaluation of evidence.** The exit-level evaluation of evidence mean score was 2.4. It represented task averages for (1) interpretation of evidence (2.8), (2) evaluation of overdetermined evidence (2.2), and (3) evaluation of underdetermined evidence (2.3). Table 11 depicts topic average scores that contributed to these task averages.

The interpretation of evidence task average scores were similar for both topics (Return to Crime, 2.8; School Failure, 2.7). Education students responded to interpretation of overdetermined evidence with partial or full success in avoiding both selectivity and selective bias. Two students were successful across both topics, 5 students were successful with 1 topic, and 1 student was partially successful across both topics. While 81.3% of the University of Victoria students avoided selectivity, 34.5% of Kuhn's (1991) main sample participants avoided selectivity.

The task average score for evaluation of overdetermined evidence (2.2) represented topic average scores of 2.3 for Return to Crime and 2.1 for School Failure. The Return to Crime topic received 4 successful responses, 2 partially
Table 11

Exit-Level Evaluation of Evidence: Cross-Student, Cross-Category, and Cross-Topic Level of Success

<table>
<thead>
<tr>
<th>Students</th>
<th>Evaluation of evidence categories</th>
<th>Evaluation of overdetermined evidence</th>
<th>Evaluation of underdetermined evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interpretation of evidence</td>
<td>RC</td>
<td>SF</td>
</tr>
<tr>
<td>A</td>
<td>2.5</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>3.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>2.0</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>3.0</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>2.5</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>3.0</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>3.0</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>3.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Topic Average</td>
<td>2.8</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Note. Topic abbreviations included Return to Crime (RC), School Failure (SF), and Unemployment (U). Level of success scores ranged from 1 to 3, 3 denoting success.

successful responses, and 2 unsuccessful responses. The School Failure topic had 4 successful responses, 1 partially successful response, and 3 unsuccessful responses. Consistency across topics was displayed by 4 students (B and H were successful; K and J were unsuccessful). Student variability was observed in the range of cross-topic average scores (1 to 3) that represented unsuccessful evaluation of overdetermined evidence for 2 students, success with 1 topic for 4 students, and success with both topics for 2 students. The percentages of
University of Victoria students who unsuccessfully expressed high certainty towards the overdetermined evidence as grounds for proof were 25% for the Return to Crime topic and 37% for the School Failure topic. The percentages of Kuhn’s (1991) college sample participants who unsuccessfully expressed high certainty towards the overdetermined evidence as grounds for proof were 59% for the Return to Crime topic and 44% for the School Failure topic.

The mean score (2.3) for evaluation of underdetermined evidence was comprised of topic average scores (Return to Crime=2.1; School Failure=2.4). The Return to Crime topic received 4 successful responses, 1 partially successful response, and 3 unsuccessful responses. The School Failure topic had 5 successful responses, 1 partially successful response, and 2 unsuccessful responses. Cross-topic consistency was evident for 5 students (3 students were successful and 2 students were unsuccessful). Student variability in topic averages (ranged from 1 to 3) represented a lack of success for both topics for 2 students, success with 1 topic for 3 students, and cross-topic success for 3 students. The percentages of the University of Victoria students who unsuccessfully expressed high certainty in the underdetermined evidence were 37.5% for the Return to Crime topic and 25% for the School Failure topic. The percentages of Kuhn’s (1991) college sample participants who unsuccessfully
expressed high certainty in the underdetermined evidence were 33% for the Return to Crime topic and 18% for the School Failure topic.

**Epistemological reasoning.** Students exit-level epistemological reasoning mean score of 1.9 represented average scores of 1.9 for both topics (see Table 12).

**Table 12**

**Exit-Level Epistemological Reasoning: Cross-Student and Cross-Topic Level of Success**

<table>
<thead>
<tr>
<th>Students</th>
<th>Return to crime</th>
<th>School failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>B</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td>C</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>E</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>G</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>H</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>J</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>K</td>
<td>2.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

**Topic Average**

<table>
<thead>
<tr>
<th></th>
<th>Return to crime</th>
<th>School failure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.9</strong></td>
<td><strong>1.9</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** Level of success scores ranged from 1 to 3, 3 denoting success.

Individual cross-topic average scores ranged from 1.5 to 2.5, reflecting a majority of partially successful responses (Return to Crime—1 student was successful, 5 students were partially successful, and 2 students were unsuccessful; School Failure—7 students were partially successful and 1 student was unsuccessful).

Partial success, characterized by justified openness to the modification of beliefs
and critical consideration of the adequacy of sources and research methodology, was limited from full success by some unjustified claims or an absence of evaluation of evidence. Four students' epistemological responses to overdetermined evidence revealed consistent partial success across both topics.

In summary, the University of Victoria education students' exit-level mean scores revealed moderate dispositions toward critical thinking, student variability in views of knowledge, and either partial or mixed success in proficiency in argument. Epistemological reasoning was partially successful overall with some unjustified responses but little, if any, critical evaluation of evidence.

**Pretest-Posttest Differences**

Comparison of the pretest and posttest scores revealed small but positive change in mean scores for disposition for critical thinking, view of knowledge, and proficiency in argument. Closer inspection revealed more dynamic changes for individual students than indicated by the mean scores.

**Dispositions for critical thinking.** The pretest-posttest difference in CCTDI mean scores was +7.3 (see Table 13). This change represented an increase in 5 subscale mean scores (truth-seeking, +2.4; analyticity, +2.6; self-confidence, +1.4; inquisitiveness, +1.5; cognitive maturity, +1.2) and a decrease in 2 subscale mean scores (open-mindedness, -.3; systematicity, -1.6). Mean gain (+ or -) scores did not fully reflect the variability across students and subscales. The
sum of the 7 CCTDI subscale change scores for individual students ranged from -16 to +36 while individual subscale change scores ranged from -7 to +16.

Variation among students within subscales ranged up to 22 points (-6 to +16).

Table 13

CCTDI—Pretest-Posttest Difference

<table>
<thead>
<tr>
<th>Students</th>
<th>T</th>
<th>OM</th>
<th>A</th>
<th>S</th>
<th>SC</th>
<th>I</th>
<th>M</th>
<th>Total change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>+2</td>
<td>-6</td>
<td>-3</td>
<td>-4</td>
<td>-2</td>
<td>0</td>
<td>+4</td>
<td>-9</td>
</tr>
<tr>
<td>B</td>
<td>-7</td>
<td>+2</td>
<td>+9</td>
<td>+2</td>
<td>+4</td>
<td>+6</td>
<td>+2</td>
<td>+18</td>
</tr>
<tr>
<td>C</td>
<td>+3</td>
<td>+3</td>
<td>+2</td>
<td>-2</td>
<td>0</td>
<td>+3</td>
<td>0</td>
<td>+9</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>+3</td>
<td>+2</td>
<td>+5</td>
<td>-1</td>
<td>-2</td>
<td>+3</td>
<td>+10</td>
</tr>
<tr>
<td>G</td>
<td>+6</td>
<td>-4</td>
<td>+10</td>
<td>-1</td>
<td>+16</td>
<td>+6</td>
<td>+3</td>
<td>+36</td>
</tr>
<tr>
<td>H</td>
<td>+4</td>
<td>-5</td>
<td>-2</td>
<td>-5</td>
<td>0</td>
<td>-1</td>
<td>-7</td>
<td>-16</td>
</tr>
<tr>
<td>J</td>
<td>+5</td>
<td>+2</td>
<td>-1</td>
<td>-7</td>
<td>-6</td>
<td>0</td>
<td>+4</td>
<td>-3</td>
</tr>
<tr>
<td>K</td>
<td>+6</td>
<td>+3</td>
<td>+4</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>+1</td>
<td>+13</td>
</tr>
</tbody>
</table>

Note. Abbreviations included California Critical Thinking Disposition Inventory (CCTDI) and subscales of truth (T), open-mindedness (OM), analyticity (A), systematicity (S), critical thinking self-confidence (SC), inquisitiveness (I), and maturity (M).

The pretest-posttest difference in CCTDI scores for these 3\textsuperscript{rd} and 5\textsuperscript{th} year University of Victoria education students can be viewed in light of Bers et al.’s (1996) data for first year college students. Bers et al.’s study involved two sections of five different first-year courses (2 different English courses and 1 mathematics, natural science, and psychology course). One section from each course formed
the control group and a second section from each course formed the experimental group. The instructors of the 5 classes in the experimental groups had voluntarily participated in a seminar on critical thinking. The seminar had been conducted over 32 weeks during which instructors met weekly during the first semester and often but less regularly during the second semester. The specific nature of the seminar and of one 16-week intervention period was not described, except to state that at least 4 of the 5 courses apparently emphasized application of analytical inquiry.

Ber et al.'s pretest results showed some difference among variables. The only experimental group difference was its higher mean score for the analyticity subscale. Results also showed that in all cases, females, older students, and more educated students had higher CCTDI scores than did males, younger students, and less educated students.

Pretest-posttest change scores for Bers et al.'s experimental groups differed from change scores for the University of Victoria students. Bers et al. found no significant pretest-posttest difference in scores. However, Bers et al. did find considerable positive and negative movement within subscale scores. Some students' scores declined on the posttest. Bers et al.'s data revealed positive change for analyticity (+1.5), truth-seeking (+.8), self-confidence (+.8), and cognitive maturity (+.2). Negative change was found for inquisitiveness (-1.0),
open-mindedness (-.1), and systematicity (-.2). After one 13-week semester of intervention, the University of Victoria students posttest scores revealed positive change for all subscales except for open-mindedness and systematicity subscales for which both Bers et al.' students and the University of Victoria students' scores declined (see Table 14).

Table 14

Cross-Sample Pretest-Posttest Comparison of CCTDI Subscale Mean Scores

<p>| Bers et al. (1996) 1st year students (N=134) | University of Victoria 3rd &amp; 5th year students (N=8) |</p>
<table>
<thead>
<tr>
<th>Scale</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
<th>Scale</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>33.0</td>
<td>33.8</td>
<td>+.8</td>
<td>T</td>
<td>40.0</td>
<td>42.4</td>
<td>+2.4</td>
</tr>
<tr>
<td>OM</td>
<td>41.0</td>
<td>40.9</td>
<td>-.1</td>
<td>OM</td>
<td>50.3</td>
<td>50.0</td>
<td>-.3</td>
</tr>
<tr>
<td>A</td>
<td>42.6</td>
<td>43.7</td>
<td>+1.1</td>
<td>A</td>
<td>44.3</td>
<td>46.9</td>
<td>+2.6</td>
</tr>
<tr>
<td>S</td>
<td>38.6</td>
<td>38.4</td>
<td>-.2</td>
<td>S</td>
<td>46.0</td>
<td>43.9</td>
<td>-2.1</td>
</tr>
<tr>
<td>SC</td>
<td>43.0</td>
<td>43.8</td>
<td>+.8</td>
<td>SC</td>
<td>43.9</td>
<td>45.3</td>
<td>+1.4</td>
</tr>
<tr>
<td>I</td>
<td>46.1</td>
<td>45.1</td>
<td>-1.0</td>
<td>I</td>
<td>49.9</td>
<td>51.4</td>
<td>+1.5</td>
</tr>
<tr>
<td>CM</td>
<td>41.0</td>
<td>41.1</td>
<td>+.1</td>
<td>CM</td>
<td>48.8</td>
<td>50.0</td>
<td>+1.2</td>
</tr>
</tbody>
</table>

Note. Abbreviations included California Critical Thinking Disposition Inventory (CCTDI), and subscales for truth (T), open-mindedness (OM), analyticity (A), systematicity (S), critical thinking self-confidence (SC), inquisitiveness (I), and cognitive maturity (CM). Scores have been rounded off to the nearest tenth.
View of knowledge. Pretest-posttest responses from University of Victoria education students exhibited little change in mean view of knowledge (see Table 15). Five students did not change in view of knowledge (B, C, E, G, K), 1 student maintained the same mean score but displayed some cross-topic inconsistency (A), and 2 students exhibited change (H, J). One student (H) appeared to make a change from absolutism to evaluativism across all topics.

Table 15

Cross-Student, Cross-Topic Pretest-Posttest Difference in View of Knowledge

Scores

<table>
<thead>
<tr>
<th>Students</th>
<th>RC</th>
<th>SF</th>
<th>U</th>
<th>Average Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-2</td>
<td>+2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>H</td>
<td>+2</td>
<td>+2</td>
<td>+2</td>
<td>+2</td>
</tr>
<tr>
<td>J</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-0.7</td>
</tr>
<tr>
<td>K</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note. Abbreviations included RC (Return to Crime), SF (School Failure), and U (Unemployment). Positive change reflected movement toward evaluativism whereas negative change suggested movement toward absolutism.
Another student's (J) responses indicated a change from an evaluativist to an absolutist position for the Return to Crime topic while maintaining an absolutist position for the other two topics.

**Proficiency in Argument.** The pretest-posttest difference in proficiency in argument mean scores (+.1) represented small but positive change to each of the 3 proficiency in argument category mean scores (argumentation, +.2; evaluation of evidence, +.04; and epistemological reasoning, +.1). All 4 argumentation dimensions (see Table 16) exhibited small but positive change in mean scores (evidence, +.1; alternative theories, +.4; counterarguments, +.1; and rebuttals, +.2).

Change scores varied across students, topics, and categories. The sum of the argumentation change scores for individual students ranged from -2 to +9. The mean topic change score across argumentation dimensions ranged from +.8 to +2.3, the highest gain representing the School Failure topic. The mean of topic change scores for each argumentation dimension ranged from +.3 to +2.7 with the alternative theory dimension exhibiting the highest mean change score.

Overall, change in argumentation was small but positive. The most positive difference was apparent for the alternative theories dimension, the School Failure topic, and for 2 students (B, +9; C, +8). The least positive change was apparent for the evidence and counterargument dimensions and for the Return to Crime
Negative change exceeded positive change for 3 students (G, -1; H, -2; K, -2).

Table 16

**Cross-Student, Cross-Dimension, Cross-Topic Pretest-Posttest Difference for Argumentation**

<table>
<thead>
<tr>
<th>Students</th>
<th>Evidence</th>
<th>Alternative theory</th>
<th>Counterargument</th>
<th>Rebuttal</th>
<th>Sum of student change scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>SF</td>
<td>U</td>
<td>RC</td>
<td>SF</td>
</tr>
<tr>
<td>A</td>
<td>-2</td>
<td>0</td>
<td>+2</td>
<td>+1</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>+1</td>
<td>+1</td>
<td>+2</td>
<td>+1</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>+2</td>
<td>+2</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>+2</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>+1</td>
<td>+2</td>
<td>-2</td>
<td>NA</td>
<td>+1</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>0</td>
<td>+1</td>
</tr>
<tr>
<td>J</td>
<td>+1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>K</td>
<td>-1</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Composite Change</td>
<td>+1</td>
<td>+1</td>
<td>0</td>
<td>+2</td>
<td>+5</td>
</tr>
</tbody>
</table>

**Note.** NA indicated that data for that topic was not available. Topic abbreviations included Return to Crime (RC), School Failure (SF), and Unemployment (U). Overall change scores indicated the sum of positive change scores and negative change scores.

Pretest-posttest change in evaluation of evidence mean scores was negligible (+.04), representative of a slight decrease in mean dimension score for interpretation of evidence (-.05) and slight positive increases in mean dimension.
scores for evaluation of overdetermined evidence (+.1) and evaluation of underdetermined evidence (+.05). Examination revealed negative sum change for both topics for interpretation of evidence, and mixed sum change for evaluation of both overdetermined and underdetermined evidence, favoring the School Failure topic (see Table 17). The sum of change scores for students ranged from -2.5 to +3.0. Overall, pretest-posttest mean change for evaluation of evidence was

Table 17

Cross-Student, Cross-Dimension, Cross-Topic Pretest-Posttest Difference for Evaluation of Evidence

<table>
<thead>
<tr>
<th>Students</th>
<th>Interpretation of evidence</th>
<th>Evaluation of overdetermined evidence</th>
<th>Evaluation of underdetermined evidence</th>
<th>Sum of student change scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>SF</td>
<td>RC</td>
<td>SF</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
<td>+1.0</td>
<td>2.0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+1.0</td>
</tr>
<tr>
<td>C</td>
<td>-1.0</td>
<td>-0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>+0.5</td>
<td>0</td>
<td>0</td>
<td>-2.0</td>
</tr>
<tr>
<td>G</td>
<td>-0.5</td>
<td>+0.5</td>
<td>0</td>
<td>+1.0</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
<td>-0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>J</td>
<td>+0.5</td>
<td>-0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K</td>
<td>0</td>
<td>0</td>
<td>-1.0</td>
<td>0</td>
</tr>
<tr>
<td>Composite change</td>
<td>-0.5</td>
<td>-1.0</td>
<td>0</td>
<td>+2</td>
</tr>
</tbody>
</table>

Note. Topic abbreviations included Return to Crime (RC), School Failure (SF), and Unemployment (U).
minimal but differences across dimensions, topics and individuals revealed the most positive change for evaluation of both overdetermined and underdetermined evidence for the School Failure topic and 2 students (A, G), and the most negative change for 1 student (C).

The epistemological reasoning mean score of 1.9 remained relatively unchanged from pretest to posttest for both topics (see Table 18). The pretest-posttest differences were stable (+.0) for 5 students (A, C, E, H, J), increased (+1) for 2 students (B, K), and decreased (-1) for 1 student (G).

Table 18

<table>
<thead>
<tr>
<th>Students</th>
<th>Return to Crime</th>
<th>School Failure</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>+1</td>
<td>0</td>
<td>+1</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>-1</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>J</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K</td>
<td>0</td>
<td>+1</td>
<td>+1</td>
</tr>
</tbody>
</table>

In summary, pretest-posttest difference in mean scores for critical thinking dispositions, view of knowledge, and argumentation proficiency was small but positive. CCTDI change scores ranged from -16 to +36, illustrating individual
variation in exit-level responses. Students’ views of knowledge remained relatively stable, the only positive change found in 1 student’s apparent movement from absolutism to evaluativism. The most apparent changes for proficiency in argument were found in the increased score for alternative theories, increased success for the school failure topic, increased scores for 2 students and declined scores for 2 students. Evaluation of evidence displayed slight increased success for the school failure topic.

**Written Argument Evidence:**

**Consistency With Interview Evidence**

Between week 11 and week 13, the University of Victoria preservice elementary social studies teachers completed a written argument about the social issue of school violence that paralleled Kuhn’s (1991) interview format. The written arguments provided evidence that paralleled 2 of Kuhn’s proficiency in argument categories: (1) argumentation dimensions of evidence generated to support theories, alternative theories, and rebuttals; and (2) epistemological reasoning. Written arguments also provided evidence of or insights about the students’ view of knowledge.

The written argument task asked students to describe their positions about the cause of school violence with supporting reasons, consider alternative views with supporting reasons, evaluate the strength of their arguments relative to
alternative views (including the view of an expert), and justify why their argument deserved serious consideration. Students were advised that the use of sources was optional, as long as their conclusions reflected the quality of their support.

Evaluation of written arguments adhered to Kuhn’s criteria for argumentation claims about evidence, alternative theories, and rebuttals. Epistemological reasoning evidence that paralleled evidence in Kuhn’s study was found in students’ claims about the degree of confidence that one should have in their conclusions, given the quality of support they provided. Successful epistemological responses (score=3) evaluated the quality of support provided as well as its capacity to sustain the conclusion. Partially successful responses (score=2) were inconsistent in identifying and evaluating support or related evaluations inconsistently reflected the degree of confidence that was warranted by the quality of evidence and credibility of source. Unsuccessful responses (score=1) expressed more confidence in conclusions than the support warranted, accepted source claims as fact without evaluation, or they did not justify the adequacy of support for causal claims. Kuhn’s view of knowledge criteria for absolutism, multiplism, and evaluativism did not align closely with students’ written arguments since the written arguments did not explicitly address potential for expert certainty nor students’ confidence level relative to that of an expert as requested. However, students’ explicit claims about the nature of their causal
theories, experts and evidence, and expressions of degree of certainty provided an indication of students' view of knowledge. Evaluativism (score=3) was suggested by claims that denied both certainty and the universal generalizations implied by single causal theories, expressed a level of confidence that did not go beyond the quality of evidence, recognized the need for evidence to support claims, and accepted the relative merit of evidence. Multiplism (score=2) was implied by claims that denied both expert certainty and universal generalizations but unlike the evaluativist, accepted all causal theories as equally viable or refused to accept any specific cause as more likely than another. Absolutism (score=1) was interpreted from claims that promoted one single broadly applicable causal theory about school violence and accepted expert certainty or expressed a high level of confidence in experts' conclusions without consideration of the quality of their evidence.

**Written Argument: Exit-Level Description of Proficiency in Argument and View of Knowledge**

Evidence from written arguments revealed a mean score of 2.4 for proficiency in argument and 2.0 for view of knowledge. The view of knowledge mean score reflected 3 evaluativist positions (B, C, G), 1 multiplist perspective (A), and 3 absolutist viewpoints (E, J, K). Evidence was inadequate to determine 1 student's (H) view of knowledge. The proficiency in argument mean score
masked variation across categories, dimensions, and students (see Table 19).

Three argumentation dimensions (evidence, alternative theories, and rebuttals) had mean scores of 2.6 and epistemological reasoning had a mean score of 1.7.

Students' average proficiency in argument scores ranged from 2 to 2.9.

Table 19

Exit-Level Cross-Student, Cross-Category, Cross-Dimension Written Argument Data

<table>
<thead>
<tr>
<th>Argument Proficiency</th>
<th>University of Victoria education students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Evidence</td>
<td>3.0</td>
</tr>
<tr>
<td>Alternative Theory</td>
<td>3.0</td>
</tr>
<tr>
<td>Rebuttal</td>
<td>3.0</td>
</tr>
<tr>
<td>Epistemological Reasoning</td>
<td>2.5</td>
</tr>
<tr>
<td>Average</td>
<td>2.9</td>
</tr>
<tr>
<td>View of Knowledge</td>
<td>2</td>
</tr>
</tbody>
</table>

Note. NA indicated that data were inadequate for that item.

Comparison of Written Argument and Exit-Level Interview Data

Scores for relevant dimensions of argumentation, epistemological reasoning, and view of knowledge were compared to similar scores from posttest
interview data. Comparison scores are provided in Table 20. Difference scores are provided in Table 21.

Table: 20

University of Victoria Students: Comparison of Exit-Level Interview and Written Argument Data for View of Knowledge and Proficiency in Argument

<table>
<thead>
<tr>
<th>Students</th>
<th>View of knowledge</th>
<th>Evidence</th>
<th>Alternative theories</th>
<th>Rebuttals</th>
<th>Epistemological reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>WA</td>
<td>I</td>
<td>WA</td>
<td>I</td>
</tr>
<tr>
<td>A</td>
<td>1.7</td>
<td>2.0</td>
<td>1.7</td>
<td>3.0</td>
<td>2.7</td>
</tr>
<tr>
<td>B</td>
<td>3.0</td>
<td>3.0</td>
<td>2.7</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>C</td>
<td>2.3</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>E</td>
<td>3.0</td>
<td>1.0</td>
<td>3.0</td>
<td>3.0</td>
<td>2.3</td>
</tr>
<tr>
<td>G</td>
<td>1.0</td>
<td>3.0</td>
<td>2.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>H</td>
<td>(3.0)</td>
<td>NA</td>
<td>2.3</td>
<td>3.0</td>
<td>1.7</td>
</tr>
<tr>
<td>J</td>
<td>1.0</td>
<td>1.0</td>
<td>3.0</td>
<td>2.0</td>
<td>2.3</td>
</tr>
<tr>
<td>K</td>
<td>3.0</td>
<td>1.0</td>
<td>1.3</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Mean</td>
<td>2.1*</td>
<td>2.0*</td>
<td>2.4</td>
<td>2.6</td>
<td>2.6</td>
</tr>
</tbody>
</table>

Note. I denoted Interview data and WA indicated Written Argument data. NA indicated inadequate data for that item. * Base N=7. The mean scores calculated for view of knowledge did not include H's interview score (3) to allow comparison. A discrepancy of .1 in mean difference scores for epistemological reasoning are due to rounding error.
### Table 21

**The Difference of Written Argument Scores Relative to Interview Scores**

<table>
<thead>
<tr>
<th>Students</th>
<th>View of knowledge</th>
<th>Evidence</th>
<th>Alternative theories</th>
<th>Rebuttals</th>
<th>Epistemological reasoning</th>
<th>Argument proficiency sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>+.3</td>
<td>+1.3</td>
<td>+.3</td>
<td>+1.7</td>
<td>+1.0</td>
<td>+4.3</td>
</tr>
<tr>
<td>B</td>
<td>.0</td>
<td>+.3</td>
<td>.0</td>
<td>+2.0</td>
<td>-.5</td>
<td>+1.8</td>
</tr>
<tr>
<td>C</td>
<td>+.7</td>
<td>.0</td>
<td>.0</td>
<td>-1.0</td>
<td>.0</td>
<td>-1.0</td>
</tr>
<tr>
<td>E</td>
<td>-2.0</td>
<td>.0</td>
<td>-3</td>
<td>+3</td>
<td>-.5</td>
<td>-.5</td>
</tr>
<tr>
<td>G</td>
<td>+2.0</td>
<td>+1.0</td>
<td>.0</td>
<td>+1.0</td>
<td>+.5</td>
<td>+2.5</td>
</tr>
<tr>
<td>H</td>
<td>NA</td>
<td>+.7</td>
<td>-.7</td>
<td>+1.0</td>
<td>.0</td>
<td>+1.0</td>
</tr>
<tr>
<td>J</td>
<td>.0</td>
<td>-1.0</td>
<td>+.7</td>
<td>+1.2</td>
<td>-1.0</td>
<td>-.1</td>
</tr>
<tr>
<td>K</td>
<td>-2.0</td>
<td>-3</td>
<td>.0</td>
<td>+.7</td>
<td>-1.0</td>
<td>-.6</td>
</tr>
<tr>
<td>Mean</td>
<td>-.1</td>
<td>+.3</td>
<td>+.0</td>
<td>+.9</td>
<td>-.3</td>
<td>+.9</td>
</tr>
</tbody>
</table>

**Note.** The abbreviation Epistem. Reasoning denoted Epistemological Reasoning. NA indicated that data were inadequate for that item.

Written argument evidence revealed positive and negative differences compared with interview evidence. Positive mean differences were apparent for evidence (+.3) and rebuttals (+.9) and negative mean differences were found for epistemological reasoning (-.3) and view of knowledge (-.1). Alternative theories showed no mean difference (+.0). There was variation across students and across dimensions. Students’ written proficiency in argument difference scores ranged
from -1.0 to +1.3 for evidence, -0.7 to +0.7 for alternative theories, -1.0 to +2.0 for rebuttals, and -1.0 to +1.0 for epistemological reasoning. Students' view of knowledge difference scores ranged from +2.0 to -2.0.

Three assertions summarize the evidence for the third research question:

1. **Written argument evidence revealed mixed views of knowledge, most which were inconsistent with interview evidence.** Compared with interview evidence, views of knowledge in written arguments were stronger for 3 students (A, C, G), the same for 2 students (B, J), and less for 2 students (E, K). Data were inadequate to judge H's view of knowledge.

2. **Written arguments exhibited stronger proficiency in some argumentation dimensions than did interview evidence.** More specifically,
   - written argument evidence to support theories displayed stronger scores than the interview evidence average scores for 4 students (A, B, G, H), the same for 2 students (C, E), and weaker for 2 students (J, K);
   - alternative theories were stronger for 2 students (A, J), remained the same for 4 students (B, C, G, K), and were weaker for 2 students (E, H); and
   - rebuttals were stronger for 7 students (A, B, E, G, H, K, J) and weaker for 1 student (C).
3. **The quality of epistemological reasoning evidenced in written arguments was somewhat weaker than corresponding interview evidence.** The mean score for epistemological reasoning decreased slightly from interview (1.9) to written argument (1.7). Compared with interviews, written arguments displayed epistemological reasoning scores that were stronger for 2 students (A, +1.0; G, +.5), the same for 2 students (C, H), and weaker for 4 students (B, -.5; E, -.5; J, -1.0; K, -1.0).

Overall, consistency between interview and written argument findings was weak. Higher scores in written arguments for proficiency in argument suggested that evidence drawn from the two mediums was not closely associated. Discrepancies in view of knowledge could reflect variation in evidence and criteria. However, both forms of evidence displayed limited epistemological reasoning success, suggestive of more congruency for this dimension. It should be noted that the written argument scores were based on only 1 topic—school violence; whereas the interview data represented 2 or 3 topics—school failure, return to crime, and unemployment, some for which students responded successfully (e.g., six students generated successful evidence for the interview school failure topic as well as for the written argument). Integrative rebuttals that addressed the relative merit of positions about school violence were higher in quality in the written medium, suggesting that Kuhn’s interview format did not
necessarily reflect students’ capacity for rebuttals. Even so, some variability across students (e.g., C’s stronger rebuttals in the interview medium) suggested that one instrument may not adequately reflect argument competency. Students’ epistemological reasoning was not enhanced by the additional time for reflection afforded by the written argument. Evidence indicated that the 8 University of Victoria education students’ mixed success in generating quality arguments did not necessarily align with their capacity to reflect on the quality of their own evidence and its capacity to sustain an argument.

**Qualitative Evidence: Perspectives of the Effectiveness of Intervention**

What sense can be made of the apparent overall minimal positive change in dispositions for critical thinking, views of knowledge, and proficiency in argument? Were students’ dispositions and perspectives resistant to change across all forms of evidence? Did qualitative evidence depict change that was not reflected in the pretest-posttest instrument findings? Did evidence of change reflect real change? What, if any, impediments might have influenced pretest-posttest difference scores? Were ideals of critical thinking, instruments, and instruction congruent? These questions are central to the 4th research question that addresses participants’ perspectives of the effectiveness of intervention.

Examination of qualitative evidence of participants’ perspectives of the course helps provide some insights into the quantitative results for this study.
Relevant evidence was gathered from both students’ and the instructor’s journals, the in-class survey, the two course evaluations, two in-class group assignments, and students’ individual assignments. These artifacts suggested that (1) the preparatory phase of intervention corresponded with evidence that students acquired a foundational understanding of and rationale compatible with teaching for critical thinking and an evaluative view of knowledge in elementary social studies, (2) some instruction during the application phase corresponded with increased levels of concern about clarity of conceptual distinctions for both students and the instructor, and (3) in the consolidation phase, most students expressed positive dispositions toward teaching for critical thinking in elementary social studies, indicated that they had acquired beginning understandings of what was involved in teaching for critical thinking and an evaluativist view of knowledge in elementary social studies upon which they could build, and demonstrated largely appropriate conceptual distinctions and relationships in their assignments. It was also clear that students’ understandings of the distinctions among dimensions of critical thinking were not readily accessible. These findings are organized according to phases of instruction (preparatory, application, and consolidation) to illustrate clear distinctions in perspectives for each phase.

Qualitative evidence were judged in light of instructional objectives, pedagogical decisions, and more formal measures. As with quantitative evidence,
judgments do not suggest that instructional decisions and actions are necessary or sufficient to realizing objectives; nor do judgments apply beyond the context of the study. Moreover, claims about dispositions are not to be interpreted as actual dispositions or actions. However, evidence do relate to attainments that are important for implementation beyond the study. That is, while positive change may not predict further implementation beyond the context of the study, it does represent a foundation important to such implementation.

The Preparatory Phase

The preparatory phase of the course (January 8-January 29) (1) fostered a warm and supportive context conducive to an honest exchange of ideas; (2) immersed students in critical thinking to facilitate some basic sense of critical thinking and to establish connections for later reference; (3) focused on the connections among students’ and curriculum ideals of an educated citizen, views of knowledge, society, and instruction; (4) provided experiences to facilitate an understanding of and disposition toward objective, fair-minded interpretation of intent and supporting interpretations with evidence before judging what to believe; and (5) developed with students a compelling reason for teaching for critical thinking and an evaluative view of knowledge in elementary social studies.

Evidence from the preparatory phase of intervention revealed the following:
1. Climate: All 8 students were comfortable enough to provide critical responses to intervention.

2. Dispositions: Most students indicated positive dispositions toward teaching for critical thinking and an evaluative view of knowledge in elementary social studies. Five students suggested that intervention had made a difference to their perspective.

3. View of knowledge: Some basic distinctions among views of knowledge were demonstrated by 6 students.

4. Judgment of meaning—sustainable fair-minded interpretation of intent: Most students (N=5) used evidence to support their interpretations of authors’ intent.

5. Effectiveness of Instructional Strategies: Students and the instructor found key strategies used in this preparatory phase of intervention were effective.

6. Alignment of Evidence: Alignment of evidence was mixed—consistently positive about intervention, inconsistent regarding view of knowledge, and roughly consistent for dispositions for critical thinking.

Climate. Participants’ responses indicated that the instructional climate during the preparatory phase of the course was supportive and comfortable. For example:

- C (journal, January 10): “Thank you for the warm welcome to your class. I feel fortunate my first class back in many years is with you.”

- G (journal, January 10): “I find this technique very helpful because some people (like myself) may find going right into the discussion as a class a bit uncomfortable, especially if you do not know much about the concept....I like it when you record our responses....By recording our thoughts it validates what we say....”

- Instructor (journal, January 10): “I feel quite comfortable with the group. I think the interviews are partly responsible. It seems like I know these individuals quite well.”
• K (journal, January 17): “I really like the interviews. I end up talking to you about so many other things. I’m really happy that I am doing it.”

• C (journal, January 17): “I found the... article very wordy, & hard to stay focused on.”

• E (journal, February 5): “I’m not sure if I have any criteria to add right now.... As a group, we seemed to have enough trouble with the six we have already!”

Dispositions. Seven students indicated positive dispositions toward teaching for critical thinking and an evaluative view of knowledge in elementary social studies, attributes that they had agreed were important to their ideal educated citizen. Relevant responses are included to illustrate the range of perspectives toward critical thinking and view of knowledge, variables central to this study:

• B (journal, January 17): “Paul & Binker.... I consider their approach ‘ideal.’ I am a firm believer in the values of critical thinking....” On January 22, B wrote, “As an individual who embraces critical thinking you should be prepared to analyze and justify your views in all areas of your life.”

• C (journal, January 17): “Concepts I especially agreed with were: points of view need to be critically analyzed & assessed; emphasis on the diversity of human perspectives should not be covered in a way that implies that all points of view are valid.”

• E (journal, January 22): “The more of an evaluativist that I can be, the better it will be for my teaching in social studies....”

• G (journal, January 10): “As a soon to be student teacher, I feel it is important that I take on these characteristics because if I want my students to become educated citizens then I will need to model these characteristics....”
On January 17, G recalled, "I was never given a chance to critically reflect on events and texts. I think it is important...."

- H (journal, January 16): "I enjoy the style of teaching social studies and employing Critical Thinking."

- J (journal, January 17): "We want our students to analyze and criticize information including their texts...Students need to be taught that their view is not the one right answer and neither is their teacher's."

- K (journal, January 22): "Teachers have a responsibility to move towards an evaluative view of knowledge."

It is possible that students already held these dispositions toward and understanding of the ideal educated citizen, characterized by an evaluative view of knowledge and the capacity and disposition for critical thinking. However, journal comments from 5 students (B, C, E, G, J) did suggest that the preparatory phase of intervention contributed to either change in or increased awareness of related values:

- B (journal, January 10): "Although I participated in the creation of the list (on what characteristics the class thought an ideal citizen should possess) it was the forming of the summary statement that really helped to clarify things for me. It forced me to choose from contributions on the list, evaluate them, and decide which are important to me.” In addition, B indicated, “[The Paul and Binker reading] helped me to rethink my attitude towards teaching Social Studies.”

- C (journal, January 16): "I feel studying the ideal Educated Citizen is very important for someone wanting to teach social studies and I know it will impact how I would teach a class."
• E (journal, January 15): "...We have been discussing the ideal Educated Citizen. I think that discussions such as these can have a great impact on how we teach social studies....[They help] to magnify and clarify my approach...."

• G (journal, January 15): "After examining and developing the criteria of the ideal Educated Citizen, I realize these traits are important for each person to have or at least be aware of the importance if we want to have a fair and just society."

• J (journal, January 10): "I enjoyed finding these characteristics because the educated citizen is used a lot without necessarily knowing what it means—now I have a clear idea what it means....It is very important that these characteristics are the goal for the social studies teacher personally—one can’t expect their students to act a particular way if they are not trying it themselves...."

**View of knowledge.** Students’ applications of their understandings of views of knowledge to sample lesson plans during in-class activities, their own perspectives (B, C, E, G, K), and assigned readings (A, B, C, E, G), suggested that some basic distinctions among views of knowledge were understood. H’s response indicated inadequate interpretations and J did not provide relevant information. Five students (B, C, E, G, K) expressed a disposition toward evaluativism. However, these students also indicated that a disposition towards evaluativism did not necessarily equate with possessing an evaluative epistemological position. For example:

• B (journal, January 22): "I would like to believe that I hold an evaluative view on knowledge, or at least it is something that I’m working towards. I suppose it should be viewed as a journey rather than a destination....I do explore alternative perspectives...however, I’m not sure how fairly I evaluate them....I’m really trying to be honest with you here. I wish I
could say that I hold an evaluativist view and be able to justify it. But I am trying to be willing and open to move in that direction...."

Five students’ (B, C, E, G, K) comments about the implications of view of knowledge for education also reflected some basic understanding of views of knowledge. For instance:

• G (journal, January 22): "I believe it would be deadly if an Absolutist taught Social Studies because he or she believes that there is only one interpretation to an event."

**Proficiency in argument: Sustainable, fair-minded interpretation of intent.** Instruction encouraged accurate or sustainable fair-minded interpretation of intent before judging what to believe. Evidence indicated that five students (A, B, C, E, G) supported their interpretations of the authors’ intent with evidence from the authors’ claims, two students (B, C) applied strategies of summarization or critical reading to class readings with reported success, and 3 students (A, E, G) acknowledged the difficulty they experienced in achieving fair-mindedness. For instance:

• C (journal—supporting interpretation with evidence, January 17): "There does not seem to be a lot of emphasis on establishing criteria to evaluate diverse points of view in the Sears & Parsons report. It tends to emphasize the necessity for a multitude of ideas and asking lots of questions but not a great deal of emphasis on evaluation, or that various points of view are not necessarily equal...seems like more emphasis is on the multiplist."

• B (journal—regarding application of strategies, January 29): "Well, I'm pleased to say, one of the strategies you mentioned in class has worked for me. By re-writing that quote in my own words, yet trying to keep the author's true intent, I think I came to an understanding."
• E (journal—regarding fair-mindedness, January 29): “It is hard to ignore ideas and language which are in direct contrast to some of my beliefs to see through to the valid arguments beneath. However, I think we are all concerned for the same reasons, the sake of the students and because of this we need to try meeting in the middle if possible.”

**Strategies for instruction.** Both students and the instructor found that key instructional strategies used in this preparatory phase of intervention were effective. Strategies that students indicated were especially effective in promoting understandings included generating criteria for examples, non-examples, and borderline examples (N=8) and for justifying related decisions (N=3); individual reflection followed by pair or small-group interaction followed by whole class sharing (N=7); the use of multiple modalities, particularly visual reinforcement (N=5); and summarization (N=4). For example,

• J (journal, January 15): “We have done a lot of discussing on the subject of the ideal citizen and I enjoy putting the information into a more practical perspective for us to use in the classroom....The strategy used with the 3 different lesson plans put into practice characteristics of an educated citizen and most of us probably didn’t realize it. We worked together to evaluate and analyze the lessons. We had also previously created criteria that we could test the lesson plans against. I found the lesson plan strategy effective for me.”

Seven students’ journal entries indicated that interactive constructivism, as employed, had been effective. For example:

• A: “By talking with someone else, I was able to see whether my understanding appeared to be correct or whether the other person’s viewpoint was more valid.”
• E: "I find that the value of group work cannot be overestimated simply because we never really know what other people are thinking. With group work comes the necessity of judging and coming to a clearer understanding of others and your own opinion."

Five students (C, G, H, K, J) indicated that visual reinforcement and attention to a range of modalities were effective. For instance:

• G: "I also like it when you record our responses, most due to the fact that I am a visual learner and so I need to look at things in order for me to fully comprehend the concept."

Four students (B, C, G, I) found that generating criteria for summarization and the task of summarization were effective. For example:

• B: "[The task was] meaningful for me....It was the forming of the summary statement that really helped clarify things for me."

• C: "Through analyzing the summary statements we were able to find the criteria-teaching us ways to find the answers to questions rather than being spoon-fed which would have lacked our involvement...therefore held less meaning and soon to be forgotten."

Overall, the instructor’s perception of the effectiveness of intervention during the preparatory phase was positive, as were the perceptions of most students. Evidence suggested that most preparatory phase objectives had been realized by the majority of students. It was not clear from the evidence that intervention decisions regarding classroom climate were responsible for the evaluative feedback that students were willing to provide. However, evidence did support a correspondence between instruction and the basic conceptual understandings that students exhibited regarding the educated citizen and view of
knowledge. Most students indicated that their rationale for teaching for critical thinking had been impacted in some way by the instruction.

Students' responses indicated that instructional decisions were congruent with meaningful links among students' prior knowledge, the curriculum, and important ideas relevant to views of knowledge and the ideal educated citizen as they applied to practice. Students' comments and applications of understandings suggested that some strategies for judging intent were acquired. Immersion was a positive experience for the instructor and provided referents for later explicit attention to concepts and strategies that had already been employed. It was not clear, however, to what degree students benefited from these connections.

**Preparatory phase: Alignment of evidence.** Participants' perspectives were largely positive about the instruction during the preparatory phase of intervention, were inconsistent with interview evidence of their view of knowledge, and roughly consistent with CCTDI evidence of dispositions for critical thinking. Pedagogical decisions and action and most students' responses were consistent with objectives toward an open climate, and a compelling case for teaching for critical thinking and an evaluative view of knowledge in elementary social studies. Immersion in interactive constructivist experiences involved key concepts and strategies central to critical thinking.
It is difficult to compare qualitative evidence relevant to students’ views of knowledge with Kuhn’s data. The nature of the instruction and qualitative evidence largely concerned interpretation of multiple perspectives and evaluating the relative merit of positions whereas Kuhn’s criteria for evaluativism focused on acceptance of multiple perspectives, denial of the potential for certainty, and less confidence in conclusions compared with an expert. G’s identification with evaluativism and denigration of an absolute truth in the qualitative data were not supported by G’s pretest interview results. According to Kuhn’s criteria, G’s responses aligned with absolutism across all 3 topics. This discrepancy could reflect differences between Kuhn’s criteria and the instructional focus or it could reflect G’s inadequate understanding of what evaluativism actually entails. For example, the following comment aligns with consensual-based judgment rather than evaluation of the quality of reasons or evidence: “I think I am an evaluativist. When I hear or see different points of view, I tend to take whatever is similar or overlaps and walk away with that view.” Some students’ (B, C, E, K) admitted that their disposition toward evaluativism did not necessarily align with the view of knowledge they actually demonstrated. This variable could account, in part, for discrepancies between evidence of students’ predominant view of knowledge (across 2 or more topics) in Kuhn’s pretest interview results (Absolutist—A, G, H,
J; Multiplist—C; Evaluativist—B, E, K) and qualitative evidence in which 5 students’ (B, C, E, G, K) supported the ideal of evaluativism.

The moderate strength found in most CCTDI results paralleled the positive dispositions toward critical thinking expressed in students’ journals (CCTDI pretest scores for individuals ranged from 279 to 359: A, 315; B, 319; C, 279; E, 359; G, 308; H, 336; J, 328, and K, 336). There was some discrepancy for at least two students (C, B) that merits consideration. For instance, in spite of a CCTDI entry-level score of 279 that apparently borders on weak disposition toward critical thinking, C expressed strong commitment towards a pedagogy for a conception of critical thinking that aligned with evaluativism:

- C: (journal, January 17) “Their premise that ‘students must think their way to knowledge’... in itself is a reason for ‘all’ to read it.... Concepts I especially agreed with were: points of view need to be critically analyzed and assessed; emphasis on the diversity of human perspectives should not be covered in a way that implies that all points of view are equally valid.... discussions & activities should develop the student’s use of critical reading, writing, speaking, & listening; include skills of identifying & evaluating various viewpoints....”

B also expressed and exhibited a stronger disposition toward critical thinking than her modest CCDDI score of 319 implied. For example, on January 17, B wrote:

I feel very much in agreement with the article, and its theories reflect our class description of an ‘ideal educated citizen.’ Some key issues/points I agree with: want to be careful, reflective seekers of truth; the concept of ‘point of view;’ want students to be able to identify P.O.V. and consider alternatives; ... all points of view are not equally valid.... multiple sources should be examined critically....I have much to learn. I feel it takes both knowledge of subject matter and confidence in that knowledge before I can
I take notable risks with my students. I am a firm believer in the values of critical thinking. I now need to learn how I can apply appropriate activities to foster it within my classroom.

B went beyond claims of disposition to accounts, which if true, demonstrated that she acted upon expressed disposition toward critical thinking. Such claims were apparent throughout B’s journal:

1. openness to self-correction or change in belief if the evidence warranted it:
   - January 10: “Paul & Binker made me rethink my attitude towards teaching social studies;”

2. intellectual curiosity:
   - January 17: “I enjoy learning and I love discovering new ways of looking at things. It is exciting to read or hear something and you think to yourself ‘I never thought of it that way before’;”

3. intellectual honesty:
   - January 22: “I do explore alternatives but I’m not sure how fairly I evaluate them;”

4. an active search for understanding and an intellectual work ethic:
   - January 29: “I’ve tried to re-word it to clarify it more for myself….I do hope that we get a chance in class to discuss this reading. The language/wording forced me to have to re-read several times for understanding. I particularly would like some clarification on relativism of values and open-mindedness;”

5. recognizes a need for sustainable evidence:
   - January 22: “I wish I could say that I hold an evaluativist view and be able to justify it;”
January 22: "I then went back and re-read the article to find evidence to support the claim... [For example...]."

6. application of ideas to new situations:

- January 15: "It's theories reflect our class description of an 'ideal educated citizen';"

- January 29: "I'm pleased to say, one of the strategies you mentioned in class has worked for me... [For example...]."

It is not clear why two students who expressed strong dispositions toward critical thinking in qualitative evidence did not attain as high a CCTDI score as did other students whose qualitative evidence of dispositional strength was less apparent (H, J). Four explanations come to mind. First, claims of disposition are not necessarily congruent with actual disposition or action. Second, students' responses could represent inconsistencies that reflected transition between views of knowledge and dispositions toward critical thinking. Third, claims of disposition could be based on students' inadequate understanding of what is represented by an evaluative view of knowledge and the conception of critical thinking it implies. Fourth, it could be that the CCTDI did not adequately align with the different critical thinking dispositions than were apparent in the qualitative evidence in this study. Preparatory phase evidence were inadequate to judge to what extent the first 2 explanations apply to results. Yet one could argue that while J's neglect of relevant assigned journal topics cannot be construed as an absence of disposition toward critical thinking, her inattention to assigned journal
topics did not reflect the intellectual work ethic of a critically thoughtful person that was consistently exhibited by B and C. That is, J’s CCTDI responses that supported disposition for critical thinking did not align with the behavior that a critically thoughtful person might exhibit. There was some evidence that suggested H did not adequately understand distinctions among the views of knowledge, corresponding with the third explanation. There was also some evidence to support the 4th explanation. Some of the dispositions toward critical thinking that were evidenced by B and C were not salient in the CCTDI inventory. The reverse may also be the case.

Overall, qualitative evidence was consistent in positive dispositions toward instruction. Qualitative evidence was inconsistent with interview evidence of students’ views of knowledge, in part because of some differences in constructs, but roughly aligned with the CCTDI evidence of dispositions for critical thinking.

**The Application Phase**

The application phase of intervention (February 5 - March 18) involved a more explicit engagement with critical thinking as it applied to teaching elementary social studies. Instruction in the preparatory phase had emphasized an immersion approach to critical thinking (students engaged in tasks of generating criteria, evaluating quality, and justifying decisions against shared criteria without explicit attention to the concepts of criteria and justification). In the application
phase, instruction integrated immersion and infusion (Ennis, 1992). Infusion involved explicit attention to the distinctions and associations among concepts relevant to epistemology (e.g., evidence—observation versus inference; justification with evidence versus mere assertion; views of knowledge) and the instructional framework (e.g., kinds of judgments and relevant criteria, the vision of the educated citizen). The intent was to (1) enhance students' capacity to justify their instructional decisions and (2) build students' commitment toward teaching for critical thinking and an evaluative view of knowledge in elementary social studies. Lessons focused on a critical challenge central to the curriculum and associated judgments (e.g., about authors' intent, the meaning of a concept, interpretive representations of reality, the significance of a claim and what should be done about unacceptable behavior). Students then applied related understandings to the development of instructional plans and justified how their plans could help their pupils move toward the class profile of the educated citizen.

Evidence regarding students' perspectives of instruction during the application phase were found in students' journals, the instructor's journal, and in the in-class survey. The instructor's perspective of this phase of intervention was based on the alignment among pedagogical objectives, instructional decisions, and both students' and the instructors' responses (participants' journals, an in-class survey, and instructional plan assignments).
Evidence indicated that both students and the instructor struggled with an excessive concept load and with inadequate clarity of some related conceptual distinctions (e.g., judgments about concept referents and judgments about empirical facts; evidence and source). Confusions were slowly unraveled and some clarification was achieved through remedial efforts that emphasized systematic conceptual development and having students justify how examples, non-examples, and sometimes borderline examples of concepts did or did not share concept attributes. Both the students’ and the instructor’s perspectives were largely congruent about instructional decisions that helped or hindered clarity. The following claims summarize the findings:

1. The instructor found that the instructional decisions associated with the most fruitful responses included:
   - an interactive constructivist approach that connected with students’ prior knowledge, involved students in the generation and evaluation of criteria and their application to examples, non-examples, and borderline examples; and challenging and justifying related decisions;
   - posing counter-examples to encourage consideration of other options;
   - seeking feedback from students and responding to concerns where feasible;
   - engaging students in practical applications; and
   - providing constructive feedback on assignments and inviting revisions.

2. The instructor found that the instructional decisions that corresponded with increased level of concern included:
   - a structural organization that embedded instruction and assignments within the framework of a unit plan that was not adequately understood by students;
   - an excessive concept load that included an unwieldy instructional framework and inadequate prior background readings;
• conceptual confusion exacerbated by the instructor's own evolving conceptual clarity;
• an imbalance between teacher direction and student problem solving;
• matters of pacing and dispersal of assignments, compounded by adjustments to students' readiness; and
• cessation of required journal submissions.

3. Students found effective aspects of instruction included:
• instructor feedback that aided comprehension;
• interactive strategies that included generating and applying criteria to examples, non-examples, and borderline examples; and
• justifications and practical applications that were essential to clarity.

4. Some students indicated that the least effective instructional decisions involved:
• an overwhelming instructional framework;
• a concept overload with inadequate prior background reading;
• a lack of initial clarity in structural organization of assignments; and
• matters of pacing and assignment dispersal.

Application phase: Instructor's perspective of effective instructional decisions. Evidence of strategies implemented that corresponded with positive responses were found in the instructor's journal:

• Instructor's journal (February 5): "I distributed sorting strips of critical challenge criteria and asked students to organize [sample challenges] into examples and non-examples for each criterion. This captured students' full attention and had them discussing where samples best fit. Some were quite critical of the samples and I responded that [this] suggested they had some depth of [understanding of] the criteria...."

• Instructor's journal (February 12): "Responses were quite appropriate from all three groups. They saw justifying relative merit as the most difficult [task to demonstrate conceptual understanding]."

• Instructor's journal (February 12): "In small groups, students now created definitions of what 'discrimination' is, displayed them, and we discussed them to gain clarity and relate them to each other. The quality of these
summaries had improved considerably from initial attributes [based on students prior knowledge] that were shared and listed on the board."

- Instructor's journal (March 18): “Sample arguments were provided to students who described what attributes each argument exhibited relative to a quality argument and then [students were] to identify which view of knowledge each represented. Responses indicated understanding. Students' justifications of views [that were] represented [in the arguments] provided additional criteria that were recorded on the board. These notes provided a fairly good visual representation of what makes a quality argument and how criteria fit their assignments....”

The instructor's journal indicated that for the most part, when students engaged in model lessons, classes went well. Students were also responsive when the instructor modeled the use of counter-examples to test decisions and encouraged students to consider alternatives. The strategy, described as it was conducted in one class, evoked considerable student engagement and a positive comment from one student:

- Instructor's journal (February 26): “Then I asked class to find evidence in text of racial discrimination, to record on slips of paper, to post with tape on board, and to be able to justify why they are examples of racial discrimination...I had each group share their evidence and justify it. I questioned them as to how sure they were that these were examples/evidence and if there might be other explanations for the behaviors, offering possibilities. I emphasized that while their evidence might indeed be representative of racial discrimination, they needed to be able to convince [others] that other factors were not responsible. I had them recall their definition....We used this definition to compare to evidence. I also asked if the treatment was unique to the Chinese. I then noted this activity related to deciding about meaning—one way to test understanding of the concept at a deeper level....C commented after class and during class that my questions made her doubt the evidence had been adequately justified as racial discrimination. She appreciated my role as
devil's advocate and agreed that we accept things too readily and need to do more such challenging of conclusions."

Seeking feedback from students was fruitful. It provided some insight into students' ideas and concerns, allowing for adjustment of instruction to better address students' needs. For example:

- Instructor's journal (February 5): "The journal responses sure provided depth of insight into students' understandings and dispositions, unlikely possible otherwise....Some students are not maintaining journal responses, some are handing them in intermittently."

- Instructor's journal (March 4): "At the end I asked how things went and J said I was trying to do too much...."

- Instructor's journal (March 18): "Re-critical thinking framework, I noted most were initially confused but as they got feedback from assignment-it made more sense to them."

The instructor's feedback on students' assignments (A, B, E, G, J), through personal meetings (G, H, J), in journals, and during class was utilized profitably by 6 students who revised a total of 12 assignments with much improved conceptual distinctions. Personal circumstances and absenteeism corresponded with inadequate evidence of comprehension for 2 students (H, K).

**Application phase: Instructor's perspective of instruction that corresponded with the increased level of concern.** The following assignment expectations corresponded with the noted increase in anxiety:

1. Students were now accountable by grades for assignments (instructional plans and justifications). They were expected to adapt the teaching, learning, and
assessment activities they had engaged in during class to different but parallel instructional plans for their prospective pupils. This involved decisions about how they might help their pupils acquire the intellectual resources (background knowledge, critical thinking concepts, criteria and standards, habits of mind, and thinking strategies) important to decisions about what to believe or what to do about judgments associated with a central critical challenge.

2. Each lesson plan focused on a different kind of decision that followed logically from a central critical challenge. This organization was compatible with the sequence of lesson plans for a unit of study and formed the 'unit plan' structural organization for instruction and assignments. As such, lessons were connected and built upon each other, imposing connections among students' instructional plan assignments that was limiting for some students.

3. Students were expected to justify their instructional plans for its relative merit for moving their pupils towards their ideal of the educated citizen. To do this successfully, students had to grasp the alignment among their instructional plans, their ideal educated citizen, the curriculum, and the intellectual tools from the instructional framework that would facilitate their pupils' movement toward the ideal.

The impact of new demands for student accountability could not be determined from the evidence. Nor was the instructor initially aware of the
limitations that the 'unit-plan' framework imposed on students through expectations for connections among lesson plan assignments. Immediately, however, the instructor was aware of increased student concern following elaboration of the instructional framework that was done in a hurried manner at the end of two classes. The instructor was also concerned about the concept overload imposed by concepts central to the lesson plans that students engaged in and additional concepts relevant to social studies, the instructional framework, and assignment expectations. The following excerpt marks February 12th as the class in which level of concern visibly increased for both instructor and students:

- Instructor’s journal (February 12): “I have concerns about how fast we are NOT moving and the amount of absenteeism as it relates to what we are hoping to achieve....Virtually every concept that comprises part of the framework and lesson criteria needs to be taught. Modeling in all classes is insufficient. One must explicitly address how each [lesson] relates to the assignments. Perhaps I have too many expectations. It seems that each [assignment] not only is a test of understanding of how to teach something but also [requires] an understanding of each of the parts [of the instructional framework] and how they relate to criteria. I must examine this more closely and see what my problem is. There is not enough time to do it all. I will limit lesson plans to three...There is only so much feedback I can expect from students. I will limit journal responses during assignments (that will mean most of the balance of the term) and I will limit readings.”

This entry represented the first of a few adjustments to students’ assignment expectations, a response to perceived students’ needs that had already placed assignments later than desired in the course.
On February 19, student inquiries led the instructor to question the wisdom of trying to develop concepts through practical application without adequate background reading. Handouts were provided with basic descriptions of some framework dimensions (e.g., habits of mind, cognitive tasks). The instructor also was becoming concerned about too much teacher direction and inadequate student-directed engagement in connecting lessons experienced during class to the instructional framework and assignments. For example:

- Instructor's journal (February 19): Some students inquired "about the lesson framework titles of intellectual tools and cognitive tasks. I took time then to try to clarify their intent. I believe I had addressed [these] superficially... before and said I hoped to develop an understanding of them gradually. I gave examples of what we had done so far.... I noted there seemed to be considerable overlap among dimensions of the framework such as cognitive tasks and strategies and student activities and forms of assessment. I had not dwelt much on cognitive tasks and find that I too am having trouble rationalizing them. There seems to be too many dimensions and there is overlap and I am having doubts about the framework.... about the complexity, the required conceptual foundation, and I feel that much of the excitement that is social studies is being lost because of my perceived need to develop understandings of needed concepts.... (e.g., justification, criteria, standards, fit, effectiveness).... [In addition to] concepts [that] are... in the lesson being taught/modelled (e.g., racial discrimination.) Perhaps I should not be providing any criteria but having students evolve their own.... I must also work on limiting criteria. How to work on criteria... [within the lessons] and also criteria for course [assignment] expectations...?"

- Also on February 19: "The time needed to develop what seem to be necessary concepts (justification) may be beyond what can be accomplished in such a course. To do this for each assignment is not possible but as we move along, they should build on each other. How can I make this more appealing and yet retain substance and facilitate necessary understandings? My sample lesson plans and justification were a lot of
work... I think they were helpful and the handouts will hopefully help the four who were absent understand related concepts. But teachers with many classes do not have sufficient time for this. The time that I am expending is made necessary, in part, because I do not have a collection of actual students' work. Doing my own samples has the advantage that I can structure them to show what I want. I hope I am not fighting a losing battle."

Inadvertent illogical development and the emergence of conceptual questions by the instructor added to the confusion on February 26. The citation below revealed these concerns:

- Instructor's journal (February 26): "During the session... I realized something was amiss; was [the] current focus on justifying cases of discrimination becoming confused with justifying whether the cases actually occurred? .... Was the term evidence confusing? Was it the logical development that was confusing? .... I hope students are able to give me feedback. I asked them to but did not assign it as they have other readings and some have lesson plans due, as well as preparing Jigsaw lesson. I am not convinced that I am clear enough about the conception myself."

The instructor found that concept load, the instructional framework, pacing, and the pressure of assignments that were backing up, in part because they were dependent on students first participating in model lessons in class, continued to be a concern. A sense of the desirability of building instruction around needs that students identify was also emerging. Concepts relevant to the structure of knowledge in social studies, the instructional framework, and assignment expectations continued to compete for attention. Acceleration of concern was apparent on March 4th following the instructor's second and last elaboration of the instructional framework for the upcoming assignment. The concern that led to the
instructor seeking written feedback from students in the next class is reflected in
the following entry:

- Instructor's journal (March 4): "I did not move into [decisions about] generalizations for fear of confusing students but have some reservations that in order to understand facts one needs to understand related concepts. I wish it was not necessary to address these concepts as it adds to the load for students and me. I had students gather in home groups...and work together to flesh out parts [of the instructional framework]...Then I asked them to help me...[elaborate] a chart on the board that would meet all the criteria for Ass. 2. This was too much for the time and I realize that I cannot continue with this. The framework is unwieldy and there is too much to expect of the students. At the end I asked how things went for them and I said I was trying to do too much. I wondered if not having Ed. E 746 made it more difficult. She said she had thought this might be the case but discussed it with others and that others felt the same way. I said we would discuss it next week. My expectations for the course are too much; the assignments too many....There are too many concepts...to develop adequately with [the] class to make them comfortable with their application....I guess I should share a paper [what paper?] with them that provided a more comprehensive account. I find myself asking if I am providing students with too much detailed guidance and not letting them do enough for themselves. I also feel I am rushing over some things superficially but feel a race against the clock. I am not sure what else to do but must find a way to do less and do it well....Before I got into the critical thinking decisions, I felt things were going quite well. I was comfortable with the strategies, the readings, discussions, and related activities and seemed to be getting fairly positive responses. This is not so now. Time is rushing toward the end of the term and [preparation for assignments is]... needing more class time than allotted....I wish now I could find a better synthesis task."

This period of heightened concern spanned from February 12 to March 5 during which time lessons focused mainly on the development of intellectual tools that students could use to justify their own decisions about what to believe or do about the meaning of concepts and the truth of factual claims and how these
experiences related to instructional plans (assignments) and their justification.

Evidence of students' perspectives supported the instructor's judgments about the instructional framework and revealed some additional problems related to the organization of assignments.

**Application phase: Students' perspectives of effective intervention.** Four students (B, C, E, J) indicated that feedback from the instructor enhanced their understandings:

- **B** (journal, January 29): "Speaking of feedback, thank-you for all the information you included in the journal for me about concepts....Your explanations clarified the definition for me."

- **B** (journal, March 21): "When I first handed my 'concept' assignment I received 6/10 which I feel was a fair mark because I missed the whole point of the assignment. this reflects my confusion! [B had been absent for the class in which the assignment was clarified]. However, the feedback on the first assignment was good and you clarified some points for me in class which I really appreciate. I may not get perfect marks on the rewrite of that assignment, but I do feel that I now understand what is going on."

- **C** (in-class survey, March 11): "I now feel after doing the 2nd assignment that I have a handle on the Critical Thinking Overview so would like to continue using them. I think last class you really helped to clarify the parts. Prior to that, I think it probably seemed a bit overwhelming...."

- **E** (in-class survey, March 11): "At first I had problems with it [the instructional framework] but after I got Lesson #1 back, I understood it more fully."

- **J** (journal, February 26): "I really enjoyed this private class that I had with you and thank-you for spending the time with me! The critical thinking visual organizer really helped me understand all these concepts. It makes it much easier to tell how they interact and relate to each other."
Positive responses from all 8 students in the early application phase paralleled those in the preparatory phase for the interactive strategies of sorting examples and non-examples in partners, justifying responses to another pair, testing criteria and justifying them to the class, and practical application of criteria. For example:

• A (journal, February 5): [Sorting examples, non-examples, and borderline examples] "tested our understanding of the criteria....Taking that even further, when you have to justify your responses, you think them through more clearly, so that you are able to explain them to someone else."

• B (journal, February 5): "I think this is an excellent activity. Doing it really helps to clarify the concept in your mind. As my group started sorting the criteria we kept on disagreeing with each others' choices. However, after discussion, and justifying our choices to each other, we began to realize how it was all supposed to fit together. Between us we had then come to some sort of understanding about the concept of criteria in a critical challenge....By trying to rank the criteria, and see if anything could be deleted or added, you really had to have some sort of understanding about the criteria. Having to justify your choices is a vital part of this process...by talking it through, you all end up with a better understanding. applying critical challenge criteria...was by far the most difficult activity for me. It took careful thought and consideration to find a critical challenge to fit all the criteria we had previously been working with. I personally think the application part of the process is vital to trying understanding the concept. Without this step I wouldn't have the same depth of understanding. I wish that we could have had more time in class for the peer evaluation."

• E (journal, February 5): "I find that the value of group work and interaction cannot be overestimated....With group work comes the necessity of judging and coming to a clear understanding of others and your own opinion. This was evidence[d] today when we had different views of what the criteria for critical challenges actually meant. By sharing our understandings with partners and the group as a whole, we were all able to
come to a better and clearer view of how we all feel and think on such subjects....Examples and non-examples remains as one of my favorite strategies for exploring concepts.”

- B (journal, February 12): “Since we have been talking in class about the importance of teaching conceptually, I think it is really valuable that we have been given the opportunity to learn that way as well. By actually having to grapple with the concept of discrimination, not only did I come to a better understanding of the concept, but also the process. Having to analyze examples, non-examples, and borderline examples really helps to clarify the concept in your mind.”

Three students (A, C, E) indicated that a course overview on March 11th helped clarify connections for them (B, H, & J missed this overview). For example:

- C (in-class survey, March 11): “Your overview of the course—what you’re teaching and how it ties into the course description and the sequencing of what we’re struggling with was good—helps to show how it all ties together.”

**Application phase: Students’ perspectives of ineffective intervention.**

Some students indicated difficulty comprehending the ‘unit-plan’ structure that linked lesson plan assignments. The inter-dependency among model lesson plans and related assignments (judging concept meaning, factual and general claims about reality, and judging the value of behavior and what should be done about unacceptable behavior) and inadequate early clarity about the relationship between these plans and a unit-plan framework imposed constraints on early assignments. For example, if the critical challenge students selected related to immigration, and if the embedded or associated concept they chose to teach was immigrant, this
meant the following lesson plan about an empirical judgment claim and then the judgment about what should be done had to be connected in some logical way to the previous plans. In some cases, the challenge and concept the students chose did not adapt well to the kinds of decisions that instruction focused on. For example:

- B (journal, March 21): "...During last weeks’ class (March 18th) I really felt like things started to come together for me. I had considerable difficulty with the assignment #4 (Deciding what to do). I think I clearly understood the procedure—it makes so much sense, but I struggled in fitting it into my original concept of ‘immigrant.’ I think if I were actually doing my unit on ‘immigrants’ with students that decision would be left out—I felt like I was forcing the concept where it didn’t belong."

- A (in-class survey, March 11): “I found the assignment very confusing initially [A had been absent for the clarification lesson] because I didn’t understand where we were going with the unit and how the lessons related. I would have found it more helpful to go over the 4 pts. today before I started writing the lessons."

The critical thinking instructional framework had been initially overwhelming for most students. Five students indicated that it had been confusing but that it was becoming much clearer now. One student (K) still found it problematic. For example:

- K (in-class survey, March 11): “I do find the overviews overwhelming. I’m still having difficulty with the cognitive tasks and Intellectual Resources... I think that we find this hard because we are not used to critical challenges, etc. and we have to learn to think differently. I think that what you are trying to do with the Overview is good but we need more time and practice with it. (Unfortunately, [K added,] we don’t have the time which makes things very hard.)"
• J (in-class survey, March 11): "C. T. overview is starting to make more sense. But it could have worked better if it was fully explained right off."

All 6 students who completed the in-class survey expressed interest in receiving background reading relevant to the components of the instructional framework. Additional comments suggested that pacing and overloaded lessons were problematic:

• A (in-class survey, March 11): "The reason for the confusion in the class is because we fly through everything and touch on so many things."

• J (journal, February 26): "I'm hoping these new methods will start coming easier to me and then I won't be so frustrated. I do understand the meaning lesson and hopefully you'll agree when you mark it....The lesson plans are becoming more clear to me but at times I wonder how realistic it is to go into such detail. Most classes only get about 2-3 classes a week and this method of unit planning seems to take a long time to get into the material. I think it would be more effective for us to see more of a variety of activities to use in the classroom. We've done a lot of sorting and listing and they seem to be effective for us, but they're not real activity based. I don't think kids would stay interested for long enough. They need more variety and excitement. Not saying I wouldn't use sorting but just not on a regular basis. I feel very rushed in the class. The methods are new (parts of them) to use and I need time to be able to figure it out. I think the speed and fullness of the class almost stifles ideas at times. Maybe the curriculum needs to be altered. I'd rather learn several things well instead of half-learning a lot."

The application phase of instruction revealed heightened level of concern for all participants. Later reflection by the instructor suggested that impediments to effectiveness during this intervention phase involved concept overload and organization of instruction and assignments that impacted pacing, insufficient
student problem-solving and too much teacher direction, and evolving conceptual clarity for the instructor.

**The Consolidation Phase**

The objectives of the consolidation level phase of intervention (March 25-April 1) were to help students synthesize course ideas and to gather evidence of intervention effectiveness. These last two classes included 3 consolidation activities that provided evidence to supplement written assignments: students were asked to predict unit plan criteria, to match kinds of decisions with relevant criteria and to justify related judgments, and to both demonstrate relationships and elaborate dimensions of critical thinking by building and evaluating resulting concepts maps. Exit-level qualitative evidence was provided by the instructor’s journal, final unit plan assignments with accompanying justifications (N=8), one instructor-created course evaluation (N=7), and one formal university instructor evaluation (N=9). Consolidation phase evidence corresponded with the following findings:

1. **Instructor’s perception of the effectiveness of instruction**: Most students were successful with most basic conceptual distinctions relative to the instructional framework for teaching for critical thinking in elementary social studies.

2. **Instructor’s perception of evidence of view of knowledge**: Evidence of views of knowledge were mixed.

3. **Students’ perspectives of the adequacy of their understanding to implement course ideals**: Most students indicated that they had acquired adequate
understandings from their experience in the course to attempt to implement some aspects of teaching for critical thinking in elementary social studies.

4. Students' perspectives toward teaching for critical thinking and the nature of intervention: Students were largely positive toward teaching for critical thinking in elementary social studies and toward the nature of instruction.

5. Alignment of evidence: Alignment across qualitative and quantitative evidence was mixed.

Consolidation phase: Instructor's perspective of the effectiveness of instruction. Students (N=6) were largely successful in demonstrating appropriate relationships among dimensions of critical thinking and the instructional framework with evidence of some uncertainty about conceptual distinctions (e.g., accuracy of claim and credibility of source). Most students were responsive to critical thinking and an evaluative view of knowledge but some inconsistencies, vagueness, or inadequate internalization of ideas across responses suggested that some students were insecure in perspectives, understandings, and proficiencies.

The following excerpt from the instructor's journal reveals that students' predictions about what unit plan criteria should include did not indicate priority for both intellectual tools and justification, ideas that had been central to the course:

- Instructor's journal (March 25): "I had them first brainstorm with partner(s) what they anticipated would be expected on [the] unit plan. Responses were somewhat disappointing in that intellectual tools and justification were not mentioned."
A sorting activity also revealed that students’ conceptual distinctions among kinds of judgments and criteria were insecure. However, the concept mapping activity was more fruitful. Students’ concept maps and related comments exhibited appropriate relationships among dimensions of critical thinking, varying degrees of elaboration, and alignment with concept map rules (see Appendix C for related concept map summaries). The more successful elaborations were developed by the few students who had brought their course notes to class (e.g., B, H). H noted on the instructor’s course evaluation, “I liked the concept map—this cleared up a lot.”

Final unit plan assignments with accompanying justifications provided evidence of the effectiveness of the instruction for individuals. Lesson plans embedded in this unit plan assignment did not necessarily represent application of learning to new situations (some revised lesson plans were included). However, a culminating lesson plan that was to represent summative assessment ideas for prospective pupils and a justification for the unit plan did allow for new applications and demonstrated some basic understandings of teaching for critical thinking in elementary social studies.

Lessons embedded in students’ unit plans were designed to help prospective pupils acquire the intellectual tools to justify 3 different kinds of decisions (e.g., about meaning, fact, and action) and as such, reflected a decision
that aimed for alignment among dimensions and attributes of the instructional framework. For example, a lesson plan directed to a particular decision about empirical claims would include teaching, learning, and assessment activities that would contribute to pupils’ capacity to evaluate the quality of evidence and the credibility of sources in order to decide what conclusion is best supported. Six students (A, B, C, E, G, J) exhibited predominantly successful conceptual distinctions and demonstrated largely successful decisions relative to the instructional framework.

Evidence of inadequate conceptual understandings included the distinction between accuracy of evidence and credibility of source (N=5), kinds of judgments (N=2), and the nature of a concept (N=1). While B exhibited appropriate criteria for the summative evaluation lesson for pupil’s work, 5 students (A, C, E, H, J) were either inaccurate or vague. For example:

- B: Criteria for evaluating her prospective pupils’ included “Written argument—clarity & fit of evaluation of degree of accuracy of info & credibility of source with class-developed criteria;”

- J: (vague regarding distinction between accuracy of claim and credibility of source): Pupils were to “create a list of criteria of what makes a claim accurate and credible;”

- E: (vague-lacks specifics): “The class will develop a list of criteria which is accepted by the whole class.”

Two students exhibited evidence of unclear distinctions among kinds of judgments:
• C referred to an empirical decision as "...the value decision on if there still is conflict and who is involved"). This confusion was not carried through to lesson plans.

• K: (unclear about what a concept is and related instructional strategies, relative to teaching about facts): "I had difficulty with the concept of meaning in this unit since I did not have a philosophical theme such as power or racial discrimination. As a result, I chose the idea of defining the C.P.R. (a form of transportation) using the 5 W's as the criteria....I feel that by getting the students to look at half truths, it will assist them in looking for information....

Justifications that were to provide arguments of the relative potential for plans to help move pupils toward the ideal educated citizen varied in quality. Six students (A, B, C, E, G, J) provided evidential support for how their plans might facilitate the acquisition of intellectual tools important to realizing the ideal educated citizen. Five students (A, C, E, H, J) argued that their plans were directed toward helping students gain the intellectual tools (or as E, H and J noted, skills or abilities) to justify their decisions, implying that other approaches would not do so. C provided the strongest evidence that discounted alternative approaches but did not show whether the tools developed were necessary or sufficient to the ideal (see C’s justification below). While K’s claims did not represent appropriate integration in her unit plan, they did show some sense of the importance of attributes relevant to the ideal educated citizen.

Justifications represented the essence of course ideas, as perceived by students. For example, B’s justification depicts an evaluativists’ consideration for
multiple views and the intellectual resources for judging relative merit, emphasis on the pupil as decision maker, and evidence of conceptual distinctions and relationships among kinds of judgments, criteria, and teaching and learning activities. A and C who did not exhibit evaluativism in their interviews and who both scored relatively low in the moderate range of CCTDI mean scores reveal positive dispositions for critical thinking and exhibit some attributes of evaluativism in their justifications, suggestive of insecure positions. B's strong dispositional response and clear evidence of evaluativism also belie her less than exceptional CCTDI posttest response score. (See Appendix D for remaining justifications.)

- B: "This unit builds connections with what students know....This personal perspective may make the learning more meaningful for the students....This unit helps students in class evolve accurate definitions for concepts used in the unit....through sorting examples and non-examples, writing definitions in own words and checking with accepted usage, and justifying decisions every step of the way....Informed decision-making and critical thinking skills are integral components of the Ideal Educated Citizen. Such skills are developed in this unit when students are guided through a systematic process where they develop criteria to examine a source for credibility, and accuracy of information. This criteria [sic] is then applied to test the source to see the degree of support that can be found. The students are continually reading critically, and comparing what they read to the class-developed [criteria]....Students are guided to recognize a problem, come up with and examine solutions, weigh the evidence considering the pros and cons for all parties involved, and come to some sort of informed decision (one that is most acceptable to most people). Examples of these kinds of strategies are shown in Lesson #5....The IEC takes an active part in making the world a better place. This is addressed in Lesson #5 where the students are making decisions to positively influence the situation in their classroom. This will in turn affect other people's lives beyond the
classroom as well....Students get used to working in partners, and small groups. This requires respect for different levels of knowledge and sensitivity to feelings of others in the group deliberations.... Students will explore values involved and see how they affect decisions and how these attitudes etc. change over time. The IEC would need to recognize the complexity of relationships and be able to take that into consideration when making a decision."

- C: "This unit plan should help the student become an informed decision maker and educated citizen who will behave in a socially responsible way....From these activities the students build empathy and a sense of what conflict feels like. Through sorting examples and non-examples that they justify in small groups the students are able to evolve their own definition [of the concept]. This will be far more meaningful than having a teacher start the unit by simply telling the students the challenge, the concept, and the meaning. Rather than the teacher and textbook being the unquestioned source...this unit plan encourages the students to challenge the teacher and the text-book, to question numerous sources and to learn about the issues through deciding what to believe. Obviously this is a better way to achieve our goal of turning out critical thinkers rather than students who can't think but can memorize. One cannot memorize answers to new problems and once a problem is solved it is not a problem any longer....The activities nurture their understanding and competent use of the intellectual tools necessary for critical thinking...."

Overall, students exhibited basic understandings of the relationship among some key dimensions of critical thinking. Such capacity was adequate for 6 students to build largely successful relational arguments that integrated critical thinking into teaching, learning, and assessment activities in instructional plans directed toward teaching for critical thinking in elementary social studies. Evidence suggested that distinctions among kinds of judgments, evidence, and criteria were not adequately internalized for ready access.
Consolidation phase: Instructor’s perspective of evidence of view of knowledge. Evidence of views of knowledge was mixed. Two students’ (B, E) claims clearly aligned with evaluativism and the other 6 students were either inconsistent or evidence was inadequate to ascertain their view of knowledge. Overall, students’ responses provided some evidence of evaluativism. Inconsistencies represented students’ insecurity that could indicate a readiness for change and consolidation.

A summary analysis of students’ views of knowledge follows (for related evidence, see students’ justifications above and in Appendix D):

- A’s suggestion that students evaluate and justify conclusions about reality could represent either absolutism or evaluativism. A’s advocacy that students evaluate options and consider the relative acceptability of options to those who might be impacted by decisions was representative of an evaluativist view about value issues but it was not clear what A’s position was about empirical claims.

- B clearly exhibited an evaluativist view of knowledge in an emphasis on the degree of support provided by evidence and the need to weigh evidence.

- C’s justification opposed absolutism in promoting the questioning of authoritative sources but did not clearly exemplify evaluativism.
E advocated the need to consider multiple points of view and to strive for the best and most accurate version of the truth, a clear acknowledgment of the relative merit of conclusions about reality and an exemplar of evaluativism.

G’s justification aligned with evaluativism with its focus on the adequacy of evidence to support conclusions but her claim supporting a potential for true evidence aligned more with absolutism.

H’s suggestion that ideas should be evaluated for flaws does not clearly distinguish between absolutism and evaluativism but does exclude multiplicity.

J’s advocacy for weighing options and evaluating the impact of decisions did not align with absolutism but it was not clear which view of knowledge her justification did represent.

K’s support for different points of view and the use of arguments to justify conclusions is suggestive of evaluativism.

**Consolidation phase: Students’ perspectives of the adequacy of their understanding to implement course ideals.** Six (A, B, C, E, G, K) of the 7 students who completed the instructor-created course evaluation indicated that they had acquired through the course adequate understanding to try to implement some aspects of teaching for critical thinking in their prospective elementary social studies classes. The 7th student (H) was unsure about the adequacy of understandings but indicated a disposition to learn by doing. For example:
• B: "I would definitely try—decisions about meaning (concept attainment), testing empirical/factual claims I consider to be an invaluable aspect in terms of research and probably one of the biggest gifts you could give students would be how to construct and present an informed argument—something I will use with my students."

• C: "I think I have a base to build on—to try things, an awareness of the importance of it all. I would like to use concepts and probably a novel the way we have here. The United Nations charter and its values would be one I might try."

• E: "I think I feel confidence enough to try them all—probably one at a time at first—and then put them all together."

• H: "Yes & no. I'm still a bit foggy in some areas but I think that trying it (even if it flops) will help to straighten it out. On the whole I am a lot more confident with the idea of teaching social studies than I was before taking this class."

**Consolidation phase: Students' perspectives about teaching for critical thinking in elementary social studies and about intervention.** To varying degrees, all 7 respondents to the instructor-created course evaluation and the 9 anonymous respondents to the university course evaluation (included 7 of the 8 study participants and 2 non-participants) expressed positive disposition toward teaching for critical thinking in elementary social studies and about the nature of intervention (e.g., the relationship among social studies, the ideal educated citizen, and critical thinking; the role of cognitive tasks in the instructional framework; the range of critical thinking judgments and their logical sequencing; teaching strategies; and course difficulty).
Six (A, B, C, E, G, K) of the 7 students who responded to the instructor-created course evaluation agreed with the importance of teaching for critical thinking in elementary social studies. For example:

- C: "I think these are life skills. I think everyone needs them and we fail students if they leave school not learning them. I think they need to be taught and practiced in order to become an automatic way of thinking about things. Our future depends on it."

- E: "I think it is extremely important but that the teacher must have a clear view of it to teach it successfully."

- K: "I think it is essential. It is what we as teachers strive for in ourselves as well as in our students. It is a learning process on both sides. It gives students ownership, accountability, and knowledge in their learning."

H, on the other hand, addressed the complexity of critical thinking. Her comment corresponded with considerable absences and insecurity:

- "Seeing how difficult [critical thinking is] to grasp as an adult, I think you need an early entry into it in order to become proficient at recognizing, understanding, and using the intellectual tools."

Seven students indicated that the course had increased their clarity about the nature of social studies or facilitated a greater sense of the role of critical thinking relative to the ideal educated citizen. For example:

- A: "I didn’t have any real knowledge before this class. I found it very useful....I haven’t taken the social studies methods course yet, so I don’t know how many strategies are taught there. From what I’ve heard, it doesn’t involve critical thinking to the same degree—if that’s the case, I think this will be a great foundation for the methods course."
• B said that she was "more conscious of the need to evaluate and justify! (your arguments, etc.)-evaluate sources (is info accurate, are sources credible)-these are things I hadn't recognized as being priorities-this course has changed that."

• H: "It has helped a lot for me, especially in regards to the textbook vs. the curriculum guide. The concept of the Big Idea vs the little stuff in the curric guide helped me a lot!!"

Students were asked about the importance of the cognitive task dimension to the instructional framework. Five students (A, B, E, G, K) indicated that cognitive tasks were important to the framework whereas 2 students (C, H) thought that cognitive tasks detracted from intellectual tools. For example:

• A stated, "I think it is important to leave it in because by making students deliberate and justify their decisions it makes it more meaningful to them."

• E: "I think it is essential to how critical thinking works."

• B: "I think it [the cognitive task dimension] needs to be covered in more depth."

Two students disagreed, indicating that cognitive tasks detracted from the attention that might be more effectively applied to intellectual tools. For example:

• C: "The intellectual tools are more important and in order to develop them they're probably using cognitive tasks. So rather than adding one more thing to possibly confuse the issue-why bother?"

• H added that "other classes have tons of cognitive task information."

Students were mixed in their views about the adequacy of the range of decisions but only one student disagreed with how the kinds of decisions were
sequenced. One student (A) thought more kinds of decisions with less examples would be preferable; 2 students (G, K) thought the full range of decisions was important; and 4 students (B, C, E, H) indicated that depth of understanding might be better realized by reducing the range of decisions and increasing opportunities to explore more examples. A, B, and E indicated that the sequencing of decisions made sense to them.

Students responded to a list of 17 teaching strategies, indicating which had been personally effective:

- All 7 students found the following strategies effective: sorting examples, non-examples, and borderline examples; the use of justification as a means of building students' understanding of relevant criteria; testing claims to try to show they are wrong; the t-shirt strategy; relating criteria of the educated citizen to lesson plans and relating teachers' decisions in each plan to their views of knowledge and the kind of citizen each facilitated; and the use of a novel as a springboard for learning.

- Six of 7 students agreed that the following strategies were effective: personal reflection, followed by pair or small group sharing and then whole class sharing (A, B, C, E, G, K); having students evolve criteria to guide evaluations of different decisions (A, B, C, E, G, K); sorting ideas into categories and justifying their fit (A, B, C, E, G, K), and using journals to
gage students' understandings and to help tailor the course to students' needs (A, B, C, E, G, H).

- Five of 7 students found effective strategies in summarization (A, B, C, E, K), the Request Strategy (A, C, G, H, K), cooperative conflict resolution (B, C, E, H, K) and testing for causal necessity (A, B, E, H, K).

- Four of 7 students (A, B, C, G) indicated that the critical thinking instructional framework worked for them. H indicated that both the instructional framework and the generation of criteria to guide decisions were problematic: "I found this extremely confusing. I was never sure how to write, how to phrase, etc."

- Two of 7 students added to the list of strategies that the concept map (C, H) and the use of various visual aids (C, G) had been helpful.

- Six of 7 students (B, C, E, G, H, K) indicated that the instructor's attempt to facilitate understandings of the attributes of the dimensions of critical thinking gradually through practical application with subsequent elaboration of the critical thinking instructional framework needed to be supplemented with prior background readings. Students were particularly interested in elaborations of the intellectual tools.
Students were invited to pose questions that remained for them from the course and to offer relevant suggestions and comments, 3 of which included concerns about practicality. Sample responses include:

- A: "[Use] more condensed information; teach lesson plans earlier in unit."

- B: "I really appreciate how such an incredible variety of teaching strategies were used in the class—most teachers don’t do this—but it helped me considerably."

- C: "Sometimes I felt there was too much talk about an ideas—we go over & over it until it became unclear; [have] more confidence the students have grasped the idea & give them a task to prove they have, then go over it by reviewing their finished product, instead of beating it to death (muddling it up)."

- C: "How will I use it—will I really be able to put it together esp. in the role of student teacher or subs?"

- E: "How does this work in a real classroom?"

- K: "Where can I get more actual activities for teaching units via critical thinking? PRACTICAL STUFF!"

Responses to the university evaluation (April 1, 1996) displayed largely positive perspectives of the nature of intervention. The identities of the 9 students who responded were not indicated. One study participant (J) was absent and 2 students who had not participated in the study were included in this evaluation. Relevant Likert scale responses are included here as they aligned closely with each other and with other consolidation evidence:
• clarity of objectives (agree-2; strongly agree-6; missing response-1);
• instructor perceived student difficulty (agree-2; strongly agree-7);
• instructor’s comments on assignments were helpful (agree-2; strongly agree-7);
• instructor’s openness to alternative views (agree-1; strongly agree-8);
• instructor made good use of examples (agree-2; strongly agree-7);
• students perception of course difficulty (very difficult-1; somewhat difficult-3; about right -5); and
• effectiveness of instructor (average-3; top 30%-4; top 10%-2).

Related written comments aligned with the Likert scale responses. Students’ written comments supported the instructor’s attention to an open and responsive climate, particularly the adjustments made to overall pacing and remedial efforts to enhance clarity. Students’ responses corresponded with other qualitative evidence about course difficulty, indicating that by the end of the course, the majority of students (N=5) judged that the level of difficulty was about right. Overall, responses were positive about both the nature and effectiveness of intervention, with sufficient evidence of constructive criticism to warrant some degree of confidence in the claims. Some relevant comments from various students included:

• Very self-reflective... Very... approachable.... Need to stop talking so fast and saying the same thing over in three different ways. Assignments must
be spaced out to be more meaningful. Activities presented were excellent, above par, and beyond expectations….Appreciated and taught to the different abilities and difficulties in the class. Clear criteria for assignments….Great class—challenging but I learned a lot…. She was always asking for our input and feedback to help guide the pace and direction of the course….Carole is knowledgeable about her area and could express herself well. She was able to take a complex issue like critical thinking, and break it down into manageable and useful components that we could apply in our future classroom. Carole is a first rate teacher and I have benefited immensely from this course.

Both the exit-level instructor-created course evaluation and the university teacher evaluation indicated that in spite of the challenge presented by the course, and the mid-semester conceptual confusion with its parallel increase in anxiety, students’ perceptions of the overall effectiveness of intervention were positive. This represented a noteworthy improvement over similar student evaluations from earlier informal pilot classes.

Consolidation phase: Alignment of evidence. Consolidation phase evidence indicated mixed alignment across forms of evidence. More specifically:

1. The instructor’s and the students’ perspectives of the effectiveness of intervention were largely congruent with some inconsistencies.

2. Dispositions toward critical thinking, views of knowledge, and argument proficiency were mixed in consistency within and across some forms of evidence and some students.
3. There was no clear indication of congruency among students' dispositions toward critical thinking, their views of knowledge, and their proficiency in argument.

**Congruence across participants' perspectives.** Participants' perspectives of the effectiveness of instruction were largely congruent. Discrepant perspectives involved the instructional framework and the structure of the application phase of the course. Contrary to the instructor's growing negative perception of the instructional framework, 4 students indicated that it had been personally effective. In addition, the instructor had not realized that the unit plan framework had not been clearly communicated and how much this structure had constrained some students' assignments.

**Congruency of each variable within and across some forms of evidence and some students.** Alignment was mixed within and across some forms of evidence in the consolidation phase of intervention for dispositions for critical thinking, views of knowledge, and proficiency in argument.

**Dispositions for critical thinking.** Evidence were congruent in depicting overall moderate positive disposition for critical thinking. However, closer examination revealed inconsistencies that merited investigation: CCTDI scores for 3 students (A, J, H) declined; some CCTDI scores for 2 students (B, C) were inconsistent with qualitative evidence; and the subscale for systematicity
dropped substantially. Some of these results aligned with particular CCTDI subscale items that:

- did not align with a conception of critical thinking supported by an evaluative view of knowledge;
- invited responses that could represent dispositions other than intended;
- invited responses that could vary depending on the situation;
- were vague in intent; and
- expected agreement with absolute generalizations.

Some results corresponded with incongruencies between instruction and instrument dispositional emphases. Some students qualified responses, suggesting that relevant dispositions were stronger than scores indicated. Some evidence suggested that the nature of instruction corresponded with the drop in the mean score for systematicity while some disparity remained a puzzle.

CCTDI scale items did not always represent a conception of critical thinking compatible with an evaluative view of knowledge. Some expected responses reflected multiplicity—not critical thinking. For instance, strong disagreement was expected for the item, “Open-mindedness has limits when it comes to right and wrong.” Yet disagreement implies that all views are equally acceptable, regardless of their potential for harm. This is not a sustainable position for an evaluativist who recognizes relative merit and the status accorded
to moral principles in judging belief or action. H indicated a lack of clarity about
the intent of this open-mindedness item and her posttest score for this item
dropped 4 points. A declined 1 point while J increased 2 points for this item.

The pretest-posttest drop in the mean score for the CCTDI subscale of
open-mindedness involved 3 students (A, -6; G, -4; H, -5). Instruction aimed to
facilitate a compelling case for an evaluative view of knowledge that, if
successful, was anticipated to result in reduced scores for CCTDI items that
reflected multiplism. However, for related open-mindedness items, this appeared
to be the case for only 3 responses, accounting for only a 6-point reduction in raw
scores (e.g., A declined 1 point for 6 items, only 2 that aligned with multiplism; G
dropped 4 points for 2 items and 1 point for 2 items, none which were closely
aligned with multiplism; and H declined 1 point for each of 4 items and 4 points
for 1 item, the latter which appeared to represent multiplism).

Multiplism and vagueness were apparent in expectations for strong
disagreement with “You are not entitled to your opinion if you are obviously
mistaken.” If beliefs do not pose harmful consequences for anyone, there is little
justification for such a judgment. If, however, belief is made public and has
potential negative consequences for someone, it is not clear how one would
rationalize 'entitlement' for such beliefs. This item appears to represent multiplist intent and the notion of entitlement is sufficiently vague for misinterpretation. J's score declined 2 points and A declined 1 point for this open-mindedness item.

Similarly, strong disagreement is expected for the item, "We can never really learn the truth about most things." Yet the absolutist would disagree with this item, stating that we can and do objectively represent reality about many things. The evaluativist would also likely disagree but for a different reason, saying that while it is possible to learn the truth, we might never know how close we get. The multiplist would likely agree, indicating that truth is unique to individuals and thus does not allow for sustainable universal generalizations. Five students' (B, E, G, H, K) scores either decreased or remained low for this item.

Other items also had potential for responses that did not necessarily represent the intended disposition. For example, positive agreement was expected for the following analyticity subscale items:

- "You could describe me as logical" (A, -1; H, -1; J, -1): Actual logical proficiency or one's confidence in what others think does not necessarily represent disposition to be logical; and

- "Frequently I find myself evaluating other peoples' arguments" (A, -1; H, -3): Given the pressures of the end of a school term, it is not surprising that some students were not disposed to evaluate other people's arguments.
Interpretations of some items appeared to be situation-dependent. For example, a response to "Considering all the alternatives is a luxury I can't afford" could depend on whether the consequences of not making a decision have potential for more harm than waiting to gain a full sense of the range of options available. H declined 1 point for this open-mindedness item. The item "If there are four reasons in favor and one against, I'd go with the four" is another situation-dependent example. It could be that one lacks adequate knowledge to judge reason quality and yet is compelled to make a decision, in which case one might be justified to side with 4 reasons in accord than with 1 discrepant reason, particularly if the 4 reasons are from credible sources. This decision does not necessarily indicate that one is closed to evaluating evidence quality. H only scored 2 points in both pretest and posttest responses for this item and J declined 1 point for this item.

Other items were also vague and subject to different interpretations. For example, "It's impossible to know what standards to apply to most items" is open to misinterpretation as 'standards' and 'criteria' are often interchanged. Instruction had addressed such potential and had specifically defined criteria as the quality or ideal desired and standards had denoted the amount such quality was evident. B recorded "unclear about intent" beside this item on the inventory.
The use of absolute generalizations was problematic. For example, expected agreement with the analyticity item, “I must have grounds for all my beliefs” does not acknowledge that some beliefs are too trivial to require grounds. Four students’ (A, B, C, J) scores decreased by 1 point for this item.

Disparity between the CCTDI and intervention emphases on dispositions for critical thinking also corresponded with some cross-evidence inconsistencies. Intervention emphasized:

- an active pursuit of truth, even while recognizing one might not ever know how close one is;
- fair-mindedness, an intellectual responsibility that is requisite to other judgments, involving the active seeking out of alternative positions and striving to understand them and their supporting reasons as fully as possible;
- open-mindedness, remaining open to the idea that there might be evidence that counts against one’s beliefs and when faced with sufficient counter evidence, being willing to changing one’s position; and
- respect for reason, basing judgment on the deliberative evaluation of evidence or support for a claim and its adequacy to sustain a conclusion.

These dispositions for critical thinking seemed most aligned with the CCTDI subscales of truth-seeking, analyticity, and maturity, dispositions for
which most students' scores increased. Truth-seeking scores increased for 6 students (A, C, G, H, K, J), remained the same for 1 student (E), and declined for 1 student (B). Analyticity scores increased for 5 students (B, C, E, G, K) and decreased for 3 students (A, H, J). Some possible reasons for the lower scores for analyticity (discussed above) included vague subscale items for which responses could be interpreted in more than one way.

Maturity, the scale intended to represent epistemological disposition, increased for 6 students (A, B, E, G, K, J), remained the same for 1 student (C), and declined for 1 student (H). H's score declined 3 points for one item that she had circled as being unclear (e.g., “The truth always depends on your point of view”). H also declined one point for “We can never really learn the truth about most things.” (-1). The other two items for which H’s score declined a total of 4 points did appear to reflect weaker maturity.

In spite of posttest decline in CCTDI scores, inspection of specific scale scores showed that A and J both increased in truth-seeking and in maturity scores, scales central to intervention. H, whose CCTDI score dropped -16, gained +4 in truth-seeking but dropped -7 in maturity. It is possible that H’s illness and high absenteeism contributed to insecurity and this decline.

Some responses were qualified, suggesting that students’ scores did not reflect their actual disposition toward critical thinking. B’s pretest-posttest decline
of 2 points for the CCTDI truth-seeking item, 'If there are four reasons in favor and one against, I'd go with the four' did not reflect her written qualification that agreement "depends on strength of argument." For the item, "Even if the evidence is against me, I'll hold firm to my beliefs,' B noted that it "depends on strength of evidence against me." B's pretest and posttest agreement with the item, 'I believe what I want to believe,' could be interpreted as evidence that B had the courage to resist coercion or manipulation. B's 3-point decline for agreeing with 'It’s never easy to decide between competing points of view' could reflect inattention to the absolute 'never.' Agreement did not, however, necessarily imply that relative merit was impossible to judge. Yet B's 3-point decline for agreeing with "To get people to agree with me I would give any reason that worked" and her strong agreement that everyone, including herself, argued from their own self-interest, while intellectually honest, did not align with a disposition toward sustainable judgment or with qualitative evidence that supported B's expression of strong commitment to striving for truth.

One puzzlement was the negative pretest-posttest difference in the CCTDI mean score for systematicity. Scores for 3 to 5 students declined for items related to being easily distracted (N=5), ease of organizing thoughts (N=4), care in keeping records for personal finances (N=3), rushing decisions too quickly (N=3), and procrastinating decisions (N=4). One might attribute such reduction to end-
of-year pressure, had not students in this study declined so much more than did Ber et al.'s students who were similarly affected (1996). Overall, instruction had been conducted systematically, with one exception where a logical flow was inadvertently interrupted. Most students had agreed with the logical sequence of decisions that evolved from a critical challenge. Yet systematicity was not emphasized explicitly during the instruction except in one class near the end of the course. It could be that the complexity or unfamiliarity of instruction ideas and the lack of simple recipes to follow contributed to the drop in systematicity. This possibility corresponds with A's comment on the instructor's course evaluation about a lesson that employed the t-shirt strategy for systematic deliberation about points of view regarding options about what should be done about an issue:

A: "I liked the structure of the last lesson plan best: It was very easy to read, flowed logically, while still maintaining all the component parts. I think that would be a better example to give when first completing the lesson plans."

The notable discrepancy in C's quantitative and qualitative evidence for her relative strength of disposition towards critical thinking remains a puzzle. C's posttest gain of 9 points for a score of 288 reflected the lower end of the range of scores for moderate disposition toward critical thinking (280 - 350), not compatible with qualitative evidence more aligned with strong disposition toward critical thinking.
View of knowledge. Alignment among forms of evidence for students' views of knowledge was mixed (see Table 22). Four students (B, E, H, K) who exhibited consistent evaluativist views of knowledge in posttest interviews also exhibited evaluativist positions in justifications of unit plans. Written argument evidence of their views of knowledge did not necessarily align with other evidence. B was unique in cross-evidence evaluativism. Both E and K's written arguments aligned with absolutism but were inconsistent with their interview and justification evidence. Data were inadequate to make any claims about H.

Inconsistency in or absence of evaluativism in interview responses (A, C, G, J) aligned with unclear evidence of students' view of knowledge in unit plan justifications. A's acceptance of all pupils' views in her unit plan was compatible with multiphism in her written argument, her support for evaluating options and justifying conclusions was more representative of evaluativism, and her inattention to the relative merit of evidence in her unit plans aligned with posttest interview evidence of predominantly absolutist responses. C's justification that questioned authoritative sources clearly opposed absolutism and was compatible with evidence of predominant multiphism in her posttest interviews. However, C's written argument exhibited an evaluativist perspective. G's justification aligned with evaluativism with its focus on the adequacy of evidence to support
Table 22

Cross-Student, Cross-Evidence Comparison of View of Knowledge

<table>
<thead>
<tr>
<th>Student</th>
<th>Kuhn's interview argument</th>
<th>Written argument</th>
<th>Unit plan justification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Return to crime topic</td>
<td>School failure topic</td>
<td>Unemployment topic</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>H</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>K</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Note: Numbers denoted view of knowledge—absolutist, 1; multiplist, 2; evaluativist, 3; letters denoted students in study; ? denoted position unclear.

conclusions but her claim supporting the potential for true evidence aligned more with her consistent absolutist responses for her posttest interview. On the other hand, G’s written argument depicted evidence of evaluativism. J’s advocacy for weighing options and evaluating the impact of decisions did not align with absolutism but absolutism was most apparent in her expression of certainty in both her written argument and in her posttest interviews.

Inconsistencies among forms of evidence might, in part, reflect the nature of the evidence that did not necessarily allow the use of Kuhn’s criteria for
distinguishing students' views of knowledge. Dispositions toward evaluativism, expressed in 5 students' journals (B, C, E, G, K), were not necessarily compatible with their practice (e.g., C and G's inconsistency). It is plausible that inconsistency across forms of evidence reflected insecure epistemological positions. Increased consistency in students' epistemological responses likely reflected increased security in that position.

*Proficiency in argument.* Students' inadequate conceptual distinctions during class activities contrasted with their more successful arguments in their concept maps, unit plans, and justifications for unit plans. Compared with interview data, written arguments exhibited similar success with alternative theories; stronger rebuttals for most students; similar success in generating genuine evidence; and a slight decrease in epistemological reasoning. Compared with interview data, unit plans revealed more cases of genuine evidence whereas justifications for unit plans exhibited slightly fewer students generating genuine evidence. Table 23 depicts cross-evidence proficiency in argument.

In-class activities and the students' instructional plans provided evidence of argument competencies other than for empirical arguments about cause and effect, the foci of interviews and the written argument. Each activity, instructional plan, and justification represented a judgment of meaning relations that aimed for alignment with the instructional framework. Two in-class activities revealed
Table 23

Cross-Evidence Comparison for Proficiency in Argument

<table>
<thead>
<tr>
<th>Student</th>
<th>Kuhn interview - evidence quality</th>
<th>Kuhn interview- rebuttal quality (RQ)</th>
<th>WA</th>
<th>WA</th>
<th>UP</th>
<th>J</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>EQ</td>
<td>RQ</td>
<td>EQ</td>
<td>EQ</td>
<td>RQ</td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>SF</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>J</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>K</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Letters A, B, C, E, G, H, J, and K denoted students in this study; RC—the return to crime topic; SF—the school failure topic; U—the unemployment topic; WA—written arguments; UP—unit plans; EQ—evidence quality, and RQ—rebuttal quality.

NA depicted inadequate relevant evidence.

inadequate understandings while the concept map exhibited basic conceptual distinctions among key dimensions of critical thinking. Six of 8 students (A, B, C, E, G, J) demonstrated largely successful relational arguments in their unit plans and in their justifications.
Five students (B, C, E, H, J) successfully generated genuine evidence for 2 to 3 interview topics. Except for H, these students were also successful across 2 or 3 of 3 forms of qualitative evidence. A, who was unsuccessful generating genuine evidence for 2 of 3 interview topics was consistently successful across all 3 forms of qualitative evidence.

Rebuttal success was limited in interview responses (1 to 3 students for each topic). More students generated successful integrative rebuttals in their written argument (A, B, H, K and J who offered 2 rebuttals in her written argument--1 successful and 1 partially successful). Only 1 student (C) generated a successful rebuttal in the justification. H and K, each who had generated 1 successful rebuttal during interviews, were successful in written argument rebuttals. A, B, and J were unsuccessful generating integrative rebuttals during interviews but offered successful rebuttals in their written arguments. Alignment was close between the interview and the written argument for successful generation of evidence whereas consistency was not the case for rebuttals. C’s unique, consistent success in the interview rebuttals contrasted with only partial success in the written argument.

Students’ justifications represented arguments that were to provide evidence of the relationship between their unit and instructional plans and the ideal educated citizen and arguments that supported the potential effectiveness of
their plan in contrast to more absolutist approaches. The success of students’
arguments in their justifications depended on the adequacy of the evidence that
they provided to illustrate alignment with the intellectual tools of the ideal
educated citizen. Six students (A, B, C, E, G, J) provided appropriate evidence in
their justifications for their relational arguments. The success of students’
argument for relative merit depended on providing evidence for a cause-effect
argument. Five students (A, C, E, H, J) provided assertions about the potential for
relative effectiveness but only 1 student (C) supported assertions with evidence
that discounted alternative approaches. Even so, C’s causal argument did not
address whether the intellectual tools facilitated by her plan were necessary or
sufficient to the ideal citizen.

Some discrepancies among forms of evidence were better understood in
light of the nature of instruction, evidence, and criteria. Insecurity with some
conceptual distinctions was not surprising, given the limited experiences that
students had in selecting appropriate tools for a range of judgments. The
instructor’s decision to limit prior background reading relative to an elaboration of
the dimensions of the instructional framework might also have contributed to such
insecurities. It is plausible that the use of the instructional framework contributed
to students’ sense of alignment among dimensions for the concept map activity.
The performance demands imposed by in-class tasks did not apply to the written assignments that allowed time for reflection and access to relevant course notes. Many of A's interview responses had been unsuccessful because of inadequate specific information. Increased success in the written argument suggested that A benefited from additional time to access background information not afforded by the interview. Discrepancy in rebuttal quality suggested that students' capacity for rebuttals was also not adequately reflected by the interview medium. However, C's consistent rebuttal success across interview topics that was not replicated in the written argument suggests that caution is needed in using just one form of assessment to make important judgment about the capacity of all students.

H and K's inadequate demonstration of comprehension on written assignments corresponded with considerable absenteeism (H missed 14 hours and K missed 7.5 hours). K did not benefit from personal instructor guidance or assignment feedback as all assignments were submitted late after the course.

Instruction did not necessarily align closely with the formal instruments, as in the case of the specific dispositions toward critical thinking, their view of knowledge, and the kinds of arguments that were emphasized during intervention. The instructor had not tried to match the instruction to the instruments. Rather, the instructor had focused on the conceptual framework for teaching for critical
thinking in elementary social studies. The duration of the course necessarily limited the range of attributes that were emphasized. Comparison of dispositions toward critical thinking that were emphasized in the course with similar CCTDI subscales revealed increased consistency.

The nature of some subscale items was problematic. Some subscale items did not clearly align with an evaluative view of knowledge, the epistemological foundation for the conception of critical thinking embedded in the instructional framework. The intent of some CCTDI items was not clear and the expected responses for some items did not necessarily reflect the intended disposition.

As for students’ views of knowledge, instruction did not focus just on Kuhn’s criteria. Intervention emphasized the evaluative ideal citizen who accepted multiple ways of viewing the world but recognized that some positions are better supported than others and the need to evaluate the adequacy of support to sustain decisions. Students were guided toward the notion that certainty was suspect. While students who indicated potential for certainty were challenged with counter-examples, their positions were not dismissed.

The arguments that students engaged in during the course encompassed more than the cause-effect focus of Kuhn’s interviews. Intervention involved students in deliberations that evolved from a critical challenge, resulting in arguments for conclusions about concepts, intent, discrete facts and
generalizations, and the value of actions. Only one class was devoted to a cause-effect argument. Students’ instructional plans formed arguments relative to the instructional framework and justifications of unit plans relative to the ideal of the educated citizen. While expectations that students justify the relative merit of their instructional plans for facilitating their pupils’ progress towards the ideal educated citizen were appropriate to a cause-effect argument, the instructor did not clarify such intent. This could have contributed to only 1 student providing a successful rebuttal for their unit plan justifications.

Five students’ low scores for 1 CCTDI item corresponded to the nature of the instruction. Students had been involved in generating criteria judging source credibility (that included overlap among sources). Low points were received for agreeing with 4 reasons in favor of a claim over one reason against the claim. Although the instruction had also addressed the quality of evidence to support claims, it is possible that the distinctions between the accuracy or plausibility of source evidence and credibility of source were not adequately explored. This explanation is supported by related insecurities in students’ unit plans. In retrospect, these distinctions were central to the evaluativist view of knowledge and required more adequate emphasis and clarification by the instructor.

3. Congruency among variables. Comparison of consolidation phase evidence did not reveal any clear pattern of alignment among variables of
students’ dispositions for critical thinking, their views of knowledge, and their proficiency in argument. For example, moderate strength in CCTDI scores did not predict the view of knowledge displayed in interviews (absolutism: G, J). Nor did evidence of evaluativism in interview responses (H) or qualitative evidence (B, G, J) necessarily correspond with stronger CCTDI scores. Proficiency in argument did not predict strength in CCTDI scores or view of knowledge (C). Such inconsistencies reflected an apparent lack of congruency among constructs.

In summary, alignment among consolidation phase forms of evidence was mixed. Perspectives of instruction effectiveness were largely positive. Discrepancies were found within instruments, among students, and across forms of evidence, most which corresponded with differences in emphases between the instruction and forms of evidence, differences in criteria for constructs, and some instrument and intervention limitations. In areas of alignment between intervention and forms of evidence, cross-evidence consistency and evidence of success increased. The medium of written arguments provided evidence of stronger proficiency in cause-effect arguments, particularly for rebuttals, than did the interview context. Epistemological reasoning did not apparently benefit from the written medium.

These results suggested that students who did not appear to change in their view of knowledge were disposed to change and but still insecure in related
understandings. When interpreting the results of formal instruments, one should keep in mind that students' consolidation phase responses about dispositions and perspectives were made with more understanding of the complexities and issues about teaching for critical thinking and an evaluative view of knowledge than was the case in the preparatory phase of the course so increases could be more important than they appear. Some lack of parallelism between formal instruments and instruction helps clarify the mixed success in students' argument proficiency that did not always relate to students' moderately successful responses in instructional plans and justifications.
CHAPTER FIVE

SUMMARY AND CONCLUSIONS

This pretest-posttest case study explored the responses of 8 female 3rd and 5th year University of Victoria preservice teachers to a 13-week course intended to facilitate dispositions, understandings, and proficiencies relevant to teaching for critical thinking and an evaluative view of knowledge in elementary social studies. This chapter provides an overview of methodology and key research results; discusses instruction, instrumentation, and the interpretive framework; and delineates some challenges to teaching for critical thinking in elementary social studies, strengths and limitations of the study, and implications for future research.

Methodology

This case study documented a small and non-random intact group of 8 of 10 female university students' dispositions toward critical thinking, their views of knowledge, and their argument proficiency over 13 weeks of an advanced curriculum and instruction elementary social studies course. Prior to this study a conception of critical thinking and an instructional approach in social studies were developed and pilot tested. Pilot study results were used to modify the instructional approach and to identify related dispositions, views, and proficiencies.
The instructional framework used in this study integrated critical thinking dimensions, an evaluative view of knowledge, and instructional plan elements as they applied to the elementary social studies curriculum. The epistemic foundation of an evaluative view of knowledge was implicit in the ideal educated citizen that provided the vision for instructional plans (Siegel, 1992). Critical thinking dimensions included (1) judgments associated with a critical challenge about what to believe or what to do; (2) cognitive tasks of self-regulation, deliberation, judgment, and justification; and (3) the intellectual resources of the ideal critical thinker: background knowledge of the problematic situation in context, critical thinking concepts, criteria and standards for judging relative merit, habits of mind, and thinking strategies (Bailin et al., 1993; Facione, 1991). Instructional plan elements included learning outcomes; associated teaching, learning, and assessment activities; and relevant evaluation criteria. This instructional framework was used to plan the course, provide a visual organizer as scaffolding to be elaborated with students following related practical engagement, and to judge students' instructional plans and justifications.

Triangulation of quantitative and qualitative evidence facilitated stronger conclusions about the preservice teachers' entry-level and exit-level critical thinking dispositions, their views of knowledge, and their argument proficiency; pretest-posttest changes; and the participants' responses to instruction (Miles &
Huberman, 1994). Formal instruments included The California Critical Thinking Disposition Inventory (CCTDI, Facione & Facione, 1992) and Kuhn's interview protocol (1991). The CCTDI included measures of truth-seeking, open-mindedness, analyticity, systematicity, inquisitiveness, self-confidence in critical thinking, and cognitive or epistemological maturity. Kuhn’s interview protocol provided evidence of students’ views of knowledge (absolutism, multiplism, and evaluativism) and evidence of their argument proficiency about 3 social issues. Other evidence sources included a written argument, participants’ journals, an in-class survey, 2 course evaluations, instructional plan assignments, student-generated artifacts, and the instructor’s lesson plans. The written argument that paralleled Kuhn’s interview provided evidence that was removed from instructor-interviewer influence and the oral language medium.

The use of the CCTDI and Kuhn’s interview protocol as a pretest and posttest served to measure entry-level and exit-level critical thinking dispositions, views of knowledge, and argument proficiency for the 13-week course. Qualitative comments and artifacts were collected during the course and a written argument was completed near the end of the course. These data sources were scored and analyzed to address the 4 research questions.
Results

What Were Entry-Level and Exit-Level Dispositions for Critical Thinking, View of Knowledge, and Proficiency in Argument?

Preservice 3rd and 5th year elementary social studies teachers’ pretest mean scores revealed moderate positive critical thinking dispositions (322.5), mixed views of knowledge (2.1), and partially successful argument proficiency (2.1). Mean scores masked variation across categories, dimensions, topics, and students. When viewed in light of the mixed-gender American samples, the University of Victoria students’ pretest CCTDI overall scores were higher than first year students’ pretest scores and lower than for graduate students (Bers et al., 1996; Facione et al. 1994). Caution is required when the single-gender CCTDI data from the present study are compared with mixed-gender data from other research. Facione et al. (1995) reported no gender differences in 1st year college students’ overall disposition toward critical thinking but small significant differences were observed for some subscales. Females were more inclined than males toward open-mindedness and maturity and less inclined toward analyticity in one study. Similar results were found in a second study for open-mindedness and maturity but not for analyticity. Ber et al.’s (1996) pretest scores for 1st year college students indicated that compared with males, females displayed stronger scores in truth-seeking, open-mindedness, and maturity. In the present study, the 3rd and 5th
year female students’ mean scores for truth-seeking, open-mindedness, and maturity were higher than either Ber et al.’s male or female students’ scores. As for epistemological perspectives, compared with Kuhn’s (1991) main sample and college sample, the University of Victoria students exhibited more evidence of evaluative views of knowledge across social issues with one exception (School Failure). The University of Victoria students’ weaker performance for the School Failure topic was of interest, considering that students’ career choice was elementary teacher education. The University of Victoria students exhibited partial success in argument proficiency, somewhat less than for Kuhn’s subjects. These entry-level data did not support Kuhn’s suggestion that argument proficiency for one topic predicted success for other topics. Kuhn did not find significant gender differences in either view of knowledge or argument proficiency data. Nor did reflective judgment research find significant gender influence for epistemological perspectives (Woods, 1997).

The University of Victoria students’ exit-level CCTDI and interview group mean scores aligned with modest positive critical thinking dispositions (329.8), mixed views of knowledge (2.3), and partial but mixed success in argument proficiency (2.2). These 3rd and 5th year female education students’ exit-level CCTDI scores exceeded mixed-gender first year college students in critical thinking disposition scores (Bers et al., 1996). The percentage of University of
Victoria students who scored less than 40 were similar to or less than Facione et al.'s (1994) mixed-gender graduate students in 5 of 7 subscale scores; the percentage of education students who scored more than 50 were similar to or exceeded the graduate students in 4 of 7 subscales. Percentage-wise, compared with Kuhn's (1991) samples, the University of Victoria students exhibited more evidence of evaluative views of knowledge, genuine evidence, objective interpretation of evidence, and appropriate evaluations of overdetermined evidence. On the other hand, the University of Victoria students generated fewer successful alternative theories, counterarguments, rebuttals, and evaluations of underdetermined evidence.

**What Pretest-Posttest Change was Evident in Students' Dispositions for Critical Thinking, View of Knowledge, and Proficiency in Argument?**

Changes in mean scores were small but positive. Closer examination revealed relative stability for individual students’ views of knowledge and epistemological reasoning but substantial positive change for alternative theories and more sizable changes in dispositions than reflected in the mean scores. Data were compared with similar research. The percentage of the University of Victoria students who had pretest CCTDI subscale scores under 40 was more than Facione et al.'s (1994) mixed-gender graduate students’ scores and the percentage of their scores over 50 was lower than for the graduate students. On the posttest,
the percentage of the University of Victoria students’ posttest scores that were less than 40 were similar to or less than the graduate students for 5 of 7 subscales and the percentage of their scores over 50 exceeded the graduate students’ scores for 4 of 7 subscales. Gender differences were observed by Ber et al. (1996) who found that compared with female students, more male students increased and fewer males decreased in self-confidence scores. The percentage of students in the present study who increased and decreased in self-confidence scores were similar to Ber et al.’s females. As for truth-seeking, Bers et al. (1996) reported that more males’ scores increased and decreased than did females’ more stable scores.

The present study findings parallel reports that epistemological beliefs are slow to change (King & Kitchener, 1994; Kurfiss, 1988; Perry, 1968; Woods, 1997). Students who demonstrated consistent evaluativist responses across pretest topics did not regress in posttest responses. Evidence did not support a linear progression from absolutism through multiplicity to evaluativism (Kitchener & King, 1981; Woods, 1997). Kuhn (1991) did not find significant gender differences in epistemological change; nor did advancing age predict an evaluative view of knowledge. Rather, level of education has been consistently associated with views of knowledge (Kitchener & King, 1981; Kuhn, 1991; Perry, 1968; Woods, 1997).
Mean change in argument proficiencies was also minimal, aligning with research that suggested more than one semester is required for substantial change (Hatcher & Price, 1998). Alternative theories were distinguished by substantial positive change scores, particularly for the School Failure topic. In spite of a slight increase in instances of genuine evidence, the relative quality of genuine evidence was both inconsistent for individual students across topics and declined from pretest to posttest.

**Was the Written Argument Evidence Consistent With the Interview Protocol Evidence?**

Overall, consistency between interview and written argument findings was weak. Views of knowledge were inconsistent across these 2 mediums. Argument rebuttal proficiency was stronger and epistemological reasoning was slightly weaker in the written arguments.

**What did Participants’ Informal Responses Suggest About the Effectiveness of Intervention?**

Qualitative responses revealed the following findings.

1. There was a lack of clear evidence of alignment among student critical thinking dispositions, their views of knowledge, and their argument proficiency.

2. Written arguments produced stronger rebuttals than did the interview protocol.
3. Epistemological reasoning did not benefit from the written argument medium.

4. Evidence of rough overall alignment across qualitative and formal measures in moderate dispositional strength, mixed views of knowledge, and mixed proficiency in argument was accompanied by internal inconsistencies.

5. Students indicated that increased level of concern during the application phase had dissipated considerably by the end of the course. Students reported that the early introduction of a comprehensive instructional framework and inadequate prior readings to elaborate the framework dimensions contributed their level of concern. Most students indicated that effective instructional experiences included: the initial focus on attributes of the ideal educated citizen, an interactive constructivist approach, concept attainment strategies, expectations that students justify responses against shared criteria, challenging conclusions with alternative explanations, and specific instructor feedback designed to guide revisions and enhance success.

6. The instructor found that seeking students’ responses to instruction allowed for more informed instructional decisions to better align with students’ readiness.

7. Positive change was more apparent in measures that aligned with instruction.

Discussion

Some of the more interesting results from this study involved the adequacy of results relative to students’ responsibilities as educators, the apparent resistance
to change, some promising practice for change, the challenges to clarity, and the adequacy of instrumentation and the instructional framework for guiding practice and assessment. The following discussion addresses these results.

**Results Relative to Ideals**

Study results suggest that the University of Victoria 3rd and 5th year preservice elementary social studies teachers’ critical thinking dispositions were moderately strong for all students. Yet the posttest mean score range of 81 points placed some students at the lower end of the range for moderate dispositional strength. Only half of the education students displayed consistent evaluativism in interviews. Qualitative responses were not always congruent with interview data. Although students in the present study interpreted information objectively and were quite successful generating evidence, alternative theories, and rebuttals, they were less successful generating counterarguments and in epistemological reasoning. Inconsistencies in genuine evidence quality suggested that students were not cognizant of the quality of evidence generated. This could account, in part, for limited epistemological reasoning. Limitations in personal dispositions, perspectives, and proficiencies would seriously impede prospective teachers’ preparedness to help their pupils acquire these attributes.

Some research suggests that teachers’ views of knowledge and critical thinking dispositions influence how they teach (Lyons, 1990; Orton & Lawrenz,
If these results are generalizable, half of the students in the present study will lack adequate resources for teaching toward the educated citizen. If the implications of ill-equipped citizens are untenable, and if educational institutions' major goal is to help students fulfill the responsibilities of citizenship, then it is reasonable to expect that teachers demonstrate these qualities themselves. If preservice elementary social studies teachers lack the qualities of the educated citizen, prospects for facilitating their pupils' journey toward this ideal are dim.

**Resistance to Change**

Instruction in a 13-week course was not adequate to substantially impact students' critical thinking dispositions, their views of knowledge, and their argument proficiencies. These results were aligned with related research that reported peoples' critical thinking dispositions, their views of knowledge, and their argument proficiencies are resistant or slow to change and may evidence instability during transition. Intervention duration, education level, the nature of intervention, and gender have been identified as variables relevant to change in dispositions, perspectives, and proficiencies that do not appear to evolve as a naturally occurring phenomenon (Kuhn, 1991). Woods' (1997) secondary analysis of reflective judgment research found little evidence of change toward evaluativism in pretest-posttest intervention studies of short duration. Hatcher and Price (1998) reported data that 2 semesters made a statistically significant positive
difference to students’ critical thinking in written composition whereas 1 semester did not make much impact. Research has reported a consistent relationship between educational level and epistemological perspectives (Brabeck, 1980; King & Kitchener, 1994; Kuhn, 1991; Woods, 1997). Even so, few graduate students demonstrate the upper levels of epistemological reasoning, suggesting that education level may be important but insufficient for an evaluative view of knowledge. It may be the nature of education rather than the level of education that impacts students’ views of knowledge and reasoning proficiency. Lehman, Lempert, and Nisbett (1988) found that differences in reasoning proficiencies aligned with the kind of reasoning central to different disciplines. Kuhn (1991) reported that 5 advanced philosophy doctoral students uniformly displayed evaluative views of knowledge and excelled in argument proficiencies. It could be that the nature of reasoning central to advanced philosophical studies facilitated students’ reasoning about grounds for belief. Whereas Ber et al. (1996) found little evidence of change in critical thinking dispositions following 16 weeks of intervention, the University of Victoria students evidenced small changes in critical thinking disposition scores over 13 weeks of intervention. Much dispositional change aligned with the intervention emphases. Such results suggested that researchers needs to consider carefully the nature of instruction and its alignment with measurement instruments if conclusions about students’
resistance to critical thinking are to be warranted. It is also possible that the nature of some University of Victoria students' prior education made them receptive to change, although their comments about their prior education suggest otherwise.

As for gender influence on responsiveness to change, students in the present study exceeded Bers et al.'s (1996) males and females in increased scores for truth-seeking but responded similarly to other females with less inclination toward self-confidence than males and strong disposition toward open-mindedness. Ber et al. (1996) discussed their data in light of Belenky et al. (1986) who identified women's priorities for understanding and caring relationships as well as some evidence of women's silence or perceived lack of voice that were attributed to life experiences and feelings of disempowerment. Ber et al. expressed concern about the gender differences in self-confidence. While such concern should not be dismissed, it is not clear that all CCTDI self-confidence subscale items represent the responses of a critically thoughtful person. Belenky et al.'s study has been recognized for its contributions to understanding female epistemological perspectives. However, Belenky et al.'s interview protocol was not clearly epistemologically oriented. Clinchy (1996), a co-author of the Belenky et al. research, admitted that some interview responses more appropriately identified the basis for women's beliefs rather than the source of women's
knowledge. Other epistemological research has not identified significant gender differences (Kitchener & King, 1981; Kuhn, 1991; Woods, 1997). Kuhn (1991) also reported no significant gender differences in argument proficiencies.

Judgments about the instructional impact on critical thinking dispositions, views of knowledge, and critical thinking proficiencies also need to be considered in light of anecdotal and research evidence that suggest until conceptual distinctions are internalized, students' judgments may be impeded. If this is so, empirical testing immediately following a critical thinking course could underestimate instructional impact since this may not allow time for internalization (Blair, 1992; Scriven, 1988). For example, Beck et al. (1982) found a performance lag for vocabulary comprehension and Shymansky, Yore, Treagust, Thiele, Harrison, Waldrip, Stocklmayer, and Venville (1997) found that complex ideas that students have demonstrated, until secure, may appear to regress.

Promising Practice for Change

Positive change towards the ideal of evaluativism is possible when favorable conditions coalesce. Promising practice reported in the literature includes: instruction that is incremental; conceptual demands that are limited; and explicit criteria for sustainable judgment (Bangert-Drowns & Bankert, 1990; Hatcher, 1995; Leshowitz & Yoshikawa, 1996; Nelson, 1996; Paul, 1996).
Students appear to benefit from epistemological knowledge to help them see the point of critical thinking (Nelson, 1996).

Instruction in the present study that corresponded with these ideas met with positive participant responses. Some positive components of instruction included interactive constructivism, concept attainment strategies, specific assignment feedback from the instructor with revision opportunities, seeking students' responses to instruction to guide future instruction, and having students build an epistemological rationale for teaching for critical thinking and an evaluative view of knowledge in elementary social studies. The global approach of interactive constructivism built on students' prior knowledge, utilized students' ideas, and provided time for individual reflection, small group sharing, and large group interaction. Interactions involved discussion, evaluation, challenging ideas, and in particular, justification of responses against shared criteria. Students' responses involved individual reflection and group engagement, culminating in individual constructions and justifications of arguments for belief and action. Students evolved criteria to distinguish examples, non-examples, and sometimes borderline examples of key concepts and of assignment expectations; evaluated and revised criteria; applied criteria to other samples, and justified judgments. Students utilized specific instructor feedback to guide successful assignment revisions. In addition, seeking students' responses to instruction allowed the instructor to
monitor and adjust instruction to reflect students' needs and to help address students' concerns. Nelson (1996) found similar benefits from interacting with students to monitor their comprehension. Students were also positive about the early epistemological focus on constructing a compelling case to teach for critical thinking and an evaluative view of knowledge. Students' prior knowledge provided the basis for their vision of the ideal educated citizen while the draft curriculum (British Columbia Ministry of Education and Ministry Responsible for Multiculturalism and Human Rights, 1993) provided a referent against which students' vision could be compared and refined. Practical connections were made between students' vision of the educated citizen and 3 different lesson plans, each of which represented instruction to facilitate a separate view of knowledge—absolutist, multiplist, or evaluativist. Students inferred and justified the kind of citizen each lesson plan would facilitate. These lesson plans were linked explicitly to views of knowledge and critical thinking instruction. Students' epistemological understandings were then applied to articles that represented different views of knowledge. Students decided that an evaluative view of knowledge was most sustainable for public schools that serve a complex citizenry.

These preparatory phase activities immersed students in critical thinking toward judgments about meaning and value; development, evaluation, and application of relevant criteria; and justification of judgments against criteria. It
could be legitimately argued that the epistemological emphasis posed limitations on what might reasonably be accomplished toward teaching for critical thinking. Yet the vision of the ideal citizen provided criterial attributes against which application phase instructional plan assignments were justified. Further justification for the epistemological focus was found in research that had documented resistance to change (Perry, 1968) and pilot studies that had documented such resistance. Furthermore, the importance of dispositions for change was reported in implementation literature and research speculation (Facione et al., 1995; Fullan & Stiegelbauer, 1991) and researchers and educators have suggested that epistemological connections and critical thinking concepts need to be explicit (Bailin, 1995; Bangert-Drowns & Bankert, 1990; Leshowitz & Yoshikawa, 1996; Nelson, 1996; Norris & Phillips, 1994). Thus the option of building an epistemological rationale to teach for critical thinking in elementary social studies was more sustainable than simply attending to the resources important to constructing and teaching toward sustainable judgments. Given the short duration of the course and the long time apparently required for full comprehension of critical thinking, the instructor decided that an early focus on having students build a compelling case for an evaluative view of knowledge and critical thinking might provide the foundation for future pedagogical impact. Empirical research is needed to judge whether this decision was justified.
Challenges to Clarity

This study involved several challenges to clarity. Some instructional decisions; an overwhelming concept load; the unique demands of a curriculum and instruction course for preservice teachers; limited congruence among instruction, students’ prior educational experiences, and the elementary social studies curriculum; the relative complexity of social studies; the constraints of a one-semester course on assimilating ideas; and the time required for an interactive constructivist approach towards deep understanding corresponded with increased level of concern during the study’s application phase.

Instruction

Less effective instruction related to clarity concerns. These included (1) inadequate early clarification of instructional organization, (2) the decision to use the instructional framework as a visual organizer and scaffolding for elaborating lesson plans following practical engagement with each kind of judgment (Ausubel, 1963), and (3) the decision not to share comprehensive readings prior to practical engagement with pertinent instructional framework elements. The plan to work with one critical challenge and associated judgments allowed incremental investigations into each kind of judgment and was intended to demonstrate the logical connections among judgments. Unfortunately the sequential organization of the course that emphasized the use of model lesson plans within a potential unit
plan was not adequately clarified. The connections among model lesson plans also imposed unanticipated assignment constraints, particularly problematic for students who missed classes. The instructional framework complexity combined with rushed elaborations of the framework toward the end of two classes further impeded clarity. Notwithstanding the desirability of increased clarity, it may be that some anxiety during knowledge construction is characteristic of an interactive constructivist approach.

**Concept Load**

Overall, the most prevailing impediment was the heavy concept load for both the instructor and the students. Concepts related to epistemological knowledge (absolutism, multiplism, and evaluativism); content knowledge (nature of social studies, social studies curriculum; critical thinking), and pedagogical knowledge (teaching, learning, and assessment ideas and evaluation criteria; elements of and criteria for instructional plans) are important components of a teacher education course for elementary social studies. Unlike courses that have reported successful integration of subject content, critical thinking content, and perhaps epistemological knowledge (Hatcher & Price, 1998; Leshowitz & Yoshikawa, 1996; Nelson, 1996), teacher-education courses must also integrate pedagogical knowledge and domain knowledge (the nature of the subject area and the school curriculum). These responsibilities add to the conceptual and logistical
demands for one 13-week, 39-hour course—quite a challenge given research that has suggested the need to limit concepts. Such conceptual demands lead educators to the unrealistic fragmentation of learning in behaviorism. If authentic learning is to be promoted, then these conceptual demands will need to be addressed.

**Incompatible Prior Education**

Students' reported that evaluating grounds for belief had not been a salient part of their prior education; some could not recall questioning the basis of claims at all. The problematic nature of knowledge was not emphasized. Assumptions that a liberal education prepares students for critically thoughtful judgments do not correspond with available evidence. For example, knowledge of concepts of evidence, conclusion, and justification do not appear to have been prominent in students' prior education (Leshowitz & Yoshikawa, 1996; Norris & Phillips, 1994; Nelson, 1996; Wright, 1995b). If these concepts are part of the conceptual background knowledge of educators, they may be tacit knowledge that educators have trouble articulating (Nelson, 1996; Paul, 1996). Evaluativism and a compatible conception of critical thinking are not an apparent priority of undergraduate education (Goodlad, 1990; Paul et al., 1997).

**Complexity of Social Studies**

The complexity of social studies added to these impediments. Distinguished as an area of study by its teleological citizenship emphasis and its
focus on relationships among people and between people and their environment (Bliss, 1991-1992; Bender, 1986-1987), social studies encompasses many social sciences with different goals, issues, questions, theories, methods, evidence, and criteria (Fullinwider, 1991; Paul, 1992). There is even disparity within each social science discipline. For example, Furet (1983) describes history as a discipline in which consensus is lacking about criteria for effective practice. Novick (1988) claims that the specification of history into sub-disciplines has led to new questions, innovative methods, higher standards, and a proliferation of knowledge, but at the expense of a discipline-wide conversation. The implications for the social studies educator who is responsible for the criteria and methods that guide this integrated subject area are substantial.

Social studies teachers who would supplement or replace the social studies textbook are further hindered by epistemological disparity and vagueness in social scientists’ accounts. For example, social science evidence varies in form and in strength, ranging from archeological radiocarbon testing of charcoal evidence, that can be replicated, to Davis’ (1987) more speculative evidence—pardon tales or letters written by or for people accused of crimes to persuade the king that there was adequate reason for pardon. Davis read between the lines and across the grain of the letters to better comprehend the interface between the tellers’ motivations and the structures in place in 16th century France that helped shape the
tales. Interpretations were then evaluated for their "fidelity to 'real events'" (p. 5) as represented by others in order to comprehend the truth-value of the tales as perceived within their historical context and to construct plausible representations of the past. In the social sciences, evidence is often indirect—physical artifacts or testimony—that must be interpreted and represented by observers with different perceptual filters in the construction of sustainable conclusions. How evidence is used also varies. Evidence may be represented indirectly (Fisher, 1977), explicitly described and evaluated for quality and adequacy to sustain conclusions (e.g., Arthur, 1974; Davis, 1987), or sometimes no evidence is offered—only unsupported assertions (Barbeau, 1960).

**Incompatible Curricula**

Historical claims in social studies are compatible with "best explanation" arguments that involve evidence that cannot be experienced directly and conclusions that remain open to new evidence (Ennis, 1996, p. 222). This evaluativist epistemological position has not been apparent in textbooks and curricula. Indeed, traditional curricula have exhibited an absolutist perspective for the right answer—historical facts that are revealed (British Columbia Department of Education, 1936; British Columbia Department of Education, 1934). More recently, multiplism has been reflected in an openness to different perspectives without much consideration of the relative merit of support for diverse beliefs as
long as conclusions are internally logical (Ford et al., 1997). This criterion neglects the fact that arguments can be logical but the grounds unsustainable. The curricular rhetoric for evaluativism and critical thinking is increasing but there is limited evidence that these ideals have been adequately translated and integrated for practice (Ford, 1988; Ford et al., 1997). Social studies textbooks, have tended toward unsupported assertions—conclusions devoid of evidence (Ford, 1988; Paul, 1992). Such inadequate curriculum resources add to the challenge of critical thinking education, particularly considering that social studies embraces a full range of judgments, unlike other disciplines that emphasize more limited and usually less contentious judgments (Ford et al., 1997).

**Clarity: An Elusive Ideal**

Limited congruency among students' prior education, intervention, and existing curriculum and supporting resources, combined with the complexities of teaching for critical thinking and an evaluative view of knowledge in the context of an elementary social studies curriculum and instruction course for preservice teachers, contributed to the challenge to clarity. Yet the conceptual distinctions associated with critical thinking are important to discriminating evidential bases for judgment (Case & Wright, 1997). Research reveals a need for such conceptual understandings for students, teachers, and teacher educators (Court & Francis, 1993; Norris & Phillips, 1994; Paul, 1996; Unks, 1985; Wright, 1995b).
Ideally, models of critical thinking dispositions, an evaluative view of knowledge, and critical thinking argument proficiencies should be clear and the instructor should have adequate understandings to avoid confusing students. Yet large scale American studies have found a lack of such understandings pose a dilemma for change (Goodlad et al., 1990; Paul et al., 1997). A basic understanding of critical thinking and an evaluative view of knowledge prior to instruction is important to successful implementation, but even more important is the teachers’ desire to learn. Teacher-student interaction is necessary to move the teacher and students closer to full clarity.

Given the complexity of critical thinking, the challenge remains to find ways to satisfy a novice’s need for understanding while retaining the essence of critical thinking. Unfortunately, a simple unified conception of critical thinking is problematic. What is comprehensible and useful to the expert is not comprehensible or useful to the novice. Paul (1996) argues that while there is no one common definition of critical thinking accepted by all those who have seriously studied critical thinking, there is a common core of meaning reflected both in the multiplicity of definitions and in the history of the concept....Given the complexity of critical thinking—its rootedness in 2500 years of intellectual history as well as the wide range of application—it is unwise to put too much weight on any one ‘definition’ of critical thinking. Any brief formulation of what critical thinking is is bound to have important limitations....There is no one way to define what critical thinking is, nor one way to explain it” (pp. 14-15; italics in original).
Full understanding of complex concepts can take many years to realize and conceptions continue to evolve through different applications (Scriven, 1988). This was the case among the 40 critical thinking experts who contributed to the Delphi Consensus critical thinking conception. Weinstein (M. Weinstein, personal communication, January 5, 1996) indicated that the 1991 report did not necessarily reflect a current consensus. Anderson and Freebody's (1981, p. 341) citation from Labov is germane to the elusiveness of clarity and the problematic nature of simple definitions: "Words have often been called slippery customers, and many scholars have been distressed by their tendency to shift their meanings and slide out from any simple definition."

Yet simplicity is desirable and likely even necessary for early stages of critical thinking education. But simplicity is not justifiable at the expense of concept integrity (Scriven, 1988). Reductionism that divests a concept of essential distinctions disassociates the idea with the phenomenon it was intended to represent. Consider, for example, a critical thinking conception that merely itemizes discrete skills or delineates cognitive tasks that are embedded in decision making. Such simplicity may be useful for a skills and drill model of teaching but it is not very useful for the educator who needs to know what counts as evidence or reasons and whether the evidence or reasons are adequate to sustain a conclusion in order to plan for teaching and assessment (Court & Francis, 1993).
Without the claims, evidence, and warrants, it is difficult for educators to help their students construct, evaluate, and justify decision quality.

If simplistic notions of critical thinking lack the essence of critical thinking, they are not morally sustainable. They do not help students acquire the knowledge and proficiency to make good judgments about belief and action, leaving them vulnerable to manipulation. Siegel (1980) argues that educators have a moral mandate to honor students' rights as persons to reasons and justification for what is taught. Teaching in a critical manner demonstrates, facilitates, and expects students to question, challenge, reason, and justify. Sustainable critical thinking conceptions need to satisfy both the essence of critical thinking and educators' moral mandate to teach in a critical manner. Self-regulation toward sustainable judgments that are guided by criteria and standards of quality arguments is necessary to sound critical thinking conceptions. The task facing teacher educators is to identify and integrate significant aspects of critical thinking into a compelling, meaningful, and practical learning experience that allows for gradual progression of dispositions, understandings, and proficiencies.

**Adequacy of Instrumentation and Interpretive Framework**

The adequacy of instruments used in this study for measuring critical thinking dispositions, views of knowledge, and argument proficiency was mixed. The initial intent had been to utilize the CCTDI as a broad indicator of critical
thinking disposition. However, 4 students displayed inconsistent dispositions between their qualitative evidence and their CCTDI scores. This provided reason for a more in-depth analysis to gain insight into the discrepancies. CCTDI inadequacies that had been identified before the study were found to be more problematic than had been anticipated. Some subscale items invited answers that did not necessarily reflect the intended disposition. Additionally, individual test items did not consistently represent an evaluative view of knowledge, important to an adequate inventory for critical thinking dispositions. There was some correspondence between the specific dispositions that were emphasized during intervention and positive change in specific CCTDI subscale scores. Some dispositions that were emphasized in the course were not comprehensively represented by a separate subscale. Perhaps a more comprehensive representation of specific dispositions would have detected more change. Or put another way, it may be that any general scale for critical thinking disposition is inadequate for intervention research. Furthermore, the disparity across measures suggested that performance compatible with critical thinking dispositions (as found in the qualitative evidence in this study) should be compared with inventory assertions. Performance evidence better represents habits of mind than do expressions of disposition (Norris & Ennis, 1989).
Students’ responses to Kuhn’s interview protocol provided evidence of views of knowledge and cause-effect argument proficiency. The simplicity of Kuhn’s 3 epistemological categories was useful for integration into instruction. The simplicity that benefited instruction had less utility for evaluating performance evidence of views of knowledge. More elaborated view of knowledge criteria would facilitate more refined evidence of epistemological change and application to mediums that have more utility in an educational intervention study (e.g., Christen, Angemeyer, Davison, & Anderson, 1994). A major drawback to Kuhn’s protocol was the substantial time that it took the researcher to attain interpretation consistency (Hofer & Pintrich, 1997). Additional concerns that emerged included the discrepant nature of evidence in natural science and social science and students’ interpretation of argument genre. Judgments about social issues involve people who are not amenable to scientific experimental manipulation (Gardiner, 1952). It may be, as Keefer (1996) argues, that some students did not perceive the form of argument as one requiring scientific evidence and justification. Rather, they may have viewed the genre as a practical argument for which their personal experiences were appropriate evidence. Kuhn (1991) would classify experiential narrations as pseudo-evidence with limited value for generalized causal explanations. Yet the inconsistency in the kind of evidence that students in the present study offered within and across
pretests and posttests suggests that students were not cognizant of what counts as quality evidence for either a practical reasoning or scientific argument. After examining students' responses, it became apparent that respondents' pseudo-evidence were inadequate for confidence in general causal conclusions about human behavior, regardless of how students interpreted the genre.

As for the critical thinking conception and the instructional framework in which it was embedded, study participants found it initially overwhelming. The cognitive task dimension received mixed responses but most students favored its inclusion. Two students thought that the cognitive tasks detracted from understanding intellectual resources. Yet as integrated, the overlap among categories (cognitive tasks, judgments, and some intellectual resources) was recognized by some students and likely contributed to some lack of clarity. Also, adding the cognitive dimension to the instructional framework contributed to its unwieldiness.

The main reason for including the cognitive dimension was to help clarify the connections between cognitive tasks and intellectual resources and in so doing, also demonstrate a compelling case for the necessary role of intellectual resources in social studies education. Implementation literature has addressed the need to illustrate the connections between the innovation (in this case critical thinking) and existing beliefs, curricula, and practice (Fullan & Steigelbauer, 1991).
Curricula continues to emphasize cognitive processes and activities. Missing, however, has been adequate guidance for teachers and students to judge what counts as quality cognitive engagement (Ford, 1988). Moreover, curricula infrequently has expected students to decide what to believe about empirical claims. The resources for judgments and expectations for students to make such judgments have been inadequately integrated to facilitate implementation. This could account for some of the inadequate evidence of critical thinking in elementary classrooms and in teacher education classrooms (Bognar, Cassidy, & Clarke, 1996; Ford et al., 1997; Goodlad et al., 1990; Paul et al., 1997).

It has been argued that cognitive tasks are presumed in the application of intellectual resources. For example, if the intellectual resources are applied to judging what to believe or what to do about a problematic situation, one would be immersed in deliberation and judgment. Self-regulation may also be part of such engagement as one checks for adequate integration of resources, but such self-regulation may not be ongoing. It is not clear, however, that justification is necessarily implicit in the Bailin et al. (1993) conception upon which this study is based. For example, one can think through a problematic situation without there being any expectation that related judgments are justified. Justifications involve explicit arguments that show how well reasons and evidence align with relevant criteria and also provide evidence of the strength of relationship between support
and conclusions. In the present study, explicit justifications in conjunction with the use of examples and non-examples were perceived by both teacher and students as a powerful means to enhancing understanding. Explicit attention to the relative merit of reasons and decisions requires that students make conscious the adequacy of their use of intellectual resources. Such consciousness is essential to evaluating one’s own thinking and thus to self-regulation and to education toward the ideal educated citizen.

The relationship between cognitive tasks and intellectual resources could be illustrated briefly early in a course, without making cognitive tasks an explicit part of a framework, which would reduce complexity. This does not address the concern that preservice teachers might find utility in some scaffolding to help with the systematic integration of intellectual resources. The thinking strategy dimension of intellectual resources could be elaborated to provide such a structure. However, it is the view of this researcher that justification is a necessary dimension of an adequate conception of critical thinking, central to its epistemological foundation and the nature of knowledge. As such, it strengthens a conception of critical thinking in an educational context. Justification is also important to the educated citizen who must decide what belief or action is most justified. Empirical research is needed to judge whether an attempt to clarify the connections between cognitive tasks and intellectual resources by including both
in an instructional framework has more utility for implementation than a conceptual framework that excludes cognitive tasks.

The conceptual framework in this study requires further deliberation and adjustments to increase clarity, reduce dimension overlap and to increase the salience of significant concepts such as evidence as they relate to judgment, criteria, and justification in order to better facilitate translation of theory into practice. Conceptual theorists who are interested in implementation of critical thinking face substantial challenges if misinterpretations are to be reduced. Critical thinking and an evaluative view of knowledge need to be integrated more comprehensively across and within different subject areas to facilitate clarity. Ironically, as the instructor-researcher's clarity seemed to increase, some conceptual distinctions became less clear—particularly the distinctions among kinds of judgments. It sometimes seemed impossible to isolate one kind of judgment since real problems involve so many kinds of judgments. Moreover, each kind of judgment is also an judgment of relative value and it appears that most judgments involve relationships (e.g., fit between interpretation and reality; fit between evidence and conclusion; internal fit; fit between action and moral principles, fit between values and higher values). The Bailin et al. (1993) conception delineated a range of judgments that was helpful for distinguishing, elaborating, and connecting judgments and resources. The apparent fuzziness
experienced was problematic from an instructional perspective. Yet the critical thinking conception’s inclusiveness is a substantial strength. Many concerns about the gender and cultural bias of critical thinking are addressed by the range of judgments and habits of mind which include concern for fair-minded consideration of relevant points of view and underlying reasons, implications, and also sensitivity to others during group deliberations (Alston, 1995; Bailin, 1995b; Gilbert, 1994; Miller, 1995; Norris, 1995; Orr, 1989; Phelan & Garrison, 1994; Wheary & Ennis, 1995).

Qualitative artifacts provided evidence that aligned with instruction and demonstrated students’ practical applications relevant to teaching elementary social studies, that was the instructional focus. However, inconsistencies across forms of evidence limited conclusions. Together, results from formal and qualitative evidence illustrate that students’ claims about critical thinking dispositions and their views of knowledge were not necessarily congruent with their behavior and that formal instruments did not necessarily reflect student capacity. Congruency among dispositions, views of knowledge, and argument proficiency likely indicated students who were more secure in these attributes.

Yet such congruency was limited. One could conclude from this limited congruency among evidence of dispositional strength, epistemological perspectives, and argument proficiencies that dispositions and perspectives are not
adequate for critical thinking proficiencies and that proficiencies do not necessarily imply compatible dispositions and an evaluative view of knowledge.

Brabeck (1980) reported that critical thinking was necessary to but insufficient for her female subjects to evidence advanced levels of the Reflective Judgment model. She concluded that critical thinking and epistemological positions represent different constructs. Brabeck's conclusion suggested that the technical competencies of critical thinking exhibited on critical thinking tests do not represent a comprehensive array of critical thinking attributes (Ennis & Norris, 1989). One who attains successful critical thinking test results might use these competencies in a self-serving weak sense of critical thinking (Paul, 1992). An evaluativist, on the other hand, could be disposed to sustainable judgments, even when it differs from existing beliefs, but may not yet have the critical thinking concepts and proficiencies to achieve or recognize sustainable judgments. An evaluative view of knowledge is a necessary attribute of a critical thinker who demonstrates a comprehensive conception of critical thinking. Instruments do not yet embrace a full range of critical thinking resources that would include the more advanced qualities of an evaluativist epistemological position.

The foundation against which critical thinking instruments should be judged is a comprehensive critical thinking conception that includes critical thinking dispositions or perhaps more appropriate, habits of mind that reflect an
evaluative view of knowledge, and argument proficiencies that clearly depict the range of arguments. The critical thinking framework for this study, even with its flaws, more closely represented a comprehensive critical thinking conception than did the formal instruments, accounting for some evidence incongruencies.

The relationship between epistemology and critical thinking is not clear in the research or in this study. Siegel (1992) suggests that "a theoretical understanding of the goodness of reasons, and of related issues concerning truth, fallibilism, rationality and the like...[represents the epistemology underlying] a coherent conceptualization of critical thinking" (p. 107). Bailin (1995) claims that critical thinking is essentially an epistemological enterprise. It is tempting to suggest that critical thinking involves bringing the epistemological resources of the critical thinker to bear on a problematic situation in pursuit of knowledge—judgments that can be sustained through a justificatory argument. But this definition does not make clear the distinctions between moral judgments and judgments of knowledge that are both part of the critical thinking enterprise. Nor does defining critical thinking as a means to knowledge distinguish critical thinking as an educational goal. To suggest that critical thinking is only a means to an end is to neglect its duality as a means and an end—reflected in the ideal of the educated citizen.
Conclusions

The challenge of teaching for critical thinking and an evaluative view of knowledge is full of complexities but progress has and continues to be made. The challenge of educating preservice teachers to teach for these ideals is even more complex, particularly given research that shows neither students or teacher educators have experienced or value related pedagogy (Goodlad et al., 1990; Paul et al., 1997). Additionally, social studies is comprised of multiple disciplines that lack cross-domain and internal coherence. Moreover, the interdependence of critical thinking concepts makes it difficult to isolate and work with one concept. Finally, if students do not see related ideals and pedagogy reinforced across disciplines, they are more likely to attribute them to one instructors’ idiosyncratic whims, and thus not worthy of serious attention.

In light of such challenges, the small but positive change that students realized in the present study were encouraging. The instructor found intellectual resources that were personally important included: increasing clarity about critical thinking concepts as they applied to elementary social studies; an Internet community of interested educators who responded to queries, shared resources, challenged assumptions, and justified positions; strong personal commitment; sensitivity to students’ responses; and the courage that it took to seek feedback from students, particularly when their level of concern peaked. In particular,
knowledge of potential resistance enhanced the instructors’ understanding of and sensitivity towards increased level of concern and made it easier to seek to understand reasons for it, and with the students, to work through it. Perceived resistance was recognized as possible manifestations of insecurity imposed by unfamiliar ideas and lack of clarity (Kurfiss, 1988). None the less, the discomfort was real. But so was the satisfaction earned in working through the discomfort with the students toward sustainable ideals.

A knowledgeable citizenry whose decisions have potential for profound impact on all of humanity require the understandings, dispositions, and proficiencies to think critically about what to behave or what to do. To this end, educators need to understand and make explicit both the problematic nature of knowledge and the resources for judging the adequacy of arguments for these beliefs. Teaching in a critical manner needs to be infused throughout the whole educational experience, not left to one course near the end of an undergraduate degree.

**Strengths of the Study**

This study is a response to scarce research about educating preservice teachers to teach for critical thinking in elementary social studies (Armento, 1991). Strengths of the study include context, description, triangulation, and some positive steps toward reliability. The study was conducted in a natural
classroom context (except for interviews) over a 13 week semester. The normality of the classroom contributes to the study's ecological validity. The passage of 13 weeks reduced potential for pretest influence. Instruction, the critical thinking conceptual framework utilized in the study, and responses to instruction were described in depth to facilitate comparison across situations. Description also facilitated insights into some challenges and promises for educating preservice elementary social studies teachers to teach for critical thinking and an evaluative view of knowledge. Triangulation of evidence across both formal and qualitative measures of critical thinking dispositions, views of knowledge, and argument proficiencies enhanced the richness of the results. Written arguments provided evidence in a medium apart from the interview and addressed interviewer influence. As for reliability, data were retained for recursive analysis and judgment to attain consistency. Reliability of interpretations and conclusions were strengthened by sharing one-complete analysis with Kuhn and responding to her suggestions. In addition, the substantial time lapse between the end of the course and study completion allowed multiple recursive analyses of the evidence and interaction with committee members that contributed to consistent, consolidated judgments.
Limitations of the Study

This study should be interpreted and judged in light of limitations in participants' representativeness, study design, some reliability constraints, potential for researcher bias and extraneous variables, a short intervention time line, and the instructor's understandings at the time of intervention. A small unrepresentative intact convenience sample comprised of 8 females from 3rd (N=3) and 5th year (N=5) education programs in a comprehensive university and a pretest-posttest design did not allow generalization of findings beyond the female sample and the context of study. Nor was the researcher-instructor likely representative of elementary social studies teacher educators. Strong commitment helped the instructor endure the considerable effort and time that was required to translate theory into a compelling, meaningful, and practical form, an unrealistic expectation for most educators with a similar knowledge base and competing professional commitments. Formal inter-rater reliability data were not obtained. The complexity of the protocol categories and criteria made training raters impractical in the context of this study. The dual researcher-instructor role introduced a potential for influence and bias. It was also possible that pretest contamination, participant maturation, and instrument decay could have impacted internal validity. The short time line was recognized as inadequate for intervention to effect substantial change, particularly given the complexity of the
critical thinking conceptual framework. Finally, the critical thinking framework foundational to this study necessarily reflected the instructor’s understanding at the time of intervention.

**Possibilities for Research**

Possibilities for research have emerged from this study. They involve conceptual research, instrumentation, curricula and instruction, and ethical considerations.

**Conceptual Research**

A conceptual analysis to clarify the relationship among critical thinking dispositions, an evaluative view of knowledge, and the other intellectual resources that contribute to critical thinking proficiency would have utility for curricula, instruction, and assessment. Conceptual research is also needed to facilitate integration of critical thinking, an evaluative view of knowledge, and elementary social studies curriculum, guided by criteria of simplicity, practicality, flexibility, integrity, and potential effectiveness.

**Instrumentation Research**

Instrumentation for assessing critical thinking dispositions, views of knowledge, and critical thinking proficiency relevant to elementary social studies need to be developed that are congruent with a comprehensive conception of critical thinking and an evaluative view of knowledge. These instruments need to
be practical to administer and to interpret, aligned with forms of assessment relevant to elementary social studies education, and sufficiently elaborated to detect changes.

**Curricula and Instruction: Research Designs and Research Questions**

Questions and hypotheses that emerge from this study are compatible with varied research designs—exploratory pretest-posttest case studies, experimental-control group studies, and research and development. Research is needed to explore promising critical thinking conceptions and compatible instruction toward critical thinking and an evaluative view of knowledge in elementary social studies. For example, which critical thinking concepts are most significant and what minimal number of the concepts identified can be integrated into teacher preparation courses with good prospects of successful implementation of teaching for critical thinking in elementary social studies classrooms? What is an effective balance of teacher guidance and student problem solving?

If we agree with the ideals of evaluativism, and if students' prior education has not aligned with teaching toward an evaluative view of knowledge, then exploratory research into the impact of a preliminary course for preservice elementary social studies teachers is recommended. A preliminary course would explicitly integrate social studies ideas, epistemology and critical thinking with opportunities for students to reflect metacognitively on their own intellectual
products and those of others. This course would provide the epistemological justification and background knowledge for a curriculum and instruction course that could then focus on planning for related instruction.

Experimental control-group studies are compatible for testing promising teaching ideas toward critical thinking and an evaluative view of knowledge in elementary social studies. For example, does an interactive constructivist approach that engages preservice teachers in critical inquiry into and justification of the attributes and application of critical thinking concepts impact their critical thinking dispositions and proficiencies? Does emphasis on the epistemological connections between knowledge, critical thinking, and instruction impact students’ responses to teaching for critical thinking in elementary social studies? Does seeking student feedback to monitor students’ level of concern and guide instruction influence students’ responses to instruction? Is students’ clarity of teaching for critical thinking in elementary social studies facilitated more through (1) experiences with numerous examples that represent a limited range of disconnected judgments, (2) fewer examples that represent a broader range of judgments associated with one central challenge, or (3) beginning with a range of associated judgments followed by multiple disconnected judgments that provide experience discriminating resources? Do such experiences in preservice instruction impact whether inservice instruction focuses on fragmented bits of
information or addresses important connections among ideas? Does the duration of relevant preservice education impact successful inservice implementation teaching for critical thinking and an evaluative view of knowledge in elementary social studies? Is there evidence of delayed internalization? Does cross-domain support make a difference to students’ responses to teaching for critical thinking and an evaluative view of knowledge? Follow-up longitudinal research is needed to determine the impact of such instructional variables on preservice teachers’ inservice practice.

Research-based development is required to integrate in a conspicuous way teaching toward critical thinking and an evaluative view of knowledge in elementary social studies curriculum, instructional materials, and forms of assessment. Possibilities include: (1) an elementary social studies curriculum that integrates and emphasizes student judgments and the resources that students need for success; (2) curricula that embed a range of evidence or reasons relevant to sample judgments to demonstrate the problematic nature of knowledge and provide the materials that teachers require to engage students in judging relative merit; and (3) audio-visual demonstrations of teaching for critical thinking and an evaluative view of knowledge in elementary social studies to support teacher education and inservice practice.
Ethical Research

Kurfiss (1988) claim that teachers have a responsibility to move students toward an evaluative view of knowledge raises questions that are significant to the educational enterprise: What are teachers’ responsibilities to teaching for an evaluative view of knowledge? Does Siegel’s (1980) assertion that teachers have a moral mandate to teach for an evaluative view of knowledge justify expectations that preservice teachers demonstrate a disposition toward an evaluative view of knowledge? Or is any epistemological perspective acceptable for educators in a public school?

Epilogue

"What would life be if we had no courage to attempt anything?"

Vincent Van Gough (in Murphy, 1981, p. 53)
References


King, H. B. (1939, Oct. 11). Letter to H. B. Houlder. Available at B. C. Archives & Record Service, Victoria, B. C., GR 452, Box 1, File 2, #676.


Appendix A

Sample Letter of Informed Consent

Date: __________________________

To Whom it Concerns:

I agree to contribute to a program evaluation for Ed. E 446 by participating in four separate interviews of about 30 to 60 minutes each, two near the beginning of the course and two at the end of the term. These will take place out of regular class time and will be audio-taped. After completing the fourth interview, I will receive payment of $40 for my input.

I give permission for copies to be made of my responses to class assignments which may be included with inventory responses as part of the data being gathered for the evaluation. I have been assured that my anonymity will be protected by the use of code numbers on related documentation, storing data in locked facilities, and burning written data and scrambling audio-tapes electronically on project completion. To ensure my participation will not affect my course standing, analysis of data from the inventories and interviews will be left for after final course grades have been assigned. I have also been assured that if I choose not to contribute to this program evaluation, this will not affect my course grades. In any reports of these data, individuals will be unidentifiable, referenced only by code numbers or group values.

Signed: _______________________________________________________________

Instructor/Interviewer: __________________________________________________
Appendix B

Phases of Intervention: An Elaboration

Preparation Phase

The preparation phase comprised the first third of the course, from January 8 through to January 29 for a total of 12 hours. This phase of intervention immersed students in critical thinking and epistemology, while building on and extending students' current beliefs about an evaluative view of knowledge and a conception of critical thinking it implies. Judgments of meaning and judgments of value were emphasized. Assignments during the preparatory phase emphasized journal responses and course readings. Accountability was limited to journal completion, in-class participation, and attendance, comprising 10% of the course grade. Feedback to students was informal, through the instructor's responses to journals and through in-class comments from peers and the instructor.

Lesson 1 on January 8th involved an overview of the course and assignments; administration of the CCTDI; completion of biographical questionnaires; inviting students to participate in an evaluation of the course for the purpose of course improvement; considering alternative class schedules and office hours; introducing the novel that would be a central application; and arranging for pretest interviews. The approach was teacher directed, with opportunity for student questions and discussion. A draft curriculum excerpt
about the educated citizen was the assigned reading. Students were asked to list attributes of the educated citizen in their journal to facilitate future discussion.

Lesson 2 on January 10th focused on the educated citizen, utilizing an inductive approach with pair and whole class interaction. Students developed a profile of the ideal educated citizen, a person able to satisfy the responsibilities facing citizens where decisions have potential for serious global consequences; shared and justified attributes, and discussed how the profile might be strengthened by additions or deletions. Partners summarized the profile, shared and evaluated summaries, and identified summarization criteria from qualities captured by the whole class. Students were assigned 3 brief prepared social studies lesson plans to consider in light of the goal of the educated citizen and in relationship to "Critical Thinking and Social Studies" (Paul & Binker, 1992). Journal assignments specified that students record their own view of the essential characteristics of the ideal educated citizen and what the profile would imply for them as social studies teachers.

Lesson 3 on January 15th focused on the implications of the ideal educated citizen for teaching elementary socials studies. Small group and whole class interaction facilitated extension and application of understandings. The class reflected on why the attributes of the ideal educated citizen should be of concern to them, the adequacy of their criteria for summarizations, and the utility of
summarizations for social studies instruction. Partners compiled profiles of the kind of citizen facilitated by each of the 3 lesson plans. Implicit in each plan was an example, non-example, or borderline example of pedagogy for the ideal educated citizen and either an absolutist, multiplist, or evaluativist view of knowledge. Students shared, justified, and related profiles to the ideal educated citizen, and used them to generate criteria for teaching for this ideal. The lesson concluded with an overview and discussion of the purpose of strategies modeled so far in class. Journal assignments included reflections about the ideal educated citizen, what the ideal implied for instruction, the personal impact these ideas might have for teaching elementary social studies, and the strategies used in class.

Lesson 4 on January 17th focused on extension of criteria for teaching for the educated citizen and on a strategy to facilitate interpretation of meaning. The lesson involved whole class discussion, teacher modeling and student practice, as well as direct instruction. Students evaluated “Critical Thinking in Social Studies” (Paul & Binker, 1992) for its utility for prospective social studies teachers and for its alignment with their experiences as students or prospective teachers. Partners collaborated, shared, and justified ideas from the reading, adding to criteria for teaching social studies for the ideal educated citizen. Then the instructor modeled the ReQuest Strategy (Brownlie, Close, & Wingren, 1989), identifying on-the-
line, between-the-line, and beyond-the-line questions and students supported their answers with evidence from the text. Students then collaborated to develop and identify a similar range of questions. An inductive strategy was employed to teach for greater clarity. Students examined a selection of examples for each kind of question on an overhead transparency and identified distinguishing characteristics. Benefits, applications, and ideas relevant to the technique were discussed. Assignments included reading and summarizing “Toward Critical Thinking as an Ethic” (Sears & Parsons, 1991) and rationalizing responses to “Critical Thinking in Social Studies” and the ReQuest Strategy.

Lesson 5 on January 22nd was the first 3-hour session. It focused on a critical reading strategy to interpret the author’s intent, critical thinking as an ethic, and explication of views of knowledge and their implications for education, society, and teaching elementary social studies. Instruction included small group and whole class interaction, reciprocal modeling and practice, application of understandings, and direct instruction. The instructor modeled Paul’s (1993) reciprocal critical reading strategy, identifying which task of questioning, elaborating, and rephrasing was being demonstrated. Partners took turns applying the strategy to new text, after which the instructor modeled the same passage. Whole class reflection addressed the value of the strategy, its potential application, the benefits of cooperation, and the importance of understanding intent prior to
judging claims. Then partners collaborated on summaries of “Critical Thinking as an Ethic” and the class judged the adequacy of summaries against class criteria. The authors’ ideas were related to students’ prior understandings of critical thinking. A time-on-task strategy experienced and evaluated by students involved recording of perspectives of critical thinking in preparation for later discussion.

The instructor now explicitly described three views of knowledge, asked partners to match three familiar lesson plans to a compatible view of knowledge and to judge and justify which view of knowledge aligned with their profile of the ideal educated citizen. Students were alerted to research that identified implications of disparate views of knowledge for teacher-student relations. The instructor provided support for Kurfiss’ (1988) position that teachers have a responsibility to teach for an evaluative view of knowledge by relating them to claims of what it is to be educated and the nature of knowledge (Hamm, 1989; Scheffler, 1965; Paul, 1992) and to the potential that teachers have to impact the intellectual resources citizens acquire and thereby future educational decisions.

Two readings representing opposing world views about critical thinking (Marzano, 1993/1994; Simonds, 1993/1994) were assigned. Fair-minded reading was encouraged to allow students to reach beyond emotionally-laden language to comprehend reasons for such discrepant world views before deciding what they believed about the issue. The journal assignment specified that students (1)
identify which view of knowledge two earlier readings represented (Sears & Parson, 1991; Paul & Binker, 1992); (2) react to Sears and Parson's claims that teachers' lack of ethic for critical thinking accounts for the paucity of critical thinking in schools and that universities impart understanding of critical thinking to teacher education students; (3) respond to Kurfiss' (1988) claim that absolutism and multiplism were core misconceptions and teachers had a responsibility to foster an evaluative view of knowledge; (4) identify their own view of knowledge and if possible, justify why they hold this view; and (5) describe what impact their perspective might have on how they teach social studies.

Lesson 6 on January 29th provided transition between the preparation and application phases, focusing on a context for critical thinking, a defensible view of knowledge and critical thinking for public schools, and an introduction to critical thinking—a definition, a framework, and critical challenge criteria. Instruction involved small group and whole class discussion, building on and extending students' existing knowledge, and direct instruction. Students shared interpretations of readings, the reasons underlying opposing positions, and the view of knowledge each represented; speculated about prospects for resolution; and discussed which position was the most defensible in a public school system. Justifications were recorded, providing a list of moral principles that supported an evaluative view of knowledge.
Critical thinking was introduced by having partners infer and share ideas of how critical thinking relates to each view of knowledge. The instructor provided examples and non-examples of critical thinking, illustrated curriculum emphasis on critical thinking and problem solving—vehicles for realizing curriculum goals (British Columbia Ministry of Education, 1983), and drew on personal experiences to demonstrate the limitations of steps for problem solving and for decision making that neglect specifics to help teachers and students judge the adequacy of related tasks or the final judgment. The instructor illustrated dimensions of the critical thinking framework and then asked partners to develop criteria for a critical challenge. Ideas were shared with the class, recorded by the instructor, and linked to criteria displayed on an overhead transparency. The instructor noted explicitly that the strategy being used built on what students knew, showed how their ideas related to other ideas, and facilitated extending ideas. Relevant quotations and examples were provided to illustrate intent of each criterion. Partners applied criteria to examples and non-examples of critical challenges, evaluating and justifying decisions about the adequacy of categorization. Some responses were shared with the class and students were asked to study the remaining samples in preparation for next class. Three readings were assigned: "Enhancing Social Studies Through Literacy Strategies" (Irvin, Lunstrum, Lynch-Brown, & Shepard, 1995); an excerpt from Hare (1993) that
examined critical thinking concepts of relativism, bias, and open-mindedness, as well as issues related to political correctness and implications for rewriting history; and *White Jade Tiger* (Lawson, 1993). Students were asked to generate sample challenges from the novel as a journal assignment.

**Application Phase**

The application phase of the course spanned seven weeks from February 5 through March 18 for a total of 21 hours. The instructional focus of these weekly 3-hour classes evolved from an immersion approach to an infusion emphasis where distinctions among dimensions of critical thinking were made explicit. The main critical challenge and associated judgments allowed students to experience collaboratively a potential unit of study and the range of judgments it afforded. Assignment emphases shifted from background reading and journal responses to a gradual elaboration of the critical thinking instructional framework in the context of producing instructional materials for critical judgments relevant to the *Social Studies Curriculum Guide, Grade One—Grade Seven* (British Columbia Ministry of Education, 1983) for evaluation (see table 3).

Lesson 7 on February 5th focused on epistemological concepts, critical challenges, the organization of the Elementary Social Studies Curriculum, and the nature of concepts. Instruction involved partner and whole class interaction to extend and apply ideas, inter-group liaison, and direct instruction. Hare's (1993)
account of open-mindedness, bias, relativism, and issues of political correctness and re-writing history was discussed as it related to habits of mind, views of knowledge, and teaching social studies. The Irvin et al., (1995) reading was also discussed, the instructor querying students about cautions in using novels to teach history. Students evaluated each reading for their benefits for pre-service social studies educators. The instructor shared a few novels that could be used as a springboard for or supplement to teaching history.

Small groups sorted examples and non-examples of each critical challenge criteria and interacted with another group to challenge and justify categories. Whole class discussion considered disputed or unresolved responses, evaluation of the strategy, and evaluation of the critical challenge criteria. Groups prioritized criteria and discussed how criteria might be strengthened by additions and deletions. The instructor illustrated the relationship among dimensions of the curriculum (e.g., goals, understandings, and issues that represented core curriculum ideas, and the concepts and facts that were necessary to understanding and judge the adequacy of the big ideas) to address lack of clarity about core curriculum. Non-examples of challenges reflecting core curriculum ideas were discussed in terms of critical challenge criteria and then transformed and justified as examples that captured the social essence of social studies. The instructor defined concepts directly, what they are and where they come from, offered
examples as well as illustrations of how concepts may develop from immature over-generalizations to more mature comprehensive mental images, and related the idea to concepts developed so far in the course and judgments of meaning.

Small groups developed a critical challenge and displayed challenges for peer evaluation of positive attributes. The class discussed what criteria each challenge best exemplified and evaluated lesson strategies. The journal assignment specified that students develop a critical challenge and identify a concept that aligned with the novel and the curriculum that they might use for a lesson plan assignments, justify how the challenge satisfies criteria, and respond evaluatively to the strategies used in class.

Lesson 8 on February 12th introduced building a unit plan around a critical challenge; addressed the integral relationship of social studies, curriculum and critical thinking; engaged students in a lesson for judging the meaning of racial discrimination; and modeled an elaboration of the critical thinking framework for judgments about meaning. Approaches included direct instruction, whole class discussion, and partner deliberation with whole class sharing.

Review discussion addressed students' expressed interest in distinctions between fair-mindedness and open-mindedness, the implications of relativism and political correctness for rewriting history and teaching social studies, and addressed unit plan organization that trivializes cultures and limits inter-cultural
understanding. The integral relationships among social studies' declarative and procedural knowledge, the curriculum, and critical thinking were illustrated, the instructor suggesting that a well-developed critical challenge might encompass broad curriculum goals, grade understandings and inquiries, integral concepts, supporting facts, and skills.

The curriculum-related general claim for the proposed unit plan was presented: "Early Chinese workers/migrants who came to British Columbia in the 1880s were victims of racial discrimination." Students discussed whether what one believed about the claim warranted critical inquiry, shared ideas for making the claim provocative and meaningful for pupils (visual imagery, role play, first hand experiences, pertinent stories), and assigned a value judgment to racial discrimination. The meaning of discrimination was explored via examples of acceptable and unacceptable discrimination, non-examples of discrimination, and borderline examples, students generating and justifying attributes distinguishing each. Partners created and shared definitions of discrimination from which the class constructed a common definition. Students predicted what racial discrimination means; considered predictions in light of examples, non-examples, and borderline examples; appended the class definition to reflect discrimination based on race; and revisited their positions on the value of racial discrimination. Small groups arranged strips depicting ideas for evaluating concept understanding
in order of complexity, justifying to the class how they reflected increasing depth of understanding. Alignment of forms of assessment with teaching and learning activities was discussed.

The instructor elaborated the critical thinking framework as it applied to judgments about meaning. This closure was rushed and criteria for assignment one had not yet been discussed. Expectations were adjusted. Students were asked to prepare a draft lesson plan about judging the meaning of a concept embedded in a critical challenge, to be the focus of peer evaluation next class when assignment criteria and justification would be addressed.

Lesson 9 on February 19th had students infer from prepared sample lesson plans criteria and standards for judging the value of lesson plans and justifications of plans, and then apply criteria to peer evaluations. The approach was mainly inductive, with partner and whole group interaction. Partners evaluated a prepared example and non-example of a lesson plan for teaching concepts for depth of understanding. Reasons for judgments were recorded on the board, serving as examples and non-examples of criteria for lesson plans. Students then assigned each plan a letter grade. The instructor distinguished criteria and standards, discussed examples for each with students, and alerted students to the fact that the terms were used inconsistently. Students decided which of two prepared justifications for each plan was the best, supported judgments with reasons that
provided a list of attributes for a quality justification, graded each justification to indicate which standard (beginning, satisfactory, and powerful) was satisfied, and justified assigned standards with evidence. Queries from two students indicated concern about distinctions between intellectual tools and cognitive tasks. The instructor reviewed descriptions of each dimension, illustrated how judgments and pertinent elaborations addressed so far aligned with the framework, and reiterated the intent to address other specifics through application gradually in class.

Peer evaluation criteria were discussed and partners collaborated to identify where plans satisfied criteria, respond to requests for specific input, and offer ideas that might enhance alignment between plans and criteria. Debriefing discussion concerned peer evaluation—insights facilitated, negative emotional experiences, suggestions for reducing intimidation, and possible classroom applications; perspectives of having pupils develop criteria and standards and justify their responses; and addressed questions about the lesson plan and justification. Assignments included completion of the lesson plan and justification, optional journal responses about experiences of strategies used and concerns about clarity, and gathering examples of racial discrimination from the novel.

Lesson 10 on February 26th focused on judgments of meaning, empirical judgments about facts, and developing criteria for judging the credibility of
historical novels. Instructional approaches included small group and whole class
discussion and application of the Jigsaw Expert strategy. A visit to the British
Columbia Archives was planned and students’ justification criteria were reviewed.
The relationship among dimensions of critical thinking were illustrated on an
overhead transparency and students were provided a similar handout to evaluate
for utility and adequacy.

The critical challenge about racial discrimination was recorded on the
blackboard to focus discussion, progress with judgments related to the challenge
were briefly reviewed, and empirical judgments introduced. The instructor
demonstrated through role play the concept, evidence, showing that what seems to
be the case may not be so. Small groups gathered evidence from the novel of
racial discrimination, displayed examples of evidence of racial discrimination,
justified them to the class, and where applicable, responded to the instructor’s
challenge to consider other possible explanations for behavior. The instructor
noticed in the midst of this discussion that something was amiss. The problem
was identified, the instructor noting that judgments of meaning and judgments
about fact had become confused.

Focus shifted to judgments of empirical claims and how one could be sure
that each bit of evidence or claim about racial discrimination, that seemed to align
with the class’ definition of racial discrimination, actually occurred. Some
problems for judging past events were discussed and students were asked to
consider the credibility of novels as historical sources. Students’ ideas were
shared and supporting reasons were recorded. Novels were then examined for
additional clues about credibility, responses were discussed and those indicating
that students believed what they read were challenged, and the list of criteria for
judging credibility was extended.

One student’s suggestion that other sources be examined for alignment with
the novel provided entry into the Jigsaw Expert strategy of inquiry (Aronson,
Blaney, Stephan, Sikes, & Snapp, 1978). Students gathered in home groups and
then each joined an expert group to examine a prepared package of materials for
evidence of racial discrimination, judge to what degree related claims warranted
belief, and then consider together ideas for sharing conclusions and evaluating
understandings of peers in their home group. Third-year students expressed
concern about a lack of method courses and the instructor suggested a few ideas.
Closure involved discussion of the benefits and cautions for using the Jigsaw
Expert strategy. Assignments included preparation for instruction in home groups,
and two readings for opposing views about affirmative action (Bissoondath, 1994;
Strike, 1982).

Lesson 11 on March 4th concentrated on distinctions between meaning and
empirical judgments, completing the Jigsaw home group teaching, judging if
sources supported racial discrimination actions, judging source credibility, discussing assignment 2 criteria, and applying experiences to the critical thinking instructional framework. Small group activities, whole class interaction, application, and direct instruction were utilized. Jigsaw experts finalized plans and then taught home groups about evidence of racial discrimination and why it should or should not be believed. Home groups were provided critical thinking instructional frameworks to elaborate with instructional ideas shared by group members for teaching about empirical judgments. Some ideas were shared with the class and rationalized for compatibility with the ideal of the educated citizen. Sample justifications about empirical claims were discussed.

Meaning judgments and empirical judgments were distinguished in an attempt to address any confusion from the last class. The instructor used two examples of factual claims as a review focus on how one might examine the accuracy of claims and the credibility of sources in order to judge if belief was warranted. Assignment 2 criteria was discussed and students were advised to restrict the lesson plan to judging only one single event, whether the instance really happened and why it should or should not be believed. Together, students and instructor elaborated the critical thinking instructional framework for judgments about empirical factual claims. The instructor asked students for feedback about how lessons were progressing for them. One student suggested
too much was being attempted and that other students she had talked with concurred. Students were advised that concerns would be discussed the following class. The lesson plan about an empirical factual judgment was the only assignment.

Lesson 12 on March 11th focused on consolidation, evaluative feedback from students, and empirical judgments—distinctions between facts and generalizations, limitations of primary sources, testing for cause and effect, and developing criteria for strong arguments. Instructional approaches included direct instruction, small group and whole class interactions, application, and demonstrations. Consolidation involved a brief overview of course objectives as they related to the University course calendar and an explicit account of the early emphasis on a rationale for critical thinking and immersion in concepts and strategies aligned with critical thinking; the integration of critical thinking with curriculum and instruction; and plans for the remainder of the course. Written evaluative input from students was sought about the adequacy of background reading, the utility and complexity of the critical thinking instructional framework, and suggestions about a final assignment that would have utility for students and demonstrate students’ capacity to distinguish intellectual resources as they apply to a range of judgments.
Small groups then sorted and displayed bits of evidence from earlier investigations of racial discrimination, justifying why each was true, likely true (plausible), not sure, or not true. The instructor challenged some responses with counter explanations and requests for verification. A demonstration that facts are not a matter of opinion was conducted and discussed. Students were referred to the original challenge to judge the adequacy of evidence to support the racial discrimination claim. They considered whether evidence might be added or deleted to strengthen support for the claim, whether the claim adequately reflected the nature of evidence, and whether the claim should be adjusted to better reflect the evidence. The instructor incorrectly identified this as a judgment of logical relations rather than a judgment of inferential relations. Students decided the claim should be qualified to better reflect the range of evidence and alternative explanations. The adjusted generalization was related to the broader curriculum claim about the challenges early settlers faced and students suggested what additional inquiry was needed to decide what to believe about the claim.

The instructor demonstrated the limitations of primary sources. Class discussion addressed implications if attention is not focused on an event, if only part of the event is recorded or observed, and if reports are removed in time from the event. Students shared additional demonstration ideas. Small groups recorded reasons for racial discrimination, displayed them on the board, discussed the
instructor's manual manipulation of expressed causes, and considered the possible impact this would have on racial discrimination. Students were invited to identify a situation in which neither the proposed cause nor racial discrimination was evident. A list of criteria for strong arguments was initiated. Students who wished to begin assignment 3, a written argument about school violence, were offered tentative criteria that would be finalized next class. Readings were postponed.

Lesson 13 on March 18th focused on empirical judgments--curriculum generalizations and pedagogical options that foster the ideal educated citizen, developing criteria and standards for evaluating arguments, relating past and present; and on judgments about value including generation of moral principles to guide action. Instructional approaches included small and whole group interaction, application of the t-shirt strategy (Case & Daniels, 1995), modeling, and direct instruction. Students' evaluative feedback was shared. The unit plan would be adjusted in the number of detailed lesson plans required. Students reported that the critical thinking instructional framework had initially been confusing but that it was becoming clearer now. Additional background readings were provided students and relevant references were put on hold for students' use in the curriculum library. The instructor offered to meet with individuals at their convenience, providing advanced arrangements were made.
Claims of empirical generalizations were related to the Curriculum, the instructor emphasizing teachers' responsibility to share with pupils the evidence upon which the claims rest, preferably by teaching students how to evaluate the evidence so that they may better judge what to believe about the claims and decide whether the claims need adjusting to better reflect the evidence. Arguments were defined and relevant criteria were reviewed, discussed, and extended. Past and present were linked in a discussion of evidence of racial discrimination, including recent cases of citizens discriminated by newcomers of their own race. Students were asked to judge the acceptability of racial discrimination, provide reasons for responses, and then consider what might be done to alleviate such behavior. A topic removed from racial discrimination was selected for classroom application of judgments about action, a response to students' concerns that the good ideas were being used in class, depleting ideas for students using the novel for their unit plan. The problem was identified, alternative solutions were brainstormed, the least acceptable solutions were eliminated, and pros and cons of each alternative were examined from each perspective. Responses were recorded in a t-shirt format, the proposed solution across the top of the page, pros listed below on one half of the page and cons listed on the other half of the page. Students evaluated alternatives, chose the option with the least hardships and most benefits for all, and justified their decision, providing criteria for judgments about action. Forms of
assessment were illustrated and the benefits of the t-shirt strategy were discussed, particularly its systematic attention to alternatives as they relate to relevant points of view. Criteria for the related lesson plan assignment were discussed and students were to review Bissoondath and Strike's positions on affirmative action for next class.

**Consolidation Phase**

Consolidation of course understandings was the emphasis of the last two 3-hour sessions and unit plan criteria was established. Lesson 14 on March 25th involved students in small group interaction, whole class sharing, application of conflict resolution strategy (Johnson & Johnson, 1995), and direct instruction. Students worked in small groups, sorted kinds of judgments and relevant criteria and attempted to justify decisions. Then students were asked to apply their understandings of course priorities to predicting unit plan criteria. Predictions were recorded on the blackboard and related to actual criteria. Two unit plan overviews had been provided to students in an earlier class, one an example that fit criteria and one that did not. Each overview was examined separately in conjunction with two accompanying justifications, one an adequate sample justification and one that was inadequate. Partners evaluated the relative merit of justifications for each plan, justifying their decisions to the class. Discussion
followed, the instructor responding to students' questions about unit plan assignments.

The strategy of cooperative conflict resolution (Johnson & Johnson, 1995) was introduced and applied to the issue of affirmative action. Groups of 4 students divided into partners, each pair constructing an argument to support the opposing perspective. Arguments were presented by each pair while opposing partners listened and recorded key ideas. Then roles were reversed, each pair presenting a supportive argument for the other point of view. Time did not allow for the final stage where advocacy was to be dropped and students were to build the strongest argument possible, using the best ideas from each perspective. Instead, this stage was discussed and students reflected on benefits of the strategy for fair-minded consideration of views, the impact that examination of both positions had made to their own perspective, and the relative merit of the cooperative conflict strategy and the t-shirt strategy for decisions about what to do. The instructor noted that both strategies required attention to criteria of acceptability if judgments were to provide effective and lasting solutions. Unit plans were assigned.

Lesson 15 on April 1st included administration of the CCTDI, consolidation of course ideas, and two separate course evaluations. Instructional approaches including small group and whole class interaction, application of concept maps,
and direct instruction. After the CCTDI was administered, the instructor used sample concept maps to illustrate attributes of hierarchical order, relationship among parts, and grammar (Novak & Gowin, 1984). Small groups generated a concept map depicting and elaborating dimensions of critical thinking, displayed finished products, and evaluated maps, recording how they satisfied concept map rules. Students justified meaning relations, judged conceptual relationships, considered what dimensions might be added or deleted to strengthen maps without sacrificing the essence of critical thinking, and reflected on uses for concept maps in the classroom. The benefits of integrating concept maps with interviews and multiple forms of assessment were discussed. Key course ideas were summarized by the instructor who encouraged students to have patience with complexity, to strive for understanding, and to have the courage to take risks where the potential benefits to their pupils warrant it. The course ended with students’ contributions to course evaluations, interviews were completed by mid-week, and assignments were due April 8th.
Text from concept maps follows, with students’ elaborations of dimensions indicated in brackets. Evaluative comments denoted students’ responses which they recorded on their peers’ concept maps.

- A, G, and one other student: “Critical challenge involves critical thinking utilizing cognitive tasks (self-regulate, thinking strategies, deliberate, evaluate) and intellectual tools (habits of mind [fair mindedness, open mindedness], criteria & standards, CT vocabulary [justify, evaluate]) leading to a decision which satisfies the challenge.” An evaluative comment on this map by a student identified the “good connection (loop)” between the decision which was linked back to the challenge.

- C, E, and K: “Critical thinking requires and involves a critical challenge [that] requires cognitive tasks (involves self-and peer regulation, involves deliberating, involves justifying and evaluating, involves deciding on a conclusion) and intellectual tools (includes background knowledge, includes critical thinking vocabulary, includes habits of mind, includes criteria and standards, includes thinking strategies) [that] work together to create [a] decision.” C, E, and K also linked both cognitive tasks and intellectual tools back to critical thinking, stating that they were “necessary components” of critical thinking. A student’s evaluative comment on this map observed, “I like how this shows a 2-way relationship—shows more depth of understanding of concept.”

- B, H, and another student used one student’s class notes to elaborate tasks and tools: “Critical thinking involves the development of (a) critical challenge(s) which are clarified by cognitive tasks (consisting of self-regulation, planning, monitoring, evaluating, adjusting, deliberating, concluding and justifying) [that] fit with intellectual tools (consisting of background knowledge, critical thinking vocabulary, criteria fit standards [that include empirical claims, value concerns, and logical and meaning relations], habits of mind [that] include open-mindedness, fair-mindedness, respect for truth and reason, inquiring
spirit, independent mindedness, intellectual work ethic, respect for different levels of knowledge) which can help people develop more informed decisions made about critical challenges." Students' evaluative comments for this map included "good connection—fit" for the link leading back from decision(s) to the critical challenge; "vertical & horizontal hierarchy" with an arrow pointing to elaborations for habits of mind; "elaborating" pointing to the kinds of judgments relevant to criteria and standards, and "rules of grammar followed throughout."
Appendix D

Justifications for Unit Plans

Justifications (A, E, G, H, J, K) exhibit a range of quality. See chapter 4 for B and C’s justifications.

A: "This unit is based on the active participation of students in making decisions, researching, sorting, and understanding. Students are developing an understanding of the intellectual tools needed to make judgments and decisions. It is important that students learn how to formulate an argument, how to evaluate options and their effect on significant stakeholders and how to justify their conclusions. Lessons where students are listening to lectures, accepting research done by other people, regurgitating information or merely learning about cultural celebrations rather than getting a feel for the 'big' ideas are unlikely to stimulate thought or provide meaning to their life as a whole."

E: "The purpose of deciding about meaning is to engage students in questioning and expanding their definitions, along with creating new definitions and ensuring understanding of concepts. Put simply, deciding on meaning is a building block for making bigger decisions. If students do not fully understand the subject/concept of study, how can they make informed, rational, justified, decisions on what to do about these decisions? Finally, by creating criteria and defining meaning, the students are able to create a link between commonly accepted definitions of concepts and their views; a so-called 'fit' between their worlds and the world outside the classroom. The decision about validity and accuracy builds on the decision about meaning. This decision is intended to develop in the students a need to question what they see, hear, read, etc....Inherent in this thought is the fact that students will understand that all sources, whether valid, invalid, accurate, or inaccurate, are presented from the point of view of one person and that each person will see a problem or event in a different way. In this way, students learn how to search for the best and most accurate version of the truth as we know it from which they can base further decisions. The final arguments, deciding what to do, develops students' ability to evaluate and justify their decisions based on results which are a compromise between being the best for everyone involved as possible while creating the least harm possible...; to consider all alternatives fairly, ensure accuracy and strength of arguments, and justify thinking with rational reasoning. The inherent purpose behind all of the above decisions...is to create citizens of the world who will
attempt to deliberate, justify, and evaluate their decisions for the best possible good of all involved.... Throughout this whole procedure, the students are not simply passive observers, but integral portions of the decision-making...."

G: "This unit helps students to develop and practice traits of the Educated Citizen..., develop an understanding of the concept..., justify how the examples fit with the attributes or definition of an immigrant; provides students with hand-on experience on how to test a claim; the teacher and class develop a set of criteria of a good justification which was used to test the claim in the lesson but which may also be used to evaluate sources outside the classroom and in society. In the end, students will learn to question how what they read—Is it true? Is the source credible? The third lesson questions students’ understanding by having them question further—do I have enough evidence to believe the claim? How can I change the claim to reflect the evidence I have collected? This lesson will also help the student to develop two Educated citizen traits. One, to look at all the evidence before deciding to support/believe a claim; and two, to be open to new evidence before deciding to change his/her view if evidence warrants....open-mindedness....They will understand and give a fair or equal consideration to alternative views in a problematic situation. This lesson gives students power as to what should be done but they must consider the criteria—Which option is most acceptable and does the least harm to the stakeholders in the situation? In every lesson each student activity has students justify their decisions..., support their conclusion with relative merit (or evidence that it is true)....The intellectual tools in each lesson provides students the opportunity to make their own decisions, and the evaluate meaning and claims. These tools in turn, will help them in society to become an ideal Educated Citizen...—an informed decision maker, socially responsible, fair-minded, open-minded, global awareness, and pursue the truth."

H: "I think that this unit plan has a good chance to build an ideal educated citizen because it uses a variety of activities that have students extending their thinking and building their habits of mind. They engage in a lot of activities that have them not only evaluating the things they read but also their own ideas and those around them for flaws as well as keeping an open mind about the ideas around them. The students also engage in a variety of role play type activities that have them practice their empathy and understanding of the points of view of both 'sides' of the story. This unit plan has students for the most part exploring and developing their own ideas on different subjects and analyzing those ideas rather than having it spoon fed to them. These are skills and qualities that will transfer over well because they have been so thoroughly practiced."
J: “The lessons start with what the students know and build from that.... The students learn and practise evaluating... accuracy and credibility rather than just regurgitating information. This allows them to think for themselves and decide what to believe about a claim on their own.... Unlike other lessons, the meaning lesson allowed the children to discover what the meaning was rather than the teacher having them memorize what the definition of ‘immigrant’ was.... The role-play lesson focuses on many habits of mind as the students take on different perspectives they become more open-minded which will lead to fair-minded behavior.... The fourth lesson allows the students to... view all the perspectives involved in the situation and evaluate how the situation will affect them. They are able to weigh each option and evaluate the effects it would have on the situation.... This portfolio format [for evaluation] allows the student(s) to continue their self-regulation and justification of the unit.... Throughout the unit the students must constantly justify their evaluations and decisions. This produces strong decisions, critical thinking, and independence. By justifying their work they have to fully understand their decision and what it involves. This critical thinking is core to this unit, as the students are constantly solving problems, making decisions, and resolving issues, and are able to justify each action they take, unlike others [in] which students learn information, do not know why they learned it or what it really means.... They will have a good grasp on quality independent thinking skills which together with their new knowledge about immigration contributes to forming ideal educated citizens.”

K: “I would hope that the students realize that not all things can be seen as black or white and that there is usually a large gray area [and that they take]... a closer look at and questioning statements which they might judge to have fallacies.... I feel that when students have the opportunity to discuss ideas and make decisions collaboratively they are gaining more knowledge as well as an essential lifelong skill.... I have also chosen to use the 5 W’s as an underlying concept/strategy to teaching this unit. I think it has benefits in questioning, decision making, building understanding, etc.... I feel that this Social Studies unit promotes cooperative learning [and].... the “Ideal Educated Citizen” in a number of ways: cooperative work ethic; socially responsible to peers; respectful towards others and self; appreciate self and others in a variety of situations; incorporates responsibility and accountability in the decision making process; motivational activities.... critical thinker—questions meaning, claims, values, etc.; looks at different perspectives...; uses sound logic and reasoning in order to come to conclusions; can justify conclusions with valid arguments; Open/Fair Minded...; informed decision maker; uses previous knowledge and various thinking strategies to further own understanding and make decisions.”