

2009

The thought processes of administrators as they review and make evaluative judgment of a second-year teacher portfolio: A qualitative study

Terri Anne Lasswell
University of Northern Iowa

Let us know how access to this document benefits you

Copyright ©2009 Terri Anne Lasswell

Follow this and additional works at: <https://scholarworks.uni.edu/etd>



Part of the [Curriculum and Instruction Commons](#)

Recommended Citation

Lasswell, Terri Anne, "The thought processes of administrators as they review and make evaluative judgment of a second-year teacher portfolio: A qualitative study" (2009). *Dissertations and Theses @ UNI*. 322.

<https://scholarworks.uni.edu/etd/322>

This Open Access Dissertation is brought to you for free and open access by the Student Work at UNI ScholarWorks. It has been accepted for inclusion in Dissertations and Theses @ UNI by an authorized administrator of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.

Copyright by
TERRI ANNE LASSWELL
2009
All Rights Reserved

THE THOUGHT PROCESSES OF ADMINISTRATORS AS THEY REVIEW AND
MAKE EVALUATIVE JUDGMENT OF A SECOND-YEAR
TEACHER PORFOLIO: A QUALITATIVE STUDY

An Abstract of a Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

Dr. John Henning, Co-Chair

Dr. Mary Herring, Co-Chair

Dr. Sue A. Joseph
Interim Dean of the Graduate College

Terri Anne Lasswell

University of Northern Iowa

July 2009

ABSTRACT

This qualitative study explored the thought processes of administrators as they reviewed and judged second-year teacher artifacts (a portfolio) relative to the Iowa teaching standards and criteria (ITS/criteria). In addition, data was collected pertaining to the tools principals used as they conducted portfolio evaluation and the amount of bearing the portfolio had on a licensure decision.

Data for the study was gathered via a think-aloud process in combination with guided interview questions. Nine principals participated in the study; three each from elementary, middle, and high school. The nine principals were also representative of rural, suburban, and urban geographic/demographic regions. The think-alouds and ensuing interviews were audio-recorded and then transcribed. The resulting verbal reports (comments) were analyzed and categorized using the constant comparative method. The comment counts were used to report the accumulated data and make comparisons between academic level and between geographic/demographic regions.

The verbal reporting data revealed that the thought processes of the principals were similar. Each review consisted of three distinct phases. Within in each phase, the principals attended to processing activities, judgment activities, and coaching activities. In addition, the principals identified two critical pieces of teacher evaluation as teacher reflection and principal's observation of teacher.

Findings also made clear the impact of the Iowa Evaluator Training Program (IEATP) on the consistency of evaluation. Principals across academic level and geographic/demographic region used a similar four-step rhythm as they judged artifacts.

In addition, a distinct consistency existed in the kinds of artifacts the principals identified as valid evidence of the ITS/criteria. Further, the leadership style of the principals was indicative of the formative nature of the portfolio.

Six distinct tools that principals used while they evaluated were identified and described in the study. In addition, it was evident that, while value was placed on the portfolio, the principals put more emphasis on observation. Principals indicated that the portfolio review would account for roughly 30% of a licensure decision.

The findings from this study were relevant to consistency in evaluation across academic level and geographic/demographic region. The information may help inform continuing efforts relative to teacher evaluation across the state.

THE THOUGHT PROCESSES OF ADMINISTRATORS AS THEY REVIEW AND
MAKE EVALUATIVE JUDGMENT OF A SECOND-YEAR
TEACHER PORFOLIO: A QUALITATIVE STUDY

A Dissertation

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

Approved:

Dr. John Henning, Co-Chair

Dr. Mary Herring, Co-Chair

Dr. Rebecca Burkhardt, Committee Member

Dr. Victoria Robinson, Committee Member

Dr. Rick Traw, Committee Member

Terri Anne Lasswell

University of Northern Iowa

July 2009

DEDICATION

This dissertation is dedicated to my partner Pam for enduring all the stops and starts and in memory of my parents, Ann and Gene, who taught me well.

ACKNOWLEDGMENTS

Completion of this dissertation was possible due to the willing participation of nine principals and three teachers from Iowa. Without their involvement, this research would have remained only an idea. I am grateful for their time, their commitment, and their enthusiasm.

I would like to thank my family, my friends, and my colleagues for believing in me. Their unyielding support, encouragement, and love were essential as I balanced my health, my work, and my writing.

I am extremely grateful to my co-chairs, Dr. John Henning and Dr. Mary Herring. Their patience and guidance elicited from me a passion and desire to question, to think deeply, to write well, and to feel my work had value. I deeply appreciated their high expectations, their mentorship, their understanding, and their advocacy. I will continually strive to parallel their professionalism, leadership, and compassion throughout my career. I offer my best to Dr. Henning as he moves to Ohio University. His future students and colleagues are indeed fortunate.

In addition, I would like to thank the members of my committee, Dr. Rebecca Burkhardt, Dr. Victoria Robinson, and Dr. Rick Traw for their time and commitment. Each of them provided me with a unique perspective of my work. Their willingness to listen, refine, and challenge was remarkable.

I also want to thank Joyce Peters, the copy editor for my dissertation. Joyce spent long hours proofing, refining tables, and listening throughout my writing process. Her work ethic and positive nature were inspiring. She provided me with consistent support

and confidence. As with others in my life, Joyce is a constant reminder of unconditional friendship.

Finally, I want to thank my partner Pam for her patience, understanding, love, and support throughout my doctoral program and my dissertation process. It has been a journey and I am fortunate to have traveled it with a faithful companion.

TABLE OF CONTENTS

	PAGE
LIST OF TABLES	x
LIST OF FIGURES	xiii
CHAPTER 1: INTRODUCTION	1
Setting Standards and Defining Teacher Quality	2
Implementation	3
The Evolution of Evaluation	4
Portfolio Use and Evaluation	7
Statement of the Problem	10
Purpose of the Study	10
Research Questions	11
Definition of Terms	11
Significance of Study	12
CHAPTER 2: REVIEW OF THE LITERATURE	14
Assessment Systems	14
The Role of the Principal	18
Portfolios as a Link	22
Evaluating Portfolios	24
Scoring Portfolios	27
Shortcomings	30
Summary	32
CHAPTER 3: METHODOLOGY	33
Research Design	34
Participants	37
Gaining Entry	39
Data Collection	44
Portfolios	45
Think-aloud/Interview Process	48
Think-aloud/Interview Pilot	50
Data Analysis	51
Reliability	60
Validity	61
The Role of the Researcher	63
Limitations of the Study	64
Summary	64

CHAPTER 4: RESULTS	66
Research Question 1	67
Overall Comment Counts per Academic Level and Geographic/ Demographic Region	68
Three “Phases” of Thinking	72
Phase I: Pre-Assessment Thinking per Academic Level	75
Overview	75
Elementary Principals’ Thinking in Phase I	77
Elementary Principals and Processing in Phase I	77
Elementary Principals and Judging Activities in Phase I	80
Elementary Principals and Coaching Activities in Phase I	81
Elementary Principals and Critical Pieces in Phase I	81
Summary of Elementary Principals’ Thinking in Phase I	82
Middle School Principals’ Thinking in Phase I	83
Middle School Principals and Processing in Phase I	83
Middle School Principals and Judging Activities in Phase I	87
Middle School Principals and Coaching Activities in Phase I	89
Middle School Principals and Critical Pieces in Phase I	91
Summary of Middle School Principals’ Thinking in Phase I	93
High School Principals’ Thinking in Phase I	93
High School Principals and Processing in Phase I	94
High School Principals and Judging Activities in Phase I	97
High School Principals and Coaching Activities in Phase I	99
High School Principals and Critical Pieces in Phase I	100
Summary of High School Principals’ Thinking in Phase I	100
Phase I: Pre-Assessment Thinking per Geographic/Demographic Region	101
Geographic/Demographic Overview	102
Rural Principals’ Thinking in Phase I	103
Suburban Principals’ Thinking in Phase I	104
Urban Principals’ Thinking in Phase I	106
Summary: Findings for Phase I Thinking	108
Phase II: Judgment Thinking per Academic Level	109
Overview	110
Four-step Rhythm	112
Elementary Principals’ Thinking in Phase II	114
Elementary Principals and Processing in Phase II	115
Elementary Principals and Judging Activities in Phase II	120
Elementary Principals and Coaching Activities in Phase II	141
Elementary Principals and Critical Pieces in Phase II	147
Summary of Elementary Principals’ Thinking in Phase II	151
Middle School Principals’ Thinking in Phase II	153
Middle School Principals and Processing in Phase II	154
Middle School Principals and Judging Activities in Phase II	160
Middle School Principals and Coaching Activities in Phase II	186

Middle School Principals and Critical Pieces in Phase II	198
Summary of Middle School Principals' Thinking in Phase II	201
High School Principals' Thinking in Phase II	204
High School Principals and Processing in Phase II.....	205
High School Principals and Judging Activities in Phase II.....	210
High School Principals and Coaching Activities in Phase II.....	232
High School Principals and Critical Pieces in Phase II.....	243
Summary of High School Principals' Thinking in Phase II.....	246
Phase II: Judgment Thinking per Geographic/Demographic Region.....	250
Geographic/Demographic Overview	250
Rural Principals' Thinking in Phase II	252
Suburban Principals' Thinking in Phase II.....	254
Urban Principals Thinking in Phase II.....	256
Summary: Findings for Phase II Thinking	259
Phase III: Reflection via Guided Interview Questions per Academic Level.....	263
Academic Level Overview	264
Elementary Principals' Thinking in Phase III	266
Elementary Principals and Processing in Phase III	267
Elementary Principals and Judging Activities in Phase III	269
Elementary Principals and Coaching Activities in Phase III.....	272
Elementary Principals and Critical Pieces in Phase III	274
Summary of Elementary Principals' Thinking in Phase III	275
Middle School Principals' Thinking in Phase III	275
Middle School Principals and Processing in Phase III.....	276
Middle School Principals and Judging Activities in Phase III.....	277
Middle School Principals and Coaching Activities in Phase III	279
Middle School Principals and Critical Pieces in Phase III.....	280
Summary of Middle School Principals' Thinking in Phase III	281
High School Principals' Thinking in Phase III.....	281
High School Principals and Processing in Phase III	282
High School Principals and Judging Activities in Phase III.....	283
High School Principals and Coaching Activities in Phase III.....	286
High School Principals and Critical Pieces in Phase III.....	289
Summary of High School Principals' Thinking in Phase III.....	290
Phase III: Reflection via Guided Interview per Geographic/Demographic Region	290
Geographic/Demographic Overview	291
Rural Principals' Thinking in Phase III	293
Suburban Principals' Thinking in Phase III.....	294
Urban Principals' Thinking in Phase III.....	295
Summary: Findings for Phase III Thinking	296
Summary: Research Question 1.....	298
Research Question 2	303
Summary: Research Question 2	307

Research Question 3	308
Academic Level Responses	309
Geographic/Demographic Region Responses	312
Summary: Research Question 3	312
CHAPTER 5: DISCUSSION AND IMPLICATIONS	314
Research Question 1: What do principals think during portfolio review?	315
Processing	315
Judging	319
Critical Pieces	328
Research Question 2: What tools do principals use?	330
Research Question 3: Bearing of portfolio on licensure decision	332
Implications for Further Research	333
REFERENCES	336
APPENDIX A: IOWA TEACHING STANDARDS AND MODEL CRITERIA	342
APPENDIX B: IOWA DEPARTMENT OF EDUCATION SAMPLE COMPREHENSIVE EVALUATION	347
APPENDIX C: UNIVERSITY OF NORTHERN IOWA HUMAN PARTICIPANTS REVIEW INFORMED CONSENT—PRINCIPAL	350
APPENDIX D: DEMOGRAPHIC INFORMATION	353
APPENDIX E: UNIVERSITY OF NORTHERN IOWA HUMAN PARTICIPANTS REVIEW INFORMED CONSENT—TEACHER	355
APPENDIX F: THINK ALOUD INSTRUCTIONS TO PARTICIPANTS	358
APPENDIX G: EXPERT PANEL RUBRIC	361
APPENDIX H: ELEMENTARY TEACHER PORTFOLIO CROSS-REFERENCE INFORMATION	363
APPENDIX I: HIGH SCHOOL TEACHER PORTFOLIO TABLE OF CONTENTS AND CROSS-REFERENCE INFORMATION	365
APPENDIX J: MIDDLE SCHOOL TEACHER PORTFOLIO ARTIFACT COVER PAGE	368
APPENDIX K: GUIDED INTERVIEW QUESTIONS	370
APPENDIX L: SAMPLE OF CODED TRANSCRIPT	372
APPENDIX M: KWIC WORKSHEET	374
APPENDIX N: INITIAL CODING INSTRUCTIONS	376
APPENDIX O: CATEGORY DEFINITIONS	378
APPENDIX P: REVISED CODING INSTRUCTIONS	382
APPENDIX Q: SAMPLE OF CATEGORY HOLDING PAGE	384
APPENDIX R: IVAN’S ITS/CRITERIA CROSS-REFERENCE TOOL	386
APPENDIX S: MIKE’S MODEL DESCRIPTORS SAMPLE	388

APPENDIX T: MIDDLE SCHOOL TEACHER’S ARTIFACT:
FOUR-QUADRANT VOCABULARY FOR ITS 4B.....390
APPENDIX U: GAVIN’S EVALUATION GUIDE: ITS 2.....392

LIST OF TABLES

TABLE	PAGE
1 Administrative Demographics by Academic Level.....	42
2 Administrative Demographics by Geographic Location	43
3 Emergent Categories.....	58
4 Summary of Phase I Comment Counts by Academic Level.....	76
5 Summary of Phase I Comment Counts made by Elementary Principals.....	77
6 Summary of Phase I Comment Counts made by Middle School Principals.....	84
7 Summary of Phase I Comment Counts by High School Principals.....	94
8 Summary of Phase I Comment Counts per Geographic/Demographic Region.....	103
9 Summary of Phase I Comment Counts made by Rural Principals	104
10 Summary of Phase I Comment Counts made by Suburban Principals.....	105
11 Summary of Phase I Comment Counts made by Urban Principals	107
12 Summary of Phase II Comment Counts by Academic Level	110
13 Summary of Phase II Comment Counts made by Elementary Principals	115
14 Judgment (J) Comment Count Breakdowns for Elementary Principals in Phase II.....	121
15 Judgment of Artifact (J/A) Comments per Artifact by Elementary Principals in Phase II	122
16 Classification of Judgment Comments per Artifact made by Elementary Principals in Phase II	124
17 Coaching (C) Comment Count Breakdowns for Elementary Principals in in Phase II	142

18	Suggested Alternative (C/SA) Comments per Artifact by Elementary Principals in Phase II	144
19	Critical Pieces (CP) Comment Count Breakdowns for Elementary Principals.....	148
20	Summary of Phase II Comment Counts made by Academic Level.....	154
21	Judgment (J) Comment Count Breakdowns for Middle School Principals In Phase II	161
22	Judgment of Artifact (J/A) Comments per Standard by Middle School Principals in Phase II	164
23	Classification of Artifact Judgment Comments per Standard made by Middle School Principals in Phase II.....	167
24	Principals' Opinion Comment Count Breakdowns for Middle School Principals in Phase II	178
25	Coaching (C) Comment Count Breakdowns for Middle School Principals in Phase II	187
26	Suggested Alternative (C/SA) Comments per Eight ITS* by Middle School Principals in Phase II	192
27	Critical Pieces (CP) Comment Count Breakdowns for Middle School Principals in Phase II	199
28	Summary of Phase II Comments made by High School Principals	205
29	Judgment (J) Comment Count Breakdowns for High School Principals in Phase II	211
30	Judgment of Artifact (J/A) Comments per Artifact by High School Principals in Phase II	213
31	Classification of Judgment Comments (J/A) per Artifact made by High School Principals in Phase II	215
32	Principals' Opinion Comment Count Breakdowns for High School Principals in Phase II	225

33	Coaching (C) Comment Count Breakdowns for High School Principals in Phase II	233
34	Suggested Alternative (C/SA) Comments per Artifact by High School Principals in Phase II	236
35	Critical Pieces (CP) Comment Count Breakdowns for High School Principals in Phase II	244
36	Summary of Phase II Comment Counts per Geographic/Demographic Region.....	251
37	Summary of Phase II Comment Counts made by Rural Principals	252
38	Summary of Phase II Comment Counts made by Suburban Principals	254
39	Summary of Phase II Comments made by Urban Principals.....	257
40	Summary of Phase III Comment Counts by Academic Level.....	265
41	Summary of Phase III Comment Counts made by Elementary Principals	267
42	Summary of Phase III Comment Counts made by Middle School Principals.....	276
43	Summary of Phase III Comments made by High School Principals	282
44	Summary of Phase III Comment Counts per Geographic/Demographic Region.....	292
45	Summary of Phase III Comment Counts made by Rural Principals.....	293
46	Summary of Phase III Comment Counts made by Suburban Principals	294
47	Summary of Phase III Comment Counts made by Urban Principals.....	295

LIST OF FIGURES

FIGURE		PAGE
1	Total Comment Counts by Academic Level Per Category.....	69
2	Total Comment Counts by Geographic Demographic Region per Category	71
3	The Thinking Phases of Principals	72

CHAPTER 1

INTRODUCTION

The dilemma of renewing education is constantly evolving as a result of ongoing concerns that students are not being adequately prepared for the demands of the 21st Century (Beyer, 2002; Schlechty, 1997). Concerns over lagging achievement, a perceived need for order, common content coverage, and overwhelming pressure from business and higher education has given rise to standards-based education practices which are designed to increase student literacy levels.

While the standards movement has existed for some time, the piece that separates the most recent resurgence of school reform is the accompanying accountability sanctions (Ellis, 2001). Standards today not only address what students should know and be able to do, but also hold students to higher standards of performance and improved test scores (Tellez, 2003; Wasley & McDiarmid, 2003). “Most states have implemented assessment programs that are being used for high-stakes purposes such as holding schools accountable to improved instruction and student learning as well as for grade promotion and certification” (Lane & Stone, 2002, p. 24).

The current movement illuminates the relationship between teacher quality and student achievement (Wasley & McDiarmid, 2003). “Consistent with the movement for standards for students, this reform [teacher quality] has been called standards-based teacher evaluation” (Henneman & Milanowski, 2003, p. 174). At the heart of reform regarding student achievement and the associated teacher quality issue is the question of teacher effectiveness; how it looks, and how it is measured. Teaching standards provide a

framework for this measure of effectiveness. “Standards of teaching state what teachers should know and be able to do in the exercise of their profession” (Danielson & McGreal, 2000, p. 32). Danielson and McGreal say that school districts need to ensure that their teachers can help students achieve these higher standards and point out that this makes every level of education concerned with teacher performance. As Costantino and De Lorenzo (2006) explain,

The national focus on performance standards for teachers is grounded in the proposition that high standards for student achievement can best be reached if teachers have the knowledge and skills necessary to prepare students to meet these standards (p. 9).

Setting Standards and Defining Teacher Quality

Iowa legislators, cognizant of the critical relationship between student achievement and teacher quality, developed and passed legislation mandating a teacher quality program. Nearly a year later, in January 2001, the federal government reinforced this legislative mandate by emphasizing teacher quality and the measurement thereof as part of the No Child Left Behind (NCLB): Reauthorization of the Elementary and Secondary Education Act. NCLB legislation addressing teacher quality requires that states develop plans and annual measurable objectives regarding the assurance of teacher quality.

The four major elements of the Iowa teacher quality program are (a) quality instruction to all students, (b) closing the achievement gap, (c) recruitment and retention of quality teachers, and (d) the development of quality teachers (Iowa Evaluator Training Manual, Training Module 1, 2005). The overarching focus of Iowa’s teacher quality program is to improve student achievement via improving classroom instruction. Key to

this focus is how quality teaching is defined. To this end, the Iowa Department of Education (DE) established eight teaching standards and 42 corresponding criteria (Appendix A).

The Iowa teaching standards and supporting criteria represent a set of knowledge and skills that reflects the best evidence available regarding effective teaching. The purpose of the standards and supporting criteria is to provide Iowa school districts with a consistent representation of the complexity and the possibilities of the qualities of teaching (Iowa Code, Chapter 284, 2001).

Beginning teachers in Iowa complete a two-year initial licensure period. Near the end of the two-year period, they are evaluated by trained administrators against the established teaching standards using a comprehensive evaluation form developed by the Iowa DE (Appendix B). Based on this evaluation, second year teachers are recommended by their administrator for one of three things: (a) a standard license, (b) a third year of mentoring and induction, or (c) nonlicensure.

Implementation

Implementation of the teacher quality program called for change to occur on a system wide basis in Iowa schools. Implementation began with a commitment by the state to train every principal in Iowa regarding the eight Iowa Teaching Standards and the accompanying method of evaluating teachers. This process has been accomplished and is fully implemented in administrator certification programs at the university level.

Principals in Iowa who evaluate teachers must complete the Iowa Evaluator Approval Training Program (IEATP). The training program is intended to develop skills in the following areas: (a) knowledge and understanding of the eight Iowa teaching standards and criteria, (b) data collection and management skills, and (c) feedback and

conferencing techniques. In addition, it specifically prepares principals to make licensure recommendations at the end of a teacher's two year initial licensure period.

The Evolution of Evaluation

Historically, teacher evaluation has been accomplished using checklists and rating scales that describe teacher behavior and its relation to student achievement. Danielson and McGreal (2000) remark that, "These rating scales and checklists explicitly encouraged a single view of teaching" (p. 14). The authors continue by saying that the simplicity of this type of teacher assessment has established a summative atmosphere with regard to evaluation; one that has been challenged as new insights are gained (or old ones acknowledged) concerning how knowledge is constructed.

Danielson and McGreal (2000) suggest that learning and teaching have shifted away from a behaviorist view and towards a constructivist view. The active construction of knowledge by learners is a basic tenet of constructivism (Gallini & Barron, 2001/2002; Pedersen & Liu, 2003). Constructivism is a theory that, "...assumes that knowledge is individually constructed and socially coconstructed by learners based on their interpretations of experiences in the world" (Jonassen, 1999, p. 217).

In a constructivist setting, students develop skills that include critical thinking, collaborative learning, problem solving, and lifelong learning. The role of the teacher in a constructivist classroom is different from that in a traditional setting. "Teachers serve as a guide, engaging students by helping to organize and assist them as they move towards taking the initiative in their own self-directed explorations" (Herring, 1997, p. 30).

McNelly (2002) says that the role of the teacher has shifted from a “provider of knowledge” to a “learner and instructor of knowledge” (p. 56).

Nolan and Hoover (2004) point out that while teachers have many characteristics in common, each is still individually unique. Like their students, teachers possess different learning styles, motivation levels, cognitive abilities, and personal lives. The authors proclaim, “A one-size-fits-all approach makes no more sense than does a one-size-fits-all approach to teaching children and adolescents. Yet remarkably, many school districts that advocate differentiated instruction for children take a one-size-fits-all approach to supervision and evaluation” (p. 7).

The use of teacher portfolios has been suggested as a means of not only evaluating teachers with more accuracy and depth than previous means but for also providing formative, individualized professional development as well (Danielson, 2001; Henneman & Milanowski, 2003; Nolan & Hoover, 2004; Peterson, 2004). St. Maurice and Shaw (2004) contend that the use of portfolios may promise vast changes in the study and practice of teacher assessment. They remark that, “Many educators say that portfolios promise improved documentation and reflection on professional development as well as rich data from authentic and localized assessments of teaching aligned with state and national standards” (p. 17). Peterson (2000) maintains, “One way to make educational evaluation more authentic is to gather representative artifacts and products into a portfolio” (p. 237).

Wolf, Lichtenstein, and Stevenson (1997) describe three types of portfolios, each constructed for a different purpose. The authors say that portfolios used for evaluation

need to be well structured and have systems of evaluation that are apparent, consistent, and fair. The second type of portfolio, constructed to advance professional growth, is more individually customized than those designed for evaluative purposes. As opposed to the evaluation portfolio, more latitude is given regarding content and structure. The authors say that teachers often design a professional portfolio to fit their personal needs and goals. The third type of portfolio is used in job searches. The authors point out that those in hiring positions may not have a great deal of time to review the portfolio prompting candidates to pay closer attention to details of the portfolio such as presentation, attractiveness, and accessibility. The authors emphasized that, "...a single portfolio can advance all three goals if the person responsible for conceptualizing the portfolio is clear about his or her purposes and thoughtful in design" (p. 196).

Dietz as cited in Danielson and McGreal (2000) says that "A professional development portfolio provides teachers with a framework for initiating, planning, and facilitating their personal/professional growth while building connections between their interests and goals and those of the school" (p. 110). Peterson (2000) cites work by Wolf that describes, "...portfolios more as an attitude of teacher behavior than as a container of information" and are "...strong for capturing the complexities of teaching" (p. 239). Wolf et al. (1997) say that four key features must be present to make a portfolio an effective tool:

1. A portfolio should be structured around sound professional teaching standards and individual and school goals.

2. A portfolio should contain carefully selected examples of both student and teacher work that illustrates key features of a teacher's practice.

3. The content of the portfolio should be framed by captions and written commentaries that explain and reflect on the contents of the portfolio.

4. A portfolio is a mentored or coached experience, in which the portfolio is used as a basis for ongoing professional conversations with colleagues and supervisors (195).

Portfolio Use and Evaluation

In the evaluation process, the use of portfolios must work in concert with observation; not replace it. Per Danielson and McGreal (2000), "Classroom observation is a critical evaluation methodology for those aspects of teaching that may be directly observed" (p. 47). Stronge and Tucker (2003) say that classroom observation is only one piece of the comprehensive puzzle that is teaching. They contend that, "Another important source of obtaining documentation of a teacher's performance is analysis of artifacts (i.e., the collection of written records and documents produced by the teacher as a part of his or her job responsibilities)" (p. 58).

First and second-year teachers in Iowa are not required by the state to complete a portfolio, per se. The language in the legislation, i.e., Chapter 284 of Iowa Code (2001), does not specifically mandate a portfolio. Warren Weber, an evaluation consultant for the DE Teacher Quality Team, says that local districts really don't have to tell the DE how they work through their collections of information and the DE has not asked them to provide their procedures of doing such (personal communication, November 13, 2005).

The summative evaluation form required by the state sets forth expectations of evidence, as judged by trained evaluators, that beginning teachers are meeting the eight Iowa teaching standards. Teachers are required to collect and provide artifacts, as defined by local districts, representative of the teaching standards established by the state. Furthermore, trained evaluators are required to examine evidence as it relates to the eight Iowa teaching standards and criteria as they make licensure decisions at the end of the initial two year period. The consistency across the state is established by the eight teaching standards. However, due to local control, methods of evaluation vary.

Stronge and Tucker (2003) define a portfolio as, "...a formalized process of organizing and reviewing artifacts" (p. 58). Local districts in Iowa may choose to have teachers display their collection of artifacts in portfolio form. Research was not found documenting how many Iowa school districts use the term "portfolio"; however, a review of sample case studies regarding implementation of professional development on the DE website revealed that two schools used the word "portfolio" in their professional development plans (Iowa Department of Education, n.d.1, Iowa Professional Development Model section). For purposes of this study, a portfolio will be defined as the artifacts an Iowa teacher is expected to collect to illustrate that they have sufficiently met the ITS/criteria.

Portfolios are used in combination with formal and informal observations. The DE in Iowa does not mandate the number of observations, the formative process, the length of observations, etc. Once again, these parameters are determined by the local districts.

Perhaps the biggest challenge in any type of evaluation is the element of judgment. The constructivist nature of portfolios amplifies the need for quality evaluative criteria. Tigelaar, Domans, Wolfhagen, and van der Vleuten, (2005) contend that, “Unambiguous, objective rating of portfolios is difficult to achieve, because the richness and uniqueness of the contents of the portfolio necessitate interpretation and taking account of the context before judgment can be passed” (p. 595). The credibility of the process and the evaluator is increased by the strength and clarity of the assessment policies (Shinkfield & Stufflebeam, 1995). It is imperative that those being evaluated trust that they will be evaluated fairly and consistently. Creating this trust means making sure that instructions to teachers for creating a portfolio are explicit and that both the teacher and the evaluator understand the rubric to be used for evaluation. (Green & Smyser, 1996).

Aside from using the eight Iowa teaching standards to guide the construction of a portfolio, a state-wide method/rubric for analyzing portfolios is currently not available and may be impossible to create due to the element of local control. The document most representative of a state-wide evaluation tool might be the comprehensive evaluation form (Appendix B). Principals indicate on the form, based on various data sources, whether the teacher has met or not met the prescribed standard. There is space for narrative under each standard where the evaluator is encouraged to incorporate and address each criterion. There is no delineation as to the level of proficiency that the teacher has met. They either meet the standards or not.

Statement of the Problem

The value of using portfolios as a component of teacher evaluation has been and continues to be advocated, and, at least in Iowa, the use of portfolios as an integral piece of teacher evaluation and professional development is a reality. However, little is presently known about (a) how principals in Iowa critique the contents of portfolios, (b) how principals make judgments concerning the contents, and (c) what bearing the portfolio contents might have on licensure decisions.

Have principals developed methods and tools at the local level that represent consistent, fair portfolio evaluation? Do administrator thought processes bear any similarities across demographic and academic levels? Could it be that administrators, in the interest of time, have established yet another checklist to evaluate teacher portfolios negating the potential for reflective assessment and constructive growth? Do they simply make sure that a “piece” of evidence exists in the teacher’s portfolio? Does this evidence indicate that a standard has been sufficiently met? Or, have principals, in fact, developed thoughtful processes and tools for portfolio evaluation that are conducive to the growth of beginning teachers and the assurance that quality is being proliferated?

Purpose of the Study

The purpose of this qualitative study was to discover and describe the thinking and methods used by principals as they evaluate second-year teacher portfolios. In addition, the study was designed to ascertain how much bearing the portfolio evaluation has on decisions to move second-year teachers beyond the initial licensure stage.

Research Questions

The central question guiding this study is, “How do principals evaluate second-year teacher portfolios?” Three research questions will direct the study:

1. What do principals verbally report they are thinking as they review second-year teacher portfolios for purposes of evaluative judgment?
2. What tools have principals developed to assist them in evaluating second-year teacher portfolios?
3. How much bearing does portfolio evaluation have on the judgment the administrator makes regarding licensure?

Definition of Terms

Terms used in this study include:

Artifacts: The products and by-products of teaching that demonstrate a teacher’s performance (Tucker, Stronge, & Gareis, 2002, p. 25).

Beginning teacher: An individual serving under an initial license, issued by the board of educational examiners under Iowa Code chapter 272, who is assuming a position as a classroom teacher (Iowa Code, Chapter 284, 2001). First and second year teachers are beginning teachers.

Comprehensive evaluation: A summative evaluation of a second year teacher conducted by an evaluator for purposes of determining levels of competency relative to teaching standards and for recommendation for licensure (Iowa Code, Chapter 284, 2001).

Evaluator: An administrator who successfully completes an evaluator training program (Iowa Code, Chapter 284, 2001). Used interchangeably with principal and administrator.

Initial licensure: The license issued to 1st and 2nd year teachers in Iowa.

Teacher evaluation: Any of a variety of formal and informal programs for assessing the competence and effectiveness of an instructor (Danielson & McGreal, 2000).

Portfolio: For purposes of this study, a portfolio will be defined as the artifacts an Iowa teacher is expected to collect to illustrate that they have sufficiently met the ITS/criteria.

Rubric: A scoring guide to assess subjective exercises (Green and Smyser, 1996).

Standards: Expected outcomes that delineate the key aspects of professional performance (Campbell, Cignetti, Melenyzer, Nettles, & Wyman, 2001).

Significance of Study

St. Maurice and Shaw (2004) assert that, "...teacher portfolios may be on the verge of bringing enormous changes to the study and practice of teacher assessment, a field which heretofore has been dominated by standardized multiple-choice tests and checklists devised outside of the classroom " (p. 17). They continue by saying that the effects of the use of portfolios are still unmeasured. The authors contend that administrators who use portfolios need research-based information to properly assess portfolios.

The information gleaned from this study will help ascertain if teacher evaluation has truly evolved beyond the checklists described by St. Maurice and Shaw (2004). The resulting information may provide guidance to DE and university-level administrator preparation programs regarding current practice in regard to portfolio evaluation. The strategies and processes used by the participants may provide a broader basis for accurately and consistently assessing teacher portfolios across the state.

CHAPTER 2

REVIEW OF THE LITERATURE

The purpose of this qualitative study was to discover and describe the thinking and methods used by principals as they make evaluative judgments regarding second-year teacher portfolios. In addition, the study was designed to ascertain how much bearing the portfolio has on licensure decisions.

The information in this chapter will provide further background concerning the function of teacher evaluation, the role of the principal in the evaluation process, the use of portfolios as an instrument in evaluation of teachers, and suggested methods of portfolio evaluation.

Assessment Systems

In 1996, The National Commission on Teaching and America's Future (NCTAF) set the following goal: "Within a decade—by the year 2006—we will provide every student in America with what should be his or her educational birthright; access to competent, caring, qualified teaching in schools organized for success" (p. 21). The impact of this goal was a new intensity surrounding the purpose and process of teacher evaluation. Reauthorization of NCLB has moved schools towards data-based decision-making processes (Marshall, 2004). As a result, protocol pertaining to teacher quality and teacher assessment has become an integral piece of school-based accountability systems with an increased emphasis on the process of assessment.

The concept of teacher assessment is not new. However, the system with which teachers are evaluated has evolved as a result of current reform. Reeves (2004) notes that,

“...the assessment of teachers in some schools has been transformed from a superficial checklist and hasty observation to deep reflection by teachers, colleagues, and administrators, all with a view toward improving professional practice rather than merely rendering an evaluation” (p. x). The transformation of evaluative practices has been an effort to move evaluation and supervision towards the common goal of improvement.

Stronge and Tucker (2003) indicated that teacher evaluation serves a dual purpose of improving teacher performance and documenting accountability. Some have pointed out that these two purposes have been considered incompatible (Beerens, 2000; Danielson & McGreal, 2000; Stronge & Tucker, 2003). However, this incompatibility may have more to do with how an evaluation system is structured rather than irreconcilable differences (Danielson & McGreal, 2000).

Single dimensional systems may be the culprit. “An evaluation system should recognize, cultivate, and develop good teaching” (Danielson, 2001, p. 13). An example of a poorly constructed system might consist of a one-shot observation that is often perceived by both teachers and principals in terms of efficiency rather than effectiveness. “Neither the teacher nor the principal has any misconceptions about the process. Both may be highly motivated, dedicated, and skilled professionals, but both see the observation process as a formality to be dispensed with as painlessly as possible” (Blake, Bachman, Frys, Holbert, Tamara, & Sellitto, 1995).

Nolan and Hoover (2004) agree with Danielson (2001) concerning the critical nature of a sound evaluation system.

An effective teacher supervision and evaluation system must be capable of remediating or eliminating poor performance as well as nurturing excellent

performance. Its teacher evaluation process must be robust enough to differentiate between the two. A comprehensive system of supervision and evaluation also leads to greater clarity for all educators concerning the purposes and the procedures that are employed for accountability and for professional growth (p. 7).

The authors differentiate teacher evaluation from teacher supervision. They describe teacher evaluation as a summative measure that ascertains the level of all teachers using given standards as judged by an appropriately trained expert. They assert that the purpose of teacher supervision is to “[promote] teacher growth, which in turn leads to improvement in teaching performance and greater student learning” (p. 26). They contend that supervision is not concerned with judgment. Nolan and Hoover say that supervision and evaluation complement each other by ensuring that acceptable levels of performance exist as well as do opportunities for growth. In short, evaluation should be intended to support teacher growth.

Davis, Ellett, and Annunziata (2002) posit that a well-developed teaching and learning assessment system can support concepts such as collegiality and collaboration and identify professional growth needs. Nowhere is teacher growth more apparent or more critical than in the first two years of teaching. Evaluation during this period is of vital importance due to licensing requirements and successful induction.

Peterson (2000) advises that evaluation of new teachers consists of two major functions: (a) reassurance, and (b) an affirmative introduction to the evaluation system including data collection and documentation. The author suggests that new teachers need to experience proactive support, the use of multiple and varied data sources, feedback tied to in-service education, and teacher control. He states that, “...the key for beginners

is to develop sound data and attitudes” and that without this kind of foundation, “...teachers become poor consumers of evaluation, permitting disastrous practices and failing to demand good ones” (p. 287). Peterson points out that assistance and assessment during the first year are not merely to make it more pleasant. The goal is to, “promote positive career-long attitudes and development” (p. 287).

System components: Recognizing that teaching is a complex activity is vital to a teacher evaluation system designed to make judgments and perpetuate growth (Beerens, 2000; Danielson & McGreal, 2000; Good & Mulryan, 1990; Nolan & Hoover, 2004; Peterson, 2000; Stronge & Tucker, 2003). Nolan and Hoover contend that “Teaching by nature is recursive, multifaceted, and nonlinear” (p. 17). Based on this recognition, Danielson and McGreal suggested that an effective teacher evaluation system contain three essential elements:

1. A coherent definition of the domain of teaching (the “What”), including decisions concerning the standard for acceptable performance (“How good is good enough?”).

2. Techniques and procedures for assessing all aspects of teaching (the “How”).

3. Trained evaluators who can make consistent judgments about performance, based on evidence of the teaching as manifested in the procedures (p. 21).

The teacher quality program implemented in Iowa was discussed in Chapter 1. A review of the components reveals that the three essential elements to which Danielson and McGreal (2000) refer are reflected in the program. The standards of performance are clearly outlined in the 8 Iowa teaching standards and 42 model criteria (Appendix A).

Suggested techniques and procedures for assessment consist of both formative and summative measures including artifact collection, dialogue before and after observation, and a summative review. In addition, principals are required to complete an evaluator training program that is designed to increase their knowledge and understanding of the eight Iowa teaching standards and criteria, coach them on data collection and analysis, management skills, and improve their skills in feedback and conferencing techniques.

The Role of the Principal

In essence, the principal has two roles in a teacher quality program. In one role, the principal is the facilitator of teacher evaluation. In the second, the principal is the evaluator. The roles eclipse at the point of teacher evaluation. Peterson (2004) observes that research over the past 25 years identifies, "... the principal as the central person in school teacher evaluation" (p. 70). Zimmerman and Deckert-Pelton (2003) echo this when they say, "Many stakeholders and educational researchers would also agree that principals are key players in the success of an effective teacher evaluation, and any subsequent teacher improvement and increased student achievement" (p. 28).

Although principals have long been in evaluator roles, the evaluation process and the high stakes surrounding teacher evaluation have put increased emphasis on this familiar role, especially as it relates to beginning (1st and 2nd year) teachers. Shinkfield and Stufflebeam (1995) write that staff evaluation is one of the most important responsibilities of a principal. They say that, "... the school principal must examine the performance of staff members in order to provide constructive feedback and to make decisions that affect individual teachers and the school itself" (p. 303).

Davis et al. (2002) contend that, "...leadership makes the difference between perfunctory and summative teacher evaluation and meaningful assessment of the teaching and learning process that has the potential to enhance the quality of teaching and student learning" (p. 288). The authors describe two case studies that accentuate the critical role that principals' beliefs and behaviors play in the acceptance of an evaluation system by their teachers. In each case, the leader of the organization was implementing a new teacher evaluation system. In the first case, the leader acted as a "knight in shining armor" and believed he was protecting his teachers from an unfair judgment system that was an insult to the integrity of his teachers. In doing so, he alienated his staff from the system before they gave it a chance to work. There was little, if any, understanding of any facet of the evaluation system or how it might have enhanced student learning.

In the second case, the leadership reflected a "small jazz combo (SJC)" style. Everyone played a leadership role in the implementation of the new evaluation system. "The principal of the SJC school was enthusiastically supportive of the opportunity provided by the new evaluation system to focus attention on teaching and learning" (Davis, et al., 2002, p. 296). The activities undertaken in the school were reflective of the initiatives of the new system. Consequently, the staff in the second school reported feeling positive about the change and attributed it to the enthusiastic support the principal exhibited.

A study by Brock and Grady (1998) examined the perceptions of first year teachers regarding the role of their principals. Results of the study indicated that beginning teachers want principals to communicate with them regarding expectations of

good teaching. Furthermore, beginning teachers identified the school principal as the major source of support and guidance and, “stressed the importance of classroom visits, feedback, and affirmation” (p. 180).

A study by Zimmerman and Deckert-Pelton (2003) ascertained teachers’ perceptions of principals as evaluators. Specifically, the authors attempted to establish how teachers viewed their principals as primary evaluators, how they perceived the principal’s role in the evaluation process, and what they thought made their principal a good evaluator. The participants in the study were practicing K-12 teachers who were enrolled in educational leadership graduate classes. The years of experience of the participants were not evident from the study.

The findings of Zimmerman and Deckert-Pelton (2003) parallel those of Brock and Grady (1998). Most notable was the expressed desire of 89% of the teachers for feedback via a bidirectional process, “...the educators consistently expressed both a desire to have a reciprocal, communicative relationship with their evaluators and a need for the evaluation process to contain constructive feedback about their professional strengths and weaknesses” (p. 32). The respondents also indicated that they perceived the principal’s commitment to the process as pivotal to the success of the teacher evaluation system. Zimmerman and Deckert-Pelton summarized that with commitment from principals, “Teachers seem to view the process [evaluation] as holding great potential for improving their pedagogical knowledge, skills, and abilities” (p. 34).

Protheroe (2002) notes that, “To do teacher evaluation well, the principal needs an understanding of standards for student learning, an in-depth sense of what good teaching

looks like, and a strong ability to communicate and provide constructive feedback” (p. 48). She also explains that principals must have an understanding of the differing philosophies of the teachers whom they observe. Without this understanding, the possibility of misinterpreted instructional practice exists.

Burke (1997) emphasizes that, “...conducting a summative evaluation of a professional portfolio requires a great deal of thought, planning, and organization” (p. 118). Credibility and fairness lie in the balance. Teachers perceive fairness in terms of consistent, acceptable application of evaluation standards and procedures (Kimball, 2002). Kimball conducted a qualitative study in three school districts that had each implemented a new standards-based evaluation system similar to that used in Iowa. Each school’s evaluation system made use of teacher portfolios as a data source. Interview questions explored the knowledge and acceptance of teachers and evaluators in terms of evaluation standards and evidence requirements, the nature of feedback and support, perceptions of fairness, and impacts on teaching and professional development.

Kimball (2002) reported that teachers in each district spoke approvingly in terms of the fairness of the new system; however, he noted a tension regarding reliability and validity of the new system due to increased burden placed on teachers and evaluators. “Increased workload may have contributed to some evaluators cutting corners on evidence gathering, writing reports, and providing feedback” (p. 261). He made the following propositions based on this concern:

1. Regardless of the clarity of evaluation manuals and commitment of central office staff to the evaluation reform, without required on-going training and accountability of evaluators, evaluation consistency will suffer.

2. No matter what cautions are taken to assure sufficient validity of a teacher evaluation system, if evaluators are not consistent in their approach and teachers do not see the system as “valid” and professionally credible, it is not likely to contribute to meaningful instructional change (consequential validity; p. 262).

The critical role of the principal and his/her training in the evaluation process is clear but how the principal evaluates with an appreciation of the complexities involved in the act of teaching is not. What evidence is available to principals that may be used to make qualified judgments concerning a teacher’s ability? How can they accurately assess all aspects of teaching?

Portfolios as a Link

Reese (2004) says that, “A portfolio can convey a teacher’s beliefs, knowledge, and skills” (p. 18). Wolf (as cited in Burke, 1997) explains that, “...a portfolio contains more information than is normally available for assessing a teacher’s competence...” (p. 120). Xu (2004) describes a teaching portfolio as “an organized collection of evidence about a teacher’s best work that is selective, reflective, and collaborative” (p. 198). Xu also emphasizes that a teaching portfolio is constructed from the teacher’s perspective and is a useful means to increase communication with those outside the classroom.

Xu (2003) conducted a case study relative to the impact of teaching portfolios on professional learning and professional collaboration. A portfolio project was introduced

into an urban elementary school. Twelve teachers, their principal, and a staff developer were interviewed and portfolio artifacts were collected. The results revealed that the portfolio project had a positive impact on professional learning and professional collaboration. Specifically, teachers and administrators reported that their relationships with each other were positively enhanced. The principal felt that reading a teacher's portfolio prior to an observation provided insightful conceptual information about the teacher and his/her teaching.

The teachers felt that the portfolios they developed provided a channel of communication between them and the administrator; a venue for dialogue about teaching in general, and a means to individually discuss their personal strengths and weaknesses in a constructive way. Xu (2003) also reported that as a result of the portfolio project teachers began to view themselves as change agents in the evaluative process.

Gelfer, Xu, and Perkins (2004) write that, "...teaching portfolios can provide a practical method to document both the characteristics of the instructional environment and the outcomes of teaching" (p. 128). Green and Smyser (1996) state that, "The essential value of a teaching portfolio is its benefit to the teacher who prepares one" (p. 95). They further assert that, "A few observations by the principal do not tell the whole story" (p. 101). Portfolios, they say, are a means for teachers to explain the background, i.e., provide the context. St. Maurice and Shaw (2004) maintain that teacher portfolios can provide a rich data source for, "...authentic and localized assessments of teaching aligned with state and national standards" (p. 17). They point out that it is not yet clear how teacher portfolio assessments will be designed and validated.

Attinello (2004) assessed teacher and administrator perceptions of the value, accuracy, utility, and feasibility of teacher portfolios as part of the teacher appraisal system in a large public school system. The mixed-methodology research results indicated that teachers and administrators thought that portfolios were more accurate than one-shot observations; however, it is interesting to note that, “Administrators were significantly more supportive than teachers in their perception of portfolios as a comprehensive measure of teacher performance” (p. 111).

The teachers in Attinello’s (2004) study identified improved communication and interaction with their administrator as an advantage of using portfolios as part of the evaluation process. Lack of administrator time was an identified disadvantage. Attinello asked teachers and administrators if they thought that the portfolio process promoted good teaching practices. The administrators in the study were significantly more supportive than the teachers with respect to the level that the portfolio process promoted good teaching practice. Teachers were very concerned about the focus of the portfolio, i.e., fluff versus content.

Evaluating Portfolios

Green and Smyser (1996) write that teacher concerns regarding portfolio content have merit. The authors assert that, “...it is possible for a teaching portfolio to look better than a teacher” (p. 102). They suggest that evaluators keep the following principles in mind as they evaluate portfolios so that judgments are consistent and fair: (a) evaluate the teacher, not the portfolio; (b) establish the purpose of the evaluation; (c) develop the rubrics for the evaluation; (d) train the evaluators; and (d) validate the evaluation rubric.

Peterson (2000) voiced concern about mandated use of portfolios in summative evaluations. He indicated that evidence and the process of judgment can be distorted. He remarks, "Summative uses reward portfolio producers, not necessarily good teachers. An evaluation system that places a premium on portfolios soon creates an industry of portfolio assembly far beyond authentic samples of teacher work" (p. 242).

Danielson and McGreal (2000) said that evaluation must focus on teaching practice. Evaluators must strive to judge the quality of teaching rather than the quality of the portfolio. The portfolio should not be the object of the evaluation. The authors also indicate that the purpose of the evaluation must be apparent. In the current study, portfolios will be utilized by principals to evaluate teacher performance against the eight Iowa teaching standards for purposes of teacher licensure. This conceptual framework (established standards), or one similar to it, must be in place so all stakeholders in the system clearly understand the purpose of the evaluation system and the values that underlie it (Nolan & Hoover, 2004).

Wolf et al. (1997) suggest that, "[A] way to make the portfolio construction and evaluation process more manageable and fair is to specify the requirements for the portfolio in advance" (p. 201). The authors recommend that the following information be made available to novice teachers via a handbook:

1. Purposes of the portfolio.
2. Procedures for constructing the portfolio.
3. Timeline for completion and evaluation of the portfolio.
4. List of required and/or suggested portfolio contents.

5. Description of the evaluation process.
6. Evaluation criteria (content and performance standards).
7. Description of the feedback and appeals process. (p. 202)

Nolan and Hoover (2004) point out that evaluating a formative portfolio is much different than evaluating a summative portfolio. In other words, they agree with Stronge (2002); the purpose of the evaluation must be apparent. They believe that each contextual factor associated with good teaching practice must be represented and considered due to the high stakes for the novice teacher and for the school district that hired him/her. Green and Smyser (1996) concur:

Performance evaluation calls for a different kind of portfolio and for a different approach to evaluation. When a teacher's professional performance is being considered, specific rubrics, or rules, need to be developed and followed. In this case, the balance between uniformity and flexibility becomes delicate. The portfolio needs to include evidence of essential teaching skills. In addition, the variety of teaching situations and diversity of individual strengths must be accommodated. Typically, a portfolio that is going to be used as the summation of a teacher's professional performance will have more "required" documents than one that is going to be used for self-evaluation" (p. 103-104).

The National Board for Professional Teaching Standards (NBPTS) is recognized by many as the hallmark in the use of portfolios. A visit to their website allows visitors to view exemplar models including criteria explanation, artifact description, and scoring methods (rubrics). Teachers who are certified by NBPTS receive explicit instructions on what to include in their portfolio. Evaluators are extensively trained in making judgments based on the representative criteria.

Green and Smyser (1996) point out that the process used by NBPTS has been established for veteran teachers and that different rubrics would need to be established for tenure decisions. They add that the type and extent of artifacts would vary as well.

Scoring Portfolios

Burke (1997) asserts that evaluation of all types are effective only if the scoring instrument measures what it is supposed to measure consistently and reliably. Rubrics are thought to promote consistency and reliability. Rubrics used for summative evaluation purposes when scoring portfolios should contain three distinct features: (a) the attributes of good teaching (e.g., the eight Iowa teaching standards outlined in Appendix A), (b) the characteristics of the evidence used to reveal good teaching, and (c) the performance criteria used when the evidence is considered (e.g., exemplary, proficient, unsatisfactory (Green & Smyser, 1996). Both teachers and evaluators must be acutely aware ahead of time what values will be applied in an evaluation and what aspects of teaching are to be emphasized. If this information is not apparent, "...the evaluation of the teacher, as evidenced by the portfolio, deteriorates into a portfolio contest" (p. 104).

Burke (1997) describes two ways of judging portfolios. She says that evaluators can use a holistic scoring method where raters assign a single overall score based on the overall quality of the portfolio or they can score analytically where raters give separate ratings to different aspects of the portfolio. She also brings attention to critical issues that should be considered when developing a scoring process:

1. Whether each piece, selected pieces, combination of pieces, or the total collection will be scored.

2. Whether analytic, holistic, or a combination of scoring approaches will be used.
3. Who will be scoring and what training they will have received.
4. What scoring rubrics will be used to judge or grade each item, and who will develop them and select and/or prepare the benchmarks to go with them.
5. Who will monitor the judges and ensure fairness, accuracy and integrity of the scoring process.
6. What type of scale or system will be used to report the results of the portfolio scoring to the individual teacher and to others (e.g., mentor, teacher, evaluator) (p. 119-120).

Glatthorn (1996) describes how an evaluator should approach the process of evaluating portfolios. He makes the assumption that the evaluator will have, at the very least, several portfolios to evaluate. He divides the process into three phases. In the first phase, the evaluator reviews the portfolios to ensure that they meet minimum design standards. Portfolios not meeting the minimum are returned to the teacher for additional work.

The second phase represents a holistic review. This review provides the evaluator with a general impression of overall quality. The portfolios are sorted into as many piles as there are rating levels. In the aforementioned example, three rating levels were suggested (exemplary, proficient, and unsatisfactory); thus, three piles would be created. Glatthorn (1996) suggests that a second review in this phase will ensure that the first judgment was valid.

In the third phase of the evaluation, the evaluator makes an analytic rating of each portfolio. Each criterion is considered, the evidence is reviewed, and then a rating for each is assigned using the same rating terminology as in the second phase. A final rating is assigned but it is not the average of the analytic ratings. It is an overall assessment of the general quality of performance. The author notes that, "...the holistic rating is made first, based upon a general impression of performance; that holistic rating is then supported with analytic assessment" (p. 66).

Wolf et al. (1997) recommend a similar systematic review process including the following steps:

1. Read the entire portfolio to get a sense of the overall performance.
2. Review the portfolio in light of the content standards and teacher goals.
3. Take notes about significant pieces of information in the portfolio.
4. Assign a rating for the portfolio (if appropriate).
5. Provide feedback to the teacher. (p. 202)

Green and Smyser (1996) voiced concerns about portfolio evaluation relative to validity and reliability. "The validity of the evaluation of teaching portfolios depends upon rubrics, and the reliability depends upon the training the evaluators receive" (p. 105). They acknowledge the subjective nature of evaluating portfolios and say that training evaluators to use evaluation rubrics properly contributes to reliable ratings.

The authors emphasized that validity is also a necessary component. They say that validity is making sure that the evaluation is representative of what teachers actually do and encourage users to withhold approval of any rubric until a level of confidence in the

instrument has been established. They believe that reliable, valid evaluation tools build trust with teachers and ultimately help teachers accept change. They also point out that in cases where personnel decisions are necessary, highly reliable and valid tools provide defensibility.

Shortcomings

The use of portfolios to this point has sounded much like a panacea; the answer to effective teacher evaluation and thus, the beginning of meeting the goal regarding teacher quality set by NCTAF in 1996. Wolf et al. (1997) describe portfolios as, “exciting assessment tools because they allow teachers to represent the complexities and individuality of teaching in great detail” (p. 198). However, the authors say that portfolios have associated liabilities in that they are time consuming to construct, cumbersome to store, and difficult to score (p. 194). Others point out problems with teacher portfolios as well. As previously discussed, Peterson (2000) and Green and Smyser (1996) believe that good portfolios can make bad teachers look good and vice versa. Peterson identifies three additional problems with teacher portfolios:

1. The open ended nature and nonuniformity makes it difficult to judge overall adequacy.
2. Portfolios are bulky and present difficulty with storage.
3. Portfolios often do not include perspectives other stakeholders, i.e., parents, students, peers. (p. 241-242).

Evaluator training and judgment are elements of concern that are continually mentioned in the literature regarding teacher evaluation and portfolios. Burke (1997) has

concerns that demands on principals' time and the possible lack of depth in content knowledge skew judgments. She suggests that the review process be completed by a committee of peers. She says peer reviews ensure that evaluators recognize competencies and quality documentation more adequately. An inherent problem in using peers to evaluate is the time commitment of yet other teachers.

Glatthorn (1996), while an advocate of portfolio use, explains that, from a teacher's perspective, portfolios are time consuming and might interfere with, rather than enhance, growth activities. He adds that, "...portfolios by themselves do not always give the teacher objective feedback about performance" (p.33). The author indicates that disadvantages for the administrator exist as well. He says that portfolios used in isolation do not provide sufficient objective evidence for use in designing professional development since the artifacts are gathered and assembled by the teacher and represent highly selective evidence of teaching. He does not believe that portfolios comprise the sum of objective evidence needed to make evaluation decisions in terms of tenure.

Beerens (2000) echoes Glatthorn's (1996) concern regarding the use of portfolios as the only means of evaluation. He believes they are only a piece of the total picture and that other data sources must be considered. Tucker, Stronge, and Gareis (2002) agree. They are quick to point out that their support of portfolio use in no way suggests that classroom observations be eliminated. They advocate, "...the use of multiple data sources, with a particular focus on performance portfolios, in order to develop a fuller, more accurate picture of performance" (p. 70).

Summary

Defining teacher quality and evaluating the effectiveness of teachers is front and center in today's world. It has been pushed there because of questions regarding the achievement of our country's students. Politically, accountability is of the essence from school grounds to the nation's capital. Those that are responsible for providing it are being asked to "show us the evidence." Has this new focus changed the way teachers are evaluated or has it simply increased the verbiage surrounding it?

The intent of the current study is to get a glimpse into the reality of the portfolio evaluation process in Iowa; an important piece of the bigger picture that is teacher evaluation. The literature review provided background concerning the function of teacher evaluation, the role of the principal in the evaluation process, the use of portfolios as an instrument in the evaluation of teachers, suggested methods of portfolio evaluation, the role of judgment in decision making, and protocol analysis as a tool used in qualitative research to ascertain cognitive processes of subjects as they perform a task. This review has been an effort to not only describe these entities but to connect them in such a way that makes sense of the process of evaluation relative to teacher portfolio assessment.

CHAPTER 3

METHODOLOGY

The purpose of this study was to examine the thinking processes, methods, and materials used by principals as they made evaluative judgments regarding second-year teacher portfolios. The overarching question of the study was, "How do principals evaluate second-year teacher portfolios?" The following research questions guided the study:

1. What do principals verbally report they are thinking as they review second-year teacher portfolios for purposes of evaluative judgment?
2. What tools have principals developed to assist them in evaluating second-year teacher portfolios?
3. How much bearing does portfolio evaluation have on the judgment the administrator makes regarding licensure?

The current study was conducted using a qualitative approach. Data for the study was gathered via a think-aloud process in combination with guided interview questions. Nine principals participated in the study; three each from elementary, middle, and high school. The nine principals were also representative of rural, suburban, and urban geographic/demographic regions. The think-alouds and ensuing interviews were audio-recorded and then transcribed. The resulting verbal reports (comments) were analyzed and categorized using the constant comparative method. The comment counts were used to report the accumulated data and make comparisons between academic level and between geographic/demographic regions. In addition, any tools that the principals used

during evaluation were documented and/or photocopied. The information in this chapter provides rationale for the design of the study, the protocol used to collect and analyze data, and the definition of the sample. The chapter concludes with discussion concerning the reliability, validity, and limitations of the study.

Research Design

Because there is no established, common method for portfolio evaluation in Iowa, there exists no standard against which to measure procedures that are currently in use by individual principals. Consequently, observing and listening to principals as they evaluated a second-year teacher portfolio supplied rich context to answer the questions posed in this study.

Qualitative research provided the best window through which to view this context. Bogdan and Biklen (2003) describe qualitative research as naturalistic, descriptive, process concerned, inductive, and meaning producing. Qualitative research is naturalistic in the sense that setting and dialogue provide authentic significance to the data acquired and ultimately to the genuine nature of the study. "Qualitative data in the words and categories of participants lend themselves to exploring how and why phenomena occur"(Johnson & Onwuegbuzie, 2004, p. 20). The very nature of this study was to capture data concerning mental processes used by principals as they critiqued and made judgments concerning second-year teacher portfolios; the naturalistic feature of qualitative research provided an appropriate lens of discovery.

To facilitate the collection of said qualitative data, i.e., to document the thinking of principals, a technique known as a "think-aloud" (TA) was used. At first glance, it

might appear that simply interviewing principals could reveal their thinking and thus provide cues regarding their mental processing and physical processes and resulting judgment that occur during portfolio evaluation. However, think-aloud (TA) protocol designed by Ericsson and Simon (1993) provided deeper insight. Think-aloud protocols are commonly used in reading strategies and studies. "In think-aloud studies, subjects report their thinking as they do a task [concurrent reporting]" (Pressley and Afflerbach, 1995, p.1). Muth (1993) explains the TA process by saying that those involved in the process are "...asked to say aloud the things that they usually mumble to themselves" (p. 5).

Shavelson, Webb, and Burstein (1986) explained that cognitive research may be considered as a means to collect data on mental processes by probing the thoughts, judgments, and decisions of participants. They included TA protocol as a tool in cognitive research and described it as a verbal reporting method that "...is interpreted as a series of mental operations that the researcher infers that the subject used to reach a judgment, decision, or problem solution" (p. 79). Furthermore, they infer that a TA has the ability to produce verbal protocols that are complete, have little or no effect on process time, and do not distort the structure and course of cognitive processes. The authors were quick to point out that "...we do not claim that verbal report data reflect actual (neural) cognitive processes" (p. 82). However, they did maintain that verbal reporting can provide specific, good quality data for examining cognitive processes.

Pressley and Afflerbach (1995) say an advantage of using TA protocol is that "...verbal reports sometimes can provide access to the reasoning processes underlying

sophisticated cognition, response, and decision making...” (p.4). Since it was the intent of this study to examine the cognitive processes that lead to a judgment, it is important to access those processes via the best means possible; in this case, TA protocol was an appropriate modality to answer the research questions adequately.

The descriptive aspect of qualitative research allows for the use of quotations from the data to substantiate patterns, themes, and theories (Bogdan & Biklen, 2003). Conducting TA/interview sessions with principals to establish their methods of portfolio evaluation lent themselves well to using the words of the participants as they described the physical and mental processes they used to evaluate a second-year teacher portfolio. Emerging trends were not only identified, compared, and clarified; they were strengthened by the personal voices of the participants.

Qualitative research places an emphasis on process above outcomes. The search in the qualitative process is for meaning. Bogdan and Biklen (2003) suggest that considering the following process-oriented questions is crucial when conducting qualitative research:

1. How do people negotiate meaning?
2. How do certain terms and labels come to be applied?
3. How do particular notions come to be taken as part of what we know as “common sense”?
4. What is the natural history of the activity of events under study? (p. 6)

The task of evaluating portfolios is, in and of itself, process oriented, as is determining how it is accomplished. The meaning associated with procedures used by the participants was teased out via the think-aloud/interview process.

The goal of this study was to discover how and by what thought processes principals review and make evaluative judgments about second-year teacher portfolios. Bogdan and Biklen (2003) refer to this as an inductive process. "Theory developed this way emerges from the bottom up (rather than from the top down), from many disparate pieces of collected data that are interconnected. The theory is grounded in data" (p. 6). A quantitative study, such as a survey, makes predetermined assumptions concerning context. In the qualitative venue used in this research, assumptions were not made about data or the potential thereof. Rather, the data created a context for understanding the thought and judgment-making processes of the participants.

The intent of this study was to identify themes from the interconnected process descriptions captured during the TA/interview episodes. These themes provided a basis for answering the research questions regarding how principals evaluate portfolios and the bearing the portfolios have on tenure decisions. Bogdan and Biklen (2003) pointed out that the early portions of the research should guide latter portions. They said that, "The qualitative researcher plans to use part of the study to learn what the important questions are" (p. 6).

Participants

Nine participants were selected from school districts using maximum variation sampling strategy. This type of purposeful sampling allowed the researcher to select

participants who represent "...the range of variation in the phenomena to be studied" (Gall, Gall, & Borg, 2003, p. 179). These authors say it also aids the researcher in determining whether common themes established from the research cut across the variation.

Iowa has communities that represent a range of settings; therefore, if the information collected is to be useful to all educators in the state, it is pertinent to include schools representative of the populations served in the state. Using the maximum variation strategy of purposeful sampling (Gall, Gall, & Borg, 2003), nine principals were asked to participate in the study; three each from the elementary level, middle school level, and high school level. Principals that were representative of the varying size and community settings indigenous to Iowa were invited to participate in the study with the intent that one principal from each level (elementary, middle, and secondary) represent an urban school, one a suburban school and one a rural school. The principals invited to participate in the study had evaluated second-year teachers in the 2004-2005 school year and were not associated in any way with the teachers who contributed their portfolios for use in the current study.

For purposes of this study, an urban school was defined using the definition established in 1984 by the Urban Education Network of Iowa (UEN). According UEN by-laws,

Any duly organized and legally constituted public school district in Iowa with two or more comprehensive high school attendance centers and/or 10,000 or more students whose composition includes the major characteristics of "urbanness" including, population density, multicultural and broad and varied socio-economic and ethnic representation, may become a member of the network... (Urban Education Network of Iowa By-laws, Section 2).

The following school districts in Iowa are members of the Urban Education Network: Cedar Rapids, Council Bluffs, Davenport, Des Moines, Dubuque, Iowa City, Sioux City, and Waterloo. Three principals from three different UEN districts agreed to participate in the study.

The United States Department of Agriculture Economic Research Service (USDA-ERS; n.d., ¶ 4) indicates that "...rural areas consist of all territory located outside of urbanized areas and urban clusters" Given this definition and the established urban districts as defined by the UEN, the researcher contacted four rural districts with county populations less 20,000 and district student populations less than 1,500. Three principals elected to participate in the study.

For purposes of this study, suburban was defined as those districts in counties that are not identified by the UEN and do not have characteristics of a rural school. Three principals in schools having characteristics of said suburban definition agreed to participate in the study.

Gaining Entry

Based on the defined geographic/demographic factors and grade level considerations, potential participants whose school districts were situated within a 100-mile radius of The University of Northern Iowa were identified. Building principals representing nine different school districts were contacted via phone. The researcher was able to secure nine principals with the desired geographic/demographic and grade level considerations with only ten phone calls. The principals' names were changed for

purposes of the study. In addition, the specific school districts represented by the principals were not revealed in the study.

Potential participants were informed of the study and the criteria for participation and asked if they would be willing to participate. Principals who agreed to participate were sent a follow-up letter with a complete description of the study, including statements regarding risk and confidentiality. Participants were asked to read the information, then sign and return the informed consent (Appendix C) as required by the University of Northern Iowa's Human Participation Review. A convenient meeting time was then arranged.

The letter of entry also included a demographic information sheet for each participant to complete (Appendix D). The demographic data served as a data source and was considered during the data analysis and interpretation phase of the study. The participants mailed the demographic collection document back to me, along with the consent form, prior to the TA/interview sessions. Three female and six male principals participated in the study. The average number of teaching years for the 9 principals was 11, while the average number of years as a principal was 16. The principals had an average of 8 years of experience in their current administrative position and their average age was 50. There were seven with masters degrees, one with an Ed. S., and one with a Ph.D.

In terms of academic level (Table 1), the elementary principals, on average, were older (57) and had more years of teaching experience(13), more years as a principal (20), and had served more years in their current position (11) than those at the other two levels.

Middle school principals had an average age of 48, had taught for 10 years, been a principal for 17 years, and served 9 years in their current position. The high school principals had an average age of 44, had taught for an average of 9 years, been a principal for 9 years, and been in their current position for only 3 years.

Demographic location information (Table 2) showed that rural principals, on average, had taught for nine years, been a principal for ten years, and served in their current position for seven years. The rural principals averaged 46 years old, and they served an average district population of 815 students. Urban principals had more experience as teachers (11) and as principals (20), but had also served 7 years in their current position. They were an average of 55 years old and served an average district population of 20,000 students. Suburban principals averaged the most years of teaching with 12, had been principals for 17 years, and served in their current position for 9 years. They averaged 49 years old and served an average district population of 12, 834 students.

Table 1

Administrative Demographics by Academic Level

Name (fictitious)	E/M/HS	Gender	YT	YP	YCP	Age	HDE	U/S/R	DSP
Brenda	E	F	10	16	16	60	MS	R	900
Leo	E	M	12	32	10	62	MS	U	30,000
Norma	E	F	17	13	7	53	MS	S	12,000
Ivan	MS	M	9	6	6	37	MS	R	544
Mike	MS	M	10	27	14	55	MS	S	4,000
Rob	MS	M	11	19	8	52	MS	U	19,000
Kathy	HS	F	9	8	1	50	MS	U	10,000
Keith	HS	M	8	8	2	40	Ed.S.	R	1,400
Gavin	HS	M	9	11	6	42	Ph.D.	S	22,500

Note. Abbreviations were used and included: E/M/HS = Elementary/middle school/high school, YT = Years as a teacher, YP = Years as a principal, YCP = Years in current position, DE = Highest degree earned, U/S/R = Urban/Suburban/Rural, DSP = Total district population.

Table 2

Administrative Demographics by Geographic Location

Name (fictitious)	U/S/R	Gender	YT	YP	YCP	Age	HDE	E/M/HS	DSP
Leo	U	M	12	32	10	62	MS	E	30,000
Rob	U	M	11	19	8	52	MS	MS	19,000
Kathy	U	F	9	8	1	50	MS	HS	10,000
Norma	S	F	17	13	7	53	MS	E	12,000
Mike	S	M	10	27	14	55	MS	MS	4,000
Gavin	S	M	9	11	6	42	Ph.D.	HS	22,500
Brenda	R	F	10	16	16	60	MS	E	900
Ivan	R	M	9	6	6	37	MS	MS	544
Keith	R	M	8	8	2	40	Ed.S.	HS	1,400

Note. Abbreviations were used and included: U/S/R = Urban/Suburban/Rural, YT = Years as a teacher, YP = Years as a principal, YCP = Years in current position, DE = Highest degree earned, E/M/HS = Elementary/middle school/high school, DSP = Total district population.

Data Collection

Creswell (1994) indicates that qualitative research may incorporate four basic types of data collection: (a) observation, (b) interview, (c) documents, and (d) visual images. Data for this study was collected via combination think-aloud/interview sessions. In addition, the researcher made observation notes and collected available evaluation tools that the principal may have used. Principals were audiotaped as they thought aloud during the review of a second-year teacher portfolio.

The TA/interview episodes served as the main methods of data collection for this study. The intent of the research was to establish the thinking of the principals, not to judge the documents they used. However, documents and observation served as valuable secondary data sources. Bogdan and Biklen (2003) acknowledge that documents can be considered as supplemental information to observation and interviewing.

Documents clearly fit the criteria of using data rich in description but to what extent the researcher uses them in a manner that is naturalistic, inductive, and concerned with the process of meaning construction for those who produce them or use them has to be examined in each case. (p. 58)

Principals were not formally asked to bring any documents they used (self-generated or district provided) as they evaluated portfolios. However, as the TA/interview progressed and the principal made reference to tools he/she used for evaluation purposes, the researcher made notes about the tool and asked the principal for copies following the TA/interview session. Asking the principal to bring documents to the interview might have inferred to the principal that they *should* be using some sort of document. Consequently, they may have created something specifically for the

TA/interview session that they did not otherwise use, thus contaminating the naturalistic atmosphere of the environment.

In addition to the audiotaped sessions and the collected documents, the researcher made observational field notes throughout the TA/interview episodes. The notes were analyzed along with the two other data sources.

Portfolios

It was necessary to acquire three sample portfolios for use in the study; one from an elementary teacher, one from a middle school teacher, and one from a high school teacher so that principals from each academic level (elementary, middle, and high school) could review the same portfolio. The sample portfolios used in the study were actual (“live”) portfolios submitted by second-year teachers to principals of schools other than those participating in the study. Non-participating area principals were contacted via phone and asked to recommend teachers that they thought might be willing to allow the use of their portfolio for purposes of the study. The researcher received the names of fourteen teachers from area principals.

The recommended teachers were contacted via phone and advised about the study and how their portfolio would be used. They were advised that any reference to them, their school, their community, or specific students would be blacked out prior to use in the study. Ten teachers agreed to have their portfolios evaluated for use in the study; four elementary teachers, three middle school teachers, and three high school teachers. The participating teachers were sent letters describing the study, including confidentiality and risk statements (Appendix E). The researcher teamed with individual teachers who agreed

to participate and together the teacher and the researcher marked up the portfolios for confidentiality purposes. To further ensure confidentiality, media items, whether produced by the teacher or by his/her students, such as PowerPoint, streaming video, taped audio, and pictures of the teacher were not included in the portfolio. A table of contents was included indicating to the reviewing principals that the second-year teacher did indeed use this type of artifact.

The recommending principal's role was as a referring agent only. The relationship concerning consent existed exclusively between the teachers who elected to participate and the researcher. Teachers felt no pressure to participate from their recommending principals and their right to refuse could have been indicated by their lack of interest in participating in the study, or ultimately by them not signing the informed consent as prescribed in the University of Northern Iowa's Human Participation Review.

Once the pool of portfolios was obtained, two former principals who were current faculty members in the Department of Educational Leadership at the University of Northern Iowa, had completed the IEATP, and had experience in portfolio review, examined the portfolios to ensure their usability for the study. It was important that the sample portfolios be at neither extreme, unsatisfactory nor exemplar, if they were to produce usable data. Either extreme would have limited the amount of TA data that might be produced by a participating administrator.

Each sample portfolio was coded with an identifying number known only to the researcher for confidentiality purposes. The expert reviewers used the holistic approach described by Glatthorn (1996); a general impression of overall quality. The panel used a

rubric (Appendix G) that was developed by the researcher as a means of evaluating the pool of portfolios for purposes of the current study. The rubric was based on Glatthorn's holistic approach.

Due to the cross section of portfolios that were contributed from different local districts, the rubric was understandably general to ensure that the portfolios met basic criteria specific to the eight ITS. Portions of Burke's (1997) portfolio rubric (p. 129) were adapted for use in development of the rubric provided to the panel. However, because of the holistic approach, less emphasis was placed on the evaluation of each artifact and more emphasis on the overall quality of the portfolio.

Each reviewer independently reviewed each portfolio using the rubric and gave the portfolios an overall rating of unsatisfactory, proficient, or exemplar. Based on a suggestion from Glatthorn's (1996) work the researcher had the panel conduct a second review to ensure that the first judgment was valid. All ten portfolios were deemed proficient in both reviews; hence, a drawing was necessary to choose three portfolios (one from each academic level) to be used in the study.

A fifth grade teacher who taught content in all areas produced the elementary portfolio that was selected. An eighth grade teacher who taught math and one section of technology produced the middle school portfolio that was drawn. A tenth grade teacher who taught social studies produced the high school portfolio that was selected. The participating teachers and their recommending principals were not advised if their portfolio was chosen for use in the study.

The fifth grade teacher and the high school teacher chose to organize their portfolios by artifact. Each artifact represented one or more of the eight ITS. See Appendix H for cross-referencing information used by the elementary teacher. See Appendix I for the table of contents and cross-referencing information used by the high school teacher. The middle school teacher organized her portfolio by standard. She included sections representing each of the eight standards. Each section contained artifacts representing a respective standard. She included sections representing each of the eight standards. Each section contained artifacts representing a respective standard. It was not necessary that she use a cross-reference guide because the middle school teacher included, on an artifact cover page, the multiple standards/criteria represented by each artifact. See Appendix J for a sample of an artifact cover page.

Think-aloud/Interview Process

The motive for using TA protocol in this research was to establish what thoughts led to judgments concerning second-year teacher portfolios. Pressley and Afflerbach (1995) say that, "Think-aloud data should reflect exactly what is being thought" (p. 9). The authors contended that it is not the role of the participant to categorize his or her cognitions as they verbalize their thinking and that the directions given to participants should make this clear.

Instructions and probes were designed to elicit verbal reports from participants while they were actually performing the task of reviewing a second-year teacher portfolio. Ericsson and Simon (1993) propose that "...whenever possible, concurrent verbal reports should be collected, so that processing and verbal report would coincide in

time” (p. xiii). Rather than having participants perform a review and then report via a series of interview questions, principals were asked to virtually “think aloud” as they reviewed the sample portfolio. They were asked to verbalize everything they were thinking from the time they first saw the portfolio until they felt they had rendered a complete review. This ensured that participants were retrieving thinking that coincided with the sample portfolio they were currently reviewing. Refer to Appendix F for instructions that were read to the participants.

The researcher conducted the individual TA/interview sessions. The session began with introductions and an opportunity to build rapport. The researcher briefly restated the purpose of the research and reassured the participant that the TA/interview session and the forthcoming transcript were confidential. All nine of the TA/interview sessions were audio recorded. The sessions ranged in length from a minimum of two hours to almost three hours. The resulting transcripts ranged in length from 22 pages to 54 pages.

As the audiotaping began, the researcher read the TA instructions (Appendix F) to the participant and then handed him/her the sample portfolio. The researcher immediately began to take observational notes. The verbal role of the researcher during the session was very limited and occurred only to encourage the participant to keep talking or to ask the participant to provide clarification of terms. When the principal deemed the portfolio review complete, the researcher asked four guided interview questions (Appendix K). The questions were open-ended and for purposes of clarification and probing. The questions in Appendix K served as a preliminary guide. Additional questions surfaced

during the sessions and varied with each participant. Participants were also encouraged to ask questions during this time. Each TA/interview session was transcribed from the audiotapes. The transcriptions were returned to each participant for review and clarification.

Think-aloud/Interview Pilot

In preparation for the study, a pilot TA session was conducted with a middle school principal from the surrounding area who was not a participant in the study. The participant was included in the Human Subjects Participation Review and signed a consent letter. One of the sample portfolios was used for the pilot session. The pilot session was videotaped. In an effort to perfect the TA technique, the researcher reviewed the tape two times. For the first review, a member of the researcher's dissertation committee who is endorsed in reading and is familiar with TA protocol participated. For the second review an outside colleague who is also reading endorsed and familiar with TA protocol joined the researcher. The data from the pilot TA session was not transcribed or included in the reported results.

Valuable information was gleaned from the pilot. The reviewers and the researcher noted that the pilot principal had difficulty with the structure of the portfolio. While he was familiar with the content (math) and the grade level (middle school), he struggled with how the content of the sample portfolio was organized; it simply was not how he coached his teachers to construct a portfolio for his review. This proved to be a barrier and initially affected his ability to judge the content of the portfolio. As he settled in to the structure and began to use the table of contents and related organizational

materials provided by the teacher, he began to verbalize his thoughts about the artifacts and his judgment thereof.

Upon the recommendations of the two reviewers and information gathered from further discussion with the pilot principal, the TA instructions were revised to include a brief description of how the portfolio was structured (Appendix F).

Data Analysis

Bogdan and Biklen (2003) describe data analysis in qualitative research as, “the process of systematically searching and arranging the interview transcripts, field notes, and other materials that you accumulate to enable you to come up with findings” (p. 5). They suggest that data analysis is the process of incorporating a system that arranges information from various data sources in a way that facilitates the development of findings. They describe interpretation as “... explaining and framing your ideas in relation to theory, other scholarship, and action, as well as showing why your findings are important and making them understandable” (p. 147). Separating data analysis and data interpretation is difficult when conducting qualitative research (Bodgan & Biklen, 2003; Gay & Airasian, 2003). Gay and Airasian say that the intertwined nature of data analysis and data interpretation is an important aspect of qualitative research.

Tesch, as cited in Creswell (1994), says there is no “right way” to analyze qualitative data; that the process is eclectic. Creswell goes on to say that the researcher must “...be comfortable with developing categories and making comparisons and contrasts” (p. 153).

Ericsson and Simon (1993) describe two different processes specific to coding TA. In one method, the researcher only categorizes/codes speech signals and is not concerned with meaning. In the second method, the researcher is concerned with meaning, but the presence of an existing theory used by the researcher "...limits the coding to selected aspects and features rather than the full meaning of the verbalization" (p. 6). The presence of predefined coding schemes and/or theory limit the search for meaning that is unique to qualitative research as described for purposes of this study. Ericsson and Simon point out that a need for a less formal kind of analysis does exist. "...the encoding scheme is not defined formally a priori, but the search for interpretations proceeds in parallel with the search for an appropriate model or theory" (p. 6). With this concept in mind, the following plan was designed for the analysis of data in this study.

A synchronized approach known as constant comparative, developed by Glaser and Strauss (1967) was used. Data collection, data analysis, and data interpretation were simultaneous, ongoing activities. The constant comparative strategy is inductive in nature and consistent with the intent of qualitative research. The strategy is "...devised to assist in generating social theory" (LeCompte & Preissle, 1993). LeCompte and Preissle's description of the constant comparative method provides clarity,

...as social phenomena are recorded and classified, they also are compared across categories. Thus, the discovery of relationships, or hypothesis generation, begins with the analysis of initial observations, undergoes continuous refinement throughout the data collection and analysis process, and continuously feeds back into the process of category coding. As events are constantly compared with previous events, new typological dimensions as well as new relationships may be discovered (p. 256).

Boeije (2002) suggests a step-by-step approach relative to the constant comparative method. She says that the steps that she proposes will help remove the vagueness of providing an account of the analysis. "It is the lack of explication and account that reduces verification and therefore the credibility of qualitative reports" (p. 392).

Boeije (2002) references a study of couples coping with the effects of Multiple Sclerosis. Twenty married couples participated in the study and both partners were interviewed. Five analytical steps emerged from the study. The author is quick to point out that "...the number of steps as such is not important, because that depends on the kind of material that is involved" (p. 395). Boeije suggested the following analytical steps when using the constant comparative model:

1. Comparison within a single interview.
2. Comparison between interviews within the same group.
3. Comparison of interviews from different groups.
4. Comparison of pairs at the level of the couple.
5. Comparing couples.

Boeije's (2002) steps were adapted to more accurately fit the content of the current study and to coincide with the way the TA/interview episodes were scheduled. The adapted steps used in the current study are as follows:

1. Comparison within the first TA/interview session (an elementary principal).

Repetitive, key words and phrases were identified and isolated.

2. Comparison between the first TA/interview preliminary categories and each ensuing TA/interview with previous TA/interviews and coding categories. Coding categories were revised and updated as each transcript was read.

3. Comparison within the academic levels.

4. Comparison between academic levels.

5. Comparison between same geographic/demographically defined schools.

6. Comparison across geographic/demographically defined schools.

The objective of the first step was to develop initial coding categories and summarize the core message of the interview for comparison with the forthcoming transcripts. As soon as the first transcript was received, the researcher read it as she listened to the audiotape of the TA/interview. Satisfied that the transcription was accurate, the researcher began the work of coding.

A valuable tool in the first step was the use of Ryan and Bernard's (2003) technique of developing a list of key-words-in-context (KWIC). In this process the researcher creates a concordance "...by finding all the places in a text where a particular word or phrase appears and printing it out in the context of some number of words (say 30) before and after it." (p. 269).

As the researcher read the first transcript again, those words/phrases that seemed to recur frequently were arranged on a list for use in the first coding attempt. Key words/phrases that emerged from the first think-aloud were: portfolio structure (structure, format, lay out, laid out), process steps (first, next, finally), reflection, tools of evaluation (form, guide, checklist, model), judgment of the artifact (evidence, proof, good, bad,

lacking), recommended alternatives for an included artifact (rather than, instead, suggest, judgment of the teacher (recommend, licensure, this teacher), and concerns about the DOE evaluation model (State, DOE). The researcher established a KWIC worksheet with the words and phrases for each developing category (Appendix M). The researcher used the category worksheets to work through each transcript to ensure that each transcript was checked for all key words. As coding progressed, new categories were established, others were eliminated, while some were combined. Each time coding adjustments were made, the researcher reviewed each transcript to recode as necessary.

Using the “find” function of the word processor, the key word for each category was entered. As each occurrence appeared, the context surrounding the word was reviewed and a decision made concerning the appropriate code. Using the “copy/paste” function of the word processor, the text surrounding the key word was moved to a coding *holding page*. This process created a way to tease out and separate categories that materialized during the think-aloud. Boeije (2002) refers to this process as fragmenting. She says that fragmenting “...emphasizes the separate themes [categories] which emerge during the interview and focuses on an individual ordering process which is relevant to the research questions” (p. 394).

Creswell (1994) refers to this reviewing and sorting process as data reduction. “The researcher takes a voluminous amount of information and reduces it to certain patterns or categories, and then interprets this information by using some schema” (p. 154). The words, phrases, patterns of behavior, and subjects’ ways of thinking that were repeated were coded into categories, thus creating a coding system to separate topics

from one another (Bogdan & Biklen, 2003). However, due to the process of constant comparison, the categories had to remain flexible holding areas that were constantly updated to reflect the emerging nature of the collected information. Gay and Airasian (2003) note that this process allows the researcher to focus the study during the data collection phase. They add that developing a focus during data collection allows for updates to ensuing collection sessions leading to greater depth of data.

Once the researcher established the earlier described preliminary coding categories, Boeije's (2002) second step was incorporated. The researcher continually compared the coding categories to the most recent transcript that had been received; re-categorizing and establishing new categories as she read and reread each transcript. The researcher began to develop preliminary and primitive definitions of the coding categories. After refining the categories and definitions, the researcher took them to her committee co-chair who is an expert in TA and the coding thereof. The co-chair was provided with the preliminary definitions of the categories along with specific instructions for coding (Appendix N). The co-chair was asked to practice coding an excerpt that included five comments from one of the transcripts so that the researcher could ensure that he understood the instructions that had been provided. After the practice session, the co-chair coded a lengthier excerpt (16 comments) as the researcher kept track of the number of times he agreed with the researcher's coding and the times that he "missed." Two of the comments were thrown out because the researcher provided too much information to the co-chair; thus, making his coding decision biased. Of the 14 remaining comments, the co-chair's coding agreed with the researcher's coding on nine

of comments and he “missed” on five of them for an intercoder agreement rate of only 64 percent. Krippendorff (1980) used 80 percent intercoder agreement as a benchmark of acceptable reliability. Thus, an intercoder reliability of 80 percent was adopted as the acceptable reliability measure for the study. It was apparent that revisions to the definitions were necessary to achieve the intercoder reliability of 80 percent or higher.

Based on discussion with the co-chair, the researcher revised the category definitions (Appendix O) and restructured the coding instructions (Appendix P). The researcher recoded the data a second time using the revised definitions, and isolated the comments from the parts of the transcript that were not coded. See Table 3 for samples of emergent categories with definitions and example statements that illustrate the category.

The researcher asked two independent readers to code excerpts. One of the independent coders was a former principal who had IEATP training and was a current chair of a university teacher education department. She had her doctorate. The other independent coder had just completed her master’s program in instructional classroom leadership with a reading endorsement. The researcher provided both coders with the revised category definitions and the same instructions provided to the co-chair in the first reliability check. Each independent coder was asked to code five comments for practice and then coded 15 comments. The first coder correctly identified 14 of the 15 comments. The second coder correctly identified 15 of the 15 comments for a combined 96% coding reliability; acceptable for a reliability check for this study.

Table 3

Emergent Categories

Three sample* categories with definitions and examples of statements that illustrate each

1. Portfolio Structure (PFS) – Comments in this category made reference to how the teacher physically structured and organized his/her portfolio. Typically, the principal referred to specific pages that provided structure within the portfolio they were viewing as well as the overall structure/organization of the portfolio. KWIC used: structure, organize, lay or laid out, figure out. A comment representative of this category might begin “In our district we organize portfolios...” or “It looks like this has been organized by artifact...”

a. Brenda said, “First of all, I like to see all the structure here as far as how they organize their artifacts. So they’ve taken an artifact and then they’ve identified standards and criteria that falls under this. Overall, I look at areas of standards.”

b. Rob said, “I can see that she’s starting to be descriptive of the artifacts. I guess I like the format that’s being used. I see that one thing I like in regards to this is that it simplifies it in terms of the administrator, already telling me to kind of focus in on 1A and 1B.”

2. Process steps (PS) – Comments in this category refer to how the principal progresses through the portfolio; the steps he/she takes. KWIC used: first, next, second, last, finally, always, and usually. Comments indicative of this category include “The first thing I like to do is...” or “My next step is to...”

a. Mike said, “So first of all, as I’m sitting down thinking about this teacher, I’m thinking not only about what this will show me, but I’m gonna be thinking about what I’ve seen.”

b. Kathy said, “First off I would review. This is a high school social studies teacher so I probably would do some time thinking about our current social studies department and what the needs are, the people in that department and what the overall departmental goals might be so that when I’m looking at this, I’m looking at it in context of not just the development of this person in his content area, but ...in our school there are 12 social studies teachers.”

(table continues)

3. Principal's role (PR) – Comments in this category refer to how the principal perceives his/her role in the mentoring and induction of the teacher with emphasis on portfolio preparation. KWIC used: role, job. An example of a comment in this category; ...”my role is to make sure by the end of year two that they’ve done what it is...”

a. Kathy said, “So again, I would just constantly reflect on my own role in making sure – if this is a young talented teacher, my role is to make sure by the end of year two that they’ve done what it [portfolio construction] is. If I haven’t done my part, that’s not his fault.”

b. Gavin said, “If I’m doing my job I know this teacher inside and out before their two years are up.”

* See Appendix O to view all categories with associated definitions and examples.

Each transcript was coded using the “comment” function of the word processor. As each transcript was read, the researcher highlighted the comment or comments that related to a specific category and then labeled the highlighted category with a comment in the right margin (see Appendix L for example of coded page). Highlighting was used so that the researcher knew she had read and made a decision about a comment or group of comments. Additionally, the researcher continued to copy/paste the highlighted items into category holding pages (see Appendix Q for category holding page sample) so that the data could be more easily managed during the comparison phases of the study. Ten comment categories resulted from the transcriptions. As a way to report the data, the researcher conducted a comment count for each category. To ensure accuracy, the researcher’s copy editor conducted a comment count as well. The results of the comment counts are reported and illustrated in Chapter 4.

Reliability

Reliability is typically associated with replication (Bogdan & Biklen, 2003; Creswell, 1994; LeCompte & Preissle, 1993). “It [reliability] assumes that a researcher using the same methods can obtain the same results as those of a prior study” (LeCompte & Preissle, 1993, p. 332). Creswell points out that, “...the uniqueness of a [qualitative] study within a specific context mitigates against replicating it exactly in another context” (p. 159). Bogdan and Biklen say that, “In qualitative studies, researchers are concerned with the accuracy and comprehensiveness of their data. Qualitative researchers tend to view reliability as a fit between what they record as data and what actually occurs in the setting under study, rather than the literal consistency across different observations” (p. 36).

LeCompte and Preissle (1993) contend that the reliability of a qualitative study can be strengthened with the use of tape recorders and by providing, specifically and precisely, the procedures used in the study. The research design, the participants, and the data collection/analysis procedures, including the use of tape recorders, were clearly outlined in the current study. The use of two independent coders, as described in the Data Analysis section of this chapter, greatly enhanced the reliability of the study. The steps in the study were logically designed to answer the research questions and could be easily repeated in a follow-up study with the understanding that, “Qualitative research is a personal endeavor; no investigator does research just like another” (LeCompte & Preissle, 1993, p. 341).

In addition, the researcher asked each of the participating principals if they, in fact, conducted portfolio reviews of their own teachers in the same fashion that they had conducted the review of the sample portfolio. Eight of the nine principals indicated that the review they conducted for purposes of the study mimicked the way they reviewed portfolios from teachers in their own buildings. One principal indicated that the process he used was similar but that since his teachers produced portfolios electronically, he used the electronic template developed by his district during evaluation. The fact that the principals in the study conducted the reviews of the sample portfolios in the same fashion as they conducted reviews of portfolios developed by their own teachers solidified the fit between the recorded data and what actually occurred in the setting under study (Bogdan & Biklen, 2003).

Validity

Merriam (1998) describes internal validity as an accuracy measure; one that insures that information matches reality. Creswell (1994) suggests that member checks be used to strengthen internal validity; that is, the information (data) gathered from the participants is taken back to them for verification.

Procedures in the current study provided for the transcriptions of the TA/follow-up interview sessions to be reviewed by the participants. The researcher, per follow-up discussion, made clarifications and revisions to the transcribed text with each participant. In addition, the participants were contacted for accuracy checks as the data was analyzed and coded. LeCompte and Preissle (1993) believe that through this type of collaborative

effort, a shared, understood meaning between the researcher and each participant becomes apparent (LeCompte & Preissle, 1993).

Creswell (1994) discusses external validity in terms of generalizability. Typically, generalizability refers to "... whether the findings of a particular study hold up beyond the specific research subjects and the setting involved" (Bogdan & Biklen, 2003, p. 32). Generalizability seems to be somewhat of a gray area in qualitative research. Gall, Gall, and Borg (2003) say that claims to generalize knowledge are typically not associated with the realm of qualitative research. However, the authors do assert that efforts to randomize the sample within a given group can increase the possibility of generalizability.

The sample for this study was selected using the maximum variation sampling strategy. This type of purposeful sampling was used to provide some variation to the sample but did not allow for a great deal of randomization.

The intent of this study was not to report results that insinuate application to all principals. The researcher's commitment was to provide an accurate account of the participants' explanations of the principals' thinking and their reviewing processes as related to the evaluation of second-year teacher portfolios. Generalizing the results beyond this particular study is in and of itself another study at minimum. Perhaps Bogdan and Biklen (2003) summarize it best, "...some qualitative researchers approach generalizability [by thinking] that if they carefully document a given setting or group of subjects, it is then someone else's job to see how it fits into the general scheme of things" (p. 33).

The Role of the Researcher

The researcher was the primary instrument of data collection. LeCompte and Preissle (1993) points out that "...the researcher's identity and experience [is] critical to the scientific merit of the study" (p. 92). As such, it was necessary to acknowledge the experiences of the researcher and any potential bias. The researcher was a former high school teacher who believed strongly in the K-12 school experience. The teacher quality program in Iowa was an effort designed to enhance the quality of schools and the evaluation process was part of the initiative. Therefore, the researcher considered the commitment of the primary evaluator, the principal, to be a vital piece of beginning teacher induction and thus, the quality of education provided by all teachers.

The researcher's primary focus in the current study was to add to knowledge, not to pass judgment. Bogdan and Biklen (2003) contend that, "The worth of a study is the degree to which it generates theory, description or understanding" (p. 33). The goal of the study was not to label the data collected as good or bad; it was merely to report the findings in a reflective, conscientious, and organized way.

The role of the researcher as data interpreter also gives rise to concerns regarding subjectivity in qualitative research (Bogdan & Biklen, 2003). However, Bogdan and Biklen point out that the method used by researchers to interpret data is, in and of itself, an aid to increasing the level of acceptable subjectivity.

The researcher spends considerable time in the empirical world laboriously collecting and reviewing piles of data. The data must bear the weight of any interpretation, so the researcher must constantly confront his or her own opinions and prejudices with the data. Besides, most opinions and prejudices are rather superficial. The data that are collected provide a much more detailed rendering of

events than even the most creatively prejudiced mind might have imagined prior to the study (p. 33).

The use of Boeije's (2002) constant comparative methodology was indicative of the researcher's intent to preserve the integrity of conscientious interpretation of data as described by Bodgan and Biklen.

Limitations of the Study

Participants in the study evaluated a portfolio somewhat "out of context." Typically, the principal knows the teacher (author of the portfolio), has observed him/her teach, and has had several meetings to discuss the developing portfolio. Concerns existed that such circumstances might limit the principals' ability to accurately evaluate and judge the portfolio.

The pilot study proved to be a valuable means of determining a potential barrier. The pilot principal did indeed struggle with the structure of the portfolio; however, the adjustments the researcher made to the instructions given to the principals bridged the gap effectively. In addition, the researcher tracked and categorized those comments that were made by principals indicative of their comfort level with the structure. The potential limitation of the study proved to be minor.

Summary

The purpose of the study was to examine the thinking processes, methods, and materials used by principals as they made evaluative judgments regarding second-year teacher portfolios. The overarching question of the study was, "How do principals evaluate second-year teacher portfolios?" To that end, three sources of data were

collected from nine principals representing differing academic levels and geographic/demographic areas:

1. Think-aloud information as they evaluated a “live” portfolio.
2. Interview question responses after the think-aloud using four core questions.
3. Tools that the principal brought along to the TA/interview session.

The data gathered during this qualitative study was coded using the constant comparative method of analysis. Two independent coders were used to increase the reliability of the analysis effort. Contrasting and comparing the results of the coding produced the findings in the study.

CHAPTER 4

RESULTS

The purpose of this study was to examine the thinking of administrators as they reviewed and made evaluative judgment of a second year teacher portfolio. Qualitative data was collected using a think-aloud (TA) process and four open-ended questions. Results reported in this chapter include information relative to the three research questions that guided the study:

1. What do principals verbally report they are thinking as they review second-year teacher portfolios for purposes of evaluative judgment?
2. What tools have principals developed to assist them in evaluating second-year teacher portfolios?
3. How much bearing does portfolio evaluation have on the judgment the administrator makes regarding licensure?

The findings indicated that (a) the participating principals operated within a similar “thinking framework” as they evaluated the portfolio provided to them by the researcher, (b) the participating principals were able to successfully evaluate a foreign/sample portfolio, (c) the participating principals attended most to judging and coaching activities as they reviewed the portfolio, (d) the participating principals established a similar four-step rhythm when judging individual artifacts, (e) the participating principals varied in terms of judgment pattern across artifacts, (f) the participating principals placed a higher value on observation than on the portfolio, (g) the participating principals rated their respective sample portfolio as proficient, (h) principals

used similar tools to evaluate portfolios, and (i) portfolios accounted for roughly 30% of licensure decisions.

The three research questions and associated findings will be addressed in Chapter 4. Each research question will be attended to individually and represented in separate sections of the chapter. Findings for each research question will be structured with regard to academic level (elementary, middle school, and high school) as well as geographic/demographic region (urban, suburban, and rural). Figures and tables will be used to more clearly illustrate the data within each section. In addition, quotes from the participants will be used to support the illustrated data.

Research Question 1

What do principals verbally report they are thinking as they review second-year teacher portfolios for purposes of evaluative judgment?

Findings for research question one indicated that (a) the participating principals operated within a similar “thinking framework” as they evaluated the portfolio provided to them by the researcher, (b) the participating principals were able to successfully evaluate a foreign/sample portfolio, (c) the participating principals attended most to judging and coaching activities as they reviewed the portfolio, (d) the participating principals established a similar four-step rhythm when judging individual artifacts, (e) the participating principals varied across academic level in terms of judgment pattern across artifacts, (f) the participating principals placed a higher value on observation than on the portfolio, and (g) the participating principals rated their respective sample portfolio as proficient.

The findings for question one resulted from the accumulated comment counts that were produced from the TA/interviews. From the comment counts, ten categories emerged. They were: (a) coaching (C), (b) comfort level (CL), (c) critical pieces (CP), (d) judgment (J), (e) portfolio structure (PFS), (f) principal's opinion (PO), (g) principal's role (PR), (h) process steps (PS), (i) tools (T), and (j) time invested (TI). The ten coded categories clustered around four broader categories that provided a means of efficiently reporting data within each academic level and geographic/demographic region. The four broad categories were (a) processing activities, (b) judging activities, (c) coaching activities, and (d) critical pieces.

The findings for research question one will be reported in the following format: First, the *overall* comment counts for the entire study, by category, will be reported and illustrated per academic level and geographic/demographic region. The overall comment counts will serve as a continual reference for reporting the findings in Chapter 4. Second, the three phases of thinking that emerged from the TA/interviews will be illustrated and explained. Finally, the results within each thinking phase will be reported; first per academic level and then per geographic/demographic region.

Overall Comment Counts per Academic Level and Geographic/Demographic Region

Figure 1 illustrates the *total* verbal comment counts for each of the ten categories by academic level. Middle school principals had the highest overall coded verbal comment counts with 45% of the total coded comments. High school principals had the second highest overall verbal comment count with 31%. Elementary principals had the fewest overall verbal comments with 23 % of the total coded comments.

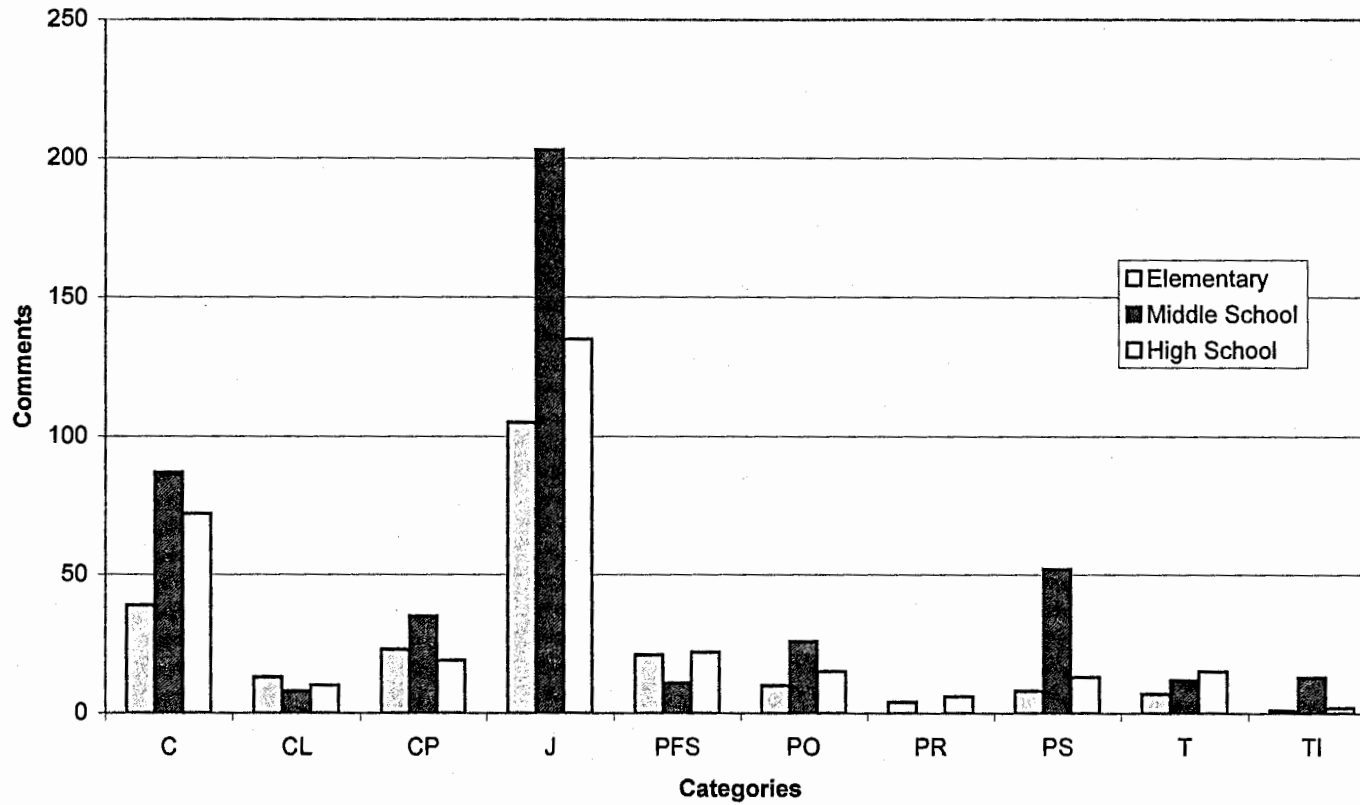


Figure 1. Total Comment Counts by Academic Level Per Category

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS = Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

Two categories represented 65% of the coded verbal comments; judgment (J) and coaching (C). The multi-level judgment category that included comments relative to judgment of artifacts, portfolio, and teacher accounted for 45 % of the total comment counts. The multi-level coaching category that included coaching suggestions regarding portfolio preparation, meetings between the principal and the teacher, the use of questioning by the principal, and suggested alternatives accounted for 20% of the total coded comments. The remainder of the categories ranged between eight percent and one percent of the total comments coded for the study.

Figure 2 illustrates the total verbal comment counts for each of the ten categories per geographic/demographic region. It is important to be reminded that while principals in each academic level reviewed the same portfolio, the geographic/demographic region verbal comment counts were representative of comments made by the same nine principals *across* academic level.

The total number of verbal comment counts per geographic/demographic region was remarkably similar. Suburban principals had the highest overall percentage of coded verbal comments with 34%. The rural and urban principals each accounted for 33% of the total verbal comment counts.

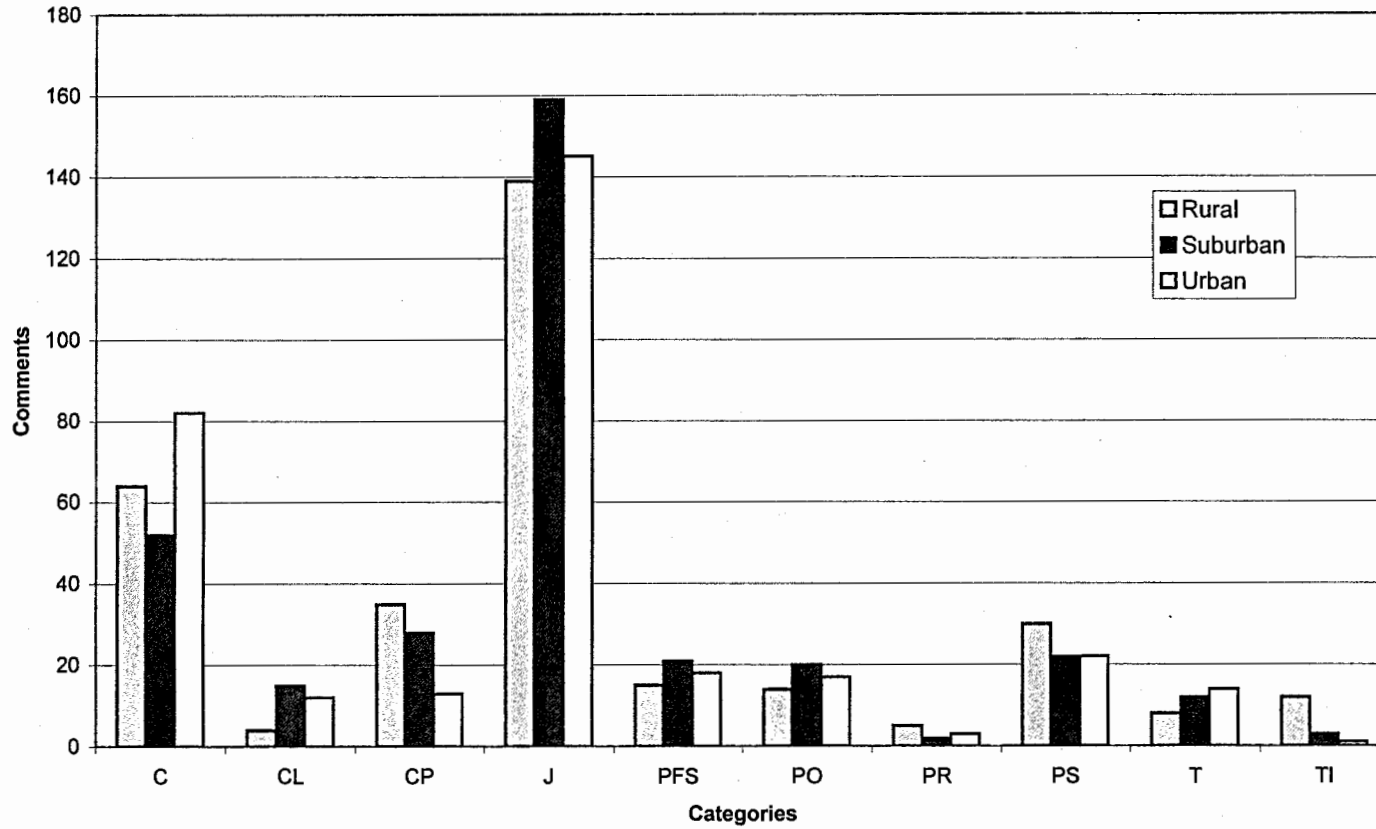


Figure 2. Total Comment Counts by Geographic Demographic Region Per Category

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS = Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

Three “Phases” of Thinking

While the coded comments provided the core of the data for answering research question one, it was the observational information that provided an entry point for the analysis. As the categories unfolded and the tapes and observational notes were revisited, a pattern of principals’ thinking developed. Figure 3 illustrates the three thinking phases that emerged. This is not to say that the thinking of the principals was completely linear and that the coded comment categories fell nicely into one phase or another. The comment categories permeated each phase.

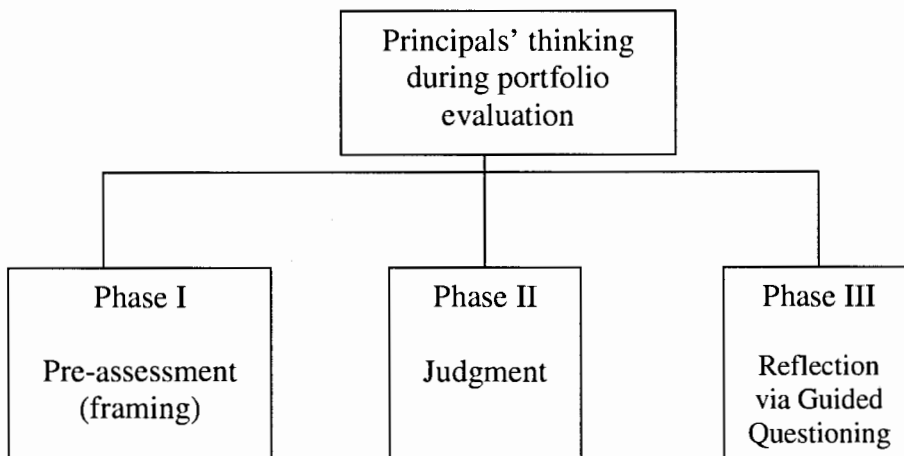


Figure 3. The thinking phases of principals.

Pre-assessment thinking (Phase I) refers to that period of time when principals prepared to judge the artifacts and ultimately, the portfolio. In short, principals *framed* their work in the pre-assessment thinking phase. The verbal comments included in the

pre-assessment thinking phase were those comments that were made prior to the first verbal comment made in the judgment phase.

Judgment thinking (Phase II) refers to that period of time after pre-assessment thinking when principals actively judged the artifacts in the portfolio that was provided to them by the researcher. The verbal comments included in Phase II (judgment) thinking were those comments that began with transition statements followed by coded comments relative to the first artifact in the portfolio. An example of a comment indicative that the principal had entered the judgment phase was made by Mike, a middle school principal, when he said "So having those in front of me [the eight ITS] I kind of go into the standard then." The comment by Mike evidenced his transition from Phase I to Phase II. After the transition statement, he began to judge the first artifact by saying, "Okay, so this teacher is showing how this particular activity does in fact impact student achievement. They're able to apply this learning into student achievement. This teacher has used a couple digital pictures to exhibit that. That one is a direct hit on a couple of the criteria as a part of standard one."

Phase II ended when it became obvious that a principal was finished with the review of the final artifact. There were two indicators that signaled when the principal had completed the review of the portfolio and was moving towards the reflective stage. One of the indicators was observation and the other was verbal. In some cases, both indicators were present.

From an observational standpoint, it became obvious when a principal completed the review because he/she physically reached the end of the portfolio. Other physical

indicators that a principal had completed, or was nearly completed with, the evaluation of the portfolio were evident when the principal moved slightly away from the desk at which he/she was seated or, in some cases, closed the portfolio. In other instances a signal that the principal was complete, or nearly complete, with the judgment phase occurred when he/she looked up and directly at the researcher rather than at the portfolio.

Verbal transition occurred when the principal was reviewing the last artifact and focused away from the artifact and towards judgment of the portfolio or the teacher. For example, Leo was transitioning when, after making a final comment on the final artifact of the elementary portfolio, he said “So if I’m thinking about this whole portfolio ...there are pieces to this portfolio that I like better than others. I think I like the way that it’s organized. She had ten artifacts and then drew out the connections of those artifacts made with each one of the standards.” At the same time he made this comment, he leaned back in his chair and put his hands behind his head.

If either observation or verbal transition indicators were present the researcher asked the principal if they deemed the review complete. The comment following the question/answer in all cases became the first comment of Phase II, thus ending the judgment phase (Phase II). Phase III began as soon as a principal deemed the review of the portfolio complete. Phase III consisted of comments made by the principal in response to the guided interview questions posed by the researcher as well as other non-solicited comments.

Phase I: Pre-Assessment Thinking per Academic Level

Findings in Phase I will establish that principals spent most of their pre-assessment thinking time in processing activities (portfolio structure, process steps, and comfort level). In this section, findings relative to research question one for Phase I will be reported as follows: First, an overview of the data will report and illustrate the total comment counts across academic level for Phase I. This overview will serve as a reference point for reporting the findings of Phase I. Then, findings for each academic level for Phase I will be reported.

Overview

Table 4 illustrates the comment counts in Phase I of the nine principals by academic level. Phase I comments accounted for eight percent of the total coded comments for the study. Middle school principals spent more time framing than did elementary or high school principals.

All ten of the coded comment categories were represented in Phase I across academic level. Categories common to all three of the academic level groups in pre-assessment thinking were coaching (C), comfort level (CL), portfolio structure (PFS), and process steps (PS).

Principals across academic level thought most about portfolio structure (32%) and process steps (18%) in Phase I. The categories of comfort level (CL) and critical pieces (CP) each accounted for roughly 11% of the comments principals made during Phase I. The categories of coaching (C), judgment (J), principal's opinions (PO), principal's role

(PR), tools (T) and time investment (TI) were all represented in Phase I thinking but at levels of less than 10% of the overall comments for the pre-assessment thinking phase.

Table 4

Summary of Phase I Comment Counts by Academic Level

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Elementary	13	0	2	1	2	7	0	0	1	0	0
Middle School	35	5	4	6	0	7	4	0	8	0	1
High School	23	1	2	0	1	9	2	1	4	3	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PR = Principal’s Role, PS = Process Steps, T = Tools, TI = Time

The ten coded categories clustered around four broader categories that provided a means of efficiently reporting data. The four broad categories were (a) processing activities, (b) judging activities, (c) coaching activities, and (d) critical pieces. Processing activities included the categories of process steps, portfolio structure, and comfort level. Judging activities included the categories of judgment, principal’s opinion, and tools. Coaching activities included the categories of principal’s role, coaching, and time investment. Critical pieces included the categories of teacher reflection and role of observation. Each academic level will be reported via the four broad clusters using data from the appropriate smaller categories.

Elementary Principals' Thinking in Phase I

The elementary portfolio used in the study was from a fifth grade teacher and was organized by artifact. See Appendix H for the cross-referencing information used by the fifth grade teacher. The elementary principals (Brenda, Leo, and Norma) made the fewest pre-assessment comments of the three academic levels. Their thinking accounted for only 24% of the overall comment comments made by all principals in Phase I.

Table 5 illustrates the breakdowns of the total pre-assessment (Phase I) comments made by the elementary principals. Brenda and Leo had the most pre-assessment comments while Norma had only one comment in Phase I.

Table 5

Summary of Phase I Comment Counts made by Elementary Principals

	Total Comments	CL	CP	J	PFS	PS
Brenda	5	1	1	1	2	0
Leo	7	1	0	1	4	1
Norma	1	0	0	0	1	0

CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PS = Process Steps

Elementary Principals and Processing in Phase I

The three coded comment categories that clustered under processing activities included the categories of portfolio structure (PFS), process steps (PS), and comfort level

(CL). Comments relative to processing accounted for 77% of the total pre-assessment comments made by elementary principals. In this section, comment counts relative to each of the three categories (portfolio structure, process steps, and comfort level) will be reported followed by a summary of elementary principals and processing activities.

Portfolio structure. Phase I processing by the elementary principals was dominated by comments regarding portfolio structure (PFS) as the principals focused on familiarizing themselves with a portfolio produced by a teacher they did not know. Comments regarding PFS accounted for 54% of the coded comments for Phase I. Comments made by Brenda, Leo, and Norma were similar in that they reflected efforts by each principal to familiarize themselves with the structure of the portfolio belonging to the fifth grade teacher. For instance, Brenda said, “First of all, I like to see all the structure here as far as how they [the fifth grade teacher] organize their artifacts. So they’ve taken an artifact and then they’ve identified a standard and criteria that falls under this. Overall, I look at areas of standards.” A comment by Leo also reflected a similar effort to become familiar with the structure of the fifth grade teacher’s portfolio. He stated that “So this one has 10 artifacts that obviously refer, going across the grid, to state standards. As I look at this, it’s a pretty good visual way to really see how the teacher is looking at connections. So that page is really pretty helpful.” Norma also looked at portfolio structure as a way of framing her assessment. She said that “I like there’s a chart as a cross-reference. This is a nice way to lay out the cross-reference, with a chart. We’ll see how it plays out.” In all three comments, a visual representation of the structure appeared to be critical.

Leo found that his thinking was assisted by the teacher's way of structuring the portfolio. He stated that "Okay. I just flipped open the book [portfolio] and I'm just looking at the state standards in detail with all the descriptors, so I'm just trying to re-familiarize myself with exactly when she wrote down." He continued his thought about the structure format he recognized when opening the portfolio by saying "So, good intro. In looking at artifact one Observation Writers Workshop, it looks like we have circled all of the criteria; all of the descriptors that she feels connect with the artifacts. Let me pull out number one here and see what she's got." The comments by each of these principals indicated that they were moving away from structure and towards content judgment.

Process steps. The category of process steps (PS) garnered only one comment from the elementary principals in Phase I. Leo provided a glimpse of his personal process steps when he said "I'm going to go back through these [artifacts] and look at them with a little bit more of an eye to detail once I get a sense here of what's been pulled out."

Comfort level. Processing was also affected by comfort level (CL). Brenda and Leo each made comments relative to comfort level. The comfort level comments they made were situated around forms that were present in the portfolio and voice. Brenda commented, "It's taking awhile to see what all these forms are. If I were doing it myself [a portfolio from one of my teachers] I'd be used to these." Leo wanted to make sure he knew what person was writing a particular portion of the portfolio when he said, "I don't know if that's the administrator or the teacher." Norma made no comfort level comments.

In summary, processing includes the categories of portfolio structure, process steps, and comfort level. While the principals spent most of their time focused on

becoming familiar with the structure of the portfolio in Phase I, their comfort level was not adversely affected. Portfolio structure and comfort level were of concern for the researcher because the middle school principal who served as the pilot participant struggled with comfort level as he worked through the structure of the portfolio. This was not the case for the elementary principals in Phase I. Because the focus of the principals in Phase I was on structure, there was little information regarding what process steps principals intended to incorporate in their review.

Elementary Principals and Judging Activities in Phase I

The three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Comments relative to judging activities accounted for 15% of the total pre-assessment comments made by elementary principals. In this section, comment counts relative to each of the three categories (judgment, principal's opinion, and tools) will be reported followed by a summary of the elementary principal and judging activities for Phase I.

As is illustrated by Table 5, elementary principals' comments in Phase I were reserved for framing the task of the portfolio review; hence, only two judgment comments occurred with no comments occurring regarding principal's opinion or tools. Phase I judgment comments made by Brenda and Leo were relative to an initial judgment of the portfolio. Brenda's initial impression was positive as is evidenced by her comment that "First of all, that I feel that they [the teacher] feel a strength in evidencing. It seems to be pretty balanced across the board. Sometimes it's more difficult for them to identify or provide evidence in standards 7 and 8. I see that they, as far as number-wise, have a

sufficient amount.” Leo was somewhat skeptical in his initial judgment of the portfolio but appeared to be open to discovery. He said, “I’m really kind of wondering here is just kind of thinking about again, kind of a cookbook look to this thing as to how does this connect with kids. What I’m seeing here are ideas for implementation but I’m wondering about the connection. Maybe there will be a sense of that as I go through.”

In summary, judging activities did not consume a large amount of the elementary principals’ thinking in Phase I. The judgment comments that were made reflected only judgment about the portfolio, not the teacher or any of the artifacts. The judgment comments were very general and did not suggest a positive or negative judgment decision about the portfolio. Norma made no comments regarding judgment. The elementary principals’ comments did not include thinking with regard to principal’s opinion or tools.

Elementary Principal and Coaching Activities in Phase I

The categories that clustered under coaching activities were coaching (C), principal’s role (PR), and time investment (TI). During Phase I, elementary principals focused on familiarizing themselves with the portfolio with very little time spent on judgment; consequently, there were no comments relative to coaching or time investment. In addition, none of the elementary principals made comments relative to their role.

Elementary Principals and Critical Pieces in Phase I

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two “pieces” carried throughout the study, the category emerged as one of the four cluster

categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. The category of critical pieces consisted of two sub-categories; critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO).

The critical piece category consumed eight percent of the elementary principals' thinking in Phase I. Brenda had one comment concerning the critical nature of teacher reflection. She was clear that she valued teacher reflection when she said "I'm going to now look at cover sheets of artifacts to see if they [the teacher] have reflections because that's what I see as being the most important thing—is what their interpretation of and that's ...umm ...I see [inaudible] artifacts sheet here. The reflection sheet is going to tell me what they got out of that particular piece of evidence."

In summary, only one elementary principal considered the critical piece of teacher reflection in Phase I. The number of comments in this category was few, but the immense value that one principal placed on teacher reflection was clearly significant.

Summary of Elementary Principals' Thinking in Phase I

Phase I (pre-assessment) thinking was described as the period of time when principals prepared to judge the portfolio. Elementary principals spent the bulk of their time in Phase I framing the work of reviewing the portfolio. The dominant thinking for elementary principals in Phase I included comments relative to the physical review of the portfolio; a means to become familiar with how the portfolio was structured. While portfolio structure was very important, it did not adversely affect the comfort level of the principals.

The elementary principals spent very little time judging in Phase I; consequently, there were no comments relative to their role as coach. Only one comment was made relative to critical pieces in Phase I; however, the value that one principal placed on teacher reflection was unmistakable.

Middle School Principals' Thinking in Phase I

The middle school teacher's portfolio was organized by standard. She included sections representing each of the eight standards. Each section contained artifacts representing a respective standard. It was not necessary that she use a cross-reference guide because the middle school teacher included, on an artifact cover page, the multiple standards/criteria represented by each artifact. See Appendix J for a sample of an artifact cover page

Table 6 illustrates the breakdowns of the pre-assessment (Phase I) comments made by the middle school principals. The middle school principals spent more time than their elementary or high school counterparts in Phase I. Their coded comments represented 49% of the total comments made by all principals in Phase I even though Ivan did not engage in Phase I thinking.

Middle School Principals and Processing in Phase I

The three coded comment categories that clustered under processing activities included the categories of portfolio structure (PFS), process steps (PS), and comfort level (CL). Comments relative to processing accounted for 54% of the total pre-assessment comments made by middle school principals. In this section, comment counts relative to each of the three categories (portfolio structure, process steps, and comfort

Table 6

Summary of Phase I Comment Counts made by Middle School Principals

	Total Comments	C	CL	CP	PFS	PO	PS	TI
Ivan	0	0	0	0	0	0	0	0
Mike	25	3	2	4	5	3	7	1
Rob	10	2	2	2	2	1	1	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, PFS – Portfolio Structure, PO = Principal’s Opinion, PS = Process Steps, TI = Time

level) will be reported followed by a summary of elementary principals and processing.

Mike and Rob both spent time in Phase I framing the task of evaluating the middle school portfolio. Ivan, however, did not engage in Phase I thinking.

Portfolio structure. The portfolio structure (PFS) category had the second highest comment count for Mike and Rob in Phase I processing. Mike’s thinking focused on comparing his district’s recommendation for portfolio structure to that of the middle school portfolio he had been provided by the researcher. He indicated that, “...within [our district] we have really given the teachers flexibility in how they organize the portfolio.” He compared the portfolio he had been provided by the researcher to portfolios he was accustomed to reviewing for his district in terms of reflection and explained it in this way.

One thing I can see in looking at this right away ...again, as I think about our process versus this process and structure, is this individual actually wrote a reflection over each and every one of the artifacts, as best I can tell. In our district,

we instead have just one reflective statement over the entire standard that encompasses the various artifacts that are used. So instead of one by one by one, I was looking for this one major reflective statement as an overview.

Mike did not negatively judge the portfolio because it was not consistent with the way his district's teachers structured their portfolios relative to reflective statements.

Rob's thinking concerning portfolio structure was centered on his feelings about organization as well, but in a slightly different sense. He liked the fact that the middle school teacher's portfolio was organized by standard and in sections. "I guess I'm saying right ahead of time that's [organized by standard] a lot better than having artifacts in a paper box or a shoe box and now we're going to sort into the categories, which would make it much more extensive. Nothing wrong with that, but a much more extensive process."

Process steps. The category of process steps (PS) surfaced with the most comment counts for Mike and Rob during Phase I. Mike's thinking revealed four process steps. First, he focused on the teacher, "So first of all, as I'm sitting down thinking about this teacher, I'm thinking not only about what this will show me, but I'm gonna be thinking about what I've seen." Then, he focused on the portfolio, "Okay, the first thing I would do as I receive the portfolio, instead of just focusing on a standard, would be to take a really quick overview just to look at how it's been organized and how it's presented. I'm gonna go cover to cover on it just to see the structure of it." His third process step was focused on himself, "The next thing I would do then is basically remind myself of each of the standards. Just in general, an overview – okay Mike, here's what you're going to be looking for as you start this process." Finally, Mike thought about the individual

standards and was clear that he had to have the eight ITSs in front of him as he prepared to evaluate.

Rob's thinking did not parallel Mike's in terms of process steps. He had only one comment relative to this category in the pre-assessment thinking phase. In essence, he provided a preview of how he commences evaluation, "I open this up and start taking a look at what's in there, which I'll do quickly. Then in the second year teachers that I've evaluated, I get to a point where I can do a lot of writing on any or all of the eight standards."

Comfort level. As was true with the elementary principals, processing was affected by comfort level (CL), but not very much. Comfort level comments represented less than 10% of the total comment counts in Phase I thinking for the middle school principals. The middle school principal who served as the pilot participant struggled with comfort level as he worked through the structure of the portfolio. While Mike and Rob had concerns, they were not detoured; in fact, they both appeared to be challenged by the "cold evaluation" of the portfolio provided by the researcher. Mike said "So to just get this as we are today would not be the norm. It'll make it a more difficult challenge." He framed his thinking in terms of the expectations of his district. He asked out loud, "Is what this teacher about –Does it match with what the district is about? Is the teacher's work consistent with the school district's expectations on student achievement? So, I know what our district goals are. I know in term what our building goals are and I know that what I'm looking for as I sit down with a [his district] instructor is."

Rob also seemed up for the challenge. His comfort level increased as the researcher described the structure of the portfolio to him in the instructions, "I'm feeling a little more comfortable because that's probably the way I would attack, open this up." Just prior to opening the portfolio he shared that "I bet though if this is all I have to go on, I feel like I probably would not do quite as complete a job or I would find that I really need to go see this teacher. But I'm ready to go!"

In summary, processing activities included the categories of portfolio structure, process steps, and comfort level. Ivan did not make any comments in Phase I. Mike and Rob made comments relative to each of the processing categories. In contrast to the elementary principals who spent the biggest share of their time thinking about portfolio structure in Phase I, the middle school principals made the most comments relative to process steps and most of those comments were made by Mike. Mike was very clear regarding the steps he used to process through a portfolio review.

Phase I comments relative to portfolio structure focused on how the portfolio was organized. Mike and Rob were comfortable with the structure of the portfolio but noted it was different than those in their own district. As was true with the elementary principals, the organization of the portfolio was not a barrier to completion of the review.

Middle School Principals and Judging Activities in Phase I

The three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Comments relative to judging activities accounted for 11% of the total pre-assessment comments made by middle school principals. The middle school principals did not make any comments relative to

judgment or tools in Phase I. Comments relative to principal's opinion were the only comments that surfaced relative to judging activities in Phase I. The principal's opinion category had two sub-categories; principals' opinions/ portfolio as evidence of good teaching (PO/PE) and principals' opinion/State Department of Education teacher assessment requirements (PO/DE).

Relative to PO/PE, Mike's thinking was focused on the recognition that the "...the portfolio is not only just a single piece, but maybe not even the most important piece in the evaluation process." He added, "The portfolio piece in and of itself doesn't necessarily show the greatness of the teacher. Sometimes that's tied to the amount of time they commit to it."

Rob's Phase I thinking regarding the portfolio as evidence of good teaching also focused on how a portfolio fit into teacher evaluation. His thinking indicated that the portfolio has a place in providing evidence of good teaching in the evaluation process but must be used alongside observation. "So I've been very pleased, I guess, with what I can look at [in the portfolio] and what I observe in class and what they can tell me they're doing that I may have missed. I'm feeling pretty comfortable that I can pretty well cover all eight standards."

Mike thought about the DE assessment requirements in Phase I thinking; specifically about the first ITS, relative to construction of classroom environment. Based on his comment, his thinking was already focused towards the structure of the standards and even though he disagreed with how the DE placed a particular criterion within the standards, he appeared to be open to its placement. "Quite honestly, I feel like 'creating a

classroom culture' fits better later on than it does up front. But, operating under the belief that culture helps drive learning and helps drive achievement in the classroom, I can see why it fits where it does too."

In summary, the three coded comment categories that clustered under judging activities were judgment, principal's opinion, and tools. The middle school principals did not consider judgment or tools in Phase I thinking. Both principals did make comments relative to their opinion of the portfolio as evidence of good teaching. Both principals considered the portfolio as only part of the teacher assessment process. Neither principal expressed dissatisfaction with the DE assessment requirements relative to teacher quality.

Middle School Principals and Coaching Activities in Phase I

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time investment (TI). Comments relative to coaching activities accounted for 17% of the total pre-assessment comments made by middle school principals. In this section, comment counts relative to two of the three categories (coaching and time investment) will be reported followed by a summary. The middle school principals did not comment on their role in Phase I thinking.

Coaching. The coaching category had four sub-categories; coaching/portfolio preparation (C/PP), coaching/meetings (C/M), coaching/questioning, and coaching/suggested alternatives for artifacts (C/SA). Only the two sub-categories of portfolio preparation and meetings were represented in Phase I thinking. Mike and Rob both thought about portfolio preparation in Phase I but only Rob thought about meetings.

In Phase I thinking, Mike and Rob both thought about the importance of being involved with their teachers as they prepared their portfolios so that no one was surprised. Rob commented, “I really tell them they need to tell me if there’s something that I can’t normally observe, or we’ve just going to dream up some things that we need to get in the portfolio to make it more complete.” Mike’s thinking paralleled that of Rob. He said, “...you gotta know that there would never ever be a time like this where a teacher walks in and hands me a portfolio that I already wouldn’t already have a pretty good understanding of what the structure is gonna be, and have helped think with them about the kinds of things that represent their teaching.”

Rob’s thinking about meetings reflected his thinking about preparation. He indicated that he provides teachers with multiple meeting opportunities so that he could be made aware of what he might look for in observations or in the portfolio.

Time Investment. Time (TI) accounted for the fewest coded comments in Phase I thinking of middle school principals. Mike’s lone comment referenced the “when and where” regarding portfolio evaluation. He indicated that he did not spend time in his office evaluating portfolios. “A lot of this ends up being evening time and quiet time. So that would be at home. A lot of it ends up being weekend time.”

In summary, the categories that clustered under the coaching activities were coaching, principal’s role, and time. Ivan did not engage in Phase I thinking. Mike and Rob did not comment on their role in Phase I thinking. Both principals indicated that they needed to be involved with teachers during portfolio preparation. Further, Rob indicated

that he felt it important to provide multiple meeting opportunities for his teachers. Mike indicated that he evaluated portfolios during quiet time away from his office.

Middle School Principals and Critical Pieces in Phase I

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two “pieces” carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. The critical piece category consumed 17% of the middle school principals’ thinking in Phase I. In this section, comment counts relative to the two sub-categories, critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO), will be reported followed by a summary.

Teacher reflection. Mike, once again, had the most comments in the CP category with a total of four. Three of his comments reflected his thinking about teacher reflection while one comment reflected his thinking about the role of observation. Once again, he framed the task of evaluating the portfolio by thinking about his own district and explained that “Particularly here in [our district], we have weighted more heavily the reflective writing that’s a part of the portfolio than the artifacts themselves. So I certainly will be putting my attention on the reflective writing piece...” Mike also shared that teachers initially struggled with the reflective writing piece and that he intentionally did not provide much direction. He explained that,

I wanted it to be about them. I felt like I could really, through their writing, understand what mattered most and more importantly how they could connect all of their artifacts through the writing. There’s no form, there’s no structure. There’s no right or wrong on length. It is very individual. I’m telling you what – it

was one of the most powerful administrative things that's ever happened to me when I ended up with the portfolios. The reflective writing.

From an observational standpoint, Mike's passion was overwhelming as he described the power of reflection. He appeared genuinely excited to explore the contents of the portfolio that had been provided to him by the researcher so that he could read the teacher's reflections.

Role of observation. Rob and Mike both considered the role of observation; the second sub-category under critical pieces (CP/RO). Both principals thought about the importance of observation as it compared to the portfolio in similar ways. Rob commented that "Before I ever open it [the portfolio] I still think it's important for the administrator to share the process with the teacher and the fact that I'm going to be able to come in before they've ever even purchased this notebook to put something in it, I will have some observations where I can start to pick out things from all eight standards." Mike thought along the same lines and indicated that he used walk-throughs as his main method of evaluation. He shared that "The in and out really matters to me. I really weigh heavily on that – what the teaching and learning looks like more than what the portfolio would show or share."

In summary, Mike was the only middle school principal to think about teacher reflection in Phase I. He did so with great passion; indicating that the reflections in the portfolio were both personal to the teacher and powerful for him to read. Both principals thought about the role of observation in Phase I. Rob viewed observation summaries as important documents that should be included in the portfolio while Mike indicated that observation was his main method of evaluation.

Summary of Middle School Principals' Thinking in Phase I

Phase I (pre-assessment) thinking was described as the period of time when principals prepared to judge the portfolio. Principals engaged in processing, judging, and coaching activities. As was true with the elementary principals, processing was most important to the middle school teachers in Phase I. Portfolio structure and process step comments accounted for over half of the middle school principals' thinking. Level of comfort was a consideration in Phase I thinking of the middle school principals but was not a barrier for the review. One of the principals indicated that he expected teacher reflection to be a critical piece of the portfolio.

Middle school principals thought their involvement with teachers throughout the portfolio process was essential. However, middle school principals viewed the portfolio as only one "piece" of teacher assessment and indicated that they valued observation above the portfolio.

High School Principals' Thinking in Phase I

The high school teacher chose to organize his portfolio by artifact. Each artifact represented one or more of the eight ITS. See Appendix I for examples of the table of contents and cross-referencing information used by the high school teacher.

The high school principals (Gavin, Kathy, Keith) spent more time in Phase I thinking than did their elementary counterparts and less time than the middle school principals. Coded comments made by the high school principals represented 32% of the total comments made by all principals in Phase I. Table 7 illustrate the breakdowns of the pre-assessment (Phase I) comments made by the high school principals.

Table 7

Summary of Phase I Comment Counts by High School Principals

	Total Comments	C	CL	J	PFS	PO	PR	PS	T
Gavin	6	0	1	0	4	0	0	1	0
Kathy	10	1	0	0	2	1	0	3	3
Keith	7	0	1	1	3	1	1	0	0

C = Coaching, CL = Comfort Level, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PR = Principal’s Role, PS = Process Steps, T = Tools

All three high school principals spent time framing the task of evaluating the high school portfolio provided to them by the researcher. Kathy spent the most time in pre-assessment thinking; her comments accounted for nearly 43% of the Phase I comments made the high school principals. Gavin and Keith were similar regarding the amount of time they spent in Phase I thinking. Gavin’s comments represented 26% of the total comments made by high school principals in Phase I while Keith’s comments represented 31% of the total.

High School Principals and Processing in Phase I

The three coded comment categories that clustered under processing activities included the categories of portfolio structure (PFS), process steps (PS), and comfort level (CL). Comments relative to processing accounted for 65% of the total pre-assessment comments made by high school principals. In this section, comment counts relative to

each of the three categories (portfolio structure, process steps, and comfort level) will be reported followed by a summary of high school principals and processing.

Portfolio structure. Phase I processing by the high school principals was dominated by comments regarding portfolio structure (PFS) as the principals focused on familiarizing themselves with a portfolio produced by a teacher they did not know. Thirty-nine percent of the total comments made by high school principals in Phase I thinking was relative to portfolio structure. Gavin accounted for 44% of the total Phase I comments made relative to portfolio structure. Keith's thinking accounted for 33% of the pre-assessment comments made by high school principals relative to portfolio structure while Kathy's thinking accounted for 23%.

The high school principals thought about the fact that one artifact could represent multiple standards. For instance, Gavin commented that "One of the first comments that I see here is that the state standards are multiple sources of artifacts. Not just from one or two areas but multiple sources." Kathy echoed Gavin's comments by saying that "I've seen a few of these and they are organized in different ways. I actually sort of like this kind of organizational structure. I like the fact that one artifact can represent lots and lots of standards."

Gavin indicated that his district used the electronic format and that the "hardcopy" version he was viewing was similar to the electronic version. "We do it on the E portfolio so it's all set up that way. Also if anybody doesn't feel comfortable with the electronic part of thing, then they can just do it hard copy. But it's basically the same structure. We break it down article by article and we just copy. If it fits three different criteria we copy

it.” At this point, Gavin seems to understand the structure of the portfolio. In fact, all three of the high school principals appeared to grasp the structure.

Process steps. The category of process steps (PS) had the next highest comment count for high school principals in Phase I. Only Gavin and Kathy made comments relative to process steps. Kathy accounted for 75% of the total comments made relative to process steps in Phase I thinking. Similar to Mike, one of the middle school principals, Kathy framed the task of evaluating the portfolio via several steps. She differed from Mike in that her thinking centered on the school and the department before she considered the teacher or his portfolio.

...so I probably would do some time thinking about our current social studies department and what the needs are, the people in that department and what the overall departmental goals might be so that when I'm looking at this, I'm looking at it in context of not just the development of this person in his content area, but ...in our school there are 12 social studies teachers.

Gavin's thinking in Phase I regarding process steps revealed the importance of reviewing the criteria of the ITS. “I need to read each of the criteria to make sure that I'm seeing what I'm seeing.”

Comfort level. Comfort level consumed very little of the high school principals' thinking in Phase I. Gavin and Keith each had one comment relative to comfort level but, similar to the middle school principals, neither appeared detoured. Gavin thought about the lack of familiarity with the portfolio. He indicated that “In our situation we have multiple meetings over this [the portfolio] so I pretty much know the flow of things – but it's very hard from the evaluators standpoint to look at this right away and say ‘I know what you're doing here’.” Keith's thinking centered on his lack of familiarity with the

teacher when he said, “Not having any background on what this teacher has done in the classroom is making this whole thing a challenge. I’m nervous about trying to hear this [overview of portfolio structure] and trying to give value to something.”

In summary, processing activities included the categories of portfolio structure, process steps, and comfort level. The high school principals spent most of their time focused on becoming familiar with the structure of the portfolio in Phase I; however, their comfort level was not adversely affected. Portfolio structure and comfort level were of concern for the researcher because the middle school principal who served as the pilot participant struggled with comfort level as he worked through the structure of the portfolio. This was not the case for the high school principals in Phase I. Two high school principals were comfortable with one artifact evidencing more than one ITS.

Some attention was given to process steps providing evidence that one high school principal considered departmental structure and goals before commencing her review. Another principal indicated that he reviewed a list of the standards/criteria prior to beginning a portfolio review.

High School Principals and Judging Activities in Phase I

The three coded comment categories that clustered under judging activities were judgment (J), principal’s opinion (PO), and tools (T). Comments relative to judging activities accounted for 26% of the total pre-assessment comments made by high school principals. In this section, comment counts relative to each of the three categories (judgment, principal’s opinion, and tools) will be reported followed by a summary.

Judgment. Keith was the only principal across all levels to think about judgment in terms of the teacher (J/T) in Phase I. He based his judgment on a quick overview of the portfolio and commented that,

Just looking at it tells me that this person is open to new ideas and new directions and isn't necessarily, at this point of observing, focused on one particular thing and that's, I think important, as you look at where teachers, especially veteran teachers, tend to find themselves in the same mold and they do things the same way year in and year out.

Principal's opinion. Kathy and Keith thought about the portfolio as evidence of good teaching (PO/PE) during Phase I thinking. Kathy thought about quantity and quality. She said, "Some of the portfolios that I get are three inches thick and some of them are one inch thick. Not necessarily in any sense that quantity means quality, but some people are collectors and they like to document every single thing that they've done. Some people think that they're going to meet some kind of minimal standards." Keith, like Rob and Mike at the middle school level, viewed the portfolio as a "...tool to help support what's taking place in the classroom."

Tools. Kathy was the only principal to think about tools (T) during Phase I. She clearly indicated that part of her pre-assessment framing was to make sure that she had a copy of the ITSS in front of her. "I'm looking at the Iowa Teaching Standards and Criteria List that I like to have in front of me. There's so many of them that I sometimes lose track which thing I'm looking at when I'm reading an artifact."

In summary, the three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). All three categories appeared in Phase I thinking of the high school principals but not much time was

committed to any of them probably due to the fact that the principals were focused on portfolio structure during Phase I.

High School Principals and Coaching Activities in Phase I

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time investment (TI). Comments relative to coaching activities accounted for nine percent of the total pre-assessment comments made by high school principals. In this section, comment counts relative to two of the three categories (principal's role and coaching) will be reported followed by a summary. No comments were made relative to time investment.

Coaching. In the coaching category, Kathy was the only principal at any level to think about coaching/questioning during Phase I. She had only one comment that was relative to clarification of a quote on the cover page of the portfolio. She simply indicated she would question the teacher about the quote.

Principal's role. Keith was the only principal across all levels to think about his role (PR) during Phase I thinking. He remained faithful to his judgment comment regarding his thinking about teachers trying new things when he said "As an administrator, I continue to try to push for my staff to try new things and to make their teaching more relevant to the students."

In summary, during Phase I, high school principals focused on familiarizing themselves with the portfolio with very little time spent on judgment; consequently, there were very few comments relative to coaching or principal's role. No comments were made by the high school principals relative to time investment in Phase I.

High School Principals and Critical Pieces in Phase I

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two “pieces” carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. The category of critical pieces consisted of two sub-categories; critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO). High school principals were the only group that did not have comments relative to critical pieces in Phase I.

Summary of High School Principals’ Thinking in Phase I

Phase I (pre-assessment) thinking was described as the period of time when principals prepared to judge the portfolio. Principals engaged in processing, coaching, and judging activities and identified teacher reflection and observation as critical pieces of teacher evaluation. As was true with the elementary and middle school principals, processing was most important to the high school principals in Phase I. Portfolio structure and process step comments accounted for 57% of the high school principals’ thinking in Phase I. Level of comfort was a consideration in Phase I thinking of the high school principals but was not a barrier for the review of the portfolio. Two high school principals were comfortable with one artifact evidencing more than one ITS.

Because Phase I thinking of the high school principals was so focused on processing (portfolio structure, process steps, and comfort level), judging activities (judgment, principal’s opinion, and tools) and coaching activities (principal’s role,

coaching, and time investment) had very low comment counts. The high school principals did not consider teacher reflection or the role of observation in Phase I thinking.

Phase I: Pre-Assessment Thinking per Geographic/Demographic Region

In this section, the findings relative to research question one for Phase I (pre-assessment) thinking will be reported per geographic/demographic region. First, a brief overview of the data collected for geographic/demographic region is provided. Total comment counts across geographic/demographic region for Phase I will be reported and illustrated. Then, findings for each geographic/demographic region for Phase I will be reported followed by a summary.

It is important to be reminded that while principals in each academic level (elementary, middle school, and high school) reviewed the same portfolio, the geographic/demographic region (urban, suburban, and rural) verbal comment counts were representative of comments made by the *same* principals *across* academic level, i.e., the principals in geographic/demographic region did not review the same portfolio because they represented differing academic levels. Therefore, data from geographic/demographic region is representative of comparison between urban, suburban and rural factors only; academic level data will not be revisited. Data will not be reported using the four broad categories as was true with the academic level reporting because it would be repetitive. Rather, data relative to the most numerically significant coded categories will be reported.

Geographic/Demographic Overview

Each geographic/demographic region represented one principal from each academic level (elementary, middle school, high school). The rural principals that participated in the study were Brenda, Ivan, and Keith. The suburban principals that participated in the study were Norma, Mike, and Gavin. The urban principals that participated in the study were Leo, Rob, and Kathy. The total number of verbal comment counts made in the study for each geographic/demographic region was remarkably similar (see Figure 2). Suburban principals had the highest overall percentage of coded verbal comments with 34%. The rural and urban principals each accounted for 33% of the total verbal comment counts made in the study. However, the similarity in the number of overall comment counts did not carry over to Phase I. Table 8 illustrates comment counts for pre-assessment thinking (Phase I) by geographic/demographic region per the ten coded categories.

All ten of the coded comment categories were represented in Phase I thinking across the three geographic/demographic groups. The four categories that were common to all three of the geographic demographic groups in pre-assessment thinking were comfort level (CL), critical pieces (CP), portfolio structure (PFS), and principal's opinion (PO).

Suburban principals spent the most time framing their work, i.e., pre-assessing, and accounted for nearly half (44%) of the total coded comments in Phase I per geographic/demographic region. Urban principals accounted for 38% of the total coded comments in Phase I. Rural principals' pre-assessment thinking accounted for less

Table 8

Summary of Phase I Comment Counts per Geographic/Demographic Region

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Rural	12	0	2	1	2	5	1	1	0	0	0
Suburban	31	3	3	4	0	10	2	0	8	0	1
Urban	27	3	3	2	1	8	2	0	5	3	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PR = Principal’s Role, PS = Process Steps, T = Tools, TI = Time

than half of that of each of the other two groups with only 18% of the total coded comments in Phase I for geographic/demographic region. As noted in academic level results, Phase I thinking comments accounted for roughly eight percent of the total coded comments for the study.

Rural Principals’ Thinking in Phase I

Table 9 illustrates the breakdowns of the Phase I (pre-assessment) comments made by the rural principals. Rural principals spent very little time framing the task of evaluating the portfolio provided to them by the researcher.

Nearly half of the pre-assessment thinking conducted by rural principals was relative to portfolio structure. The remaining comment counts were minimal and spread evenly across the five remaining categories at similarly low levels. In short, rural principals took very little time to frame their task and while they thought about six of the

ten categories, the bulk of their pre-assessment thinking was relative to portfolio structure.

Table 9

Summary of Phase I Comment Counts made by Rural Principals

	Total Comments	CL	CP	J	PFS	PO	PR
Brenda	5	1	1	1	2	0	0
Ivan	0	0	0	0	0	0	0
Keith	7	1	0	1	3	1	1

CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PR = Principal’s Role

Suburban Principals’ Thinking in Phase I

Table 10 illustrates the breakdowns of the Phase I comments made by the suburban principals. Suburban principals spent two and a half times more time in Phase I thinking than the rural principals and slightly more time than the urban principals. Their pre-assessment thinking included seven of the ten comment categories.

Suburban principals were similar to the rural principals in that they included four of the same categories; comfort level (CL), critical pieces (CP), portfolio structure (PFS), and principal’s opinions (PO). However, they also included the categories of coaching (C), process steps (PS), and time investment (TI). Suburban principals did not include judgment in their pre-assessment thinking as did the rural and urban principals. Nor did

they include principal's role as did their rural colleagues. Suburban principals were the only group to consider time investment but to a very small degree.

Table 10

Summary of Phase I Comment Counts made by Suburban Principals

	Total Comments	C	CL	CP	PFS	PO	PS	TI
Norma	1	0	0	0	1	0	0	0
Mike	25	3	2	4	5	3	7	1
Gavin	6	0	1	0	4	0	1	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, PFS – Portfolio Structure, PO = Principal's Opinion, PS = Process Steps, TI = Time

Portfolio structure accounted for 31% of the pre-assessment thinking of the suburban principals. Suburban principals spent twice as much time thinking about portfolio structure during pre-assessment as did their rural colleagues, however, two of the suburban principals accounted for the bulk of thinking in this category. Mike and Gavin accounted for 90% of the thinking concerning portfolio structure in pre-assessment. Norma made only one comment concerning portfolio structure in pre-assessment thinking.

The coded comment category of process steps accounted for the second highest percentage of pre-assessment thinking representing 26% of the coded comment counts for suburban principals. Mike's comments accounted for 88% of the total pre-assessment

comment counts for the process step category for suburban principals. As noted in the academic level reports, Mike was very clear regarding the steps he thinks about as he prepares to evaluate a second-year teacher portfolio. Gavin had only one comment relative to process step thinking in the pre-assessment phase and Norma had no comments relative to the process step category in the pre-assessment thinking of suburban principals.

The remaining pre-assessment thinking categories represented less than ten percent each of the total pre-assessment comments for suburban principals. Like their rural counterparts, portfolio structure consumed Phase I thinking of the suburban principals. However, unlike the rural principals, suburban principals thought about process steps during pre-assessment.

Urban Principals' Thinking in Phase I

Table 11 illustrates the breakdowns of the Phase I comments made by the urban principals. Urban principals spent twice as much time in Phase I thinking as did elementary principals but spent slightly less time than did their suburban counterparts. Phase I thinking of the urban principals included eight of the ten coded comment categories. The four categories that they had in common with the rural and suburban principals were comfort level (CL), critical pieces (CP), portfolio structure (PFS), and principal' opinion (PO). Urban principals thinking in pre-assessment also included the categories of coaching (C), judgment (J), process steps (PS), and tools (T). Urban principals were the only group to think about tools during Phase I. Very similar to the suburban principals, portfolio structure accounted for 30% of the total coded comments

for urban principals in pre-assessment thinking. Leo had twice as many comments concerning portfolio structure as did either Rob or Kathy.

Table 11

Summary of Phase I Comment Counts made by Urban Principals

	Total Comments	C	CL	CP	J	PFS	PO	PS	T
Leo	7	0	1	0	1	4	0	1	0
Rob	10	2	2	2	0	2	1	1	0
Kathy	10	1	0	0	0	2	1	3	3

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PS = Process Steps, T = Tools

Process steps accounted for the second highest percentage of pre-assessment thinking representing 19% of the coded comment counts for urban principals. Kathy’s thinking in this category represented 60% of the thinking. Like Mike, a middle school principal in the suburban group, Kathy was very clear regarding the steps she thinks about as she prepares to evaluate a second-year teacher portfolio. The categories of coaching, comfort level, and tools were each representative of 11% of the pre-assessment thinking of urban principals. The remaining categories of critical pieces, judgment, and principal’s opinion represented less than 10% each of the pre-assessment comments made by urban principals.

Summary: Findings for Phase I Thinking.

The findings clearly indicated that principals in the current study spent the bulk of their time during Phase I (pre-assessment) in the area of processing (portfolio structure, process steps, and comfort level). This was true across academic level as well as geographic/demographic region. The principals in the current study were anxious to understand the design of the portfolio so that they could make informed judgments about the artifacts. Also in Phase I, the principals provided some insight into the process steps used during review. These findings were supported by the comment counts. The categories with the highest percentage of overall comment counts in Phase I were portfolio structure (36%) and process steps (18%).

While comfort level was a common concern in Phase I thinking across academic level and geographic/demographic region, it was not a barrier for the principals. Although Phase I (pre-assessment) thinking consumed only eight percent of the total coded comment counts for the study, it was a significant phase in the review process because principals gained confidence, via familiarizing themselves with the portfolio, as they moved into Phase II (judgment) thinking.

Verbal reporting relative to what principals thought during Phase I was consistent in content; however, differences did exist relative to the *number* of pre-assessment comments across academic level (see Table 4) and geographic/demographic region (see Table 8). In regard to academic level, middle school principals had the highest comment count, i.e., spent the most time in pre-assessment activities. Middle school principals

made twice the comments of elementary principals and one-and-a-half times as many as the high school principals.

In regard to geographic/demographic region, rural principals spent very little time in pre-assessment thinking. The time urban and suburban principals spent in pre-assessment thinking, while comparable to one other, was nearly twice that of their rural counterparts.

Phase II: Judgment Thinking per Academic Level

Findings in Phase II established that (a) principals spent a significant amount of time in judging and coaching activities, (b) in general, principals across academic levels developed a four-step *rhythm* as they moved through the judgment phase, (c) an imbalance existed between artifact judgment comments and suggested alternatives (i.e., the artifacts were by and large judged positively), and (d) principals identified observation of the teacher by a principal and written teacher reflection as two critical pieces of teacher evaluation.

In this section, findings relative to research question one for Phase II (judgment) thinking will be reported as follows: First, an overview of the data will report and illustrate the total comment counts across academic level for Phase II. This overview will serve as a reference point for reporting the findings of Phase II. Second, a description of the four-step rhythm will be presented and discussed. Then, findings for each academic level and geographic/demographic region for Phase II will be reported.

Overview

Table 12 illustrates comment counts in Phase II (judgment thinking) of all principals by academic level. Comments made during Phase II represented 79% of the total coded comments for the study. Middle school principals' thinking, as it did in Phase I, accounted for the most comment counts during Phase II. Elementary principals had the least amount of comments in Phase II.

Table 12

Summary of Phase II Comment Counts by Academic Level

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Elementary	163	29	7	17	88	12	2	3	3	2	0
Middle School	377	78	1	21	196	4	13	0	44	11	9
High School	240	60	5	16	122	12	6	3	7	8	1

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

Just as they were in Phase I thinking, all ten of the coded comment categories were represented in Phase II (judgment) thinking. Eight categories were common to all three of the academic level groups in Phase II thinking: coaching (C), comfort level (CL), critical pieces (CP), judgment (J), portfolio structure (PFS), principal's opinion (PO), process steps (PS), and tools (T). This compares to the representation of only four common

categories across academic level in Phase I thinking. The categories that were common to both Phase I and Phase II thinking across academic level were coaching (C), comfort level (CL), portfolio structure (PS) and process steps (PS). In short, principals broadened their thinking as they moved from the pre-assessment thinking phase (Phase I) into the judgment phase (Phase II).

Phase II was dominated by judgment (J) thinking (52%). The next most dominant category during Phase II thinking was that of coaching (21%). Critical pieces (CP) and process steps (PS) each accounted for seven percent of the comments principals made during Phase II thinking. The categories of comfort level (CL), portfolio structure (PFS), principal's opinion (PO), and tools (T) were all represented in Phase II thinking but at levels less than seven percent of the overall comments for the judgment phase.

The ten coded categories clustered around four broader categories that provided a means of efficiently reporting data. The four broad categories were (a) processing activities, (b) judging activities, (c) coaching activities, and (d) critical pieces. Processing activities included the categories of process steps, portfolio structure, and comfort level. Judging activities included the categories of judgment, principal's opinion, and tools. Coaching activities included categories of principal's role, coaching, and time investment. Critical pieces included the categories of teacher reflection and role of observation. Each academic level will be reported via the four broad clusters using data from the appropriate smaller categories.

Four-step Rhythm

Seventy-three percent of the coded comments in Phase II represented judgment (52%) or coaching (21%). This was due largely to the fact that, in general, principals across academic levels developed a four-step *rhythm* as they moved through the judgment phase. First, they would identify and verbally describe an artifact. Second, they would actually read aloud as they focused on what the teacher was attempting to illustrate. Their third step was a judgment statement regarding the artifact they were judging. The judgment statement was then sometimes followed by a coaching statement; thus, the higher rate of coaching comments.

A good example of the rhythm established by the principals was demonstrated by Gavin, a high school principal, in the following portion of his transcript. He is judging an artifact used by the high school teacher to illustrate that he (the teacher) is meeting ITS 1-G. The artifact the teacher used to meet the standard was the creation of a Webpage. The italicized words indicate that Gavin was reading directly from the second-year teacher's portfolio provided to him by the researcher.

1G – Communicates with student families. That would be wonderful. He hit that. I'm sure in most cases about 80% of the students have web pages. The only thing I would want to make sure is to find out who doesn't and have them sent too, because all students need to ...go back to that 1E [creates an environment of mutual respect, rapport, and fairness] that he talks about there. If parents don't have Internet, then the fairness isn't there either. He needs to figure out how to get this home to those people too.

Gavin identifies the standard and associated criteria and then reads aloud the words the instructor used to describe the standard/criteria. He judges the artifact favorably but his

thinking moves towards coaching, specifically coaching/suggested alternative, after he makes his judgment statement.

Another example of the four-step rhythm established by principals was well illustrated by Brenda in the following excerpt. Brenda was judging an artifact titled “Diverse Learners, Ranging from a Struggling Home to ELP.” The elementary teacher indicated on her cross-reference sheet (see Appendix H) that the artifact met all or part of every ITS. In step one, Brenda clearly identified the artifact and to what it referred (diverse learners). “Okay. We have artifact number 5. This is diverse learners. I’m going to go over to the second page that’s kind of the artifact reflection tag. I’m just going to skip up here to these questions.” In the second step, Brenda read aloud the teacher’s own words, attempting to understand what the teacher thought was being illustrated.

Here are some questions that may help me in reflecting on my artifact. Why did I select the artifact? This is what I think might be missing sometimes. Why did I select this artifact? Why did I want all of my students to know or to do the result of my teaching? How did I judge the quality of my students’ work? How did my practice impact student achievement? How could I improve or strengthen my practice? I think those questions ...now I’m just keeping in mind ...over these two school years I’ve had the opportunity to work with students who have a large range of needs. Some of my students have been in SCI program, many in Title I reading, a few in ELP and two in particular with very difficult home lives. It has been through these students that I have learned to be very flexible, individualized to their needs whether it be at home or at school. Included in this artifact are an email from a parent showing their appreciation for me getting their child going to the GEI process. Another is a contract and calendar plan I made strategy I developed after taking special needs class. Where a student has struggled to come to school prepared due to home situations and finally I included a meeting syllabus that I put together for 4, 5, 6 grade teachers, administrators, counselors, and parents to discuss the development of needs. And placement of a student who is performing much beyond 5th grade.

In step three, Brenda made a judgment that acknowledged that the teacher had worked with “a lot of different students.” She said, “Okay. This is a good background into some

of the things that she's done with a lot of different students. She has explained why she's put all these things in with a cover letter." However, Brenda wanted more. Specifically, she wanted the teacher to show growth via reflection. Her fourth step, a coaching comment in which she suggested an alternative, was very clearly defined when she said, "These are all good pieces of evidence to show that she's dealt with a lot of different types of students. I think just expanding on her reflection without a lot of guided questions."

The ebb and flow of the rhythm was consistent across academic level. However, as is evidenced by the number of judgment comments as compared to the coaching comments (Table 12), the thinking was not equal. In other words, there was not a coaching comment made every time a principal made a judgment comment. Examination of the data per academic level further illustrated how principals moved through the judgment thinking phase (Phase II).

Elementary Principals' Thinking in Phase II

The elementary portfolio used in the study was from a fifth grade teacher and was organized by artifact. See Appendix H for the cross-referencing information used by the fifth grade teacher. The elementary principals (Brenda, Leo, and Norma) made the fewest Phase II comments of the three academic levels. Their thinking accounted for only 21% of the overall comments made by all principals in the judgment thinking phase as compared to 48% for middle school principals and 31% for high school principals.

Table 13 illustrates the breakdowns of Phase II comments made by elementary principals. There were five categories common to Phase I and Phase II thinking for the

elementary principals. They were comfort level (CL), critical pieces (CP), judgment (J), portfolio structure (PFS) and process steps (PS).

Four new categories emerged in Phase II for the elementary principals. They were coaching (C), principal's opinion (PO), principal's role (PR), and tools (T). Brenda and Leo, as they did in Phase I, had the most Phase II comments. Norma had significantly more comments in Phase II than she did in Phase I; however, her comments were still nearly half of those of Brenda.

Table 13

Summary of Phase II Comment Counts made by Elementary Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Brenda	71	19	0	11	30	5	2	3	0	1	0
Leo	55	7	2	3	34	6	0	0	2	1	0
Norma	37	3	5	3	24	1	0	0	1	0	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

Elementary Principals and Processing in Phase II

The three coded comment categories that clustered under processing included the categories of portfolio structure (PFS), process steps (PS), and comfort level (CL).

Comments relative to processing accounted for 13% of the total Phase II comments made

by elementary school principals. In this section, comment counts relative to each of the three categories (portfolio structure, process steps, and comfort level) will be reported followed by a summary of elementary principals and processing.

Portfolio structure. Elementary principals continued to think about portfolio structure in Phase II, but not to the same degree. In Phase II, comments in the portfolio structure category accounted for 7% of the total coded comments, a sharp drop from the 54% it captured in Phase I. Thinking about portfolio structure in Phase II moved from curiosity and information seeking to closer scrutiny of how the portfolio was structured. Brenda a Leo had slight increases in thinking relative to portfolio structure. Norma made only one comment in both Phase I and Phase II regarding portfolio structure.

Brenda became somewhat critical of the structure of the portfolio as she progressed through Phase II. Early in Phase II, Brenda liked the organization of the teacher's portfolio. "I'm thinking she's gone through and with this lesson, identified which of her data points are included in here. I'm seeing this is a good way to show me how she's gone through each of the standards. It is evidencing all eight standards. On standard number eight she's got specifically what it is either she's done or evidence from another data point." Brenda became increasingly frustrated with the organization of the portfolio when she could not make clear, immediate connections between the evidence (artifacts) and the standards. "I would like a cover sheet, a reflection sheet on those just stating 'this is why I feel this evidences these standards.' I am still doing a lot of work here to figure it out myself. I want you [the teacher] to tell me basically how this evidences the standards." Brenda diligently continued to work through the portfolio and

near the end of Phase II, she returned to the structure of the portfolio. It became clear that she would have preferred the teacher to organize the portfolio by standard rather than by artifact. “Sometimes I think starting out, getting used to the Iowa teaching standards and seeing ...sometimes organizing a portfolio according to standards helps them understand the standard better.”

Leo thought about the structure of the portfolio a bit differently than Brenda. He noted and appreciated the structure in Phase I and his thinking remained consistent in Phase II. The teacher had included a Social Studies unit as an artifact. He thought about the consistency with which the teacher presented her artifacts and noted that “Social studies is laid out in the same way and I suspect it’s going to be very similar to what we saw in the first one [artifact]. We just moved into a different content area. So again it’s laid out the same way.” His confidence and the speed with which he moved through the artifacts increased after he became familiar with the structure of the first two artifacts. However, like Brenda, he did note on one of the artifacts that clear connections between the artifact and the standard were not evident. He commented that, “The structure is the problem here. She simply circles the descriptors under the standard and I don’t know what she feels is in here that actually makes that connection.”

Process steps. The number of overall comments regarding process steps (PS) increased by only two comments from Phase I thinking. Comments in the process steps category accounted for two percent of the total coded comments in Phase II thinking for elementary principals. The thinking in Phase II regarding process steps did not reveal any

critical new information regarding how the elementary principals progressed through the portfolio.

Leo, however, early in Phase II, seemed reluctant to take the steps necessary to thoroughly investigate the links between the artifacts and the standards. He indicated his hesitation by saying “Well, what I have not done ...and it would certainly take some time to do ...but to really validate what’s going on here, a person would almost have to go through each one of the references [criteria] that she has circled here and check for continuity.” However, he did progress through the rest of the portfolio by carefully examining each artifact and comparing it to the standard/criteria that the teacher indicated she was evidencing.

Comfort level. The number of overall comments regarding comfort level (CL) increased three-fold in Phase II thinking. Comments in the comfort level category accounted for only four percent of the total coded comments in Phase II thinking for elementary principals. Comfort level thinking in Phase I was dedicated to familiarization with the portfolio and some thinking about not being able to physically observe and/or know the teacher who had produced the portfolio. Phase II thinking took on a different look.

Brenda did not dedicate any time to comfort level in Phase II. Leo’s thinking regarding comfort level increased by only one comment. His thinking centered on respecting the views of the principal of the teacher whose portfolio was used in the research.

It’s difficult to separate out what I see ...narration that he [teacher’s principal] gave to visualize what might have been going through his mind or her mind or

how I'm seeing that differently based on the artifacts that I saw. That's one of the things that makes me have to kind of look at this and try to stay away from second guessing somebody else who was looking at this same portfolio.

Interestingly, it was Norma who had the greatest increase in comment counts relative to comfort level thinking. Norma had only one Phase I comment and it was not relative to comfort level. She moved quickly from pre-assessment (Phase I) thinking to judgment (Phase II) thinking. Her confidence in jumping right into judgment was only accentuated in her comments in Phase II. Norma's Phase II comments illustrated significant confidence in being able to make quick connections between the artifact and the ITS. Her confidence was best illustrated when she said, "I know what an SCI resource is. It would appear as though she was doing some differentiation at least in the expectations because the resource student got 14 out of 15. So did some differentiation on the expectations for students. ...engaging and involving all students [ITS 4b]."

In summary, processing included the categories of portfolio structure, process steps, and comfort level. Processing was much less significant in Phase II; however, the elementary principals all dedicated thinking to portfolio structure during Phase II. While some criticism of the structure did exist, it did not detour the principals from moving through the evaluation. Two of the three principals thought about connections between the artifacts and the ITS. They indicated that the structure did not offer enough visual clarity; thus, making them (the principals) work harder to make the connections.

Phase II thinking in terms of process steps consumed very little of the elementary principals' thinking in Phase II. The comments in the process steps category for Phase II increased only slightly from the comments made in the same category in Phase I. One

principal exhibited some hesitance to linking the criteria directly to the standards but continued to evaluate the portfolio in depth; making judgments based on each artifact and its relationship to the corresponding standard (s) that the teacher indicated on her cross-reference sheet.

Comfort level thinking increased in Phase II thinking for the elementary principals. Only two principals engaged in comfort level thinking in Phase II. One principal accounted for 71% of the comments relative to comfort level in Phase II thinking. The increase was due mainly to evidence of confidence rather than apprehension as illustrated in Phase I thinking.

Elementary Principals and Judging Activities in Phase II

The three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Comments relative to judging activities accounted for 56% of the total Phase II comments made by elementary school principals. In this section, comment counts relative to each of the three categories (judgment, principal's opinion, and tools) will be reported followed by a summary of elementary principals and judging activities.

Judgment. There was a significant increase in judgment comments in Phase II as the elementary principals began to review artifacts in the portfolio (Table 13). The judgment category was a multi-level category that included judgment of artifacts (J/A), judgment of the portfolio (J/P), and judgment of teacher (J/T). Table 14 illustrates the breakdown of the judgment category comment counts for elementary principals in Phase II.

The judgment of artifacts (J/A) clearly dominated the thinking in the judgment category for the elementary principals in Phase II. The number of comment counts per each principal was comparable in the J/A category. Norma's thinking was completely focused on the judgment of the artifacts while Brenda and Leo did stray slightly towards judging the portfolio and the teacher.

Table 14

Judgment (J) Comment Count Breakdowns for Elementary Principals in Phase II

	Total Judgment Comments	J/A	J/P	J/T
Brenda	30	27	3	0
Leo	34	30	3	1
Norma	24	24	0	0

J/A = Judgment comments relative to the artifacts, J/P = Judgment comments relative to the portfolio, J/T = Judgment comments relative to the teacher

By further breaking down the judgment of artifacts category, the comparability of comments per artifact by each principal became even more clearly illustrated. Table 15 illustrates the breakdowns of the number of artifact judgment comments, per each artifact, made by the elementary principals.

As a group, the elementary principals made the most artifact judgment comments about artifacts one, two, and three. For Brenda and Norma, artifact one garnered the most

comments. Leo had the most comments in artifact one but it was artifact three that dominated his thinking. After artifact three, the number of judgment comments by each principal, relative to each ensuing artifact, tapered off.

Table 15

Judgment of Artifact (J/A) Comments per Artifact by Elementary Principals in Phase II

	Artifacts*									
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Brenda	5	5	4	1	3	1	1	3	2	2
Leo	6	2	7	2	4	2	2	2	1	2
Norma	4	3	2	2	2	3	2	3	2	1

* See Appendix H for Artifact Title and Cross-referencing with ITS

To establish the actual judgment of artifacts, it was necessary to create a rating system by which to classify the artifact judgment statements made by each principal relative to each artifact. To accomplish this, ratings of positive (+), neutral (N), or negative (-) were established. Each artifact judgment statement made by the elementary principals in Phase II was evaluated and tagged with one of the ratings.

Positive judgment statements included those statements that indicated the artifact, per the principal's judgment, had sufficiently illustrated/met the ITS the elementary teacher indicated it would. An example of a positive artifact judgment statement was made by Brenda as she worked through evaluation of the first artifact when she said,

“Now we must be into some writing examples. Which is a good thing to have in here as a result of the lesson, but actually it gives me an idea of what kids actually do when you set this [goal] and ask them to do that [assessment].”

Neutral statements included those statements that the principal made while judging the artifact, but the statements did not indicate that the principal had made a definitive judgment as to the value of the artifact. An example of a neutral judgment statement was made by Norma as she worked through evaluation of the third artifact when she said, “She intends for me to see everything here except for 6C and 6D and 8A and 8B [ITS and criteria]. Those are tough ones to show.”

Negative judgment statements included those statements that indicated the artifact, per the principal’s judgment, did not sufficiently illustrate/meet the ITS the elementary teacher had indicated it would. An example of a negative judgment statement was made by Leo as he worked through evaluation of artifact six when he said,

The documents that I’m seeing here are, again, implementation documents. These are all things that have simply indicated the fact that she has done this. So what’s going on in my mind is that she has not indicated what has happened as a result of it. She doesn’t have any student documentation here that is a follow through to indicate that yes, this has really reached out and got me. Whether it’s looking at box scores for the St. Louis Cardinals or picking up a Dear Abby column to be able to figure out relationships. I’m not seeing student connections here.

Table 16 represents the results of tagging each of the elementary principals’ judgment comments in Phase II. Clearly, the largest portion of artifact judgment comments made by the elementary principals was positive. In fact, positive comments represented 63% of the judgment comments made pertaining to the ten artifacts in the portfolio. Neutral comments represented 23% of the judgment comments made pertaining

to the ten artifacts in the portfolio. Negative comments represented only 14% of the judgment comments made pertaining to the ten artifacts in the portfolio.

Table 16

Classification of Judgment Comments per Artifact made by Elementary Principals in Phase II

	Total Judgment Comments	<u>Brenda</u>			<u>Leo</u>			<u>Norma</u>		
		(+)	(N)	(-)	(+)	(N)	(-)	(+)	(N)	(-)
Artifact 1	15	4	1	0	3	1	2	4	0	0
Artifact 2	10	2	1	2	1	1	0	2	1	0
Artifact 3	13	3	1	0	4	3	0	1	1	0
Artifact 4	5	0	0	1	1	1	0	2	0	0
Artifact 5	9	3	0	0	3	1	0	2	0	0
Artifact 6	6	1	0	0	0	0	2	0	1	2
Artifact 7	5	1	0	0	1	0	1	0	1	1
Artifact 8	8	3	0	0	1	0	1	1	2	0
Artifact 9	5	1	1	0	1	0	0	1	1	0
Artifact 10	5	1	1	0	2	0	0	1	0	0

(+) = Principal made positive statement about value of artifact, (N) = Principal made neutral statement about value of artifact, (-) = Principal made negative statement about value of artifact

In order to get a sense of how elementary principals thought about the artifacts in terms of their effectiveness in illustrating/meeting the ITS the elementary teacher indicated they would, it was valuable to look specifically at the elementary principals' thinking as they viewed artifacts one, two, three, and six. Artifacts one, two, and three were chosen because they each had the most artifact judgment counts of all ten of the artifacts. Artifact six was chosen because it was the only artifact that did not obtain an overall positive judgment.

Artifact one garnered the most artifact judgment comments from the elementary principals. It accounted for 19% of all artifact judgment comments made by elementary principals in Phase II thinking. The first artifact was labeled Observation: *Writer's Workshop*. It was, in essence, a lesson relative to *Writer's Workshop* that was observed by the teacher's principal. It included (a) a very positive formal observation summary that had been completed by the teacher's principal, (b) a comprehensive lesson plan, (c) student completed goal sheets relative to *Writer's Workshop*, (d) student writing samples, (e) task prompts, (f) student writing samples, (g) teacher reflection, (h) post observation conference notes, and (i) data tracking per RIT scores for each student. Student writing was centered on the root statement "My writing is like a donut because..." The teacher indicated on her cross-reference sheet that she met all or part of every ITS (see Appendix H).

In general, artifact one was judged as positive by the elementary principals, i.e., it was acceptable in terms of how it illustrated/met the ITS the teacher indicated it would. Positive judgment comments accounted for 73% of the total artifact judgment comments

made by the elementary principals for artifact one. Neutral and negative comments accounted for 13% each of the total artifact judgment comments.

The elementary principals identified several pieces that made artifact one acceptable. Leo and Norma liked the principal's observation and thought it was valuable evidence. Norma said, "Okay. Sounds like a good observation and positive feedback." Leo relied on the principal's observation to make a judgment. He read through the observation notes. He thought the observation allowed him to "...see how the whole thing [*Writer's Workshop* lesson] hangs together." Brenda appreciated the thoroughness of the lesson plan and said, "This is something that they [the teacher] created, the lesson that they've created, so that they can ...it shows me planning, it shows me preparation for it, it shows me how I'm assessing organized materials, so organization of the unit."

Brenda and Norma both made positive comments about the student writing samples the teacher included. Brenda commented that, "Now we must be into some writing examples. Which is a good thing to have in here as a result of the lesson, but actually it gives me an idea of what kids actually do when you set this [goal] and ask them to do that [assessment]." Norma said, "Okay. That's pretty comprehensive. Writing rubric. This is actually a student's rubric - 19 out of 20 points."

Leo liked the *Writer's Workshop* concept and particularly the focus of the writing lesson. He commented, "...the whole notion to me of a writer's workshop is a valid artifact and I would certainly say yeah, you bet, it's fine to do that. I like the donut metaphor, by the way." Norma provided a positive overview of artifact one when she

said, “Yeah, I think I saw evidence of all those things as well. I can’t think of anything that would be at least glaringly left out.”

Neither Brenda nor Norma thought negatively about artifact one. Leo had two negative comments. His first comment was relative to a piece he could not find in the artifact. He said,

I am less interested in curriculum and content and more interested in relationships. So what I have tended to ...my staff here would be very aware of that because that’s what I’ve talked as I have worked here ...is focus on relationships. Relationships have got to come first and then teaching or being able to articulate curriculum is going to come second. So what I’m seeing here in this particular development is much more content skewed as opposed to relationship skewed.

Leo did not know what was meant by RIT scores. As a consequence, it seemed to skew his evaluation of artifact one slightly to the negative side. He said, “We do have outcome information data here. Fall RIT. RIT block growth. I don’t know what a RIT is. Beats me. *Reprinted by the mid Iowa School Improvement Consortium.* We’ve got outcome data, but I can’t get a grip on what it’s telling me. So, lots of paper but just don’t have a real good sense of what that classroom looks like.”

In general, elementary principals gave artifact one a positive review. They identified positive pieces of the artifact as, (a) the principal’s observation, (b) the concept of *Writer’s Workshop*, (c) the lesson plan, (d) student samples, and (e) the thoroughness of the entire artifact. The negative thinking that was present in the judgment of artifact one stemmed from unfamiliarity with a particular term (RIT scores) and from perspective (Leo’s urban lens).

Artifact two had the third highest artifact judgment comments from the elementary principals. It accounted for 12% of all artifact judgment comments made by elementary principals in Phase II thinking. Artifact two was labeled Observation: Social Studies. The topic of the lesson being taught/observed was Pacific Northwest Native Americans. The artifact included (a) pre and post-observation forms, (b) a very positive formal observation summary that had been completed by the teacher's principal, (c) two graphic organizers with terms and definitions relating to Northwest Native Americans that was created by students using Inspiration software, (d) an example of an ELP (TAG) student's work in the form of a short answer assessment, (f) teacher reflection, and (g) a checklist of "what should be present in a classroom" given to the teacher by her principal. The teacher indicated that she met all or part of every ITS with the exception of standard seven.

Overall, artifact two was judged positively by the elementary principals, i.e., it was acceptable in terms of how it illustrated/met the ITS the elementary teacher said it would. Positive judgment comments accounted for 50% of the total artifact judgment comments made by elementary principals for artifact two. Neutral comments accounted for 30% of the total artifact judgment comments. Negative comments accounted for only 20% of the total artifact judgment comments for artifact two.

The elementary principals identified several pieces that made artifact two acceptable. Leo and Norma both relied on the observation summary completed by the elementary teacher's principal. For them, the observation provided positive evidence that the teacher had met the ITS/criteria that she indicated she would on her cross-reference

sheet. Leo read the summary and noted that the principal's remarks indicated that the elementary teacher had a strong ability to relate to her students in a positive and professional way. The comment in the observation was very significant for Leo and was consistent with his thinking about relationships in artifact one. He said, "For me, that's more of the kind of thing that I'm looking for. Again, is talking about her connection with students." Norma also trusted the principal's observation in addition to the teacher's reflection. She judged the artifact positively because she could see evidence that the principal and the teacher agreed on the ITS/criteria that were evidenced.

She has here [reflection page] how she's going to evaluate the students, both from the responses they give during the brainstorming portion as well as a final assessment that's going to be an essay. Then again lays out how the assessment will work. [Reading]*Post observation conference*. Teacher reflection and indicates the standards that she and the evaluator apparently ...so apparently the evaluator here was marking. I'd say that yes, it looks like all those things were covered.

Brenda thought the graphic organizer web created by the students using Inspiration was a good visual piece of evidence; partially because she likes the webbing concept and partially because she thought the web organizer included by the teacher was good. "Webbing is good. This one is good. This is a good web."

Brenda's thinking in Phase II was consistent with that in Phase I; she continually referred to the importance of reflection. Brenda liked the fact that the teacher reflected clearly about her goals, her instruction, and the assessment of the lesson in artifact two. "That's a good thing to always remember [lesson alignment with goals]. Just to see how her [teacher's] reflections were. She absolutely felt that they [students] learned what she

intended. It's good that she reflected on that, because I think I was questioning the connection there."

Interestingly, the two negative judgment comments by Brenda, relative to artifact two, were made prior to her reading the teacher reflection. She could not make clear connections. "Her objective here [included on pre-observation form] is kind of conflicting with her original objective of the lesson." Based on the teacher's reflection, Brenda's judgment changed regarding the artifact.

In general, elementary principals gave artifact two a positive review. They identified positive pieces of the artifact as, (a) the principal's observation, (b) use of a graphic organizer created by students using Inspiration software, and (c) teacher reflection. The negative thinking that was present in the judgment of artifact two was eliminated via teacher reflection statements. The three principals continued to be consistent in their thinking. As it did in Phase I and in her review of artifact 1, Brenda's perspective about reflection influenced her judgment. Leo's perspective concerning relationships was influenced positively due to comments the principal made in the teacher's formal observation. Norma, as she did in her review of artifact one, relied heavily on the principal's observation.

Artifact three had the second highest artifact judgment comment counts for the elementary principals. It accounted for 16% of all artifact judgment comments made by elementary principals in Phase II thinking. Artifact three was labeled Social Studies Unit: Native Americans. The teacher's written reflection best described the artifact.

This Native Americans unit is an extensive six week program that begins with the theory of Beringia nearly 15,000 years ago (possibly more) and runs through the

1400s up to Exploration. The unit includes everything from four major art projects, a field trip to Effigy Mounds and our local Ag Museum, newspaper publications, group work, categorization and analysis, writing legends, and an essay assessment.

This unit is the foundation for the curriculum in my social studies class and I developed it using several sources, one of which is participating in the Teaching American History grant. Throughout the year the students are able to remember who the 1st Americans are and how, over time, our country has expanded through all of us being part of the immigration process. This unit helps them have a better sense of continuity and change as well as time.

The artifact included (a) the teacher reflection; (b) photos of the art projects; (c) an explanation of the categorizing activity; (c) photos of the visit to the Ag Museum and Effigy Mounds; (d) announcements made by the teacher in a newsletter to parents about the visits and the schedule of activities for the unit, including homework due dates and daily lesson objectives; (e) student samples, including a graded final essay with rubric; and (f) a peer observation of the teacher for one lesson in the unit. The teacher indicated that she met all or part of every ITS with artifact three.

Artifact three received a positive review from the elementary principals, i.e., it was acceptable in terms of how it illustrated/met the ITS the elementary teacher said it would. Positive judgment comments accounted for 62% of the total artifact judgment comments made by elementary principals for artifact three. Neutral comments accounted for 38% of the total artifact judgment comments made by elementary principals. There were not any negative judgment comments for artifact three.

The elementary principals identified several pieces that made artifact three acceptable. Brenda liked the fact that the teacher challenged the students in the unit. She said, "I also like that she's using categorizing and analyzing so she's using some higher

order thinking skills there.” Brenda describes the unit as “creative” and “extending beyond the curriculum.” She liked the pictorial representations, the use of technology, the use of group work, and the connection to the student’s portion of the product. In addition, Brenda particularly liked being able to view a student writing sample that had been graded via a rubric. The writing sample included positive comments and in-depth questions from the teacher written on the work. Brenda’s thinking in Phase I and in Phase II has been consistent concerning the value of questioning. The following comment summed up her positive judgment of the artifact,

She has here put together a presentation, a booklet with visuals, by the students. So they have taken the different aspects and done a pictorial with explanation of their projects, which is a good culmination. It brings closure to their unit and something they can look back on. It gives an explanation of each one, why they did it. It gives a tie-in with the students’ part in the product, as well as why the students did it, with what role did totem poles play in Native American life. So this is a ...it took time and it shows evidence of putting technology ...I’m seeing this was done on a computer, so it’s like giving the kids the opportunity with technology in the process. Very nice project.

Leo’s thinking was extremely positive during his review of artifact three. However, he still exhibited some hesitancy to closely review each criteria in each ITS that the teacher indicated she had met. He seemed somewhat in a hurry and “thumbed” through the pages of the artifact while saying,

Uses student performance data as a guide for decision making. Okay – where is that? I’ll kind of look for that as we go through. Did she, on the fly here or somewhere along the line, did she use data to direct the teaching that was going to take place in the classroom. Don’t know, but she said she did. Creates an environment of mutual respect for unfairness. I’ll just pull out a couple of them here. So it is ...without extensively just going back through and looking here, it’s hard to know if it does or doesn’t.

This was coded as a neutral judgment comment because Leo was simply thinking aloud as he quickly scanned the beginning pieces of the portfolio. The comment illustrated Leo's thinking in terms of how he processed through making a judgment about an artifact. It was as if he was still looking for something he could value from his perspective. Then, he came across the pictures and the newsletter. He immediately slowed down. He laid the artifact on the table and began to turn the pages slowly while he commented,

She did a little newsletter. That's good. A weekly, a volume two, issue seven. Effigy Mounds trip. That's good. Pictures from the mounds. I get a real sense of community here. I get a sense of a learning community. I get a sense of these kids. It looks like they are pleased to be where they are. It looks like they are learning. It's hard ...I see she took a picture standing behind the ranger here and I see the kids all looking at the ranger. I would doubt if she set that picture up. I'm looking at kids who do seem to be focused on learning in this particular field trip.

Leo, like Brenda, spent time reading the student essay samples and paid close attention to the teacher comments on the samples. His interest in the teacher comments on the sample was further proof of his keen interest in finding evidence of the teacher's relationship with the students.

Leo was also very positive about the peer observation the teacher included. The peer reviewer had completed the observation as a sort of "running records" report. The peer drew a line down the middle of letter-sized paper, making two columns. In one column the peer observer wrote the teacher's action, words, explanations, etc. In the other column, the peer wrote the student responses. It was a non-pictorial view of the lesson. At the end of the lesson, the peer wrote notes about what she thought went well and some portions of the lesson that could be improved. Leo's thinking, as he reviewed the peer

observation in artifact three, was consistent with how he thought about the formal observations conducted by the principal in artifacts one and two; he put great value on the observation pieces. As he viewed the peer observation he said,

It [the peer observation] also gives a sense of student involvement. We can see the level of student involvement. I can pick up on the extent to which the kids are focused on the teacher. That is a great artifact. What went well, what went less well. Great artifact. That is terrific. Far better than the cookbook pages that she has frequently put in that simply indicate what could be present in a lesson, whether it's social studies or whether it was in a Writers' Workshop sort of thing.

With great exuberance, Leo was able to find, in artifact three, the pieces missing from the other artifacts.

Norma was positive about the artifact as well. She appreciated the authenticity of the unit and the fact that the teacher had provided written prompts for the evaluator. "This is the unit that she has created over the six weeks, a comprehensive unit. Pictures of projects, student work and some artsy craftsy [SIC] thing. It's authentic. She does a good job of giving captions to give the evaluator more information about what she's doing here.

Artifact three was given a positive review by the elementary principals. There were not any negative comment counts for artifact three. The elementary principals identified positive pieces of the artifact as (a) the creativity of the unit and the fact it extended beyond curriculum, (b) the use of higher order thinking skills, (c) the use of feedback relative to assessment, (d) the use of technology, (e) the pictorial accounts of the field trips, the newsletters, and (f) the authenticity of the unit. For Leo, the relationship piece missing in the other artifacts was very apparent in artifact three. He dedicated more of his thinking to artifact three than to any of the other nine artifacts.

Artifact six was the only artifact to receive a negative review from the elementary principals. In other words, it was not acceptable in terms of how it illustrated/met the ITS the elementary teacher said it would. Judgment comments relative to artifact six were relatively low accounting for only seven percent of the total artifact judgment comments made by elementary principals in Phase II thinking. The lower counts were consistent with the tapering off for all principals for all artifacts after artifact three. However, the elementary principals were very clear that artifact six did not meet expectations. Only one comment of the six comments made was positive. There was one neutral comment. Four of the six comments (67%) were negative comments.

Artifact six was labeled Newspaper in Education. The teacher's reflection once again provided the best description of the artifact.

I started using the [newspaper] in my classroom in my first year of teaching and found that the students thrived on getting a chance to actually read the paper, discuss what they read, and could even go home and talk about it. I noticed many of the students who struggled in other ways found comfort in the paper, those who were already successful using other resources were now able to use even another one, and it was great exposure for all of the kids whose families didn't get the newspaper.

I asked [name omitted] from the [newspaper] to enter pictures of my students reading the paper in a contest for the marketing papers. The papers sent out to local schools and businesses to encourage Newspaper in Education. Low and behold, my students won and become the faces in the paper and on the forms! In addition, I was asked by [name omitted] to do a radio advertisement for The [newspaper].

The artifact included (a) the teacher reflection, (b) a written note from the contact at the newspaper, (c) the newspaper and forms with pictures of the students in/on them, and (d) an email from the radio station asking the teacher to sign a release for the radio spot. The teacher indicated that she met all or part of every ITS with artifact six.

The lone positive comment came from Brenda. Once again it was the teacher reflection that influenced Brenda's judgment. Brenda did not begin her review headed in a positive direction.

It's just basically 'I used this.' It's not a lot of things that she had to create separate, but it is saying that she picked it and she had a good reason. She had a goal for why she wanted to use it. She wasn't using it because the Courier provides these free of charge to schools. It wasn't one of those things that stacked up in her corner of her room. She had a purpose to benefit the kids. I think that this artifact, while minimal in what she had to do to pull it together, took over time; it was the process that she went through that made it good. She evidenced that in her reflection, why she did it and how it was beneficial. So I think this is a good artifact and her cover [reflection sheet] is good on that.

Leo did not exhibit any positive thinking about artifact six. In fact, one of his comments shared common language with Brenda concerning the basic premise of the artifact, i.e., "I used this." Leo said,

The documents that I'm seeing here are, again, implementation documents. These are all things that have simply indicated the fact that she has done this. So what's going on in my mind is that she has not indicated what has happened as a result of it. She doesn't have any student documentation here that is a follow through to indicate that yes, this has really reached out and got me.

Interestingly, after artifact four, Leo began to systematically check the individual criteria indicated by the teacher. His review began to be more specific. In the case of artifact six, he did not find evidence that the teacher had made connections between the artifact she included and the standards/criteria she indicated were met by the artifact.

For example, standard two 'Confident in content knowledge, understands and uses key concepts' – Don't see it. 'Uses knowledge of student development to make experiences' – I would guess that that's what this does, but she doesn't give me a sense that a student who is a poor reader would find a way to maybe connect with who knows. Cartoons or sports or whatever else in there. 'Relates information with and across content areas' – Again, I don't see any evidence. She proved to me that she got the newspapers to the door. She doesn't tell me what she did with them, so I can't see that. So I don't think that this one effectively

would not meet the standards that she has indicated here. I'd want to know more about implementation. That's 6.

Norma's thinking about artifact six was similar to that of Leo. She said, "But, this doesn't show me much in its current state. I don't know what that was supposed to be showing." Norma looked at the artifact again and attempted to squeeze something credible from the evidence but was unable to do so.

The project in her class may well do all the things that she said but her artifacts certainly don't show hardly any of it. I would say that the fact that she reflected that her kids were in the paper and it showed the kids reading the paper and it helped the kids feel good about themselves, it maybe showed a few things. I think she could have done a lot more with the artifacts or she shouldn't have claimed that we were going to show all of those things.

In general, artifact six did not provide enough evidence to garner positive thinking from the elementary principals as a group. It was the only artifact where negative thinking outweighed positive thinking. The concept of the artifact appeared to be acceptable to the principals but they wanted more evidence. Brenda was consistent with her thinking regarding reflection. Leo became more specific with his review. Nancy looked for evidence to make the artifact work but ultimately had to admit that the artifact was weak.

While the artifact judgment statements dominated the thinking of the elementary principals in the judgment category, the judgment categories of portfolio (J/P) and teacher (J/T) were also present. Brenda and Leo both made comments relative to judgment of the portfolio. Brenda first judged the portfolio in Phase I saying that it looked as if the portfolio was balanced and had enough evidence to support the ITS. However, her thinking changed as she moved to Phase II thinking. She became

concerned about the one-dimensional use of one unit (Native Americans) to evidence all standards.

At this point, I'm on number 7 on the artifact review. Every lesson that is being reflected back on, although in a different way and that's fine, it's always the same lesson. I'm getting a picture that we did a really good unit here at the beginning of the year. I don't know if we're doing any other good units, because it's all based on that one unit. I would know as administrator that this is happening frequently, that we've got these well designed units that are incorporating so many different things all the way throughout. Let's show. Let's evidence that, let's show that. Get a little variety in there.

Leo's thinking, in terms of portfolio judgment in Phase II, centered on originality and perspective. He, like Brenda, wondered about the repetitiveness used by the instructor. Leo eluded to a "cookbook style" in Phase I thinking but became more concerned in Phase II. After reviewing the first artifact during Phase II he said, "Then I'm not real comfortable with these pages that look like they came out of a how-to-do-it book. The problem with that is that doesn't tell me anything other than the fact that she copied it from somewhere and put it in her book. I have no sense of what she did with it." Prior to judging artifact four he showed concern for perspective "This pretty much is simply a pretty cut and dried look at the project [unit plan]. She has a tendency to do that on all of her artifacts, is to lean on or to look pretty strongly at the lesson plan as opposed to looking at it more from a child perspective."

Leo was the only elementary principal to judge the teacher in Phase II thinking. In light of his concerns with the portfolio, his judgment comment was somewhat surprising. "I would point out to the teacher is that I'm pretty sure from what I've seen so far for these four artifacts, I'd love to have this teacher working for me. I think it's a strong

teacher.” Interestingly, the comment followed a full review of the fourth artifact, a description of a fundraiser for the Gulf Coast.

In summary, judgment comments increased significantly in Phase II thinking of the elementary principals. Artifact judgment was the most dominating sub-category in the judgment category. The number of principal comments was remarkably similar in the sub-category of artifact judgment. With the exception of artifact six, the artifacts were judged positively by the elementary principals. Elementary principals engaged in very little thinking about judgment of the portfolio or the teacher in Phase II thinking.

Principal’s opinion. The category of principal’s opinion (PO), surfaced for the first time in Phase II for the elementary principals, accounting for less than two percent of the coded comments for Phase II thinking. The category consisted of two sub-categories; principal’s opinion regarding the portfolio of evidence of good teaching and principal’s opinion regarding the DE system for teacher assessment. Brenda was the only elementary principal to have comments in the principal’s opinion category in Phase II thinking. Brenda commented relative to her opinion of the DE system for teacher assessment. She indicated that the assessment system had changed during her tenure as a principal. “I guess the difference with the teaching standards and assessments and the way it previously had been done, is the evaluator had to show, identify and prove that the teacher was doing something. Now the teacher is proving that they are doing it and proving it to me.”

Tools. The category of tools (T) was also new in Phase II thinking for the elementary principals, accounting for less than two percent of the coded comments for

Phase II thinking of elementary principals. Only two comments relative to tools were made during Phase II thinking. Brenda and Leo each made one comment relative to thinking about tools they used during evaluation. Brenda simply indicated that she used sticky notes placed on artifacts to communicate with the teacher in a written form. Leo indicated that he does not use written tools to evaluate teacher portfolios. "I have simply done this off the cuff in a verbal kind of conversation with the teacher as opposed to having any kind of a paradigm or any kind of grid form."

In summary, the coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Elementary principals spent the bulk of their time judging artifacts. Overall, the principals judged the artifacts positively with the exception of artifact six. The elementary principals identified 11 pieces that contributed to their positive judgment of the artifacts: (a) observations by the teacher's principal and a peer, (b) samples of student work, (c) pictures of activities, (d) the use of technology, (e) detailed lesson plans, (f) well-written teacher reflections, (g) evidence of incorporation of higher order thinking in lesson plans, (h) rubrics, and (i) authentic assessments. Artifact six was judged negatively because clear connections did not exist between the artifact and the ITS it was designed to evidence.

The categories of principal's opinion and tools surfaced for the first time in Phase II. The principal who commented about the DE teacher assessment system simply indicated that the system had changed during her tenure and that the burden was now on the teacher to show evidence of good teaching. One principal indicated that she used sticky notes in the portfolio as tools of communication with her teachers.

Elementary Principals and Coaching Activities in Phase II

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time (TI). Comments relative to coaching activities accounted for 20% of the total Phase II comments made by elementary school principals. In this section, comment counts relative to two of the three categories (coaching and principal's role) will be reported followed by a summary. No comments were made relative to time.

Coaching. The coaching (C) category had the second highest comment count for elementary principals in Phase II. The coaching category was not present in Phase I thinking. Comments in the coaching category represented 18% of the total coded comments in Phase II thinking for elementary principals. The coaching category was a multi-level category. Table 17 illustrates the breakdown of the coaching category comment counts for elementary principals. While the coaching category was indeed a part of the rhythm developed by the principals, it did not represent a one-to-one ratio with judgment comments. Judgment was the focus of the principals but coaching during the judgment phase was also very important.

Brenda spent time thinking in all four of the coaching categories. Leo's thinking regarding coaching did not include questioning and Norma, once again, remained focused and spent her time thinking only about suggested alternatives.

Brenda and Leo both thought about meetings with the teacher (C/M). Brenda thought about the enhanced explanation a meeting with the teacher would afford. "So a conversation, when I meet with this teacher, would be to ask how she intends to use the

information.” Leo’s thinking was similar. “I’d say [to the teacher] ‘I’m not seeing that. Help me understand this because you’ve indicated a connection. Talk to me about it.’ Maybe she could talk her way through it.”

Table 17

Coaching (C) Comment Count Breakdowns for Elementary Principals in Phase II

	Total Coaching Comments	C/M	C/PP	C/Q	C/SA
Brenda	19	4	2	4	9
Leo	7	3	1	0	3
Norma	3	0	0	0	3

C/M = Coaching relative to meetings with the teacher, C/PP = Coaching relative to portfolio preparation, C/Q = Coaching relative to the use of questioning, C/SA = Coaching relative to suggested alternatives for artifacts

Brenda and Leo both spent time thinking about coaching relative to the portfolio process (C/PP). Brenda’s coaching was intended to keep the teacher focused. “Again, a good packet. I think I’m seeing a trend here that I would like to help her focus on. That is, I would like to see her keep in mind what her goal is and not get distracted by cutesy things and ask herself why she would use that.” Leo’s coaching, on the other hand, was more specific about how to begin construction of a portfolio. He said,

The structure that I use or I’ve advocated to teachers has been to go with a folder that, instead of being divided up by artifact one, artifact two, is to simply have a folder divided up with the eight standards. Then pull an artifact that goes with

those standards and then cross reference like this [referring to the portfolio provided to him by the researcher], but the organization is different.

Brenda was the only elementary principal to think in terms of questioning as a means of coaching. Her comments relative to questioning only numbered four; however, the intensity she exhibited regarding the use of questioning merited inclusion as a category. It was almost as if the teacher that created the portfolio were sitting in the room. Brenda was particularly interested in asking, and having teachers ask themselves, about student learning and how it was illustrated in the portfolio. "A good question is always did they learn?" Examples of other questions that Brenda posed were "How are you going to know you met that goal?" and "How do you know the student met their goal?" She used questions to challenge the thinking of the teacher and the purpose of the artifacts she included when she said "I think I might ask her what makes this one [visual organizer] a better one than this one for her 5th grade students? "If she were using both, what would be the results?" "Which one would get her closer to her goal?"

All three elementary principals engaged in thinking relative to coaching/suggested alternatives (C/SA). The suggested alternative subcategory accounted for 52% of the elementary principals coaching comments in Phase II. Even though the suggested alternative category dominated the coaching category, the thinking of the principals was clearly on judgment. The low emphasis by the elementary principals on suggesting alternatives reinforced that, while they thought about potential alternatives to the artifacts the elementary teacher included, they were very focused on judgment. In addition, the suggested alternative counts were low because the judgment decisions for all but one of the ten artifacts were positive, i.e., not a great need to suggest alternatives.

Table 18 illustrates the breakdowns of the number of suggested alternative comments, per each artifact, made by the elementary principals. The table clearly illustrates the small number of suggested alternative comments made by the elementary principals in Phase II thinking. Interestingly, the only artifact that was not reviewed positively (artifact 6) garnered only one suggested alternative comment. The number of judgment comments for the elementary principals ranged between 24 and 30 (Table 14) suggesting again that, although principals developed a rhythm in their thinking, they did not make coaching comments for every judgment comment; a 1:1 ration did not exist between the two categories.

Table 18

Suggested Alternative (C/SA) Comments per Artifact by Elementary Principals in Phase II

	Artifacts*									
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Brenda	1	1	3	0	1	0	0	1	1	1
Leo	0	0	0	1	0	1	1	0	0	0
Norma	0	0	0	0	0	0	1	1	0	1

* See Appendix H for Artifact Title and Cross-referencing with ITS

Brenda's thinking continued to be consistent regarding teacher reflection as she made suggested alternatives. She also continued to focus on teacher growth and how the

teacher should illustrate her growth. She used questions as a basis for her suggested alternative; another consistent thinking trait for Brenda.

I could see that putting all these things together, going on another page and just reflecting on how ...you know, where did she come in at because I think measuring your growth is saying 'well, how did I feel when I first came into the teaching profession and dealing with these different types of students and now how do I feel? How do I handle things differently now?'

Even though Brenda judged artifact three as positive, she did include suggested alternatives. She suggested including a clearer goal for the Native American unit. She also suggested that the teacher be clear about the level of student who completed the writing samples. While she liked the comments the teacher wrote on the graded student samples, she again suggested that the teacher be more specific with why the paper, or portion of the paper, was written well, i.e., more than encouragement.

Brenda and Norma both suggested that the teacher include descriptive captions on each of the pieces of evidence so they would be more easily identifiable. Brenda also suggested that the teacher do a spell and grammar check.

Leo's suggested alternatives were consistent with his previous thinking as well. He suggested that the teacher include more student generated reflection about the projects, units, etc. Once again, he was looking for the student perspective and the relationship evidence. His only suggested alternative for artifact six, the artifact of which he was most critical was "...if it's going to be an effective artifact for me, I need to see outcomes."

Principal's role. The category of principal's role (PR) appeared for the first time in Phase II and accounted for less than two percent of the total coded comments for the

elementary principals. Brenda was the only elementary principal to have comments in the principal's role category in Phase II thinking. Brenda thought about her role in terms of teacher growth and teacher responsibility in terms of producing the portfolio. She stated clearly that "...my job is to help teachers to grow and make them better at what they're doing. If you're [the teacher] exhausted at the end of this project [portfolio], and feel that it hasn't helped, then I've missed my goal with a new teacher." She continued by identifying the responsibility of the teacher and again clarifying her role.

It's the job of the teacher to identify the criteria. If this teacher were to say, "I'm short in an area" or if I were to say, "I'm short in this standard. Are there some things that you could help me evidence or directions I could go," that's fine. It's not my job to give a lot of artifacts and then go through and mark them. I mark them only to see if I've got a good cross-reference, how many things do I have to evidence this standard. But I don't identify them.

In summary, the three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time investment (TI). Only comment counts relative to two of the three categories (coaching and principal's role) were present in Phase II. No comments were made relative to time investment. The coaching (C) category had the second highest comment count for elementary principals in Phase II. The coaching category was not present in Phase I thinking. The coaching category was a multi-level category that included coaching via meetings, coaching on the portfolio process, coaching using questioning, and coaching via suggested alternatives.

As a group, elementary principals considered all of the coaching sub-categories; however, coaching via the use of suggested alternatives was the most significant sub-category in the overall coaching category. Suggested alternatives were part of the four-step rhythm that the elementary principals developed; but not always. Suggested

alternative comment counts did not match one-to-one with the artifact judgment counts probably because the judgment counts, by and large, were positive; thus, suggested alternatives were not as frequent as judgment comments. The elementary principals suggested that the elementary teacher: (a) provide additional clarity in reflection statements, (b) provide additional clarity of goal setting and proof out outcomes, (c) provide more variety of artifacts, (d) include information regarding student level of achievement relative to each student artifact, (e) include student reflection, (f) provide additional clarity when making student comments on graded work, (g) complete a grammar and spell check on her own written work before submitting the portfolio.

Only one principal made reference to her role as a principal. She indicated that she played a significant role in the portfolio process but that it was up to the teacher, for the most part, to identify appropriate evidence for inclusion in the portfolio.

Elementary Principals and Critical Pieces in Phase II

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two “pieces” carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. In this section, comment counts relative to the two sub-categories, critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO), will be reported, followed by a summary.

The critical pieces (CP) category had the third highest comment count for elementary principals in Phase II. Comments regarding critical pieces increased in Phase

II thinking. Comments in the critical pieces category represented 10% of the total coded comments in Phase II thinking for elementary principals. Table 19 illustrates the breakdown of the critical pieces category comment counts for elementary principals. All three principals thought about the role of observation and teacher reflection as critical pieces during Phase II thinking.

Table 19

Critical Pieces (CP) Comment Count Breakdowns for Elementary Principals

	Total CP Comments	CP/RO	CP/TR
Brenda	11	3	8
Leo	3	2	1
Norma	3	2	1

CP/RO = Critical Pieces/Role of Observation, CP/TR = Critical Pieces/Teacher Reflection

Role of observation. The role of observation was considered by each of the three elementary principals. Brenda and Leo both focused on a formal observation summary conducted by the elementary teacher's principal. Brenda thought about observation as something she would need to see for herself and did not greatly value the observation performed by the teacher's principal and included in the portfolio. "I see that she was observed by the principal and his feedback on that. I look at that. I also know that that's

not what I'm going to determine the quality [of the portfolio] on, just because that's not something that they created or produced, it's somebody else's opinion and this is you proving to me why this is evidence." Leo read the entire observation summary (nearly two pages) aloud but did not comment on its significance relative to the portfolio; however, the amount of time he spent reading the summary provided evidence that the observation by the principal was of importance to him.

All three elementary principals commented on a peer observation that was included as an artifact in the portfolio. Brenda did not place great value on the peer observation but thought it was a good idea. "This [peer observation] is good feedback for her. It tells how she ran her lesson again. I don't hold these things as high. I think they should be included in there, but if ...observed by a peer ...so it ...I mean, this is your evidence from someone else other than your observer and I think that's good that they did that." Leo and Norma placed much more value on the peer observation. Leo remarked that, "I'm looking here at a script. It looks like it's an evaluation script and it looks to me like it was done ...well it says it was a peer observation. So that's good. That is really good because it gives a strong sense of actually what is taking place in the classroom."

Norma thought about peer observation as a way for everyone involved to learn. "So she also makes use of having one of her colleagues give feedback. Since it's not required, it shows nice initiative. I've always thought that teachers maybe learn more from that than they do from having the evaluator go in and watch." Norma, while she valued the peer observation, thought along the same lines as Brenda in terms of using observations by someone else to judge the portfolio. Norma explained her need to

directly observe the teacher when she said, “I know when I do my own [portfolio evaluations], I look through and say ‘Oh yeah, I was in there. I saw her doing this.’ So you know what it is. But for someone to read it [the portfolio] who doesn’t know them...”

Teacher reflection. The second sub-category of critical pieces was teacher reflection. All three principals thought about teacher reflection (refer to Table 17). However, Brenda dedicated eight times the thinking to teacher reflection as did Leo and Norma. Both Leo and Norma noted that the teacher had included some reflection.

Brenda’s Phase II thinking relative to the importance of teacher reflection was consistent with her thinking in Phase I. The value Brenda placed on reflection in Phase II was best illustrated when she said, “What I would like to see and what is the most important is ...okay, an overall description is given here. I’d like to then see a reflection. In other words, I want some more reflection. I want you to *prove* to me, rather than my figuring it out why this [artifact] fits with this [ITS/criteria]. I don’t want to have quite so many questions about the different pieces.”

In summary, the category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. The elementary principals thought that including observation summaries was a good idea. Brenda and Norma did not place as much value on the observations completed by others as they would on their own observation. Leo spent a lot of time reading the evaluations and placed higher value on them than did Brenda and Norma. Teacher reflection was noted

by Norma and Leo. Brenda thought of reflection as a way for teachers to assist her in making quick, clear connections between the artifacts and the ITS they represented.

Summary of Elementary Principals' Thinking in Phase II

Phase II (judgment) thinking was described as the period of time following pre-assessment thinking when principals actively judged the artifacts in the portfolio that was provided to them by the researcher. Processing activity drastically declined in Phase II as the elementary principals focused away from portfolio structure and towards their roles of judge and coach. The principals spent the biggest share of their time judging the first three artifacts. Their judgment thinking then tapered off.

As they began to review the individual artifacts included in the portfolio, the principals developed a four-step rhythm. First, they would identify and verbally describe the artifact. Second, they would actually read aloud as they focused on what the teacher was attempting to illustrate. Third, they would make a judgment statement regarding the artifact they were judging. The judgment statement was then sometimes followed by a coaching statement. The confidence level of the principals increased as they moved into Phase II.

Overall, the elementary principals judged the artifacts positively with the exception of artifact six. The elementary principals identified nine pieces that contributed to their positive judgment of the artifacts: (a) observations by the teacher's principal and a peer, (b) samples of student work, (c) pictures of activities, (d) the use of technology, (e) detailed lesson plans, (f) well-written teacher reflections, (g) evidence of

incorporation of higher order thinking in lesson plans, (h) rubrics, and (i) authentic assessments.

Coaching/suggested alternative comments were not at a 1:1 ratio with judgment statements. However, the elementary principals did spend time suggesting alternatives as they judged. In general, they suggested that the elementary teacher: (a) provide additional clarity in reflection statements, (b) provide additional clarity of goal setting and proof out outcomes, (c) provide more variety of artifacts, (d) include information regarding student level of achievement relative to each student artifact, (e) include student reflection, (f) provide additional clarity when making student comments on graded work, (g) complete a grammar and spell check on her own written work before submitting the portfolio.

Comment counts relative to meeting with the teacher accounted for nearly 25% of the coaching comments made by the elementary in Phase II. This percentage was second only to coaching/suggested alternatives in the coaching category. The principals emphasized that the purpose of the meetings was to gain clarification and understanding regarding the artifacts the teacher had included in the portfolio; not a time to criticize.

Comments regarding the critical pieces of teacher reflection and the role of observation increased. The principals valued the reflection comments made by the teacher; however, they indicated that they would like to have had more reflection included. It also became clear in Phase II that observation was more highly valued than a portfolio.

Middle School Principals' Thinking in Phase II

The middle school teacher's portfolio was organized by standard. She included sections representing each of the eight standards. Each section contained artifacts representing a respective standard. It was not necessary that she use a cross-reference guide because the middle school teacher included, on an artifact cover page, the multiple standards/criteria represented by each artifact. See Appendix J for a sample of an artifact cover page. The middle school principals evaluated the artifacts contained within each ITS section and made judgment statements about each artifact.

As was true in Phase I thinking, the middle school principals (Ivan, Mike, and Rob) made the most Phase II comments of the three academic levels (see Table 12). Their thinking accounted for 48% of the overall comments made by all principals in the Phase II as compared to 21% for elementary principals and 31% for high school principals.

Table 20 illustrates the breakdowns of Phase II comments made by middle school principals. There were seven categories common to Phase I and Phase II thinking for the middle school principals: (a) coaching (C), (b) comfort level (CL), (c) critical pieces (CP), (d) portfolio structure (PFS), (e) principal's opinion (PO), (f) process steps (PS), and (g) time (TI). Two new categories emerged in Phase II thinking of middle school principals. They were judgment (J) and tools (T). The middle school principals' thinking did not include principal's role in Phase I or Phase II. Although Ivan did not engage in Phase I thinking, he had the most comments in Phase II. His thinking accounted for 42% of the total Phase II made by middle school principals. Mike and Rob had significantly more Phase II comments than they did in Phase I; each accounting for roughly 29%.

Table 20

Summary of Phase II Comment Counts made by Academic Level

	Total Comments	C	CL	CP	J	PFS	PO	PS	T	TI
Ivan	157	27	0	13	73	2	6	26	1	9
Mike	109	18	0	8	65	1	3	9	5	0
Rob	111	33	1	0	58	1	4	9	5	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PS = Process Steps, T = Tools, TI = Time

Middle School Principals and Processing in Phase II

The three coded comment categories that clustered under processing included the categories of portfolio structure (PFS), process steps (PS), and comfort level (CL). Comments relative to processing accounted for 13% of the total Phase II comments made by middle school principals. In this section, comment counts relative to each of the three categories (portfolio structure, process steps, and comfort level) will be reported followed by a summary of middle school principals and processing.

Portfolio structure. The three principals made half as many comments relative to portfolio structure (PFS) in Phase II as they did in Phase I; an indication that they were becoming familiar with the structure. Portfolio structure comments in Phase II accounted for less than two percent of the total Phase II comments. Ivan wished that the teacher would have numbered the artifacts for easier cross-referencing. He also commented that,

“I can see the pattern of how she set up her portfolio and that’s personal preference.”

Rob’s comments reinforced his earlier position that using one artifact for multiple standards was acceptable. In addition, he, like Ivan, had begun to become familiar with the structure. “I can see that she’s starting to be descriptive of the artifacts. I guess I like the format that’s being used. I see that one thing I like in regards to this is that it simplifies it in terms of the administrator.”

Mike began to draw parallels between the portfolios he normally reviewed and the one presented to him for purposes of the study. He read aloud from an artifact designed to evidence ITS 1g-Communicates with students, families, colleagues, and communities effectively and accurately.

I [the teacher] use sub notes for when I’m absent so it’s easier on the sub. I was a sub at one time. I realize the pressure coming into the environment, especially math. I like to leave copies of lesson plans, answer keys, and leave a general rule sheet and a list of students.

Then, he indicated that, “Something like this would be included in most of the portfolios that I look at because it is a good communication tool.”

Process steps. The process steps (PS) category had the third highest comment count for middle school principals in Phase II. The number of comments relative to process steps increased from Phase I to Phase II. Comments relative to process steps accounted for 12% of the total coded comments in Phase II thinking for the middle school principals. The middle school principals made more comments relative to process steps in Phase II than either of the other two principal groups (see Table 12). Ivan contributed the bulk of the comments relative to process steps accounting for nearly 60% of the comments. Mike and Rob each accounted for roughly 20%.

The middle school principals' thinking in the process steps category included cross-referencing, anticipation, prediction, self-talk, and writing a formal evaluation. Ivan's process steps were mostly related to cross-referencing and his process for doing so. Ivan brought a copy of the ITS/criteria (Appendix R) with him to the review and it was beside the portfolio binder as he read aloud. His copy proved to be a significant tool as he reviewed the artifacts. For instance, as he reviewed ITS number one, he thought about teacher observations and remarked, "So I make little marks on mine [his hard copy of the ITS] and throughout the year I would have observation. I'd write 'walk through.' I'd write 'artifact' so that I would keep track so I know what the differences are." His remark concerning observation was closely followed by this remark, "I'm just going to make a little dot so I'll put Xs here for each one that they do, so when I go through in the end I can say, 'Oh yeah, we hit 1A once or twice or whatever' without having to take hours to go through these and go back and look at them." Ivan's X's were a physical form of record keeping for him. He was very methodical but was also concerned with the time he could save via cross-referencing and Xing. While still judging the first ITS, he said,

And to be honest, I'm gonna slow down on my cross referencing here and go back and cross reference later if I'm short. I've got five more of these over here to do, so I probably won't go into that much detail. What I've found is I go into a lot of detail with the first one early on and then I start bogging down and looking at my watch and saying, "Okay, if I've got enough to do it, I'm just gonna go ahead and bare bones it. Let's finish it." I would go through this with them and verbalize like I am now, though, more than the cross referencing. This is a good one. So far I'm liking that.

Ivan's quote provided some insight as to why comments pertaining to judgment and coaching/suggested artifacts clustered around the first three standards. His thinking process paralleled that of the other two middle school principals and his elementary

colleagues as well. The principals seemed to use the first several standards/artifacts as a sort of a “gauge” for the rest of the standards and, consequently, there were more comments in the judgment and coaching categories for the first three standards.

Ivan was mentally and physically “darting” from one standard to the other on his hard copy. He wanted to ensure the he “Xed off” every possibility and seemed to be challenged to do so. Again, it was quite obvious that he had the standards and associated criteria memorized. Although Ivan indicated that he was going to “slow down with the cross-referencing,” he remained consistent as he moved very methodically through the seven remaining standards. However, he needed to take a break while still working through the first standard. “Actually what’s going through my mind is that I need a break. I am starting to space off and not focus. But what I will do is, I will stumble through this last one here because I’m not gonna take a break until I get through artifact one. But I thought I’d tell you what went through my mind as I was looking at this.” At another point, Ivan got a bit anxious about cross-referencing.

2B – Uses knowledge of student development to make learning experience meaningful and accessible to every student. Certainly. We can certainly cross reference this one down to 3 and 4. ...research based instruction strategy. Well,ncing and had to coach himself to be patient and again, stay focused. it’s probably not research-based. Well, it is. The singing along. But I’ll wait. I want to see what she finds for those, so we’ll wait and see what she’s got. We’re gonna stay on task.

Ivan continued to work through the portfolio with the same diligence and excitement as illustrated in the previous quotes. His comments relative to process steps were consistent as he reviewed the remaining artifacts.

In Phase I, Mike described the four steps he used to prepare himself to review one of his own teacher's portfolios. In essence, he framed his task and anticipated how he needed to perform it. He indicated that he (a) thought about the teacher who produced the portfolio, (b) performed a brief overview of the portfolio he was reviewing, (c) did an inventory check with himself to make sure he was focused on the standards, and (d) made sure he had a copy of the standards in front of him and he reviewed those as well. Mike's processing in Phase II was similar to that in Phase I in that he framed and anticipated what he thought each artifact might and/or should represent. A good example of this anticipatory processing occurred as he began his review of ITS number two. Before he even looked at the artifacts he said aloud, "So this [ITS number 2] is going to get into the curriculum and the instruction and do you know your math stuff. The teacher understands and uses key concepts and themes and perspectives tied to content. So, I'm going to be looking for that." As he began his review of ITS number six, he anticipated and predicted. He commented that, "Number six is normally an easy one for teachers in that it deals with classroom management. So many teachers start with syllabus or have rules posted or involve students in generating expectations for the class." Mike, like Ivan, occasionally had to refocus himself. He accomplished this via self-talk, aloud. As he completed a second review of artifacts relative to ITS number 5 he said, "How does this teacher communicate the assessment and the standards? Now, let's see if in looking at this, this time, I can focus more on any kind of varied assessments or whether it offers that."

Rob also used anticipation in his processing. His thinking paralleled that of Mike's when he said, "Then, as I move along, I see an artifact...I say, 'What am I going to be seeing?'" Evidence was also present that Rob's process steps in Phase II were a continuation of his lone Phase I comment. In Phase I, Rob indicated that he reviewed and processed through the portfolio with his formal written evaluation in mind. During Phase II, he described not only how he moved through the portfolio review process, but how, what, and when he would begin to write his evaluation of the teacher.

What I do is kind of jot some notes down from 1A through 1G in terms of tipping me off about what I've seen in this notebook so that I can then eventually start to write down each one of those as we go along. Typically, and I'm pretty close to this, I would be looking at this very early in the school year to see what I could get out of it and probably the first time during the week that we have Thanksgiving break, I would begin actually writing my narrative on each of the eight standards. Typically just the technicalities of it, I would probably put it in a different font, different color ink.

In a sense, Rob used anticipation for two reasons: (a) as a guide for what he might find in an artifact and (b) in preparing to write his formal evaluation, i.e., he anticipated the task of writing the formal evaluation and kept it in mind during observation and during portfolio evaluation.

Comfort level. The number of comfort level (CL) comments decreased to one comment in Phase II; a clear indication of the growing confidence of the middle school principals. Comfort level comments in Phase II accounted for less than two percent of total Phase II comments. Comments in Phase I centered on uncertainty due to lack of familiarity with the teacher, her district, and the portfolio structure. That uncertainty dissipated in Phase II as was evidenced by Rob's comment that, "So I have a little bit of comfort here. A regular ed. math teacher. This is more normal." Rob's confidence in

Phase II was bolstered because he was evaluating a teacher with whom he was not familiar in a discipline and classroom with which he was familiar.

In summary, as the middle school principals moved into Phase II of the portfolio review, they became more familiar with the structure and began to recognize artifacts that were similar to those of their own teachers. As a result, the number of comfort level comments decreased.

Comments by the middle school principals revealed that their process steps in Phase II included cross-reference-thinking processes, anticipation, self-talk, and consideration of writing the formal evaluation. It was clear that the principals had memorized the ITS standards/criteria and could mentally align artifacts with standards/criteria rapidly. In addition, one of the middle school principals used a hard copy of the standards/criteria to track which of the standards/criteria had been met and which of them he still needed to locate.

The middle school principals were far more graphic in terms of process than the elementary principals, thus, had higher comment counts regarding processing. Both groups (elementary principals and middle school principals) used the first three standards and/or artifacts as a “gauge” for judging the remaining five standards and/or artifacts.

Middle School Principals and Judging Activities in Phase II

The three coded comment categories that clustered under judgment activities were judgment (J), principal’s opinion (PO), and tools (T). Comments relative to principal as judge accounted for 58% of the total Phase II comments made by middle school principals. In this section, comment counts relative to each of the three categories

(judgment, principal's opinion, and tools) will be reported followed by a summary.

Judgment. Judgment comments accounted for the highest comment count in Phase II. In fact, judgment comments accounted for 52% of the total coded comments made by middle school principals in Phase II thinking; similar to the 54% made by elementary principals and 51% made by the high school principals in Phase II. The middle school principals did not engage in judgment thinking in Phase I.

The judgment category was a multi-level category that represented judgment of artifact (J/A), judgment of the portfolio (J/P), and judgment of the teacher (J/T). Table 21 illustrates the breakdown of the judgment category counts for middle school principals in Phase II. The middle school principals each considered all three of the judgment sub-categories. The judgment of artifacts (J/A) clearly dominated the thinking in the judgment category for middle school principals just as it did for elementary principals.

Table 21

Judgment (J) Comment Count Breakdowns for Middle School Principals in Phase II

	Total Judgment Comments	J/A	J/P	J/T
Ivan	73	53	7	13
Mike	65	62	2	1
Rob	58	44	4	10

J/A = Judgment comments relative to the artifacts, J/P = Judgment comments relative to the portfolio, J/T = Judgment comments relative to the teacher

Middle school principals employed a similar four-step rhythm when judging artifacts as did the elementary principals. First, they would identify and verbally describe an artifact. Then, they would actually read aloud as they focused on what the teacher was attempting to illustrate. Their third step was a judgment statement regarding the artifact they were judging. The judgment statement was then sometimes followed by a coaching statement. In addition, middle school principals used *cross-referencing* as they judged artifacts. When they judged a specific artifact, they would make mental and/or physical notes that the artifact could meet more than one ITS standard/criteria.

An example of the use of cross-referencing was demonstrated by Ivan as he judged an artifact relative to ITS number one. The artifact he was reviewing was described by the middle school teacher on her cover page for the artifact when she said, “In math, we have a section on the order of operations. During this section, the students create a poster helping to describe the order of operations or PEMDAS.” The teacher included pictures of several of the posters created by the students. The teacher’s reflection indicated that, “The students enjoy making the poster because they get to use creative ways to help them memorize the order of operations. It helps with scores on tests because they can think of the idea they thought of on their own.” The teacher indicated that she was meeting ITS 1a and 1b (Appendix A). First, Ivan judged the artifact.

I can certainly see 1B. This is definitely a good strategy to support the student and the building. I’d like to see what the district goal was. I’m sure it’s part of their math standards. I’m assuming that the evidence of this to the family and staff is that they have these posters to take home and then they can show them how they did on the quiz or the test that related to this. So that’s not bad. That’s pretty good.

Then, Ivan considered another ITS that the artifact might have evidenced.

I like the fact that they were doing ...they created a poster helping describe the order of operations. Let's go down to 4d. *Engages students in a variety of experiences that meet the diverse needs and promote social, emotional and academic growth.* To me, having them create a poster and do that abstract type thinking of math certainly hits 4d, so I would add 4d to this first artifact that she's got.

Mike used the same sort of cross-referencing to evaluate ITS number two. The teacher indicated that she was attempting to show evidence of ITS 2d. The artifact was described by the middle school teacher on her cover page for the artifact when she said,

After a chapter on graphing functions, using slope, and problem solving, I have the students do a paper towel experiment. The purpose of the experiment is to find out which paper towel is best to buy. Students have to use the knowledge they learned in the chapter and some scientific knowledge to figure out which one is the best and justify their reasoning.

Mike first judged the artifact. "So, this is interdisciplinary. It's active student learning. It's relevant. She just chose to use 2d 'Understands and uses strategies that are appropriate'." Then, Mike makes a mental note that the artifact, at least for him, could be used to evidence other criteria in ITS number two.

This really, in my opinion, could be used for each one of the criteria in the second [standard]. It's across curriculum and it includes science-related stuff, learning experiences that are meaningful and its applied learning and higher level thinking and key concepts and themes. Absolutely. This one single artifact could have been used for the entire standard.

The use of cross-referencing was employed by the middle school principals consistently throughout their comments relative to artifact judgment.

By further breaking down the judgment of artifacts (J/A) category, the comparability of judgment comments, per standard, by each principal became even more clearly illustrated. Table 22 illustrates the breakdowns of the number of artifact judgment comments, per each standard, made by the middle school principals.

Table 22

Judgment of Artifact (J/A) Comments per Standard by Middle School Principals in Phase II

	Iowa Teaching Standards (ITS)*							
	#1	#2	#3	#4	#5	#6	#7	#8
Ivan	20	12	8	3	2	3	4	1
Mike	11	9	14	6	10	4	7	1
Rob	12	6	6	4	6	2	3	5

* See Appendix A for ITS and Criteria

Middle school principals exhibited the same kind of trend when judging artifacts as did the elementary principals in that the bulk of their thinking occurred as they judged artifacts that represented the first three standards. In fact, over half of each of the principals' thinking efforts occurred as they judged artifacts representing ITS one, two, and three. After artifacts representing ITS three were judged, the number of comments by each principal, relative to each ensuing ITS, tapered off.

The same rating system used to classify the artifact judgment statements of elementary principals was used to tag the artifact judgment statements of the middle school principals. Each artifact judgment statement made by the middle school principals was evaluated and tagged with a positive (+), neutral (-), or negative (-) rating.

Positive judgment statements included those statements that indicated the artifact, per the principal's judgment, had properly illustrated/met the IT'S the middle school

teacher indicated it would. An example of a positive artifact judgment statement was made by Ivan as he worked his way through the evaluation of artifacts relative to ITS one-Demonstrates ability to enhance academic performance and support for implementation of the school district's student achievement goals (see Appendix A). The artifact he was evaluating was a quarterly student summary report indicating homework completion information/scores, quiz scores, and test scores that was distributed to parents. Middle school students had to have parents sign the report and return it to the teacher. The teacher indicated that she was meeting ITS 1a-Provides evidence of student learning to students, families, and staff.

...this looks like it's printed right out of their grading system, which as would ours be. Excellent. This is really good. This is the type of thing that's easy to do. It should be done. It's definitely 1A. Now you're getting evidence to the parent. Especially the part that it's signed and returned. Even if it's not signed and returned, it would still be a good artifact. I think that raises the value of the artifact because then you know the parent is signing it. Assuming the kid doesn't forge it.

Neutral statements included those statements that the middle school principals made while judging an artifact, but the statements did not indicate that the principal had made a definitive judgment as to the value of the artifact. An example of a neutral artifact judgment statement was made by Mike as he worked his way through ITS five-Uses variety of methods to monitor student learning (see Appendix A) when he said, "We have not seen this [artifact] yet. This is being used for multiple assessments." Negative artifact judgment statements included those statements that indicated the artifact, per the principal's judgment, did not properly illustrate/meet the ITS the middle school teacher said it would. An example of a negative artifact judgment statement was made by Rob as

he worked his way through ITS number five-Uses a variety of methods to monitor student learning. The artifact that Rob was reviewing was an email the middle school teacher sent to the resource teacher in her building concerning home work for a student with special needs. The middle school teacher indicated that she was specifically meeting ITS 5f-Works with other staff and building and district leadership in analysis of student progress.

...it's more of an email regarding the homework. I guess you gotta kinda blend it to everything in this notebook. If I just take this by itself it's pretty mysterious. Not much depth. Looks like this is the opportunity for special ed to give some feedback to the teacher or to the parent. But standing alone by itself it's a little confusing or shallow I guess.

Table 23 represents the results of tagging each of the middle school principals' artifact judgment comments for each ITS in Phase II. Overwhelmingly, the middle school principals' judgment of the standards was positive. Positive comments represented 84% of the artifact judgment comments made pertaining to the eight standards evidenced by the teacher in the portfolio. Neutral and negative comments represented roughly eight percent each of the artifact judgment comments.

The principals were consistent in their positive judgment comments regarding the artifacts the teacher used to evidence the eight ITS. The artifacts used to evidence each of the standards were, by and large, judged as appropriate in meeting the eight ITS. The language the middle school principals used while judging the artifacts was remarkably similar. For instance, all three principals were exuberant when judging an artifact evidencing ITS 3b—Sets and communicates high expectations for social, behavioral, and

Table 23

Classification of Artifact Judgment Comments per Standard made by Middle School Principals in Phase II

	Total Judgment Comments	<u>Ivan</u>			<u>Mike</u>			<u>Rob</u>		
		(+)	(N)	(-)	(+)	(N)	(-)	(+)	(N)	(-)
Standard 1	43	19	0	1	10	1	0	9	1	2
Standard 2	27	9	0	3	7	2	0	6	0	0
Standard 3	28	7	0	1	7	6	1	6	0	0
Standard 4	13	3	0	0	5	0	1	4	0	0
Standard 5	18	2	0	0	7	1	2	6	0	0
Standard 6	9	3	0	0	4	0	0	2	0	0
Standard 7	14	4	0	0	6	0	1	2	1	0
Standard 8	7	1	0	0	1	0	0	3	1	1

(+) = Principal made positive statement about value of standard, (N) = Principal made neutral statement about value of standard, (-) = Principal made negative statement about value of standard

academic success of all students. The teacher described the artifact on her cover page by saying that, "When a student receives a detention, I send a note home with the student to have the parents sign. The note gives the date of the occurrence, violation, date the detention is to be served and a portion for the parents to sign." The teacher included a template of the note as further illustration of the artifact. Ivan said, "There's probably more details that can go with the routines and things in that classroom and how she sets it

up. But this certainly talks about sending the note home with the parents is certainly a piece of that. Good. I like that!” Mike had this to say about the same artifact. “Okay. This is another 3b – the communication of expectations. Social, behavioral, academic... That’s great! So that would seem appropriate. It is a notice of detention. It’s good.” Rob was also very positive in his judgment of this artifact. He said, “Then she goes into talking a little bit about detentions, 3b. I can tell already that this is accurate. She hits it on target.”

The middle school principals were not consistent in their negative artifact judgment comments. Negative comments represented only eight percent of the total artifact judgment comments made by middle school principals. The negative judgment comments were spread out across the eight ITS. In other words, there was not a particular standard (s) that was judged as negative by all three principals. The only standards that did not receive any negative comments were standard six and standard eight.

To more accurately describe the inconsistencies in the negative judgments comments made by the three middle school principals, it was valuable to specifically examine the negative judgment comments for ITS number one. ITS number one was chosen for two reasons: (a) it garnered the most overall judgment comments (27%) of the eight ITS sections, and (b) it had the highest count of negative judgment comments (three comments). It is important to keep in mind that ultimately, ITS number one was positively judged as acceptable by all three principals. In fact, 93% of the artifact judgment comments made for ITS number one were positive. There were only three negative comments relative to ITS number one; Ivan made one of them and Rob made

two. However, the negative comments made by Ivan and Rob were not relative to the same artifact.

Ivan's lone negative comment was relative to an artifact the teacher included as she sought to provide evidence of ITS 1d-Accepts and demonstrates responsibility for creating a classroom culture that supports the learning of every student. The artifact was titled "My Job, Your Job." The teacher explained the artifact on her cover sheet by saying,

I use a My Job, Your Job worksheet when I notice the students getting frustrated or when things are getting overwhelming in the classroom. Students have to list what jobs they have in the classroom, what jobs I have in the classroom, what jobs they do not have in the classroom, and what jobs I do not have in the classroom; none of which can be duplicated.

The teacher included a worksheet in the form of a graphic organizer. It had four boxes on it that were labeled with My Job Is..., Your Job Is..., My Job Is Not...and Your Job Is Not. It had been produced by School Administrators of Iowa (SAI). Ivan surmised that the teacher was having some discipline issues.

So I'll make an assumption this is a survival guide ...I'm gonna ...it's something I'm gonna say that the principal gave her, she went to them [the principal] because she was frustrated with the kids' attitudes and that they were getting moany about having to do that or this or not wanting to work. I can read a lot into that because that tells me I have someone who's willing to communicate with the administration, wants to be a good teacher. I'm going to guess there was some discipline in there and control was part of the problem here with classroom management. Whether she went to him or he went to her doesn't matter.

While Ivan appreciated the actions of the teacher, i.e., approaching the administrator and including the worksheet template, he negatively judged the artifact.

The value isn't great...because it's not there. I would have really liked to have seen a completed student piece. I think that would have been very valuable. For one thing, this doesn't prove she used it. Okay, I got it; I threw it in my portfolio.

So the value of this detracts greatly when there's no student data. I'd like to see what she wrote as her job and what her job was not. I'd like to see what the kids wrote. I'd like to see her answers to these, as well as the students. I think that really detracts from the artifact because anybody can photocopy this and turn it in.

This was the only negative comment made by Ivan relative to ITS number one. He very much liked the concept but needed more evidence of actual use of the worksheet, specifically an artifact produced by a student.

Both Mike and Rob were complimentary of the My Job, Your Job artifact. Mike said, "I'm familiar with My Job, Your Job and its good stuff." Rob was not as familiar with the concept/worksheet but he was positive about the artifact.

Next artifact – my job, your job – this is 1d. Definitely fits. It's trying to support a classroom culture in terms of learning. It looks like this is for somebody who just gets overwhelmed in the classroom. Students have lists of what jobs they have in the classroom. It's kind of an interesting thing I guess. The My Job, Your Job worksheet. This would be kind of new to me. It comes from SAI but I haven't seen this before. It certainly fits 1D.

The three principals differed in how they judged the artifact; however, the principals' familiarity, or lack of, with the concept of "My Job, Your Job" was not the deciding factor in whether or not the artifact evidenced the standard/criteria (1d).

Rob's negative comments for ITS number one centered on an artifact that was an attempt on the teacher's part to evidence ITS 1e-Creates an environment of mutual respect, rapport, and fairness. The teacher described the artifact on her cover sheet, "Last year I taught eighth grade algebra. Several students have come to me and said how much it helped this year in geometry." The teacher's reflection described the interaction. "I had three students (Tiffany, Tyler, and Mike) list all the things that they remember from class last year. When all three of them got together, it was neat to listen to them and all the

ideas they remembered from the previous year.” The artifact the teacher included was the list that the students created of the items they could remember (e.g., PEMDA poster, song for quadratic formula, Legos, math songs).

Rob was concerned about the value of the artifact the teacher included for evidence of ITS 1e. He said,

This kind of goes with the situation where people can send an email complimenting a teacher. In this case it's more of students have come back and said what they learned in Algebra was so helpful in Geometry. To be honest ...my first thought is “what kind of documentation is this? Is this just word of mouth?” I need to probably go on a little bit, but that would be the first thing that crossed my mind – is prove it.

Interestingly, Rob did find some value in the artifact but ultimately, he needed more.

This is ...Tiffany, Tyler and Mike are writing down some things. It's kind of like notes taken on how do you remember some of this. They talk about the posters. They talk about songs in here. They must have some interesting ways of trying to memorize things. Definitely a good thing. Again, my first thought is I might try to ask a little bit more about this.

Ivan and Mike did not take issue with the artifact the teacher included for ITS 1e; further illustrating the inconsistent nature of the negative judgment comments of the three principals. However, Ivan talked himself in to his positive judgment. He read from the portfolio (italicized) and made the following comment.

1e – Creates an environment of mutual respect, rapport and fairness. How does this ...maybe this is going to be some outside of the box thinking. Why is this mutual respect, rapport and fairness? Obviously there's a rapport between the kids and her if they were willing to come up in the hallway and talk to her and then come in and listed things for her. There's certainly rapport. There's no doubt about that and there has to be some respect as well. So that's actually a good one. That's a good one.

Mike's judgment of the artifact was also positive. He too read from the portfolio as he judged the artifact.

Okay. I.e, which is what was left hanging – *Creates an environment of respect, rapport and fairness*. So she's got this too. *Last year I taught 8th grade algebra. Several students have come to me and said how much it helped them in geometry.* So this looks to be a student artifact that I'm gonna be getting into. This is not only helping me on this particular artifact, but it is also helping me see some of the strategies and activities that are a part of the class.

While the principals were inconsistent about which artifacts they judged negatively in ITS number one, the thing about which they were consistent was the need for more information. Both Ivan and Rob seemed to *want* to judge the artifacts positively and indicated that they probably would have done so with more information and clarity.

While artifact judgment statements (JA) dominated the thinking of the middle school principals during judging activities for Phase II thinking, the judgment categories of portfolio (J/P) and teacher (J/T) were also present (see Table 21). Judgment of portfolio comments accounted for seven percent of the total judgment comments made by middle school principals in Phase II. The judgment of teacher category represented 12% of the total judgment comments made by the middle school principals in Phase II.

Ivan made the most comments pertaining to portfolio judgment in Phase II. In fact, he accounted for 54 % of the total portfolio judgment comments for middle school principals. Rob accounted for 31%. Mike had the least amount of portfolio judgment comments accounting for only 15% of the total.

Ivan and Rob liked the fact that the teacher included artifacts that could be cross-referenced and used to evidence more than one ITS. Ivan commented that, "That's what I like about this [portfolio]. Yeah, she has reused some of the things but it's not a four inch binder of stuff. The four-inch, three-inch binder, it just puts up a little red flag." Rob had a similar comment. He said, "Perhaps someone from the outside world might say that

some of this is redundant. But I think in some ways when you're ...particularly if you've got a really good artifact, it is going to hit more than one standard and may be kind of a repeat of something you've done before. I'm okay with that." Interestingly, two of the elementary principals, Brenda and Leo, had a very different view of the kind of cross-referencing that Ivan and Rob complimented. They were critical of the fact that the elementary teacher used one lesson or unit to evidence several standards.

Ivan and Rob developed a trust for the teacher via reviewing the portfolio. Their comments revealed their trust. Ivan provided insight into his trust for the teacher and how it influenced his judgment of the portfolio when he said,

With this one, everything is flowing well. I think there are certain portfolios, whether that is right or wrong, but there's times when you're more nit picky. If they're lacking and don't have a lot of standards Xed [cross-referenced] and a lot of things are missing, then you're not going to be as liberal probably with reading inferences into these things and you're going to sit down and say, "Explain this to me." But in the sake of time, the teachers you have observed and know are doing the quality job; you are not going to take that time. Yes, I call them in and I'm going to have a conversation with them anyway, but I'm not going to take all the notes and document it all and cross all the Ts and dot all the Is. Maybe that's not right, but I think it's the reality of the job. If there's any chance of removing a teacher who isn't doing as good, that's how you have to work with that. I'm not seeing that by any means with this one. I think we have a very quality portfolio.

Rob's comment about the portfolio produced by the teacher was similar. He, like Ivan, developed a trust for the teacher as he reviewed the portfolio.

As I get into standard three right now, I think her information on each of the sub-points that she says she has, I firmly believe it. So I would not have to scrutinize too much what she's trying to support on the eight standards. I'm feeling real comfortable. I don't have much experience with people that lie about something or are inaccurate.

Mike did not appear to develop the same kind of vicarious rapport with the teacher as he reviewed the portfolio. He generally judged the portfolio as positive but was

specific about some missing pieces. He said, "She hasn't included enough parent and student things." He continued his judgment by saying that,

This teacher has not given as many examples of parent input and community involvement as maybe need attention. Excellent work within the building. Excellent classroom. Terrific classroom. Excellent work within the building, although no examples were given, to my recollection, of working with guidance, or working with the nurse, of working with administration. It was tied a lot to resource and special ed. Maybe broaden out a little bit on the work within the building.

Mike's comments were similar to those of Leo, one of the elementary teachers. Both principals mentioned that they would like to see student produced artifacts/information.

Given the positive artifact judgment comments and the positive review comments relative to the portfolio, it was not surprising that the judgment of the teacher (J/T) comments were mostly positive as well. Ivan and Rob each accounted for 96% of the judgment of teacher comments. Ivan made the most comments relative to judgment of the teacher with 54% while Rob accounted for 42% of the total judgment of teacher comments. Mike made only one comment relative to judgment of the teacher.

The middle school principals identified three positive characteristics about the teacher during Phase II. Their comments indicated that they thought the teacher was (a) a good communicator, (b) a teacher who uses strategies and activities that make math fun and interactive, and (c) a teacher that is organized. In terms of communication, Ivan commented that "This tells me I have someone who's willing to communicate with the administration, wants to be a good teacher." Rob concurred when he said, "I already see this teacher as being very strong in terms of communication."

Ivan also commented positively about the organization skills of the teacher. He said, "I have a teacher that's probably very organized. Someone who is very meticulous about how they set things out and how they do things. A lot of math people are that way. I would guess her desks are in rows; her desk is neatly arranged and organized."

All three of the principals commented about the teacher's ability to use teaching strategies to make math fun and interactive. Ivan commented that, "I like the way this teacher is working. It's not just paper and pencil math stuff like it used to be." Rob also was impressed with the teaching ability of the teacher. He said, "I'm sure she's got people that are intrinsically motivated but I'll bet a couple people are turned on to math in that class just from this alone. I would say right now, without being in her classroom, I bet she's not boring. Also seems to make every minute count, probably." Mike was very complimentary as well. He said, "This teacher would appear to be one who makes math a fun, active setting. The examples that are shown here – a song, making a basketball court, a number machine, mind twisters, posters – I was connecting more on the strategies as I was the artifact!"

There were negative judgment comments concerning this teacher. Ivan was the only principal to make negative comments. Thirty-eight percent of his teacher judgment comments in Phase II were negative. Every one of his negative comments related to typing and grammatical errors in the portfolio. His first comment relative to these errors was that, "I think she should re-read these and she's doing them, but again, it's not going to make her not proficient. But certainly deteriorates from the quality of what I'm looking at. To send something like this to a parent would not be good." Later in the review, he

returned to the typing and grammatical errors. His frustration was obvious when he said, “We are going to talk about the grammatical stuff. Not that I haven’t sent stuff out or had to catch myself, but it’s something that we need to be very conscious of.” Neither Rob nor Mike seemed to be concerned by the typing and grammatical errors. Brenda, one of the elementary teachers commented on spelling and grammar as she reviewed the elementary portfolio.

To summarize, artifact judgment (J/A) was the most dominating sub-category in the judgment category for middle school principals, accounting for 81% of the total judgment comments in Phase II (see Table 21). Middle school principals, like their elementary counterparts, engaged in a four-step rhythm as they judged artifacts. However, the middle school principals used cross-referencing as part of their judgment routine.

The middle school principals positively judged each of the eight ITS and, in general, most of the artifacts used to evidence them. The largest percentage of the artifact judgment comments clustered around the first three artifacts and then tapered off. Positive comments concerning the artifacts used to evidence the eight ITS accounted for 84% of the total number of comments relative to artifact judgment. Neutral and negative comments each accounted for eight percent of the artifact judgment comments. The negative comments relative to artifact judgment were random and did not center on one artifact or standard.

The middle school principals’ comments about the portfolio were mostly positive. The principals appreciated the teacher’s use of one artifact to represent multiple

ITS/criteria. In addition to being positive about the artifacts and the portfolio, the principals verbally reported positive comments regarding the teacher. The principals appeared to have developed a trust for the teacher via her portfolio and felt that the teacher was a good communicator, used quality teaching strategies, and was organized. One of the middle school principals noted numerous spelling, grammatical, and typing errors made by the teacher saying that while they caused some distraction from the overall quality of the portfolio, they would not make the teacher “not-proficient.”

Principal’s opinion. The category of principal’s opinion (PO) was the second category to cluster under judging activities. It had the fifth highest comment overall count for the middle school principals in Phase II. Comments in the principal’s opinion category tripled in Phase II, but accounted for only three percent of the total comments made by the middle school principals in that phase. Only Mike and Rob commented in Phase I but all three principals made comments in Phase II.

The principal’s opinion category was a bi-level category that consisted of principals’ opinions concerning the DE requirements associated with teacher assessment and principals’ opinions concerning the portfolio as evidence of good teaching. Table 24 illustrates the breakdown of the principal’s opinion category comment counts for middle school principals in Phase II. The number of comments for each of the sub-categories was fairly even.

The middle school principals’ comments concerning the DE teacher assessment requirements focused on the system of evaluation, the standards, and the value of having

new teachers in their buildings. Ivan’s interest in, and enthusiasm about, the portfolio he was reviewing was obvious. However, he indicated some displeasure with the system when he said, “I know there’s some real value to it [the portfolio piece of the evaluation system], but boy it’s frustrating the amount of time it takes, especially when you’re always looking at three or four first and second year teachers every year, plus now putting all of them [veteran teachers] on the rotation.”

Table 24

Principals’ Opinion Comment Count Breakdowns for Middle School Principals in Phase II

	Total PO Comments	PO/DE	PO/PE
Ivan	6	1	5
Mike	3	2	1
Rob	4	3	1

PO/DE = Principals’ Opinion/State DE Teacher Assessment Requirements, PO/PE = Principals’ Opinion/Portfolio as Evidence of Good Teaching

Mike and Rob took issue with the Iowa Teaching Standards; Mike in a general way and Rob in reference to specific standards. Even though Mike, as he judged the artifacts in the portfolio, supported the use of one artifact to meet several standards, he felt that the standards themselves were repetitive. “That’s what I don’t like about the standards – you end up duplicating.” Rob took issue with two specific standards/criteria.

As he evaluated an artifact relative to ITS 3e-Uses available resources, including technologies, in the development and sequencing of instruction, he said, “I am finding the newer teacher perhaps ...if we would have done this 20 years ago with this present system, we might find people struggling with the technology end of it. That has not been an issue for me in any teachers over the last three to four years.” It seemed that Rob was thankful that his new teachers were capable of using technology and could include evidence of its use, while, at the same time, thinking about his veteran teachers who may have no background in technology.

Rob was especially expressive about ITS 8b-Demonstrates professional and ethical conduct as defined by state law and district policy.

I’d just as soon throw out 8b. I think everybody does it [is ethical] unless they violate the law. I do have a pet statement I use oftentimes. Something to the effect that, “ I have found no violations of district policy and this teacher seems to follow any expectations set up in a school board manual or employee manual.” I guess I just have a professional problem with 8b period.

In his comment, Rob referred to how he would write about ITS 8b in his formal written evaluation. The reference to the written evaluation was pervasive in each phase of Rob’s review. He seemed to constantly be thinking about how specific artifacts would be included in his formal written evaluation. Rob also indicated that ITS 8b was difficult for the teacher to evidence and the administrator to “sign off” on. His frustration was apparent when he said, “I think that’s more a fault of the system rather than the teacher or administrator.”

Mike expressed the value of having new teachers who could coach his veteran teachers about the portfolio process. “What I think is that our new teachers are stepping

in a lot more prepared on this through work on campus [teacher education preparation]. Our new people are stepping in. It's been kind of interesting. A lot of times they are able to give some tips to some of the experienced staff as to how to work on this [portfolio]."

The middle school principals also expressed opinions concerning the value of the portfolio as evidence of good teaching. The coded comments in Phase II were similar to those in Phase I. The middle school principals view the portfolio as a "piece" of teacher assessment. Mike indicated that artifacts were cues to what teachers apply in the classroom. The middle school teacher used a safety flip-chart to evidence ITS 6e-Creates a safe and purposeful learning environment.

Our teachers do exactly the same thing with that [the flip-chart]. I saw these probably 15 times last year as I opened portfolios. It's nothing the teacher created. But as long as the teacher gives evidence of use and how they're using and how they're having it available, that's what matters. So maybe that's important to think about when it comes to artifacts too. Not all of these have to be teacher created, but they have to be teacher applied. That's what I'm looking for. How does the teacher apply this?

He acknowledged that a number of teachers in his building often used identical artifacts as evidence of particular standard/criteria. The application appeared to be the key, not the artifact. In short, the artifact by itself did not provide the proof that Mike was looking for; he needed to see the chart hanging in the teachers' rooms.

Ivan seemed to have the view that "less is more" when he considered the portfolio as evidence of good teaching. "They're [teachers] trying to kill you with volumes and bore you with volumes so you don't go through it all and you're like, 'gosh, if they got all this stuff they must have done it all.' That's personal prejudice maybe, but I think there's something to that too. They can get it done in smaller amounts." He extended his thinking

to experiences he had with teachers who brought their portfolios to an interview. “I’ve had people come in for interviews and bring a beautiful portfolio. It’s really thick. I get a little nervous about that because anybody can stuff a piece of paper in a portfolio. That doesn’t prove how it went or what happened. It’s just a nice little added piece to help put it [teacher evaluation] together.” Ivan’s skepticism about the portfolio as proof of good teaching was best summed up when he said, “My personal opinion is that a teacher can put together a portfolio and I can sit here and say, ‘Wow, this is a really good portfolio. This teacher deserves her license’ when in essence there are lots of issues that they need to be worked with.”

Rob thought the portfolio provided some evidence of good teaching but that it was not a “one size fits all” product. He indicated that it was necessary for teachers in different content areas to use different ways of evidencing good teaching. “Even though I fully understand and feel pretty comfortable with all eight standards –and here you’re giving me kind of a regular ed. math teacher – it was challenging for me as an administrator to do more peculiar areas of teaching.” He cited the difference between how a language arts teacher and an autism teacher might show evidence of meeting the standards. He further indicated that trying to make some areas of teaching align with the standards is sometimes difficult. “To try to write things and give it to eight standards/42 criteria with that autism teacher versus a language arts teacher – that was a challenge for me. So every time I heard ‘well, this is good for every teacher in every subject area’ – well, sometimes that’s easier said than done.”

Tools. The category of tools (T) had the sixth highest comment count for the middle school principals in Phase II. Tools referred to any sort of instrument (paper or electronic) that a principal used as he/she evaluated the portfolio. The researcher did not prompt the principals to bring any tools with them to the portfolio review for fear of the principal creating a tool for purposes of the study. The middle school principals did not comment about the use of tools in Phase I; however, each of the three principals commented about the use of tools (see Table 19) in Phase II.

Ivan and Mike focused on tools used for cross-referencing purposes. Ivan provided a detailed description of the cross-referencing tool (Appendix R) that he used.

What I have is a copy of the Iowa Teaching Standards and Criteria. We shrunk it down so it can be just on the face front side of the sheet of paper. It has the eight standards and underneath it, it has the 42 criteria listed just like they are. Behind each one what we have done is, in parentheses, cross referenced. For instance, 1a Provides evidence of student learning to students, families and staff – we have in parentheses 1g, 5b, 5e and 8e. What we're saying is that it is more than likely that if they did 1a they probably also evidenced 1g, 5b, 5e, 8: not necessarily, but it's a quick check. I can look at those and know to quickly cross reference over here. I bet it will do that one too.

Ivan used this tool extensively as he reviewed the portfolio provided to him by the researcher. He moved around the sheet very quickly and it was obvious that he had much experience with its use.

Mike also referred to a similar tool. He called it a log. He described it this way. "Our teachers have a log that I would lay out in front of me and it shows if they are applying the artifact in many standards and criteria. They'll check mark that and they'll tell me whether or not, as I'm looking at it, that it's also gonna be found again later."

Ivan and Mike each indicated that providing their teachers with the cross-reference tools was an advantage for everyone. Ivan said, “We give this to our teachers as well to help them when they’re turning this portfolio in because it’s rare to me that you’re going to have too many artifacts that are only going to hit one or two standards. The majority of them are going to hit a vast number of standards and criteria.” Mike commented on how his teachers used the log he provided to them.

Our teachers really, really like that [the log] as far as the organization part of it. In cases like this where for instance this artifact is being used for two standards and six different criteria. It helps the teacher know ‘okay, I’m going to store it here but I’m going to reference it here and here as well.’ As an end product, they had that out in front of them as well. It helps them kind of see where the holes are, which I think is good in this process.

Mike also indicated that he provided his teachers with sample data points and sources (Appendix S). He said, “This really takes away the anxiety. Not as an end point, but as a starting point for teachers. They look at this and right away they say they can find five in every area, saying, ‘Oh gosh, I’m already doing that. It’s gonna get me multiple data sources. I’m gonna be able to get beyond just my teacher stuff into things from the principal’.”

Rob had only one comment relative to a tool for cross-referencing. He simply indicated that he used a copy of the ITS as he reviewed portfolios and cross-referenced as he proceeded. Rob’s main focus, in terms of tools in Phase II, tracked back to his formal written summary. He indicated that the summary was provided to him electronically via the DE. He referred to the summary template used by the DE when he said,

We do have a computer template that we can use. What I do is print it off initially that just has the eight standards, because I haven’t typed anything on it yet. I just use that as my scratch paper. I will write in the little boxes. When you keyboard

it, the box expands and so that's what I'll do. I'll keep expanding it during my three meetings. Then typically by the month of April or late March I will have a final product that I can then send to the DE, down to our Human Resources and then obviously the teacher and I will have a copy.

Rob's thinking concerning the formal written summary was consistent throughout every phase of his review. His comments reflected his attempt to continually revise his writing as he talked with the teacher and/or completed observations.

In summary, the three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). The judgment comments of the middle school principals increased significantly in Phase II; just as they did for the elementary principals. For middle school principals, artifact judgment (J/A) was the most dominating sub-category in the judgment category, accounting for 81% of the total judgment comments in Phase II. Middle school principals, like their elementary counterparts, engaged in a four-step rhythm as they judged artifacts. However, the middle school principals used cross-referencing as part of their judgment routine.

The middle school principals positively judged each of the eight ITS and, in general, most of the artifacts used to evidence them. The largest percentage of the artifact judgment comments clustered around the first three artifacts and then tapered off.

Positive comments concerning the artifacts used to evidence the eight ITS accounted for 84% of the total number of comments relative to artifact judgment. Neutral and negative comments each accounted for eight percent of the artifact judgment comments. The negative comments relative to artifact judgment were random and did not center on one artifact or standard.

Overall, the middle school principals judged the portfolio positively in Phase II. The middle school principals did not take issue with one artifact representing multiple ITS/criteria as did their elementary counterparts. The comments made by the elementary principals indicated that the portfolio include a variety of artifacts as evidence of the eight ITS/criteria.

The middle school principals also judged the teacher positively. One principal, Ivan, took issue with grammar and spelling errors in the portfolio. His concerns with grammar and spelling did not affect his overall judgment of the teacher. The principals developed a sense of respect and trust for the teacher via her work in the portfolio.

The middle school principals' comments in Phase II included opinions regarding the DE requirements for teacher assessment and opinions relative to the portfolio as evidence of good teaching. Mild displeasure with the system requirements was expressed, centering mainly on the amount of time required of principals for review of both new teacher and veteran teacher portfolios. In addition, the standards themselves were scrutinized. One principal thought that the standards were repetitive while another was critical of the difficulty in evidencing ITS eight. The comments of all three middle school principals made clear their feelings that the portfolio was only part of what evidences good teaching. The principals viewed the portfolio as a "piece" of the entire evaluation system and that it could provide some evidence of good teaching; however, they contended that observation provided the proof that the artifacts included in the portfolio were genuinely applied in the classroom.

A copy of the ITS proved to be the most oft-referred-to tool in Phase II. The middle school principals indicated that they used a printed copy of the ITS in some form for cross-referencing purposes. One principal brought a copy and actually used it as he reviewed the portfolio. The principals also indicated that they encouraged teachers in their buildings to use cross-referencing as they prepared their portfolios. The goal of the principals appeared to be a way of assisting teachers by providing them with some data sources for their artifacts and prompting them as to the multiple standards that one artifact might meet. Ultimately, it seemed that the checklists served to save time for the principal and for the teacher.

Middle School Principals and Coaching Activities in Phase II

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time investment (TI). Comments relative to coaching activities accounted for 23% of the total Phase II comments made by middle school principals. In this section, comment counts relative to two of the three categories (coaching and time investment) will be reported followed by a summary. The middle school principals did not make any comments relative to principal's role in Phase II.

Coaching. The coaching (C) category had the second highest comment count for the middle school principals in Phase II. Comments relative to coaching showed significant increase from Phase I to Phase II for the middle school principals. Comments in the coaching category represented 21% of the total coded comments in Phase II thinking for middle school principals. The coaching category was a multi-level category.

Table 25 illustrates the breakdown of the coaching category comment counts for the middle school principals. Coaching via meetings dominated the coaching category. The middle school principals did not use questioning as a means of coaching.

Table 25

Coaching (C) Comment Count Breakdowns for Middle School Principals in Phase II

	Total Coaching Comments	C/M	C/PP	C/Q	C/SA
Ivan	27	8	4	0	15
Mike	18	5	3	0	10
Rob	33	23	3	0	7

C/M = Coaching relative to meetings with the teacher, C/PP = Coaching relative to portfolio preparation, C/Q = Coaching relative to the use of questioning, C/SA = Coaching relative to suggested alternatives for artifacts

As was true with the elementary principals, the coaching suggested alternative (C/SA) category was indeed a part of the four-step rhythm developed by the middle school principals; however, it did not represent a one-to-one ratio with judgment of artifact comments. Judgment was the focus of the principals in Phase II but coaching during the judgment phase was also very important. The middle school principals did not spend time thinking about coaching via the use of questioning as did the elementary principals.

Meetings with the teacher had the highest comment count for the coaching category accounting for 46% of the thinking of the middle school principals. All three principals engaged in thinking in the coaching/meeting category. The principals identified clarification as the main purpose for meeting with the teacher and did include in their comments, questions they might ask in meetings. During Ivan's review of the artifact representative of ITS 1e- Creates an environment of mutual respect, rapport, and fairness (the artifact described how three former students advised the teacher of what they had learned the year before due to the many hands-on activities and strategies), he commented that,

If she [the middle school teacher] was sitting here and we were meeting, I would ask her "Why these?" I obviously see a pattern in those things [the strategies]. A lot of them are hands-on. A lot of them are singing. A lot of them are abstract types. It's tying things in and that's how kids remember. I find that interesting. I think she would probably pick up on that as well.

Although Ivan judged the artifact favorably, he felt compelled to clarify, for the teacher, the value of the strategies she had incorporated.

Another example of the principals' need to meet with teachers for clarification was illustrated during Rob's review of an artifact used by the middle school teacher to evidence ITS 4b-Uses research-based instructional strategies that address the full range of cognitive levels. The teacher described the artifact in this way. "At the beginning of the year we had a speaker come in and speak about vocabulary improvements [faculty professional development]. She gave us several different examples to use in class. It was a research-based method for helping students remember vocabulary words." The teacher used the method in her class. Her reflection best described the result. "My eighth grade

math class has to memorize eight properties of math. I split the students into groups and had them use the four quadrant vocabulary (see Appendix T) to help others be able to memorize the eight different properties.”

Rob commented that, “I would say right now that when I meet with her...I’ve gotta delve into a little bit more about how this is dealing with the full range of cognitive levels. I just need her to elaborate on it. I’m not doubting it. I just need some questions answered for my own ignorance. I would just need to ask some questions for clarification.” Rob’s need for clarification was driven by his own need to understand the use of the vocabulary quadrant and the teacher’s vision of its purpose in meeting ITS 4b. Mike also needed clarification for the artifact pertaining to 4b. “I like 4b. We’d talk a little about it. I’d say, ‘How do you see this artifact also being used in meeting 4b?’ and let the teacher think through that with me.” Mike, like Rob, approved of the artifact but needed the teacher to provide clarification.

Coaching, relative to the portfolio process (C/PP), accounted for 13% of the total coaching comments made by middle school principals in Phase II. Each principal accounted for roughly 33% of the coded comments relative to the portfolio process. All three principals spent time thinking about coaching relative to the portfolio. Two common threads emerged from this category. The first common thread was regarding the use of one artifact to meet multiple ITS/criteria. All three principals supported this concept. Ivan, however, indicated that his preference was for the artifact to be included only once. “I give them [his teachers] suggestions in the fact that I don’t need to see this in every single one. Don’t put the same artifact in there multiple times. You’ve cross

referenced it, I'll find and I'll mark it and you'll know I found it. Save them time, save me time. But they're [the teacher's district] set up a little bit different and that's okay."

Mike also supported the use of one artifact to meet multiple ITS/criteria. He indicated that he would prefer that the teacher reflection that accompanies each artifact adjust to the ITS/criteria for which it is used.

If I'm [referring to the teacher] going to use this [reflective statement] multiple places - that's kind of where the artifact tags become beneficial - it helps the teacher know, okay I'm using it here because and now I'm going to apply it here because... In my opinion it's wrong to assume that the same reflective statement would apply to both standards. I can absolutely support using the same thing in a variety of places, but I think you'd tag it differently. The benefit of it here might be different than early on. Therefore the writing here would be a little bit different. I just happen to notice as I'm seeing some of these multiple times that that seems to be the way it's been approached.

The second common thread was that of how the principal could assist the teacher if the teacher were having difficulty finding an appropriate artifact to evidence one or more of the eight ITS/criteria. Each of the middle school principals thought that providing such assistance was appropriate. Ivan indicated that including the principals' observation would be helpful. "I tell teachers, 'Use my observations and my walk-throughs in your portfolio. They should be in there and part of that'." He also said, "Mine [his teachers] are coached to put them in there because it hits so many of them [ITS/criteria]." Rob also favored assisting his teachers. "My previous experience would be with people that can't find anything on a particular sub-point. But, they may be doing it. They just can't get it orchestrated. And, if nothing else, particularly if it's a pretty decent teacher, we'll talk about it. We'll try to think it through." Mike specifically

discussed ITS number 8-Fullfills professional responsibilities established by the district, and how he might coach teachers during development of the portfolio.

Eight is the tough one for teachers and the one that oftentimes they would come to me at different times and ask for my help or ask for me to provide them with some artifacts. But I think they're underestimating. It gets into the day to day stuff. Are you being professional? Are you coming on time? Are you carrying out supervisory responsibilities? Are you attending meetings? Are you being a professional in your actions? Some things that we identified that helped them on that would be their website. Each of our teachers has a website and keeps it up to date as far as assignments and activities and things like that. So a lot of them ended up tapping into that as far as evidence that they had.

All three principals described a perspective of partnering with teachers to assist them with their portfolio and associated artifacts. This approach differed slightly from that of the elementary principals. The elementary principals' comments were directed towards focus and structure rather than assistance and partnering. Brenda, one of the elementary principals, put the responsibility of artifact inclusion on her teachers but did say that she might assist them if they asked her to do so.

The coaching/suggested alternative (C/SA) subcategory had the second highest comment count in the overall coaching category for middle school principals in Phase II. It accounted for 41% of the total coaching comments made in Phase II by the middle school principals. The suggested alternatives were part of the four-step rhythm incorporated by the middle school principals but were not a one-to-one match to the artifact judgment comments. Table 26 illustrates the breakdowns of the number of suggested alternative comments, per ITS, made by the middle school principals.

In Phase II, the middle school principals clearly focused on the judgment of the artifacts but made relatively few suggested alternative comments; probably due to the low

number of negative artifact judgment comments. The middle school principals cumulatively made only 13 negative artifact judgment comments and 32 suggested alternative comments.

Table 26

Suggested Alternative (C/SA) Comments per Eight ITS by Middle School Principals in Phase II*

	Iowa Teaching Standards (ITS)*							
	#1	#2	#3	#4	#5	#6	#7	#8
Ivan	5	1	1	3	1	4	0	0
Mike	2	1	2	2	2	0	0	1
Rob	3	2	0	0	2	0	0	0

* See Appendix A for ITS/Criteria descriptions

Distribution of the suggested alternative comments for the middle school principals bore some resemblance to the distribution of the judgment comments made relative to each ITS. The suggested alternative comments clustered around ITS number one and then tapered off, very much like the artifact judgment comments did. In fact, 30% of the total suggested alternative comments were made relative to ITS number one. Standards one, two, and five were the only standards that received comments from each of the three middle school principals.

In essence, four firm recommendations emerged in Phase II relative to coaching/suggested alternatives from the middle school principals: (a) provide additional detail in artifact description, (b) include student samples, (c) include information from other sources, and (d) use of an already-included artifact. These recommendations were not suggested for every artifact the teacher included; however, they were the recommendations that were most commonly cited by the middle school principals. Each recommendation will be discussed in the following paragraphs.

Consistent with comments made during artifact judgment, the middle school principals wanted more information included with some artifacts for clarity purposes. For instance, as Rob reviewed an artifact included in ITS number two he said, "I'd just probably want a little bit more. I'm not finding fault with it. Just may go into a little more depth." The tone of this comment was very indicative of the middle school principals' when they did make suggested alternatives. The suggestions were not critical, they were inquisitive. The comments in the coaching/suggested alternative category also aligned with the principals' indication that they would find a meeting with the teacher a valuable venue for gaining insight and depth concerning artifact selection and description.

Mike's need for more detail involved the reflective statements. This was not surprising given the number and intensity of his comments relative to the importance of reflection in Phase I. As Mike wrapped up his review of the artifacts representative of ITS number one he said,

So, it would look like, having completed standard one, that that is done well. She has included a variety of artifacts including students. Although her reflective writing could have been in more detail and could have been more tied to the

teacher part of things as opposed to the student part of things, it helped to explain [the teacher's intent].

Mike was clear that he judged the standard positively and that his preference for detail centered on the reflective statement.

In addition, the principals expressed a desire to have student work included as evidence of the standards. While reviewing ITS number one, Ivan remarked, "I would have really liked to see a student piece." Rob too, was interested in student work as he reviewed artifacts in ITS number one. "I guess I'd ask to see some kids' responses."

Mike also weighed in on student work when he said, "So, maybe a little more in the area of student."

The third recommendation relative to suggested alternatives was that of including information from other sources. Of the three middle school principals, Rob was the most interested in information from other sources. He suggested that perhaps the teacher should include evaluations from her students.

Or perhaps student evaluations of this teacher. I think back in the dark ages... before we even had artifacts...I still have some letters of support, or letters of criticism too, where they got a chance to evaluate me as a teacher at [can't include name] High School. Some of that is kind of fun to look at 15-18 years later and see what kids perceived as what was good or bad about your teaching.

He seemed intent on including student feedback in the portfolio, even if it meant that he talked with students. His purpose was not to "check up" on the teacher but rather to provide further support for teacher and the artifacts in the portfolio.

I'll tell you one thing that can be supported, and I would tend to use this more in support of this area – I've had some real good luck with calling kids in and talking with them a little bit about a teacher or perhaps some mannerisms or why they ...this is like separate from any kind of discussion I have with the teacher. Now, I'll be honest. I would hesitate using much of what I get if it's negative in nature.

But if nothing else, I can get kind of an honest ...they're going to use their own vocabulary, it's going to be their own feelings. It kind of goes back to if you really want to know something about your school or how something is going, call a kid in and ask him.

To reiterate, Rob did not, in any way, seem to indicate that he was talking with students for purposes other than to include different perspectives. He also suggested a parent survey. "I might just ask her a little bit about, 'Does she ever survey the parents?' or what kind of feedback she gets from parents; emails, etc."

The fourth coaching/suggested alternative recommendation made by the middle school principals centered on the use of artifacts already present for evidence of a standard. Ivan made this comment as he reviewed an artifact for ITS 1a-Uses student performance as a guide for decision making, "So I'm going to throw 3a [uses student achievement data, local standards, and the district curriculum in planning and instruction] on this one as well. I think that's something she can learn to do, is get a broader cross reference for herself." Ivan was physically making small X's next to ITS 1c and 3a on the hard copy of the eight standards/criteria that he brought with him. Quite obviously, Ivan had memorized the standards/criteria. He moved quickly and efficiently as he cross-referenced.

Each of the principals made reference to the use of one artifact for multiple standards; however, it was Mike's comment regarding an artifact describing small-group work relative to ITS 3b that best embodied the thinking and cross-referencing concept used by the middle school principals.

Now let's look at the artifact. *Three junior calc students taught this worksheet.* So the teacher designed the worksheet, but in small group instruction [led by the three calculus students who were visiting the 8th grade classroom], which also

...that could be used coming up here in 4. She just used this in 3. This really could have been used back in 2 regarding using instructional strategies to enhance learning. It also would be fine later on in 4 too, but she just is using it for 3. It includes a lot of good stuff. Here's one that's being used for 3D, but she also has used it earlier in 2 as well as 7 and 8, which get into the professionalism and that kind of thing.

Clearly, Mike, too, had memorized the standards/criteria. Mentally, he moved quickly and efficiently through the possible standards the artifacts might evidence. Plainly, he was attempting to get everything he could out of each artifact.

Time investment. The category of time investment (TI) was also present in the coaching activities; however, it accounted for less than two percent of the overall coded comments for Phase II thinking. Ivan was the only middle school principal to comment about time (TI) in Phase II. Ivan focused mainly on how he was using his time, but did make one comment about the amount of time it took to review a portfolio. The following comment best illustrates his concern about time.

Well, if I keep up this pace I'm going to be in this office all day. I'm thinking I've got at least five artifacts per standard. I have got seven standards to go. That's 35. A minute a piece is 35 minutes. It's probably at least two minutes a piece – that's 70 minutes. I've gotta pick up the pace or I'm never going to get done. That is what I don't like about this process. The amount of paperwork and looking at it.

The issue of time was of concern only to Ivan and his comments mostly focused on how he needed to use his time.

In summary, the three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time investment (TI). Middle school principals in Phase II addressed only two of the three categories; coaching and time investment. The middle school principals did not make any comments relative to principal's role in Phase II.

The coaching (C) category had the second highest overall comment count for middle school principals in Phase II; second only to judgment comments. Comments relative to coaching increased significantly in Phase II. The middle school principals' coaching comments in Phase II were positive and took on a tone that reflected their need for clarification and their desire to partner with the teacher to acquire said clarity. The coaching category was a multi-level category that included coaching via meetings, coaching on the portfolio process, coaching using questioning, and coaching via suggested alternative. As a group, the middle school principals attended to each subcategory except coaching using questioning.

Meeting with teachers had the highest comment count of the coaching category. Meeting with the teacher, for the middle school principals, was a venue through which they could achieve clarity concerning artifacts. The middle school principals indicated that the meetings were not designed to *tell* the teacher what the principal thought about a particular artifact. Rather, the purpose of the meetings was to *engage* the teacher in conversation about the artifact. The intent of the meetings, as described in the comments of the middle school principals, paralleled the intent of the elementary principals' intentions when meeting with teachers. Both groups, elementary principals and middle school principals, were very interested in initiating conversation with teachers that caused deeper thinking concerning quality teaching and the strategies that exemplify that quality. Relative to coaching, in terms of portfolio preparation, the middle school principals indicated that it was appropriate to assist teachers with finding appropriate artifacts to include in their portfolios.

Suggested alternative comments were part of the four-step rhythm that was incorporated by the middle school principals; however, the suggested alternative comments did not match one-to-one with the judgment comments. This was probably due to the high amount of positive judgment comments, hence, less need for suggested alternatives. The suggested alternatives were random in nature, i.e., they did not cluster around one particular standard. The middle school principals' suggested alternatives included four recommendations: (a) provide additional detail in artifact description, (b) include student samples, (c) include information from other sources, and (d) use of an already-included artifact. Two of the recommendations, additional detail and student samples, echoed sentiments of the elementary principals in Phase II thinking.

Middle School Principals and Critical Pieces in Phase II

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two "pieces" carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. In this section, comment counts relative to the two sub-categories, critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO), will be reported, followed by a summary.

The critical pieces (CP) category had the fourth highest comment count for middle school principals in Phase II thinking; three times the number of comments coded in the same category in Phase I. Critical pieces accounted for six percent of the overall comments made by the middle school principals in Phase II. The critical pieces category

was a bi-level category. Table 27 illustrates the breakdown of the critical pieces category comment counts for middle school principals. While all three middle school principals had comments in the critical pieces category in Phase I, only Ivan and Mike had comments in the same category in Phase II. Even with Rob not commenting, the comment counts increased significantly from Phase I to Phase II.

Table 27

Critical Pieces (CP) Comment Count Breakdowns for Middle School Principals in Phase II

	Total CP Comments	CP/RO	CP/TR
Ivan	13	8	5
Mike	8	1	7
Rob	0	0	0

~~CP/RO = Critical Pieces/Role of Observation, CP/TR = Critical Pieces/Teacher Reflection~~

Role of observation. Clearly, Ivan dominated the critical pieces/role of observation category. His comments indicated that observation was absolutely critical in terms of evaluating any teacher. He explained the significant nature of observation as he reviewed an artifact evidencing ITS 6d.

6d – Uses instructional time effectively to maximize student achievement. Well, the potential is there to use instructional time effectively. To be honest, it's tough to give me a paper for 6d. I need to observe 6d, that you're using instructional

time effectively because I can have the best sheet of rules on the board, but if you're not keeping kids in line and you're letting them get you off task and you've got all this wait time while you're doing attendance or while you're doing whatever...

Ivan's emphasis on observation was consistent with his thinking in regard to the coaching category. He indicated that he coached his teachers to include his observations. He left no doubt about the weight he placed on observation when he said,

To me they [written observations] should be included in it [the portfolio]. My observation is the basis, it's the foundation of whether they're going to get their license or not. This to me is the supporting evidence to help it. There are some things I can't go in there and observe and then yes, that is the base for that. But this is it. This is what says yes or no for them. Maybe I'm off base compared to others, but if my observations and walk throughs aren't in there, I just think that's missing a huge element. When a teacher comes here for an interview and they show me a portfolio, I look for that principal's observations in there.

Ivan also liked that the middle school teacher had included a peer evaluation but was clear about its value. "Peer observation is good. That's something they [teachers] want to get to and they want to go to. It's just tough to get it to work as far as on an overall realm. Plus, it can't be the evaluator piece, but it can be something they [teachers] can put in there to help them."

Mike's only comment concerning the role of observation suggested that observation reinforced what he was seeing in the portfolio. As he reviewed an artifact evidencing ITS 2d-Understands and uses instructional strategies that are appropriate to the content area, he said, "This is where all of my day to day walk-ins and walk-throughs are just gonna simply reinforce the one example she's chosen."

Teacher reflection. The comment counts in the sub-category of critical pieces/teacher reflection (CP/TR) were split nearly evenly between Ivan and Mike. Rob

did not comment in this category. While Ivan did not make any comments in Phase I relative to critical pieces, his comments in Phase II revealed the importance of the teacher including reflection in the portfolio and how it might influence his thinking about an artifact. "Because of that [information in the reflection] I really like this artifact. I like it a lot. What I like is the reflection. Just the distributive property worksheet is pretty blah. Anybody could put that together. But the reflection of that one gave that one a ton of credibility. I like that."

Mike's comments regarding teacher reflection in Phase II were very consistent with those he made in Phase I; he expects teachers to include reflective statements and he expects reflections to assist him in his review. "Particularly here in [my district], we have weighted more heavily the reflective writing that's a part of the portfolio than the artifacts themselves. So I certainly will be putting my attention on the reflective writing piece, if in fact that's been included here; more so than the artifacts themselves."

In summary, observation and teacher reflection were considered to be critical pieces of the portfolio review by middle school principals. Observation was deemed critical because it supplemented the portfolio and allowed principals to observe an artifact in action. The middle school principals also thought that reflection was critical in that it, like observation, could provide depth and clarity to an artifact.

Summary of Middle School Principals' Thinking in Phase II

Phase II (judgment) thinking was described as the period of time following pre-assessment thinking when principals actively judged the artifacts in the portfolio that was provided to them by the researcher. Processing activity drastically declined in Phase II as

the middle school principals focused away from processing activities and towards judging and coaching activities. The principals focused most on artifacts representative of the first three teaching standards. Then, similar to the elementary principals, judgment comments tapered off.

As they began to review the individual artifacts included in the portfolio, the middle school principals developed a four-step rhythm similar to that of the elementary principals. However, the middle school principals employed cross-referencing as part of their rhythm. First, they would identify and verbally describe the artifact. Second, they would actually read aloud as they focused on what the teacher was attempting to illustrate. Third, they would make a judgment statement regarding the artifact they were judging. The judgment statement was then sometimes followed by a coaching statement. Finally, the middle school principals would think about other ITS/criteria the artifact might meet, i.e., cross-referencing.

The comfort level of the middle school principals increased as they moved into Phase II. Verbal reporting indicated that the middle school principals developed a sense of rapport and trust with the teacher who produced the portfolio. The sense of rapport and trust translated into an overwhelmingly positive percentage of artifact judgment comments. Positive artifact judgment comments accounted for 84% of the total artifact judgment comments made by the middle school principals. Only eight percent of the artifact judgment comments were negative and they were very randomly distributed across the ITS/criteria. As opposed to the elementary principals, the middle school principals did not single out any one of the artifacts as being more negative than positive.

Similar to elementary principals, the middle school principals showed a proclivity towards certain artifacts. They liked artifacts that illustrated: (a) communication with parents, students, and colleagues, (b) assessment results, (c) student work samples, (d) evidence of rapport with students, (e) use of technology and (f) teacher reflection and growth.

Coaching/suggested alternative comments did not reflect a 1:1 ratio with artifact judgment comments. The suggested alternative comments made by the middle school principals centered most on the need for additional information to support artifacts. While the middle school principals verbally reported positive response to student-created work samples, they wanted more of them. Additionally, they indicated that additional sources of information such as principal observations and student/parent surveys would be beneficial.

Coaching, in terms of meetings with the teacher, was also important to the middle school principals in Phase II. The comment counts relative to meetings with the teacher accounted for nearly half of the coaching comments. The principals viewed the meetings as opportunities for formative conversation with the teacher.

Phase II verbal reporting indicated that the middle school principals placed great value on observation and teacher reflection. Both were considered critical to teacher evaluation by the middle school principals. As middle school principals thought aloud about the vital importance of observation they also commented about the value of the portfolio in the teacher evaluation process. It was clear that, for these three principals, the portfolio was only a part of the evaluation process; that observation provided the “seeing

is believing” aspect missing from written artifacts. Verbal reporting also indicated that teacher reflection was essential because it provided information to the principals about artifacts from the teacher’s perspective; thus, guiding the principal through the artifacts and eventually, the portfolio.

High School Principals’ Thinking in Phase II

The high school teacher’s portfolio was organized by artifact. Each artifact represented one or more of the eight ITS/criteria. See Appendix I for examples of the table of contents and cross-referencing information used by the high school teacher. The high school principals (Gavin, Kathy, and Keith) made fewer overall comments than the middle school principals but made more overall comments than principals at the elementary level. Their thinking accounted for 31% of the overall comments made by all principals in Phase II.

Table 28 illustrates the breakdowns of Phase II comments made by the high school principals. All ten categories were present in Phase II. Only eight categories were present in Phase I. The two new categories that emerged in Phase II were critical pieces (CP) and time investment (TI).

Comment counts for the high school principals significantly increased in Phase II. In Phase I, the high school principals made only 13 overall comments as compared to 240 comments in Phase II. Gavin had the most comments in Phase II accounting for 44% of the total Phase II comments made by high school principals. Kathy’s comments represented 37% of the total Phase II comments made by high school principals. Keith

made the fewest Phase II comments accounting for 19% of the total Phase II comments made by high school principals.

Table 28

Summary of Phase II Comments made by High School Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Gavin	106	19	2	6	59	9	5	0	2	4	0
Kathy	89	29	2	6	36	3	1	3	5	3	1
Keith	45	12	1	4	27	0	0	0	0	1	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PR = Principal’s Role, PS = Process Steps, T = Tools, TI = Time

High School Principals and Processing in Phase II

The three coded comment categories that clustered under processing activities included the categories of portfolio structure (PFS), process steps (PS), and comfort level (CL). Comments relative to processing accounted for only 10% of the total Phase II comments made by high school principals; a marked decrease from the 77% of total processing comments made in Phase I. The verbal reporting data provided evidence that, while principals remained cognizant of structural matters, they were turning their focus towards judgment. In this section, comment counts relative to each of the three

processing categories (portfolio structure, process steps, and comfort level) will be reported followed by a summary of high school principals and processing.

Portfolio structure. While the number of comment counts relative to portfolio structure increased in Phase II, the percentage of time committed to portfolio structure was much lower in Phase II than in Phase I. Kathy and Gavin were the only two high school principals to make portfolio structure comments in Phase II. Kathy began to think about the structure of the portfolio and how it compared to those of her own teachers. She indicated that she appreciated quality over quantity when she said,

If you do it the way that he's done it here, which is organized by artifacts, you do end up with a lot fewer and I think that's real good actually. Some people really do give you ...I am not joking. I had one ...we give them a 3" three ring to keep theirs in, so bigger than this. She went out and had to buy a bigger one. I'd never seen a three ring binder that was so big. Every single thing she had. It was actually to me a little bit of a sign of "Wow, what's this all about? What does this 6" three ring binder represent?" It was more of a red flag actually than it was a positive reflection on her teaching.

Kathy's portfolio structure comments in Phase II only increased by one. She appeared poised to begin making judgments of the artifacts.

It was Gavin who showed the largest increase in portfolio structure comments. He went from four comments in Phase I to nine comments in Phase II. He showed signs of understanding the structure in Phase I but became increasingly frustrated as he prepared to judge the artifacts in Phase II. Essentially, he struggled with the fact that the teacher had organized his portfolio by artifact. He was trying to understand how just seven artifacts could evidence eight standards. It was just not squaring with him. His frustration was evident when he said,

Again, a structural kind of deal for me right now is I as the evaluator, especially with tenure teachers having to have portfolios, also. I'm working too hard to get what I need to find out. It's one of those where after you do this for so long it's going to be "I can't do this anymore." That's fine. It's just with the structural thing and finding the easy ...keep it simple.

Clearly, Gavin was thinking about time and energy. He felt that the structure was making him work harder. However, he tried very hard to find positives. "Let me go back to this reference page. The way he has it set up, I do like the idea that for each one of the artifacts he does at least have the teaching standards right there so I don't have to go back and look at those too." Gavin continued to move through each of the seven artifacts. It wasn't until he was judging the seventh (last) artifact that the structure became clear to him. He said, "I am so dumb. I've just now figured out what he's doing and I'm on the seventh one!" It was as if he had just discovered the cross-reference page even though he had been referring to it all along. He appeared relieved and was noticeably less stressed.

Process steps. Comment counts relative to process steps increased from four comments in Phase I to seven in Phase II for the high school principals. Keith did not make process step comments in Phase I or Phase II. However, Kathy and Gavin both had comments in each phase. In Phase II, the principals moved away from process steps regarding the review of the portfolio to process steps relative to each artifact. Both principals needed to get a clear picture of the artifact and then compare it to the established ITS standard/criteria. The following Phase II comment revealed Kathy's thinking process as she reviewed ITS one.

Given that I now think I get all of Artifact One, I would go back and think about ...now that I kind of pretty thoroughly understand what it was that he was doing with the lesson is ...I would then go back and look at whether I think Artifact One

– does it demonstrate ...he’s saying 1D. So I would read through those and try to have some sense whether this documents all of those things or not.

Gavin, too, looked for evidence in an artifact and how it matched the intended standard.

“Let me go through here then. For example, on 2a it’s talking about communicating with evidence of student learning to student families and staff. What I would look for in something like that is how he is getting this information home or what they’re studying or how their grades are or whatever it may be.” In both cases, the process was leading the principals towards making a judgment about the artifact; the principals were identifying what they were looking for before they made a definitive judgment decision about the artifact.

Cross-referencing did not have the same sort of processing significance for the high school principals as it did for the middle school principals. Kathy made some reference to cross-referencing when she said, “I might get a better sense, looking at the whole document, of some of these standards and criteria even if he hasn’t put them under a particular one. That’s kind of a part of the process I do too.” Her statement indicated that she would peruse the entire document and find artifacts that would evidence one or more ITS.

In summary, high school principals thinking relative to process steps included self-talk about what they anticipated should be present in an illustrated ITS. Some thinking was dedicated to cross-referencing but not nearly to the degree of the middle school principals. Phase II process steps for the high school principals served as transition from familiarizing themselves with the portfolio to judging the artifacts.

Comfort level. The percentage of comfort level (CL) comments for high school principals accounted for only two percent of the overall comments made in Phase II; down from the nearly 10% of overall comments made in Phase I. The decrease was an indication that the focus of the review in Phase II was indeed shifting towards judgment. In Phase I, familiarity with the portfolio was the goal. In Phase II, all three high school principals indicated that being familiar with the teacher and his students as well as district goals and standards would be valuable as they judged the artifacts. Keith said, “As building principal, you would probably have some better ideas of exactly who the kids are.” Gavin was concerned that he would be accurate and fair in his judgments – to the point of being very hard on himself. He said, “I don’t know his kids. So it’s hard for me to make that judgment just based on what I’m seeing in front of me. I feel like a real jerk here.” Kathy’s thinking about district criteria was evident in Phase I and she remained consistent with this type of thinking in Phase II when she said, “Again, if I were his principal I would know whether that was a district or building goal.” The principals appeared poised to make judgments but wanted to make sure they were being fair.

In summary, processing activities included the categories of portfolio structure, process steps, and comfort level. The percentage of time principals spent in processing decreased significantly in Phase II as the principals shifted into judgment mode. Some anxiety still existed for the high school principals in terms of comfort during Phase II. However, comfort concerns moved away from issues of portfolio structure to concerns about familiarity with the teacher, his students, and his district. The principals conveyed

that fair judgment was important to them. Their anxiety did not detour them from completing their reviews.

High School Principals and Judging Activities in Phase II

The three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Comments relative to judging activities accounted for 57% of the total Phase II comments made by high school principals. In this section, comment counts relative to each of the three categories (judgment, principal's opinion, and tools) will be reported followed by a summary.

Judgment. Judgment comments accounted for just over half of the total comments made by the high school principals in Phase II; similar to the 54% made by elementary principals and 52% made by middle school principals. The high school principals made only one judgment comment in Phase I.

The judgment category was a multi-level category that represented judgment of artifact (J/A), judgment of portfolio (J/A), and judgment of the teacher (J/A). Table 29 illustrates the breakdown of the judgment category counts for the high school principals in Phase II. Each of the judgment sub-categories was considered by each of the principals. Similar to the elementary and middle school principals, the judgment of artifacts sub-category consumed the thinking of the high school principals in Phase II.

The high school principals incorporated the same four-step rhythm when judging artifacts as did the elementary and middle school principals. First, they would identify and verbally describe an artifact. Then, they would actually read aloud as they focused on what the teacher was attempting to illustrate. Their third step was a judgment statement

regarding the artifact they were judging. The judgment statement was then sometimes followed by a coaching statement.

Table 29

Judgment (J) Comment Count Breakdowns for High School Principals in Phase II

	Total Judgment Comments	J/A	J/P	J/T
Gavin	59	58	1	0
Kathy	36	30	1	5
Keith	27	23	0	4

J/A = Judgment comments relative to the artifacts, J/P = Judgment comments relative to the portfolio, J/T = Judgment comments relative to the teacher

The high school principals did not spend the same energy with cross-referencing that the middle school principals did. Rather, they relied on the table of contents provided by the teacher (Appendix I). After reading the specifics about the artifact, the high school principals looked at the other standards/criteria in the artifact that the teacher said he was attempting to evidence. The principals would then methodically work through the additional identified standards/criteria using the same four-step rhythm. There was not the “conjecture” about what other standards/criteria the artifact might meet as with the middle school principals. There was anticipation of what principals “should” see represented in the artifacts.

The teacher's first artifact was a lesson from a unit on United States History. The way in which Gavin moved through the first portion of artifact one provided a good illustration of the more methodical cross-referencing technique used by the high school principals. The teacher described the lesson in his artifact cover page. He said,

The following lesson is aligned with the theme "America in the World." It is from a unit I teach in U.S. History since 1877. This lesson followed discussion of Vietnam and the way people view war during different periods in American history. Discussion of popular songs provided in the lesson evidence societies' views toward war.

The teacher included in the artifact his lesson plan, a graphic organizer for comparison of different songs, and lyrics of three different songs ranging from the Vietnam era to present day. The teacher indicated that the artifact was meant to evidence criteria from ITS one, two, three, four and five.

When Gavin reviewed the first artifact, he first familiarized himself with the artifact. He quietly read the description to himself, and then he remarked, "I'm just reading through it so I can ... I like the description. It's short and sweet and tells us a little bit of what they did. His reflection does describe how he believes that he at least fit 1d. The different learning styles and things like that." Then, Gavin moved to the next standard/criteria the teacher said his first artifact evidenced. The high school teacher indicated that the first artifact also met 2c-Relates ideas within and across content areas. As he reviewed the artifact again, he said,

ITS 2c is definitely is here. He's relating ideas and information within and across content areas. It's not necessarily ...he does have them writing papers and he does have them doing some of those kinds of things, but he's also cross cultural within the pop music and the societies and things such as that and getting into the views part of things. So, I would give him credit for the 2c.

Gavin completed the four-step process by making several suggested alternatives. In general, he judged the artifact as adequately meeting the standards/criteria the high school teacher said it would. The afore-described process was indicative of the means used by the high school principals to review each artifact.

By further breaking down the judgment of artifacts (J/A) category, the comparability of judgment comments, per artifact, by each principal, became even more clearly illustrated. Table 30 illustrates the breakdowns of the number of artifact judgment comments, per each artifact, made by the high school principals.

Table 30

Judgment of Artifact (J/A) Comments per Artifact by High School Principals in Phase II

	Artifacts*						
	#1	#2	#3	#4	#5	#6	#7
Gavin	9	7	9	8	11	7	7
Kathy	6	6	3	1	5	2	7
Keith	4	3	3	2	2	3	6

* See Appendix I for Artifact Title and Cross-referencing with ITS

The distribution pattern of the comments relative to each artifact was markedly different from the distribution patterns of the elementary and middle school principals. The high school principals' comments were spread fairly consistently across all artifacts as opposed to the elementary (Table 14) and middle school (Table 20) principals'

comments that clustered around the first three comments and then tapered off. Artifact number one and artifact number seven had the highest comment counts for the high school principals while artifacts number four and six had the fewest comment counts.

To more clearly understand and illustrate the trend in artifact judgment for high school principals, it was necessary to rate the comments relative to each artifact. The same rating system used to classify the artifact judgment statements of elementary and middle school principals was used to tag the artifact judgment statements of the high school principals. Each artifact judgment statement made by the high school principals was evaluated and tagged with a positive (+), neutral (-), or negative (-) rating.

Positive judgment statements included those statements that indicated the artifact, per the principal's judgment, had properly illustrated/met the ITS the high school teacher indicated it would. An example of a positive artifact judgment statement was made by Gavin as he considered 2d—Understands and uses instructional strategies that are appropriate to the content area. "Looks like he is talking about 2d—uses instructional strategies. I'm seeing group work. I'm seeing writing, working on some content reading strategies, things such as that. So I would probably say that 2d is also there." Neutral statements included those statements that the principal made while judging the artifact, but the statements did not indicate that the principal had made a definitive judgment as to the value of the artifact. An example of a neutral judgment statement was made by Keith when he said, "This looks like one [a worksheet] that he has taken from somebody else. He doesn't necessarily have to reinvent it." Negative judgment statements included those statements that indicated the artifact, per the principal's judgment, did not sufficiently

illustrate/meet the ITS the elementary teacher had indicated it would. An example of a negative judgment statement was made by Keith when he said, "I don't see any proof that the lesson lends itself well to all learning styles."

Table 31 represents the results of tagging each of the high school principals' judgment comments in Phase II. Positive comments represented 61% of the total judgment comments made by the high school principals. Negative comments accounted for 14% of the total judgment comments while neutral comments accounted for 24%.

Table 31

Classification of Judgment Comments (J/A) per Artifact made by High School Principals in Phase II

	Total J/A Comments	<u>Gavin</u>			<u>Kathy</u>			<u>Keith</u>		
		(+)	(N)	(-)	(+)	(N)	(-)	(+)	(N)	(-)
Artifact 1	19	6	3	0	2	3	1	3	0	1
Artifact 2	16	2	3	2	2	2	2	3	0	0
Artifact 3	15	8	1	0	1	0	2	3	0	0
Artifact 4	11	5	3	0	0	0	1	1	1	0
Artifact 5	18	6	5	0	4	1	0	1	1	0
Artifact 6	12	5	1	1	1	0	1	2	1	0
Artifact 7	20	5	2	0	3	0	4	5	0	1

(+) = Principal made positive statement about value of standard, (N) = Principal made neutral statement about value of standard, (-) = Principal made negative statement about value of standard

The principals were consistent in that, individually, their total positive comments outnumbered negative and neutral comments. Positive comments made by Gavin and Keith accounted for well over half of their judgment comments in Phase II. Kathy was less positive in her judgment with only 43% of her comments being positive. Every artifact received more positive than negative comments. In short, the artifacts used to evidence each of the standards were, by and large, judged as appropriate in meeting the ITS.

The principals were somewhat consistent in how their negative comments were distributed between the artifacts. Every artifact, with the exception of artifact five, received at least one negative comment from the high school principals. Artifacts two and seven had the highest percentage of negative comments at 25% each. While the number of negative comments showed some consistency across artifacts, the percentage of negative artifact judgment comments *per principal* in Phase II showed some difference. Gavin and Keith's negative comments accounted for less than 10% of their overall artifact judgment comments. However, 37% of Kathy's total judgment comments were negative.

In order to get a better sense of how high school principals thought about the artifacts in terms of their effectiveness in illustrating/meeting the ITS the high school teacher said they would, it was valuable to specifically compare artifact two and artifact five. Artifact two was chosen because it was one of two artifacts to receive the highest percentage of negative comments (25%) and only three comments separated positive and

negative comments. Artifact five was chosen because it was the only artifact that did not receive negative comments by any of the principals.

Artifact two was titled “web page.” The teacher included a printed version of the personal web-site he had created. He indicated that he had developed the website after attending a short class on web-site development. He included the following description: “The web-site offers my students and the community access to information about classes, grading scale, expectations, and assignments. The site is also an important way for parents to keep up communication about their children’s learning.” The teacher also offered the following reflection:

The web-site is a great tool for both students and parents. It lets the community learn more about me as a teacher and is a great way to share all kinds of classroom information. I would improve my web-site by adding a questions and comments page, where students and parents could contact me directly by posting messages. I could also work on updating my page more frequently, this would allow me to add daily assignments and reduce the amount of make-up work. Overall, I have had over 150 hits to the site, but this could also be improved. I have made a poster for my room to promote the web-site, but more awareness is needed to make the most out of this great tool.

The teacher included three printed pages along with the artifact description/reflection.

The printed pages were a screen-capture of the web-site. The teacher introduced himself, provided an announcement that grade reports had been sent home recently, included a random history fact, and outlined his classroom expectations. He indicated that he was attempting to meet criteria relative to ITS one, three, five, seven and eight. (see Appendix A).

The principals made 16 total artifact judgment comments for artifact number two. Of the 16 comments, seven were positive, five were neutral, and four were negative. On

the positive side, the verbal reports of all three principals indicated that they thought the web-site was a good way to communicate with and keep parents involved (ITS 1g and 8e). In addition, the three principals agreed that the teacher effectively evidenced that he had engaged in professional growth by attending a class on web-page development and then implementing the new knowledge.

All three principals noted that the teacher posted his classroom expectations. In the posted classroom expectations, the teacher indicated that students should abide by the Four B's which were, be on time, be prepared, be teachable, and be respectful. Each expectation was followed by a short, one sentence, explanation. For Keith, posting the expectations was sufficient to meet ITS 1e. Gavin liked the idea of having the expectations posted but made only a neutral judgment statement when he said, "He does set the expectations within his classroom expectations page. So he sets them. It doesn't say anything about following through necessarily." Kathy, like Gavin, acknowledged that the teacher had posted the expectations but added that, "So the only thing in this that I see—he talks about the 4Bs of classroom expectations. In that, he talks about respect as something everyone should expect and deserve. I would say to him 'that's not doing it'." Artifact two judgment comment counts for Kathy and Gavin suggested the uncertainty of their judgment (see Table 31). Their uncertainty appeared to stem from their skepticism that a web page could sufficiently meet the multiple criteria the teacher indicated; hence, the higher amount of negative and neutral comments.

In contrast to artifact two, artifact five did not garner any negative comments. Artifact five was titled "Extracurricular work." The teacher included three documents.

The first document was a letter from the teacher's principal to coaches and parents indicating that the teacher was volunteering to conduct a strength and conditioning program for coaches and athletes in the district. The second document was an email from the local Area Education Agency (AEA) indicating that the teacher had attended a training regarding a new General Education Intervention/Problem Solving Process. The teacher indicated that he would be conducting an in-service for the teachers in his district to share his new knowledge. The third document was an email that included minutes from a committee meeting regarding class sharing with another district, on which the teacher served.

The teacher's description of artifact five was, "The following is a collection of documents that demonstrate my willingness to be involved and contribute to the school community. The letters and notes illustrate my voluntary participation on committees, professional growth, and leadership." He reflected on the artifact by saying that, "I believe in order to be a more effective teacher one must become part of the community. The best way to become part of the community is to get involved, join clubs, work on committees, and volunteer. I have been willing to do this and I firmly believe it has made me a better person." The teacher indicated that he was attempting to evidence criteria relative to ITS one, five, seven, and eight (see Appendix A).

The principals made 18 total artifact judgment comments for artifact number five. Of the 18 comments, 11 were positive and seven were neutral. The high school principals did not make any negative comments relative to artifact five. The positive comments made by the high school principals centered on the teacher's ability to appropriately

evidence *most* of the criteria for ITS one, five, seven, and eight. In cases where they questioned evidence, they did not negatively judge the artifact. Instead, the high school principals indicated that the evidence was “probably” appropriate. Because they did not know for which district the teacher worked, the high school principals were not aware of the building, school, or district goals. However, the high school principals speculated as they judged the artifact. For instance, the principals noted that the Problem Solving Process training the teacher used as evidence was most likely a building or district initiative and that the teacher’s work on the class sharing committee evidenced his collaboration efforts and his work towards meeting district goals.

The high school principals gave the letter concerning the weight lifting program neutral ratings. They stopped short of negative judgment saying that additional information was needed. In some cases, the high school principal would indicate that the ITS/criteria the teacher was attempting to evidence in artifact five had been evidenced in a previous artifact or they would look ahead to find evidence, i.e., loosely used cross-referencing. As a result, there were more neutral artifact judgment comments for artifact five.

The comparison between artifact two and artifact five demonstrates that the high school principals needed to “see to believe” relative to evidence that occurred in the classroom (i.e., classroom expectations), resulting in more negative comments about the artifact. However, in the case of artifact five, evidence relative to professional development and meeting district goals, the principals were more likely to be neutral or positive because they felt more comfortable speculating.

While artifact judgment statements (J/A) dominated the judging activities of the high school principals in Phase II, judgment comments relative to the portfolio (J/P) and the teacher (J/T) were also present (see Table 29). There were only two comments made relative to portfolio judgment; one each by Gavin and Kathy. Both principals were very skeptical in Phase II concerning the proficiency of the high school teacher's portfolio. Gavin indicated that he thought the portfolio had some "holes" in it and revealed his skepticism when he said, "I'm hoping this isn't one of your [the researcher's samples] wonderful ones." Then, he reconsidered and said, "Well, everybody holds it [the portfolio] to different standards. Again, it might just be because I'm not used to this form of organization. Overall, he's probably doing what the state wants him to do."

Kathy echoed Gavin's skepticism. However, her skepticism seemed to stem from the size of the portfolio. She said, "It's way too little for something that he thinks is going to cover eight standards." This statement paralleled her thinking as she judged the artifacts. While she seemed comfortable with one artifact meeting multiple ITS/criteria, she sometimes had difficulty with stretching one artifact to meeting as many ITS/criteria as the teacher indicated in his Table of Contents (see Appendix I). Kathy's comment evidenced that she had limits in terms of cross-referencing. A view that was different than that of her middle school colleagues.

Given the dubious nature of the portfolio judgment comments made by Gavin and Kathy, it was expected that each would make similar comments relative to judgment comments about the teacher (J/T) in Phase II. However, the judgment comments made relative to the high school teacher were mostly positive. Surprisingly, after a somewhat

critical judgment of the portfolio, Gavin made no teacher judgment comments. Both Kathy and Keith did make comments regarding their judgment of the teacher. While Kathy was not particularly impressed with the teacher's portfolio, it did not appear that it would have a negative effect on how she judged the teacher. She said, "So there would be no sense of my trying to get this person not to get their teaching license." Kathy was impressed with the teacher's research into and the use of Problem Solving Processes in his teaching. She also liked the fact that the teacher was able to communicate, via his reflections, that students need to be respected and treated differently in the classroom. She noted that the teacher made some spelling and grammar errors.

Keith did not make any portfolio judgment comments and nearly 80% of his artifact judgment comments were positive. His positive artifact judgment seemed to carry over to his judgment of the teacher. He noted that the teacher seemed to have confidence and creativity when he said,

So that's a creative way of trying to do something [a lesson] and I really appreciate the fact that this teacher has that willingness to step outside what would be comfort areas. It would be interesting to know their background in their undergraduate work, if their teacher prep program prepared them somehow to do a creative lesson like that.

Keith was impressed with the teacher's involvement in the professional development activities evidenced in several of the artifacts. While Keith was happy with the teacher's confidence to be involved in these activities and his willingness to present new information to the rest of the faculty, he was concerned about how the teacher might be perceived by other faculty.

I would agree that I would want my first and second year teachers to begin to get involved. I am a little hesitant for a young teacher that they are bringing back the

learned skills to all teachers at their district and will be presenting that. I would hope that this person would also have an experienced teacher so that they're not construed as a know-it-all, which some veterans would definitely see. But I would very much be impressed with the fact that this young person is involved.

Keith was impressed with this teacher but seemed to want to protect him and mentor him so that he could be successful on all fronts.

Keith, like Kathy, noted that the teacher made spelling errors in the portfolio and on some of the printed material on his web-site. While the spelling errors did not seem to cause Keith to negatively judge the teacher, it was apparent that, if this were Keith's teacher, he would talk with him about the gravity of the errors. "This wouldn't sell him well to his students and it's just a spelling error. Obviously he would want to correct that."

In summary, artifact judgment (J/A) was the most dominating sub-category in the judgment category for high school principals. Verbal reporting comments relative to artifact judgment accounted for 91% of the total judgment comments made by high school principals in Phase II. High school principals, like their elementary and middle school counterparts, engaged in a four-step rhythm as they judged artifacts. The high school principals used some cross-referencing, but not to the degree of the middle school principals.

It was expected that the high school principals might exhibit the same artifact judgment patterns as did the elementary and middle school principals; however, this did not bear out. The elementary and middle school principals closely scrutinized the first third of the artifacts in their respective portfolios. After early scrutiny, their artifact

judgment comments decreased (see Tables 16 and 23). This was not true of the high school principals. The high school principals closely scrutinized each artifact (Table 31).

Even with the closer scrutiny, positive artifact judgment comments were more numerous than negative judgment comments for the high school principals. Positive comments represented 61% of the total judgment comments made by the high school principals. Negative comments accounted for 14% of the total judgment comments while neutral comments accounted for 24%.

The high school principals made only two comments relative to portfolio judgment in Phase II. The judgments they made relative to the portfolio in Phase II showed skepticism of the proficiency of the portfolio. The skepticism about the portfolio did not seem to affect the principals' judgment of the teacher in Phase II. The principals were mostly positive about their judgment of the teacher. They appreciated the teacher's confidence, his willingness to teach outside the norm, and his understanding of difference. Similar to the elementary and middle school principals, the high school principals were critical of the teacher's spelling and grammar in the portfolio.

Principal's opinion. The category of principal's opinion (PO) was the second category to cluster under judging activities. It had the seventh highest comment overall count for the high school principals in Phase II. Comments in the principal's opinion category tripled in Phase II, but accounted for only three percent of the total comments made by the high school principals in that phase. The principal's opinion category was a bi-level category that consisted of principals' opinions concerning the DE requirements associated with teacher assessment and principals' opinions concerning the portfolio as

evidence of good teaching. Table 32 illustrates the breakdown of the principal's opinion category comment counts for high school principals in Phase II.

Keith commented in Phase I but not in Phase II. Kathy was the only high school principal to comment in both Phase I and Phase II. She had only one comment in each phase.

Gavin commented only in Phase II and his comments accounted for 83% of the Phase II principal's opinion comments made.

Table 32

Principals' Opinion Comment Count Breakdowns for High School Principals in Phase II

	Total PO Comments	PO/DE	PO/PE
Gavin	5	3	2
Kathy	1	0	1
Keith	0	0	0

PO/DE = Principals' Opinion/State DE Teacher Assessment Requirements, PO/PE = Principals' Opinion/Portfolio as Evidence of Good Teaching

Gavin's comments concerning DE requirements associated with teacher assessment addressed two issues. First, he indicated that the person who evaluated the portfolio should be familiar with the person who created the portfolio. He expressed opposition to an outside team that might do portfolio reviews when he said,

Hopefully the evaluator knows what's going on, but somebody who's looking at it for the first time ...and that's one of the reasons that I do believe that if we're

going to have these, I think the administrator in the building needs to look at them instead of the way they talked about to begin with bringing in three or four people from the state to go around. It would be a little more difficult from that standpoint.

The second issue Gavin addressed revealed that he valued reflective teaching but took issue with the time demands on teachers to produce a portfolio. He said,

The other thing that I see here too is I like the idea that when he [the teacher] reflects ...the very last sentence he has "*If I use this strategy in the future this is what I'm going to do differently.*" I think that is so important that we do that. I think that's one of the things that the state wants us to do with these portfolios. However, what I find is good teachers have always been doing this. What I see in all these portfolios is basically all we're doing is making teachers put stuff down on paper that they've always been doing. To me, sometimes I believe that's kind of a waste of time. Especially with 1st and 2nd year teachers who are already struggling to put a lesson in front of their kids every single day and making it the best lesson. Then we're asking them to put hours of time in to a portfolio.

The two issues that Gavin addressed were not verbally reported by any other principal in the study. Issues regarding time dedicated to the portfolio were reported by other participants; however, the comments were relative to the amount of time the principal spent reviewing the portfolio, not the amount of time the teacher spent preparing it.

Gavin and Kathy had similar opinions relative to portfolios as evidence of good teaching. Kathy thought about a teacher in her building who had produced a six-inch thick portfolio. She indicated that the thickness did not necessarily correlate with good teaching when she said, "It was more of a red flag actually than it was a positive reflection on her teaching." This was somewhat contradictory to Kathy's Phase I position where she was skeptical that the teacher had not provided enough material in his portfolio.

Gavin also questioned the portfolio as clear evidence of teaching. He, too, thought about previous situations that he had experienced. “The one thing about portfolios is ...I’ve had teachers that I have not given a license to after the second year. They can make themselves look absolutely phenomenal when it comes to a portfolio. I want to see how it’s carried out. People can put anything on paper, but I want to see the results.” The opinions expressed by Kathy and Gavin relative to the portfolio as evidence of good teaching were similar to those expressed by their elementary and middle school counterparts.

Tools. The category of tools (T) had the sixth highest comment count for the high school principals in Phase II. Tools referred to any sort of instrument (paper or electronic) that a principal used as he/she evaluated the portfolio. The researcher did not prompt the principals to bring any tools with them to the portfolio review for fear of the principal creating a tool for purposes of the study. The high school principals made only three comments regarding tools in Phase I. Their comment counts increased to eight in Phase II. Each of the principals made comments relative to tools with Kathy and Gavin accounting for all but one of the Phase II comments.

In Phase I, Kathy indicated that she used a copy of the ITS/criteria as she judged. She expanded on her thinking in Phase II as she looked more closely at the artifacts. She used the teacher’s cross-reference information (Appendix I) but also made reference to a tool that she had developed. She indicated that she used a separate piece of paper that listed the ITS/criteria. She further described the tool when she said, “I have one [a tool] that we actually go through and write notes on each of the elements that they have. So

that when I'm doing this I check it off. I say they've met that standard or I need to have a conference with them about this particular one." Kathy did not provide the researcher with a copy of the tool. She only described the tool.

Similar to Kathy, Keith and Gavin made use of the cross reference information that the teacher provided (Appendix I). The use of the teacher's cross-reference information was the only tool to which Keith referred in Phase II. Because Gavin's school has teachers submit their portfolios electronically, he described a currently and previously used tool to evaluate. He said,

It's really pretty simple. I don't even write out the standards. I just put 1A, 1B, 1C all the way down the list. I will check to see if I can find it. We used that a lot more when we had these types of portfolios. Since we've gone to the E-portfolio, then we don't use that as much. This [the teacher's cross-reference] does help. He's got it a little more organized. This is what he believes he's got. He might be right.

Gavin gave the researcher an Evaluation Guide that was used in his district (Appendix U). The guide was similar to what Rob, one of the middle school principals, described as a template that his district used. The guide that Gavin provided listed the ITS/criteria and identified possible teacher behaviors and/or written documentation that would evidence that the teacher had met the standards. Gavin, like Rob, did not use the evaluation guide as he evaluated the portfolio provided to him by the researcher.

Similar to principals at the elementary and middle school level, high school principals used a written list of the ITS/criteria in conjunction with the cross-reference information included in the teacher's portfolio. None of the high school principals physically checked off or wrote notes about the ITS/criteria during their review. Gavin was the only high school principal that provided the researcher with a tool; however, he

only indicated that it was a tool, i.e., he did not physically use the tool as he reviewed the portfolio (Appendix U).

In summary, the three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). All three categories appeared in Phase II thinking of the high school principals. High school principals' comments relative to judging activities significantly increased in Phase II. For high school principals, artifact judgment comments (J/A) was the most dominant sub-category in the judgment category, accounting for 91% of the total judgment comments in Phase II. As was true with elementary and middle school principals, the high school principals engaged in a four-step rhythm as they judged artifacts. Similar to the middle school principals, the high school principals used cross-referencing as part of their judgment routine; however, to a lesser degree than their middle school counterparts.

The high school principals made more positive artifact judgment comments than negative. This evidence suggested that the high school principals thought the artifacts included in the portfolio met, at least to some degree, the ITS/criteria. This trend was similar to that of both the elementary and middle school principals; however the high school principals had a lower positive comment count. Positive comments concerning the artifacts used to evidence the ITS/criteria accounted for 61% of the total judgment comments made by the high school principals. Negative comments accounted for 14% of the total judgment comments while neutral comments accounted for 24%.

The judgment pattern of the high school principals did not parallel that of the elementary and middle school principals. Principals at the elementary and middle school

level closely scrutinized the first third of the artifacts in the portfolio. Consequently, the greatest share of their artifact judgment comments clustered around those artifacts and then tapered off. In contrast, the high school principals' artifact judgment comments were distributed consistently across all of the artifacts suggesting higher scrutiny of each artifact (see Table 31). However, the high school principals as a group did not single out one artifact that they deemed insufficient to meet the ITS/criteria, i.e., no artifact received more negative than positive comments.

In Phase II, two of the high school principals, Kathy and Gavin, verbally reported skepticism relative to judgment of the portfolio provided to them by the researcher. Kathy took issue with too much replication, i.e., one artifact representing multiple ITS/criteria while Gavin indicated that he thought there might be some "holes" in the portfolio. In essence, at least in Phase II, both principals indicated that additional evidence would have been beneficial.

The skepticism surrounding the portfolio did not seem to negatively affect how the principals judged the teacher in Phase II. There was no indication that the principals would not recommend the teacher for licensure based on their judgment of the artifacts and consequently the portfolio during Phase II. The principals indicated that the teacher seemed to be willing to try new teaching strategies, take leadership roles, and be open to ideas. While verbal reporting indicated that the high school principals, in general, positively judged the teacher in Phase II, they did not appear to develop the same trust relationship with the teacher who produced the high school portfolio as did the middle school principals with the teacher who produced the middle school portfolio. Similar to

both the elementary and middle school principals, the high school principals noted the presence of grammatical, typing, and spelling errors in the portfolio.

The high school principals' comments in Phase II included opinions regarding the DE requirements for teacher assessment and opinions relative to the portfolio as evidence of good teaching. One principal reported that having someone, other than the teacher's building administrator, evaluate portfolios would not be a good proposition. In addition, the same principal opined that asking teachers to create a portfolio might be a poor use of their time. Two principals commented relative to the portfolio as evidence of good teaching. In both cases, the principals indicated that the portfolio was only part of the evaluation process and that a good portfolio was not always representative of good teaching. The opinions of the high school principals concerning the portfolio as evidence of good teaching paralleled the opinions of the elementary and middle school principals.

Two of the high school principals verbally reported that they used a hard copy of the ITS/criteria as they reviewed portfolios. However, neither of the principals physically checked off the ITS/criteria during the evaluation. Comments indicated that the high school principals used the hard copy as a checklist to make certain that the teacher had addressed each ITS/criteria. The high school principals indicated that they would indicate "met" or "not met" on the checklist and write notes concerning the ITS/criteria. One principal provided to the researcher an Evaluation Guide used by his district (Appendix U); however, he did not physically use the guide during the evaluation of the portfolio provided to him by the researcher. All three of the high school principals used the cross

reference information (Appendix I) provided to them by the teacher who created the portfolio.

High School Principals and Coaching Activities in Phase II

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time investment (TI). Comments relative to coaching activities accounted for 27% of the total Phase II comments made by high school principals; the highest coaching activity percentage for all three academic levels. In this section, comment counts relative to the three categories (coaching, principal's role, and time investment) will be reported followed by a summary.

Coaching. The coaching (C) category had the second highest overall comment count for the high school principals in Phase I. Comments relative to coaching showed significant increase from Phase I to Phase II for the high school principals. There was only one comment relative to coaching in Phase I as compared to 60 comments in Phase II indicating that, similar to judging activities, coaching activities were heavily emphasized in Phase II. Comments in the coaching category represented 25% of the total coded comments in Phase II thinking for high school principals. The coaching category was a multi-level category. Table 33 illustrates the breakdown of the coaching category comment counts for the high school principals. As a group, the high school principals made comments about each of the four coaching sub-categories.

All three principals made comments relative to meetings with the teacher (C/M). However, the C/M comments accounted for only 10% of the larger coaching category. In general, the high school principals viewed meetings with the teacher as a time for the

teacher to provide clarification and help the principal make connections. Kathy commented that,

...then I would conference with the teacher and say, "I don't quite get ...maybe it's here." Often that happens, that they were in there teaching the lesson so they know what they did that might have said how they related to student development. But it might not be evident to me as I look at this. I would then talk with the teacher and say, "Talk me through this. Give me some other ideas. Add to this portfolio in this way."

Table 33

Coaching (C) Comment Count Breakdowns for High School Principals in Phase II

	Total Coaching Comments	C/M	C/PP	C/Q	C/SA
Gavin	19	1	2	0	16
Kathy	29	5	5	6	13
Keith	12	1	0	0	11

C/M = Coaching relative to meetings with the teacher, C/PP = Coaching relative to portfolio preparation, C/Q = Coaching relative to the use of questioning, C/SA = Coaching relative to suggested alternatives for artifacts

Kathy's comment was representative of the verbal reports made by the high school principals relative to coaching during meetings with teachers. The comments made by the high school principals relative to meeting with the teacher were very similar to those made by their elementary and middle school counterparts; the meetings included formative coaching and created professional dialogue.

Coaching, relative to the portfolio process (C/PP), accounted for 12% of the total coaching comments made by the high school principals in Phase II. Keith did not make any comments relative to coaching the portfolio process. Gavin and Kathy once again demonstrated their desire to coach teachers through the portfolio process rather than to penalize them if evidence in the portfolio is lacking. For instance, Kathy said,

So then, once I've gone through all of that [the artifacts/portfolio] and made notes on things ...sometimes what I do too with people when it looks like I'm going to want a lot more than they've wanted to give me, is that ...I just went back and said, "Let's just do number one. Let's just get that. So rather than give me this whole book where I get to say "this isn't good enough, this isn't good enough" let's just start with standard one. You and I talk and you go out and collect things and then come back and give me standard one stuff." Then, once I've done that with a couple standards, then they know what I'm going to say. I'm going to say, "Where's the student work, where's the follow through?" I guess there would be some thought that I would have too so if it's at the end of year two and this was handed to me, what would I ...is there some things that I could have them go pull, do, collect for me that we really could end up meeting a lot of standards.

Clearly, Kathy had concerns about the teacher's portfolio at this point; however, she was also clear that her involvement up to this point would have helped shape the teacher's effort relative to what he should have included in the portfolio.

Gavin's coaching was centered on advising the teacher to supply additional information. Since the teacher was not on Gavin's faculty, he wanted more information in the portfolio that would help him make connections to the building goals under which the teacher worked. He explained how he would coach the teacher when he said,

One of my thoughts again here – even a statement in the cover page that says, "Our building goal is..." would help me identify. The one that he's saying next is 1b—Implement strategies supporting student, building and district goals. He talks about parental communication. It's very possible that communication could be a building goal, but I have no reference to that that I see.

Kathy and Gavin both seemed interested in partnering with the teacher to improve his portfolio; a perspective very similar to that of the middle school principals.

Kathy was the only principal to make comments in Phase II relative to coaching and the use of questioning (C/Q). Kathy's use of questioning was apparent when she said, "I would have talked to him about that before he even put this [artifact] in here. He seems to understand what he needs to do, but I'd like to see it." Then, she read from the artifact description, "*I also averaged the scores and display them in the classroom.* So I'd say, 'Show me. What were the scores that you averaged? What data are you basing whether this lesson was meaningful or not on?' *The students could watch their progress ...Good.* Where is it? Show me." Kathy's use of questioning in Phase II was similar to her other coaching activities in that she put herself in the role of mentor and facilitator with the teacher.

The coaching/suggested alternative (C/SA) category accounted 67% of the multi-level coaching category comments for high school principals. In contrast, the C/SA category garnered only 52% of the multi-level coaching category comments for elementary principals and 41% for the middle school principals.

The coaching/suggested alternative comments were part of the four-step rhythm employed by the high school principals. The percentage of C/SA comments was higher for high school principals than it was for elementary or middle school principals; however, similar to the elementary and middle school, C/SA comment counts, they did not represent a 1:1 ratio with artifact judgment comments. The high school principals

cumulatively made 111 artifact judgment comments (Table 31) and had 40 C/SA comments.

Given the number of artifact judgment comments made by the high school principals, it was not surprising that they dedicated a great deal of their coaching activity to making suggested alternatives. Table 34 illustrates the breakdowns of the number of suggested alternative comments, per each artifact, made by the high school principals in Phase II.

Table 34

Suggested Alternative (C/SA) Comments per Artifact by High School Principals in Phase II

	Artifacts*						
	#1	#2	#3	#4	#5	#6	#7
Gavin	3	3	1	3	1	3	2
Kathy	2	2	1	2	1	2	3
Keith	0	1	4	0	0	2	4

* See Appendix I for Table of Contents and Cross-referencing Information

The distribution of suggested alternative comments made by the high school principals showed some correlation with the negative artifact judgment comments (Table 31). For instance, artifact five did not receive any negative artifact judgment comments. As a result, the high school principals made only two suggested alternatives for artifact five. In

contrast, 25% of the artifact judgment comments made for artifacts two and seven were negative; consequently, the principals made more suggested alternative comments for those artifacts.

The suggested alternatives that the high school principals had in common were that the teacher should include (a) additional artifact description, (b) additional teacher reflection, (c) additional evidence of teacher follow-through in terms of professional development, (d) student work samples, (e) student reflection, (f) evidence of pre and post testing and how data was used to adjust teaching/instructional strategies, and (g) discipline data and records. The suggested alternatives made by the high school principals were very similar to those of the elementary and middle school principals. However, the high school principals placed more emphasis on evidence relative to professional development, assessment, and discipline.

In regard to professional development, Gavin and Kathy both indicated that additional evidence relative to implementation of professional development training needed to be included. For instance, as Gavin reviewed documentation indicating that the high school teacher had attended a Problem Solving Process Workshop, he said, “He’s definitely acquiring it [professional development] within these documents, but he might want to show how he’s implementing it.” Kathy’s request for more information relative to professional development was similar. She said,

Part of what I’d like to see is not just that he attended this professional growth activity, but then what did he do with it. Show me some documentation from the problem solving sessions that you then actually had in your school once you learned about this process at the Hawkeye Tech thing. I’d be looking for ...it even says *Please bring copies of your current problem solving forms*. I’d like to see

those. That would be helpful. Anybody could have pulled this off the Internet. It doesn't talk to me about the application or your own professional growth.

Both principals wanted more evidence of how the training made its way into the teacher's building and/or classroom, i.e., evidence of implementation and results of implementation.

In regard to assessment data, the suggested alternatives made by the high school principals in Phase II focused more heavily on assessment than did the verbal reporting in Phase II of the elementary and middle school principals. All three high school principals made comments relative to the use, value, and implementation of assessment. Gavin indicated his advocacy of pre- and post-testing when he said,

I'm a real believer in pre- and post-tests. I give the example all the time in American government that if you give a pre-test and all the kids know 90% of what you want them to know about three branches of government; I sure as heck hope that you don't want to spend three weeks on the branches of government.

He later indicated that the teacher should use data gathered from assessment to make instructional decisions. He said, "What I'd like to be able to see in something like that, how did you change what you do based on student data? Do you take extra time in class to drill and kill or do you put them in small groups because they're struggling with that or do you avoid five of the words because they know them all. I'd like to see how it actually drives decision making."

Kathy echoed Gavin's comments regarding assessment when she said, "That's what I'd like to see in your portfolio—how did the kids do on this? What were the pre-test and post-test scores on this activity?" Keith's concerns regarding assessment had a slightly different skew. His suggested alternatives indicated that he wanted the teacher to

be very clear to students about assessment criteria and that the teacher should allow students to self-assess. He further indicated that the student self-assessments should be included in the teacher's portfolio.

Relative to discipline, Kathy and Keith recommended that the teacher include conduct reports and/or discipline records. The suggestions were made in reference to artifact seven of the portfolio. The teacher indicated that he was attempting to meet all or part of ITS one, three, five, six, and eight (Appendix A). The artifact was titled "Classroom Expectations." The teacher described the artifact in the following way:

The following documents are to illustrate the positive learning environment I have in my classroom. I create a learning community that involves mutual respect and active learning. Expectations for my classroom are consistent and fair even though no two people are the same or behave the same way. I believe in keeping rules basic and try to use common sense in dealing with students in all situations.

The teacher included three documents. The first was a copy of his syllabus with the classroom expectations clearly delineated. Second, the teacher included a "My Job, Your Job" activity; the same activity included by the middle school teacher. The activity was designed to have students think about their role in the classroom as well as the role of the teacher. The teacher included a blank activity sheet that was absent of student responses.

The third document the teacher included was photographs of his classroom evidencing that he displayed his expectations and rules in the front of the classroom. The photographs also included student work that had been displayed on the walls as well, i.e., posters that students had created relative to their unit on ancient American development.

The reflection the teacher provided for Artifact seven focused on classroom management and classroom learning environment. He said,

Classroom management is an area that I pride myself in. I feel that the classroom should be a place where a student feels safe and free to learn. I keep my rules simple and encourage students to interact with one another in a positive manner. Every student is different and I treat every student differently. Some may think this is unfair but I believe just the opposite. It is more just to handle each situation, with different circumstances, as it arises than to have a set protocol that can be a hindrance in resolving conflict. Although I may handle each situation differently, I am consistent and fair. I give students the environment in which they feel comfortable and are [SIC] open to learning.

Kathy was very critical of ITS seven. In general, the three documents provided by the teacher were just a starting point from Kathy's perspective. She indicated that she wanted more than an implication that the teacher was meeting the standards. Specifically, she suggested, "I'd like to see some discipline records or conferencing that deals with students and parents or the assistant principal ...how many of the girls did you send and how did they get resolved? Keith made a similar suggestion when he said, "His classroom management I think he could also be well served to have examples of how he ...write up a little review of how he handled different situations may be beneficial to him. I don't know if they have conduct reports or once again, communications to parents about classroom expectations – that might help in there." The similarities between the two perspectives were quite obvious.

Principal's role. Comments relative to principal's role (PR) increased from one comment in Phase I (Table 7) to three comments in Phase II (Table 28). Kathy was the only principal to make PR comments in Phase II. Her commitment to fulfilling her role was evident when she said,

I always stop and think ...you really do want this to be a growing process for him [the teacher]. At this particular point, at the end of your two, to say that someone is not competent would be a horrible, horrible thing to have happening. Devastating for a young teacher. So again, I would just constantly reflect on my

own role in making sure – if this is a young talented teacher, my role is to make sure by the end of year two that they've done what it is. If I haven't done my part, that's not his fault.

Kathy's comments were very similar to those of Brenda, one of the elementary principals. Both principals assumed, and took very seriously, the responsibility of mentoring and coaching their teachers.

Time investment. The category of time investment (TI) occurred for the first time in the study for high school principals during coaching activities in Phase II. Only one comment relative to time investment was made. Kathy asked the researcher how to proceed with the review. She asked, "Do you want me to review this [criteria/evidence] for every single artifact? Like I would do it?" The researcher responded by saying, "Yes." Kathy simply acknowledged that, "When I do this [review a portfolio], it probably takes me an hour or two hours to kind of go through." Kathy's comment solidified the fact that she intended to review the sample portfolio with the same commitment with which she reviewed one from her own building.

In summary, the three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time(TI). While the high school principals made verbal comments relative to all three categories in Phase II, their thinking was dominated by coaching/suggested alternatives. This was not surprising given the number of artifact judgment comments (Table 31) made by the high school principals and the fact that they closely scrutinized all seven of the artifacts the teacher included in the portfolio. The close scrutiny of all of the artifacts was different from the judgment pattern of the elementary and middle school principals. The elementary and middle school

principals looked closely at the first one-third of the artifacts in the sample portfolio and then their scrutiny decreased (see Table 16 and Table 23). As the judgment comments made by the elementary and middle school principals decreased, so did their suggested alternative comments. For high school principals, the suggested alternative comment pattern paralleled the artifact judgment pattern; consequently, a higher percentage of suggested alternative comments existed.

In Phase II, the high school principals' comments were similar to each other relative to the suggested alternatives they made. The suggested alternatives that the high school principals had in common were that the teacher should include (a) additional artifact description, (b) additional teacher reflection, (c) additional evidence of teacher follow-through in terms of professional development, (d) student work samples, (e) student reflection, (f) evidence of pre and post testing and how data was used to adjust teaching/instructional strategies, and (g) discipline data and records. The suggested alternatives made by the high school principals were similar to those made by the elementary and middle school principals in Phase II. However, the high school principals placed more emphasis on evidence relative to professional development, assessment, and discipline.

During Phase II, high school principals did not put the same emphasis on coaching/meetings as did the middle school principals. The high school principals compared more closely to the elementary principals in terms of the number of verbal comments relative to meeting with teachers. Principals from all three academic levels verbally reported that meetings were used to coach teachers in terms of artifact quality

and the portfolio process. In Phase II, only one high school principal made comments directly related to her role as a principal. She indicated that her role was to monitor the teacher's progress closely during his first two years and to partner with him as he developed his portfolio.

High School Principals and Critical Pieces in Phase II

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two "pieces" carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. In this section, comment counts relative to the two sub-categories, critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO), will be reported, followed by a summary.

The critical pieces (CP) category had the third highest comment count for high school principals in Phase II thinking. Phase II produced significantly more comments relative to critical pieces than Phase I. The high school principals made only one comment relative to critical pieces in Phase I and sixteen comments relative to critical pieces in Phase II.

Table 35 illustrates the breakdown of the critical pieces category comment counts for high school principals. While only one high school principal had comments in the critical pieces category in Phase I, all three of them made comments in Phase II.

Table 35

Critical Pieces (CP) Comment Count Breakdowns for High School Principals in Phase II

	Total CP Comments	CP/RO	CP/TR
Gavin	6	6	0
Kathy	6	2	4
Keith	4	3	1

CP/RO = Critical Pieces/Role of Observation, CP/TR = Critical Pieces/Teacher Reflection

Role of observation. Nearly 70% of the critical piece comments were relative to the role of observation (CP/RO) for the high school principals in Phase II. This was higher than the percentages for elementary principals (41%) and the middle school principals (43%) in Phase II. While the high school principals' percentage of comments relative to observation was higher than the elementary or middle school principals, the message was the same; observation was very important.

Kathy and Gavin were both surprised the teacher did not include observations from his own principal in his portfolio. The value that Kathy placed on the administrator's observation was evident when she said,

Part of what I'm thinking about on here, too, is that none of these [artifacts] have classroom observations by the administrators as documentation. So I am wondering why there's no document here saying, "Dear [teacher], I was in your class on this day, this is what I saw." That's what I would do. At the bottom [of the observation] I'd say, "This is what I saw. Based on what I've observed in your classroom, I would say these criteria have been met." All of my teachers would

have, within their portfolios, a write-up of what I did and we would have gone through that and said which criteria/standard we think we met.

Gavin also commented on the importance of the teacher including administrator evaluations when he said, “I think if he would have some artifacts in here from evaluations that, maybe his principal had seen some of his stuff, it would make me feel better that yes, he’s doing some of the stuff. But by just looking at paper, it doesn’t do me a whole lot of good.”

All three high school principals made verbal comments suggesting that they preferred to observe the teacher themselves. For instance, Kathy commented that, “They [e.g., communication with students] might not appear in a portfolio, but would just appear in my eyes.” Gavin remarked that, “I need to see the syllabi actually being worked in the classroom.” Keith’s perspective regarding observation was more focused on his personal observation of a teacher. He was reviewing artifact one, an artifact used to evidence ITS number 2-Demonstrates competence in content knowledge appropriate for the teaching position, when he said, “That I would have had to see. I’d like to have viewed the classroom.”

Teacher reflection. Kathy made 80% of the comments relative to teacher reflection. Keith made one comment while Gavin made none. Kathy and Keith both appreciated the fact that the teacher was able to reflect on his own teaching. It was clear that both principals relied on the teacher reflection to assist them in the review of the artifacts and hence, the portfolio. Keith read each reflection statement aloud and then used the information to review the associated artifact.

Kathy, too, read each reflective statement but would have liked additional information in each statement. As she reviewed artifact number two in the high school teacher's portfolio she said, "He talks about how he would improve his web page in the future. What he's done here with the reflection about the artifact – that's good. I wish he had reflected about each of the things that he thinks that this artifact meets."

In summary, the high school principals valued observation and teacher reflection. Two of the high school principals suggested that the teacher include his principal's evaluation in the portfolio. All three high school principals expressed a preference to personally evaluate the teacher. While observation was valued as a way of corroborating the evidence included in the portfolio, teacher reflection was valued as a way for the teacher to more fully inform the principal about the purpose of each artifact. In short, the reflection statements made by the teacher were a guide, via written means, for the principals.

Summary of High School Principals' Thinking in Phase II

Phase II (judgment) thinking was described as the period of time following pre-assessment thinking when principals actively judged the artifacts in the portfolio provided to them by the researcher. Processing activity drastically declined in Phase II as the high school principals focused away from processing activities and towards judging and coaching activities.

The high school principals, like their elementary and middle school counterparts, developed a four-step rhythm as they judged artifacts. First, they would identify and verbally describe the artifact. Second, they would actually read aloud as they focused on

what the teacher was attempting to illustrate. Third, they would make a judgment statement regarding the artifact they were judging. The judgment statement was then sometimes followed by a coaching statement. Like their middle school counterparts, the high school principals used cross-referencing as they judged; however, to a lesser degree.

During Phase II, the high school principals closely scrutinized each of the seven artifacts in the portfolio; a judgment pattern much different than the elementary and middle school principals where only the first third of the artifacts were scrutinized. The high school principals made more positive than negative artifact judgment comments suggesting that they thought the artifacts, at least to some degree, met the ITS/criteria.

Additional findings in Phase II indicated that the overall percentage of negative artifact judgment comments was higher for the high school principals than for the elementary or middle school principals. This evidence suggested that the high school principals did not gain the same sense of confidence in the evidence contained in the high school teacher's portfolio that the elementary and middle school principals did for the teachers who created the sample portfolios they reviewed. Despite what appeared to be a lack of confidence in the evidence in the high school portfolio, the high school principals as a group did not single out one artifact that they deemed insufficient to meet the ITS/criteria, i.e., no artifact received more negative than positive comments.

The higher percentage of negative artifact judgment comments prompted some skepticism from the high school principals regarding the quality of the sample portfolio. Verbal reporting indicated, however, that additional information and conversation with the teacher would most likely remedy any deficiency. Neither did the skepticism

regarding the quality of the portfolio appear to negatively affect how the principals judged the teacher in Phase II. There was no indication that the principals would not recommend the teacher for licensure based on their judgment (during Phase II) of the artifacts, and consequently, the portfolio. Similar to both the elementary and middle school principals, the high school principals noted the presence of grammatical, spelling, and typing errors in the portfolio.

In addition to being highly engaged in judging activities in Phase II, the high school principals were also highly engaged in coaching activities. The greatest percentage of their coaching activity was spent making suggested alternatives relative to the artifacts the teacher used to evidence the eight ITS/criteria; however, the coaching/suggested alternative comments did not reflect a 1:1 ratio with artifact judgment comments. The suggested alternatives that the high school principals had in common were that the teacher should include (a) additional artifact description, (b) additional teacher reflection, (c) additional evidence of teacher follow-through in terms of professional development, (d) student work samples, (e) student reflection, (f) evidence of pre and post testing and how data was used to adjust teaching/instructional strategies, and (g) discipline data and records. The suggested alternatives made by the high school principals were similar to those made by the elementary and middle school principals in Phase II. However, the high school principals placed more emphasis on evidence relative to professional development, assessment, and discipline. High school principals also verbally reported that they placed importance on meetings as a means of coaching teachers. One principal

at the high school level was very specific about her role as a mentor and facilitator in the portfolio process.

During Phase II, some mild discontent was expressed by one high school principal with the DE requirements regarding teacher evaluation; current and proposed. The principal specifically indicated that having the teacher collect and present artifacts (portfolio) might be a “waste of valuable teacher time.” The same principal was also dubious about having someone other than the teacher’s administrator evaluate the portfolio. He reported concern that “... they [the DE] talked about bringing in three or four people from the state to go around and review [the portfolios].” Also during Phase II, the high school principals verbally reported that they viewed the portfolio as only a portion of the teacher evaluation process and that a good portfolio was not always representative of good teaching. This opinion was evident in the verbal reporting across academic level.

Similar to the elementary school and middle school principals in Phase II reporting, the high school principals indicated that they valued observation and teacher reflection as critical pieces of teacher evaluation. The high school principals valued observation summaries from other sources (teacher’s administrator) but more highly valued their own observation. The high school principals valued teacher reflection because of the insight it provided to the principal regarding the teacher’s thinking about each artifact included in the portfolio. Given the fact that the principals did not know the teacher, the written perspective provided unspoken guidance.

Phase II: Judgment Thinking per Geographic/Demographic Region

In this section, the findings relative to research question one for Phase II (judgment) thinking will be reported per geographic/demographic region. First, a brief overview of the data collected for geographic/demographic region is provided. Total comment counts across geographic/demographic region for Phase II will be reported and illustrated. Then, findings for each geographic/demographic region for Phase II will be reported followed by a summary.

It is important to be reminded that while principals in each academic level (elementary, middle school, and high school) reviewed the same portfolio, the geographic/demographic region (urban, suburban, and rural) verbal comment counts were representative of comments made by the *same* principals *across* academic level, i.e., the principals in geographic/demographic region did not review the same portfolio because they represented differing academic levels. Therefore, data from geographic/demographic region is representative of comparison between urban, suburban and rural factors only; academic level data will not be revisited. Data will not be reported using the four broad categories as was true with the academic level reporting because it would be repetitive. Rather, data comparing Phase I and Phase II results per each of the ten coded categories will be reported.

Geographic/Demographic Overview

Each geographic/demographic region represented one principal from each academic level (elementary, middle school, high school). The rural principals that participated in the study were Brenda, Ivan, and Keith. The suburban principals that

participated in the study were Norma, Mike, and Gavin. The urban principals that participated in the study were Leo, Rob, and Kathy. The total number of verbal comment counts made in the study for each geographic/demographic region was remarkably similar (see Figure 2). Suburban principals had the highest overall percentage of coded verbal comments with 34%. The rural and urban principals each accounted for 33% of the total verbal comment counts for the study.

Table 36 illustrates comment counts for Phase II by geographic/demographic region per the ten coded categories. As was true in Phase I, all ten of the coded comment categories were represented in Phase II.

Table 36

Summary of Phase II Comment Counts per Geographic/Demographic Region

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Rural	273	58	1	28	130	7	8	3	26	3	9
Suburban	252	40	7	17	148	11	8	0	12	9	0
Urban	255	69	5	9	128	10	5	3	16	9	1

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS = Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

The total number of comments for each geographic/demographic region was similar in Phase II. The three regions were separated by only 21 comments. The rural principals

made the most comments in Phase II accounting for 35% of the total coded comments in Phase II per geographic/demographic region. This was nearly twice as high as their comment count percentage in Phase I where they accounted for only 18% of the total coded comments per geographic/demographic region. Urban and suburban principals' comment counts in Phase II were within three comments of each other accounting for 33% and 32%, respectively, of the total coded comments per geographic/demographic region. As noted in academic level results, Phase II thinking comments accounted for 79% of the total coded comments for the study.

Rural Principals' Thinking in Phase II

Table 37 illustrates the breakdowns of Phase II (judgment) comments made by the rural principals. Rural principals' comments included all ten of the coded comment categories in Phase II.

Table 37

Summary of Phase II Comment Counts made by Rural Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Brenda	71	19	0	11	30	5	2	3	0	1	0
Ivan	157	27	0	13	73	2	6	0	26	1	9
Keith	45	12	1	4	27	0	0	0	0	1	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

In Phase I (see Table 9) only six of the coded categories were present. The categories that were present in Phase II but not in Phase I were coaching (C), process steps (PS), tools (T), and time (TI).

Not surprisingly, the percentage of comments pertaining to the structure of the portfolio (PFS) was less but still present to some degree in Phase II. Judging (J) and coaching (C) accounted for nearly 70% of the rural principals thinking activities in Phase II. This was a marked difference from Phase I where there were only two judgment comments and no coaching comments.

Two additional categories showed noticeable increase for the rural principals as they moved through Phase II. The first, critical pieces (CP), which included the sub-categories of the role of observation and teacher reflection, emerged as principals moved through artifact judgment and identified observation and teacher reflection as keys to successful teacher evaluation. The second category to show noticeable increase was principal's opinion (PO). While judging artifacts, the rural principals began to more frequently comment about the system of teacher assessment and the portfolio as evidence of good teaching.

One rural principal, Ivan, accounted for every process step (PS) comment in Phase II. Interestingly, Ivan was also the only rural principal to make comments relative to time (TI). While the category of tools (T) appeared for the first time in Phase II for the rural principals, the comment counts were minimal for the category.

In summary, the number of overall Phase II comments made by rural principals showed significant increase. The rural principals kept pace with the suburban and urban

principals in terms of the number of overall comments in Phase II (Table 36) as opposed to Phase I (Table 8) where they made half as many comments as the other two groups. Comments relative to the role of observation and teacher reflection (CP) also showed a moderate increase from Phase I to Phase II as did the category of principal's opinion. Nearly 70% of the coded comments in Phase II for rural principals were relative to judging and coaching. Clearly, the focus of the rural principals was on judgment and coaching in Phase II.

Suburban Principals' Thinking in Phase II

Table 38 illustrates the breakdowns of Phase II (judgment) comments made by the suburban principals. Only eight of the ten coded comment categories were present in Phase II for the suburban principals. The principals did not make comments relative to principal's role (PR) or time investment (TI). Principal's role was not present in either phase while time investment was present in Phase I (see Table 10) but not in Phase II.

Table 38

Summary of Phase II Comment Counts made by Suburban Principals

	Total Comments	C	CL	CP	J	PFS	PO	PS	T
Norma	37	3	5	3	24	1	0	1	0
Mike	109	18	0	8	65	1	3	9	5
Gavin	106	19	2	6	59	9	5	2	4

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal's Opinion, PS = Process Steps, T = Tools

The coded comment categories of judgment (J) and tools (T) were present in Phase II but not present in Phase I. Judgment comments accounted for nearly 60 % of the Phase II comments made by the suburban principals. Comments in the tools category emerged in Phase II as the suburban principals thought about the tools they used to “check off” which ITS/criteria each artifact had met. In all cases, the tool to which the principals referred was a one-page copy of the ITS/criteria. Coaching (C) comments also increased significantly in Phase II. The suburban principals only made three coaching comments in Phase I as compared to 40 comments in Phase II. The coaching category accounted for roughly 15% of the Phase II comments made by the suburban principals.

Additionally, critical pieces (CP) showed noticeable increase in Phase II as principals began to comment on the value of the role of observation and teacher reflection as they judged each artifact. Comfort level (CL) comments increased for the suburban principals in Phase II; however, this was not due to low comfort level. Norma’s comfort level comments increased because she was verbally reporting a high level of confidence rather than discomfort. At the same time, Mike’s comfort level comments dropped from three in Phase I to zero in Phase II. These results indicated that comfort level increased as the suburban principals moved from Phase I to Phase II.

While judgment appeared for the first time, and coaching, along with critical pieces comments increased, there was a decrease in comments relative to portfolio structure. The decrease in the percentage of comments relative to portfolio structure was consistent across academic level and geographic/demographic region in Phase II. Slight increases existed in principal’s opinion (PO) and in process steps (PS) during Phase II for

the suburban principals. One suburban principal had concerns about the amount of “teacher” time the development of a portfolio consumed and about having someone other than the building principal evaluate the portfolio, i.e., an outside group established by the DE. The suburban principals were clear that a good portfolio was not always an indicator of a good teacher. The increase in process steps was attributable to one principal at the middle school level.

In summary, the trend of the suburban principals was similar to that of their rural counterparts in Phase II in that comments regarding coaching and judging significantly increased. Noticeable increases also existed in the categories of critical pieces and comfort level. However, the comfort level increases were due to increased confidence not increased discomfort. As was the trend for the rural principals, the portfolio structure comments decreased for the suburban principals in Phase II. Also in Phase II, the suburban principals made more frequent reference to the use of a copy of the ITS/criteria as a checklist/tool than they did in Phase I. One of the suburban principals expressed mild concerns about the DE teacher assessment requirements and one about the potential for some other than the building principal conducting portfolio reviews. In addition, the suburban principals reiterated in Phase II that the portfolio was not the full picture of good teaching.

Urban Principals' Thinking in Phase II

Table 39 illustrates the breakdowns of the Phase II comments made by the urban principals. Similar to the rural principals, the urban principals attended to all ten of the coded comment categories during Phase II. Two categories that were present in Phase II

but not present in Phase I (see Table 11) were principal's role (PR) and time (TI); neither of which garnered significant verbal reporting.

Table 39

Summary of Phase II Comments made by Urban Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Leo	55	7	2	3	34	6	0	0	2	1	0
Rob	111	33	1	0	58	1	4	0	9	5	0
Kathy	89	29	2	6	36	3	1	3	5	3	1

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS = Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

The percentage of comments relative to portfolio structure (PFS) decreased in Phase II. This pattern was consistent across geographic/demographic region as well as academic region. The coded comment categories of coaching (C) and judgment (J) showed significant increase. There were only three coaching comments in Phase I as opposed to 69 comments in Phase II. Coaching comments accounted for 27% of the total Phase II comments for the urban principals. There was only one judgment comment in Phase I as opposed to 128 judgment comments in Phase II. Judgment comments accounted for half of the total Phase II comments for the urban principals. The increase in

coaching and judgment comments in Phase II was a consistent trend across geographic/demographic region and academic level.

Three coded comment categories showed moderate increase in Phase II for the urban principals. Comment counts made by urban principals relative to process steps (PS) increased from five comments in Phase I to eleven comments in Phase II. In particular, Rob, the middle school principal, became more descriptive of how he moved through artifact judgment. In addition, the coded comment category of tools (T) showed some moderate increase in Phase II. Again, it was Rob who demonstrated the greatest increase. His increase in verbal comments relative to tools correlated with his increase in process steps as he described how he used the formal summary template provided by the DE to assist him as he judged the artifacts. A moderate increase was also seen in the coded comment category of critical pieces (CP) as the urban principals commented on the significance of observation and teacher reflection. The category of critical pieces increased at a moderate to significant rate across geographic/demographic region.

Two categories showed slight increase in Phase II for the urban principals. The number of comment counts relative to comfort level (CL) increased but accounted for less than two percent of the total Phase II comments for the urban principals. A slight increase was seen in principal's opinion (PO) as well. The opinions expressed by the urban principals were similar to those expressed by the rural and suburban principals. The urban principals, like the rural and suburban principals, expressed that the portfolio was only a part of teacher evaluation.

In summary, the urban principals comment counts showed significant increase in judgment and coaching in Phase II. The significant increase in the judgment and coaching categories mirrored the trend established for both the rural principals and the suburban principals during Phase II. While some increase was also noticeable in process steps, the increase was attributable to only one urban principal. The category of critical pieces also showed some noticeable increase as the urban principals began to think more often about the value of observation and teacher reflection.

Summary: Findings for Phase II Thinking

Findings in Phase II (judgment) established that (a) principals spent a significant amount of time in judging and coaching activities, (b) in general, principals across academic levels developed a four-step *rhythm* as they moved through the judgment phase, (c) an imbalance existed between artifact judgment comments and suggested alternatives (i.e., the artifacts were by and large judged positively), and (d) principals identified observation of the teacher by a principal and written teacher reflection as two critical pieces of teacher evaluation.

The findings for Phase II indicated that the participating principals spent the bulk of their time judging artifacts and engaging in coaching activities. Verbal comments relative to critical pieces (observation and teacher reflection) also increased during Phase II. These findings were true across academic level as well as geographic/demographic region.

Across academic level during Phase II, the data revealed that middle school principals had the most Phase II comments (see Table 12). They had twice the number of

Phase II comments as did elementary principals. The high school principals also spent significant time judging but were still lower than the middle school principals. The high school principals spent one and one-half as much time engaged in judgment as did their elementary counterparts.

Total comment counts across geographic/demographic region during Phase II revealed that the groups (rural, suburban, and urban) had similar numbers of comments during Phase II (see Table 36). The three regions were separated by only 21 comments. The rural principals made the most comments in Phase II accounting for 35% of the total coded Phase II comments per geographic/demographic region. This was nearly twice as high as their comment count percentage in Phase I where they accounted for only 18% of the total coded comments per geographic/demographic region. Urban and suburban principals' comment counts in Phase II were within three comments of each other accounting for 33% and 32%, respectively, of the total coded comments per geographic/demographic region.

Verbal reporting revealed that the participating principals used a four-step rhythm as they made artifact judgment. The principals would (a) identify and verbally describe an artifact aloud, (b) read aloud from the teacher reflection and/or artifact description, (c) make a judgment statement regarding the artifact, and (d) sometimes make a coaching statement about the artifact. This type of rhythm was consistently demonstrated by all principals in the study. Cross-referencing (using one artifact to meet several ITS/criteria) was considered by all of the principals in the study as well; however, the use of cross-referencing was used most extensively by middle school principals.

Principals in the current study were consistent in the particular types of artifacts they believed to effectively evidence the ITS/criteria. The artifacts most often identified by principals as providing effective evidence were (a) observation summaries completed by the teacher's principal or a peer, (b) samples of student work, (c) pictures of activities, (d) copies of two-way email communication with parents and colleagues, (e) rubrics, (f) lesson plans, and (g) minutes from professional development sessions and committee assignments.

In general, the principals in the current study positively judged the majority of the artifacts in their respective sample portfolios; however, the elementary and middle school artifact judgment statements (see Tables 16 and 23) reflected a higher amount of positive judgment comments than did the high school principals (Table 31). The elementary principals and middle school principals exhibited similar judging patterns in that each group closely scrutinized the first one-third of the artifacts and then their judgment comments decreased. In contrast, the high school principals closely scrutinized every artifact in their sample portfolio; thus, the comment counts were distributed more consistently among the artifacts in the high school portfolio.

In regard to coaching activities in Phase II, the coded comments indicated that coaching comments relative to specific artifacts were common in the four-step judgment rhythm that was established by the principals; however, coaching comments did not represent a 1:1 ratio with judgment comments. In other words, a coaching comment was not made each time a judgment comment was made. On average, roughly 28% of the combined judging/coaching comments represented coaching. The percentage was lowest

for elementary principals (25%) and highest for high school principals (33%). The middle school principals coaching comments represented 28% of their combined coaching/judging comments. The somewhat higher percentage of coaching comments for the high school principals was expected given the higher level of negative comments surrounding the artifacts.

Principals most often made “suggested alternative” coaching comments specific to artifacts. Suggested alternatives were associated with artifacts that were judged negatively and positively. The suggested alternatives frequently reflected the principals need for clarification. Additional suggested alternatives were the use of (a) more frequent reflection (teacher and student), (b) additional proof of outcomes, (c) more variety, (d) evidence of student achievement in each artifact, (e) grammar and spell check, (f) more detailed artifact description, (g) authentic student work, and (h) data from other sources (e.g., student and/or parent surveys).

An additional coaching activity that was prevalent in Phase II was meetings with the teacher. Coaching comment counts relative to meetings with the teacher were second only to coaching/suggested alternatives. Principals indicated that the meetings with teachers were for purposes of clarification, dialogue, and coaching. The meetings appeared to be an avenue through which the principal could partner with the teacher in the portfolio process.

During Phase II, principal’s observation of the teacher and the value of written teacher reflection were identified as critical pieces of teacher evaluation. Because of the weight these two “pieces” carried throughout the study, the category of critical pieces

(CP) emerged as one of the four cluster categories; not so much based on a significant number of comments but on the importance placed on each of the sub-categories (observation and teacher reflection) via the pervasiveness of the comments throughout the study. Observation by an administrator was viewed as critical to teacher evaluation. The principals in the current study valued observation summaries made by the teachers' building principals when they were included. If written observation summaries were not included, principals highly suggested that they should be. While the written summaries from someone else were valued, the principals clearly expressed a preference to personally evaluate the teacher. Observation appeared to carry more weight than the portfolio.

The principals also indicated that they valued written teacher reflection. They suggested that teachers include more written teacher reflection with each artifact. The principals indicated that written reflections provided clarity, guidance, and depth to the artifacts included in the portfolio.

Phase III: Reflection via Guided Interview Questions per Academic Level

The findings in Phase III established that the principals who participated in the study (a) rated their respective portfolios as proficient and (b) placed great importance on observation in teacher evaluation. The findings in Phase III were, in part, a direct result of the guided interview questions posed by the researcher (Appendix K). In other words, the verbal reporting in Phase III was not a result of the less-guided think-aloud (TA) process as was the verbal reporting in Phases I and II. Because the researcher asked direct

questions, the verbal reporting data was skewed towards the categories that correlated with the interview questions. The guided interview questions were:

1. Given the rating possibilities of unsatisfactory, proficient, or exemplar, what rating would you give this portfolio? Why?

2. How much bearing would you assign to this portfolio when making a licensure decision? Explain.

3. Is this the same value that you place on portfolios you review from your own second-year teachers? What is the same/different?

4. Was today's process similar to/different than how you typically evaluate your own teachers? How?

In this section, the comments counts specific to Phase III will be reported just as they were in Phases I and II. First, an overview of the data will report and illustrate comment counts across academic level for Phase III. Then, findings for each academic level for Phase III will be reported. In addition, data relative to guided interview question number one will be reported. Data relative to guided interview questions two and three will be reported when research question three is addressed. The fourth guided interview question was a means of checking reliability of the study and was reported in Chapter 3.

Academic Level Overview

Table 40 illustrates the comment counts in Phase III of the nine principals by academic level. Each of the ten coded categories was represented in Phase III.

Elementary principals accounted for the most Phase III comments (40%) followed by high school principals (34%). The middle school principals accounted for 26% of the

Phase II comments. This was the only phase in which middle school principals did not account for the highest percentage of comments. Phase III comments accounted for 14% of the total coded comments in the study.

Table 40

Summary of Phase III Comment Counts by Academic Level

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Elementary	55	10	4	5	15	3	8	1	3	5	1
Middle School	35	4	3	7	7	0	9	0	1	1	3
High School	46	9	3	3	13	2	7	2	2	4	1

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS = Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

Given the nature of the guided interview questions, the category with the most comments was judgment (J) accounting for 26% of the coded comment counts for Phase III. Predictably, the majority of the judgment comments were made relative to the entire portfolio rather than to individual artifacts. Coaching (C) and principal's opinion (PO) each accounted for roughly 18% of the total coded comments; evidence that the principals continued to attend to coaching even when answering guided questions. In addition, the principals continued to consider the critical pieces (CP) of observation and teacher reflection. Critical pieces accounted for 11% of the total comment count in Phase

III. The remainder of the ten categories, comfort level (CL), portfolio structure (PFS), principal's role (PR), process steps (PS), tools (T), and time (TI) were all represented in Phase III but at levels of less than seven percent of the total number of comments made in the phase.

The ten coded categories clustered around four broader categories that provided a means of efficiently reporting data. The four broad categories were (a) processing activities, (b) judging activities, (c) coaching activities, and (d) critical pieces. Processing activities included the categories of process steps, portfolio structure, and comfort level. Judging activities included the categories of judgment, principal's opinion, and tools. Coaching activities included categories of principal's role, coaching, and time investment. Critical pieces included the categories of teacher reflection and role of observation. Each academic level will be reported via the four broad clusters using data from the appropriate smaller categories. In addition, data relative to guided interview question number one will be reported.

Elementary Principals' Thinking in Phase III

The elementary portfolio used in the study was from a fifth grade teacher and was organized by artifact. See Appendix H for the cross-referencing information used by the fifth grade teacher. The elementary principals (Brenda, Leo, and Norma) made the most Phase III comments of the three academic levels. Their thinking accounted for 40% of the overall comments made by all principals in the Phase III as compared to 26% for middle school principals and 34% for high school principals.

Table 41 illustrates the Phase III comment counts of the elementary principals. Norma accounted for 40% of the total Phase III comments while Brenda and Leo accounted for roughly 30% each.

Table 41

Summary of Phase III Comment Counts made by Elementary Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Brenda	16	4	0	1	2	2	1	1	2	3	0
Leo	17	1	1	1	9	1	4	0	0	0	0
Norma	22	5	3	3	4	0	3	0	1	2	1

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS = Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

Elementary Principals and Processing in Phase III

The three coded comment categories that clustered under processing were portfolio structure (PFS), process steps (PS), and comfort level (CL). Comments relative to processing accounted for 18% of the total Phase III comments made by elementary school principals compared to 77% in Phase I and 13% in Phase II. In this section, comment counts relative to each of the three categories (portfolio structure, process steps, and comfort level) will be reported followed by a summary of elementary principals and processing.

Portfolio structure. The elementary principals made few comments relative to portfolio structure in Phase III. Their comments reflected that, although the portfolio was not organized in a way to which they were accustomed, they appreciated the way the teacher had organized her portfolio. One principal indicated that handing her a box of artifacts was acceptable because the goal was to evidence the ITS/criteria in a way that works for the teacher.

Process steps. Again, the comment counts made by the elementary principals relative to process steps were not significant. Brenda simply outlined the steps that would occur after the review. She indicated that she would meet with the teacher to discuss the portfolio and begin her written evaluation based on the notes she (the principal) took during the portfolio review. Norma reiterated how she processed through each artifact, i.e., the four-step rhythm described in Phase II.

Comfort level. The comfort level comments made by the elementary principals were consistent with comments in Phase I and Phase II: Once the principals became familiar with the portfolio, they became more comfortable.

In summary, processing included the categories of portfolio structure, process steps, and comfort level. None of the three categories garnered significant comment counts in Phase III. The comments made relative to processing were consistent with the thinking of the elementary principals in Phase I and Phase II. In essence, the verbal reports indicated that as the principals became familiar with the portfolio, their comfort level increased.

Elementary Principals and Judging Activities in Phase III

The three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Comments relative to judging activities accounted for 51% of the total Phase III comments made by elementary school principals. The percentage of judgment comment counts for Phase III reflected the influence of guided interview question number one in which the principals were asked to rate the portfolio as unsatisfactory, proficient, or exemplar. In this section, comment counts relative to each of the three categories (judgment, principal's opinion, and tools) will be reported. In addition, data relative to guided interview question number one will be reported.

Judgment. The judgment category was a multi-level category that included judgment of artifacts (J/A), judgment of the portfolio (J/P), and judgment of teacher (J/T). The elementary principals made a total of 15 judgment comments of which judgment of the portfolio comments accounted for nearly 70%. This was largely due to the guided interview question that asked the principals to rate the portfolio as unsatisfactory, proficient, or exemplar. All three of the elementary principals rated the portfolio as proficient. When Brenda provided her rating she said, "I would give it proficient. I definitely would give this person the approval of going on with their teaching certificate. I guess the reason I don't consider it exemplary, and it's not that I don't ... I guess not having the opportunity to sit down and talk." Leo did not expound on his decision to rate the portfolio as proficient. Norma described her decision to rate the portfolio as proficient when she said,

Parts of it I thought were really strong; especially the observations and the unit on Native Americans, the fund raiser. Parts of it I thought were - I don't know that I'd call it unsatisfactory, but it was not as strong. It was proficient. But I think that as far as artifacts themselves, it could have had more meat to the artifact or more explanation about what the artifact was. The one portion there I thought was really pretty weak. I think that was the Newspapers in Education.

Norma was able to reflect on specific strengths and weaknesses of the portfolio as she reflected. The proficient rating assigned to the portfolio aligned with the judgment comments made in Phase II relative to the specific artifacts.

The proficient rating carried over to comments made relative to judgment of the teacher. While Brenda did not make judgment comments specific to the teacher, Leo and Norma did make positive comments relative to how they judged the teacher. Although Leo did not rate the portfolio as exemplar, he liked it enough to say, "Site unseen, if I had an opening for a fifth-grade teacher and wanted somebody to come in, that portfolio, I saw enough of it to form the opinion that she's a heck of a good teacher." Norma also seemed to like what she saw in the portfolio and felt that the teacher was strong; however, she felt that working with the teacher during the first two years when the portfolio was being developed would have made for a stronger portfolio.

Principal's opinion. The elementary principals continued verbal reporting relative to their opinions (PO) in Phase III. The category consisted of two sub-categories; principal's opinion regarding the portfolio of evidence of good teaching and principal's opinion regarding the DE system for teacher assessment. Seven of the eight Phase III comments made by the elementary principals were opinions relating to the portfolio as evidence of good teaching. It was apparent that the principals considered the portfolio, while good evidence, not the only evidence of good teaching. For Brenda and Leo, the

portfolio was not complete proof of teaching ability. For instance, Brenda reported that, “I just think that you’ve [teachers] gotta be able to prove that you can do those things. I guess that proof is partially in the concrete things [artifacts] that are in there [portfolio] and does that reflect what I see daily?” While Leo reported thinking that portfolios had value, he indicated that it was difficult for a teacher to document the intangibles such as relationships with students and positive influence. Leo showed some skepticism when he said, “As we look for ways to evaluate the success or potential success of teachers, portfolios are certainly a window to being able to do that. We just have a long ways to go. A portfolio can be something that can be in the minds of teachers a way to really over blow what might be some major deficiencies going on in the classroom.”

Tools. Brenda and Norma were the only two principals who made comments in Phase III regarding evaluation tools. Brenda described a form she used. “They [teachers] list their artifacts and then it’s kind of over here [pointing to her right] with the standards in little boxes and they ‘x’ it. Norma described using a similar form. Neither principal provided the researcher with a sample of their form. Both principals also indicated that they made use of the DE evaluation form during review. In addition, Brenda indicated that she used sticky notes.

In summary, judging activities continued to be represented in Phase III. The judgment of the portfolio consumed the largest percentage of the judgment comments. The elementary teachers judged the portfolio as proficient. Further, two of the principals judged the teacher positively. In addition the principals made comments that reflected their desire to meet with and assist teachers during portfolio development.

The elementary principals' comments indicated that, while they valued portfolios, they needed additional proof of good teaching. The elementary principals used sticky notes, a checklist coversheet, and the DE evaluation form as tools during evaluation.

Elementary Principals and Coaching Activities in Phase III

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time (TI). Comments relative to coaching activities accounted for 22 % of the total Phase III comments made by elementary school principals. Coaching activities were not present in Phase I and accounted for 20% of the total coded comments in Phase II. In this section, comment counts relative the three categories (coaching, principal's role, and time) will be reported.

Coaching. The coaching category was a multi-level category that included coaching relative to meeting with the teacher (C/M), coaching relative to portfolio preparation (C/PP), coaching via the use of questioning (C/Q) and coaching via making suggested alternatives. The elementary principals made a total of 10 coaching comments of which comments relative to the portfolio process accounted for 60%. Thirty percent of the coaching comments were relative to suggested alternatives with only one comment (10%) concerning meetings with the teacher.

Phase III comments relative to the portfolio process were consistent between Phase II and Phase III in that the elementary principals thought about how they could assist the teacher in the process. For instance, Norma said, "I tell my first year teachers, 'Just start collecting stuff now. Just even throw it in an envelope every time you think of something, every time you create something'." Norma and Brenda both indicated that

helping teachers recognize weak areas of the portfolio in the first two years was important. Brenda commented that, “When I meet with them [teacher] mid-year we kind of hit that [weak areas] and we look at that and say, ‘Are there any areas that you need some more evidence in? Let’s look at this and see if there’s anything that we might have available that you could find that would fit that’.” The use of “we” was a clear indication of Brenda’s commitment to partner with her teacher in the portfolio process.

Four firm suggested alternatives emerged from Phase III. The elementary principals suggested that the teacher might have made her portfolio stronger had she included (a) comprehensive cover sheets including more detail in the area of reflection, (b) student work throughout, (c) digital photos of students in action, and (d) more evidence of how writing was being taught within the unit.

Principal’s role. Brenda was the only principal to comment specifically about her role (PR). In essence, her comment was an extension of the comments she made relative to coaching and the portfolio process. She indicated that, “My job is not to catch them making it wrong, or that they’re insufficient. My job is to get them to turn in a portfolio that is going to reflect the ITS/criteria.”

Time Investment. Comments relative to time were low throughout the study. Phase III was no exception as only one principal commented. She simply indicated that it typically takes her about an hour to work through a portfolio.

In summary, the three coded comment categories that clustered under coaching activities were coaching (C), principal’s role (PR), and time investment (TI). The category of coaching was most prevalent for the elementary teachers in Phase III. The

elementary principals indicated that they should partner with their teachers during the portfolio process. The elementary principals made four suggestions for improvement of the sample portfolio. They suggested that the teacher include (a) comprehensive cover sheets including more detail in the area of reflection, (b) student work throughout, (c) digital photos of students in action, and (d) more evidence of how writing was being taught within the unit.

Elementary Principals and Critical Pieces in Phase III

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two “pieces” carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. The category of critical pieces consisted of two sub-categories; critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO).

The elementary principals made five comments relative to critical pieces in Phase III. Four of them were about the role of observation. The comments made by the elementary principals accentuated the importance they placed on observation. The value that principals placed on observation was best explained by Norma when she said, “In my mind that whole ‘Are you going to get your teaching license?’ is a big thing and so I’m going to put a lot of weight on what I’ve seen the person do, how I’ve seen them perform.” This comment was representative of a common thread across academic level throughout the current study.

Summary of Elementary Principals' Thinking in Phase III

The Phase III comments of the elementary principals were centered on the judgment of the portfolio and the critical nature of observation. The principals rated the sample portfolio as proficient but felt the teacher could have improved the portfolio had she included (a) more information on the cover sheets, including more detail in the area of reflection; (b) student work throughout; (c) digital photos of students in action; and (d) more evidence of how writing was being taught within the unit. In addition, the elementary principals indicated that they placed higher value on observation than they did on the portfolio.

Middle School Principals' Thinking in Phase III

The middle school teacher's portfolio was organized by standard. She included sections representing each of the eight standards. Each section contained artifacts representing a respective standard. It was not necessary that she use a cross-reference guide because the middle school teacher included, on an artifact cover page, the multiple standards/criteria represented by each artifact. See Appendix J for a sample of an artifact cover page.

The middle school principals (Ivan, Mike, and Rob) made the fewest Phase III comments of the three academic-level groups; a departure from the previous two phases where they made the most comments. Their thinking accounted for 26% of the total coded comments in Phase III as compared to 40% for the elementary principals and 34% for the high school principals.

Table 42 illustrates the breakdowns of Phase III comments made by middle school principals. Eight of the ten categories were represented in Phase III. The middle school principals did not consider portfolio structure (PFS) or principal's role (PR) in Phase III. The comments were evenly distributed between the middle school principals.

Table 42

Summary of Phase III Comment Counts made by Middle School Principals

	Total Comments	C	CL	CP	J	PO	PS	T	TI
Ivan	11	1	0	4	1	2	1	0	2
Mike	12	2	1	2	2	3	0	1	1
Rob	12	1	2	1	4	4	0	0	0

C = Coaching, CL = Comfort Level, J = Judgment, PO = Principal's Opinion, PS = Process Steps, T = Tools, TI = Time

Middle School Principals and Processing in Phase III

The three coded comment categories that clustered under processing included the categories of portfolio structure (PFS), process steps (PS), and comfort level (CL). Comments relative to processing accounted for 11% of the total Phase III comments made by middle school principals. The middle school principals did not make any comments relative to portfolio structure. In this section, comment counts relative to two of the three categories (process steps, and comfort level) will be reported.

Comfort level comments accounted for three of the four comments relative to processing for the middle school principals. The comments were representative of and consistent with earlier-voiced concerns regarding not knowing the teacher and/or being able to visually watch them perform. However, as was true in both Phase I and Phase II, the principals were able to successfully complete the portfolio review. Mike best summarized comfort level concerns when he said, "Being honest, this [evaluating a foreign portfolio] isn't a natural thing. It didn't feel right to me as I thought about it all because it undervalues and underestimates and doesn't account for the day to day stuff that is a part of it. But, it worked."

Middle School Principals and Judging Activities in Phase III

The three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Comments relative to judging activities accounted for nearly half of the total Phase III comments made by middle school principals compared to 15% in Phase I and 58% in Phase II. The percentage of judgment comment counts for Phase III reflected the influence of guided interview question number one in which the principals were asked to rate the portfolio as unsatisfactory, proficient, or exemplar. In this section, comment counts relative to only judgment and principal's opinion will be reported since only one comment was made regarding tools. In addition, data relative to guided interview question number one will be reported.

Judgment. The judgment category was a multi-level category that included judgment of artifacts (J/A), judgment of the portfolio (J/P), and judgment of teacher (J/T).

The middle school principals made a total of seven judgment comments. Five of the comments were relative to judgment of the portfolio and two of the comments were relative to judgment of the teacher. The principals' judgment comments in Phase III focused on the portfolio. The middle school principals each rated the portfolio as proficient. Interestingly, Ivan initially rated the portfolio as exemplar and then changed his mind when he recalled the typing errors. All three of the middle school principals mentioned observation as they rated the portfolio. In essence, they indicated that while the portfolio was proficient, it was only part of the bigger picture of teaching.

Both judgment comments relative to the teacher were unsolicited and indicated that, based on the portfolio, the teacher would be licensed. The favorable comments regarding the teacher in Phase III were consistent with the comments regarding the teacher in Phase I and Phase II.

Principal's opinion. The middle school principals continued verbal reporting relative to their opinions (PO) in Phase III. The category consisted of two sub-categories: principal's opinion regarding the portfolio of evidence of good teaching and principal's opinion regarding the DE system for teacher assessment. Seven of the eight Phase III comments made by the middle school principals were opinions relating to the portfolio as evidence of good teaching. Consistent with the sentiments they expressed as they rated the portfolio, the middle school principals considered the portfolio to have value but were clear that it was only a piece of teacher evaluation. They each emphasized that creating a good portfolio was possible for many teachers but that a good portfolio did not always

correlate with good teaching. These principals felt that observation carried greater value than the portfolio.

In summary, the judging activities of the middle school principals focused on portfolio judgment and the portfolio as evidence of good teaching. The middle school principals judged the portfolio as proficient and, based on the portfolio, they would probably license the teacher. However, the middle school principals were also very clear that they valued observation more than the portfolio.

Middle School Principals and Coaching Activities in Phase III

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time (TI). Comments relative to coaching activities accounted for 19 % of the total Phase III comments made by middle school principals. In this section, comment counts relative to only the coaching category and the time category will be reported.

Coaching. The middle school principals made four comments relative to coaching. The verbal reports by the middle school principals indicated their dedication to being involved in the portfolio process (C/PP) alongside the teacher and that regular meetings (C/M) were part of the portfolio process. Mike said, "We're [his district] creating time in professional development to discuss portfolio work, we're creating partnerships. I'm involved in the process along the way, so it isn't as though it's happening out there and suddenly it's here on my desk. I will have worked with them [teachers] and helped them along the way." Ivan indicated how he felt about meeting with teachers regarding the construction of their portfolios when he said, "To me, it's the

conversation that we're going to have that is the key to it. That's where the interaction is going to go."

Time Investment. Only Ivan and Mike commented about time in Phase III. In both instances, the principals reflected on the amount of time it took to complete the review of the sample portfolio. As part of the reflection, they each considered the time they invested in reviewing portfolios of their own teachers. Mike's comment summed up the sentiments of both principals. He very candidly stated that,

You get consumed by it [review] in April and May. Every year you tell yourself it doesn't have to be that way; you can do the right kinds of things along the line to space it out. But in reality, in order to give teachers the amount of time they need and to take the amount of time you need, it seems like it always hits over the final month. Like I told you, it's a lot of evenings and a whole lot of weekend time to pull it together.

Mike's concern was not only for his own time but for that of the teacher. He appeared cognizant of the fact that he needed to allow time for a quality review and that he used his personal time to conduct them.

Middle School Principals and Critical Pieces in Phase III

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two "pieces" carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. The category of critical pieces consisted of two sub-categories; critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO).

The middle school principals made seven comments regarding critical pieces in Phase III. Six of the seven comments were relative to the critical role of observation in teacher evaluation systems. Ivan was very clear about the emphasis he placed on observation when he said, "My observation is the basis, it's the foundation of whether they're going to get their license or not. This [the portfolio] to me is the supporting evidence to help it." His comment was reflective of the emphasis placed on observation by every principal in the study.

Summary of Middle School Principals' Thinking in Phase III

Like the elementary principals, the middle school comments in Phase III were generally focused on judgment of the portfolio and the value of observation. The middle school principals rated the sample portfolio as proficient. They did not make any suggestions for improvement of the portfolio during Phase III. The comments relative to the value of observation was very clear and echoed the sentiment of the elementary principals.

High School Principals' Thinking in Phase III

The high school teacher's portfolio was organized by artifact. Each artifact represented one or more of the eight ITS/criteria. See Appendix I for examples of the table of contents and cross-referencing information used by the high school teacher. The high school principals (Gavin, Kathy, and Keith) made cumulatively less comments than the elementary principals but more comments than the middle school principals in Phase III. Their thinking accounted for 34% of the overall comments made by all principals in Phase III as compared to 40% for the elementary principals and 26% for the middle

school principals. Table 43 illustrates the breakdowns of Phase III comments made by the high school principals. The total comment counts were fairly evenly distributed among the three high school principals and all ten of the coded comment categories were represented in Phase III. The total comment count for Phase III was significantly less than Phase II.

Table 43

Summary of Phase III Comments made by High School Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Gavin	16	1	1	2	5	1	3	2	1	0	0
Kathy	16	6	1	0	5	0	2	0	0	2	0
Keith	14	2	1	1	3	1	2	0	1	2	1

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal’s Opinion, PR = Principal’s Role, PS = Process Steps, T = Tools, TI = Time

High School Principals and Processing in Phase III

The three coded comment categories that clustered under processing activities included the categories of portfolio structure (PFS), process steps (PS), and comfort level (CL). Comments relative to processing accounted for only 15% of the total Phase III comments made by high school principals. In this section, comment counts relative to

each of the three processing categories (portfolio structure, process steps, and comfort level) will be reported.

The high school principals made only seven comments relative to processing activities in Phase III. During Phase III the principals were more focused on judgment of the portfolio than they were on processing activities. The structural comments that were made by the principals conveyed their understanding that all portfolios were not structured the same way. Gavin, who struggled with the structure of the portfolio early on, conveyed his understanding when he said, "I think there are easier ways to lay this out but it probably makes all kinds of sense to this teacher because they've been involved in this process their way from the very beginning." The process step comments were reflective in nature and reiterated Phase II comments made regarding cross-referencing techniques.

The high school principals made only three comfort level comments. In essence, the comments reflected that comfort level would be increased if the principal knew the teacher who produced the portfolio. The Phase III comments made by the high school principals were consistent with comments made in both Phase I and Phase II.

High School Principals and Judging Activities in Phase III

The three coded comment categories that clustered under judging activities were judgment (J), principal's opinion (PO), and tools (T). Comments relative to judging activities accounted for 43% of the total Phase III comments made by high school principals. The percentage of judgment comment counts for Phase III reflected the influence of guided interview question number one in which the principals were asked to

rate the portfolio as unsatisfactory, proficient, or exemplar. In this section, comment counts relative to each of the three categories (judgment, principal's opinion, and tools) will be reported. In addition, data relative to guided interview question number one will be reported.

Judgment. The judgment category was a multi-level category that included judgment of artifacts (JA), judgment of the portfolio (J/P), and judgment of the teacher (J/T). There were 13 judgment comments made by the high school principals in Phase III of which 70% were judgment of portfolio comments. All three of the high school principals rated the portfolio as some level of proficient and were consistent as they justified their rating. Kathy initially rated the sample portfolio on the cusp between unsatisfactory and proficient. She indicated that, "The lack of knowing the teacher affected the rating." Later, she indicated that the portfolio was proficient. Keith rated the sample portfolio as "almost exemplary" and he, too, indicated that it was difficult to make judgment not knowing the person. Gavin rated the portfolio as somewhere between proficient and exemplar. As he provided his rating, indicated his need to know the person as well. He said, "If you're sitting down with your new teachers and you're going through this on a periodic basis and you're evaluating them in their classrooms and you're talking with them through the year, you might know how this all pulls together a little bit better."

The remaining four judgment comments pertained to judgment of the teacher. Gavin indicated that the teacher appeared to be conscientious and that although he (Gavin) was sometimes critical of the portfolio, the teacher would be licensed. Kathy

commented that, "I get the sense that this is probably someone who should be a teacher, he's probably a very fine teacher." However, she was clear that the portfolio alone did not give her a clear sense of the teacher's proficiency and that she needed to be in his classroom to make that kind of determination. Keith did not make any judgment comments regarding the teacher.

Principal's opinion. The principal's opinion category consisted of two sub-categories; principal's opinion regarding the portfolio of evidence of good teaching and principal's opinion regarding the DE system for teacher assessment. All seven of the comments pertaining to principal's opinion were relative to the portfolio as evidence of good teaching. Across the board, the high school principals indicated that, while valuable, the portfolio was not the best evidence of good teaching. Returning to a familiar theme expressed by their elementary and middle school counterparts, the high school principals were insistent that creation of a good portfolio was not indicative of a good teacher.

Kathy captured the cumulative sentiment of all the principals in the study when she said,

There are teachers who shouldn't be teaching who can put together marvelous portfolios or can put together portfolios to document all of those things. I had a teacher this year that I put on track three [assistance] who had some things that ...she could give me lots of documentation to show that she was able to deal with classroom management. She could put things in a book that said, "I did this and I did this and Johnny got better." but I knew that the overall picture was that she was doing some damage to kids.

Kathy's concern for her students was quite evident in her comment. Her comment also suggested that she spent time observing in order to get a clear sense of what was occurring in the teacher's room. The fact that she put the teacher on an assistance track also suggested that she felt the teacher had potential to improve.

Tools. The high school principals made only four comments relative to tools in Phase III. The tools that the principals indicated they used were sticky notes, a blank piece of paper on which to write notes in preparation of a meeting with the teacher, and a copy of the eight ITS/criteria. A copy of the ITS/criteria was the common denominator among all of the nine principals in the study.

In summary, the high school principals judged the portfolio as proficient but felt that knowing the teacher might have caused the rating to increase. The verbal reporting also indicated that the principals placed some value on the portfolio as evidence of good teaching but that observation was the deciding factor regarding licensure. The high school principals thought that the teacher who produced the portfolio was strong but, were once again, careful not to make judgment without knowing and observing the teacher. A copy of the eight ITS/criteria continued to be a consistently used tool for the high school principals in Phase III. The principals also indicated that they took notes during their reviews on sticky notes and/or blank paper. The notes were taken in preparation for meetings with the teacher to discuss and continue to prepare the portfolio.

High School Principals and Coaching Activities in Phase III

The three coded comment categories that clustered under coaching activities were coaching (C), principal's role (PR), and time (TI). Comments relative to coaching activities accounted for 20% of the total Phase III comments made by high school principals. In this section, comment counts relative to each of the three categories (coaching, principal's role, and time) will be reported.

Coaching. The high school principals made a total of nine coaching comments. Four of the nine comments focused on coaching the portfolio process (C/PP). Three comments pertained to meetings with the teacher (C/M) and two comments were relative to suggested alternatives (C/SA).

Kathy and Keith were the only two high school principals to make comments relative to the portfolio process. Both principals indicated strong involvement with their beginning teachers during the portfolio process. Keith indicated that he and the teacher started early in the first year to talk about the portfolio and then followed up after each evaluation. He described his coaching when he said, "I don't want it to be something that people are scrambling with or become overwhelmed with because if you just work on it in bits and pieces it's nothing. You're doing the work in the classroom and you're just taking those documents that you receive and placing them in the portfolio."

Kathy, who rated the portfolio as barely proficient, indicated a clear involvement in the portfolio process when she said, "Maybe he did what his administrator wanted. He wouldn't for me, but he would have had a whole lot of different information coming from me. My guess is that if he had had that, this [the portfolio] would have been fine." Kathy's comment indirectly pointed to the influence of local control and the fact that each district may have a different perspective about what makes a portfolio proficient. Her main point was that her involvement would have shaped the portfolio contents differently and as a result made it stronger; more towards exemplar. Abundantly clear in the verbal reporting was the fact that the principals understood the value of meeting with the teacher.

There were two suggested alternative comments made by the high school principals in Phase III. The two principals who commented suggested that the teacher include additional explanation in the portfolio. Gavin explained the need for additional information well when he said,

When I'm looking at this one here, there are several times that I would have liked to have a paragraph on it [the artifact] or below it or in front of it that would have said 'this is how I believe I'm meeting this criteria. If it says something about posting rules and regulations in the front of the classroom and he's got a picture, that doesn't need an explanation. But on some of the documents it would help to see what their thinking process is behind it.

The suggested alternative relative to additional information was consistently present as all principals made judgments about the artifacts and about the portfolio. The fact that the principals in the study did not know the teacher most likely contributed to the additional emphasis on increased information.

Principal's role. Gavin was the only high school principal to comment on his role in Phase III. Gavin's perspective on his role was two fold. First, he wanted to make sure his teachers understood the portfolio process and completed a high-quality portfolio in the event that he might leave and a new principal would be reviewing the portfolios. This was a sentiment that was similar to a comment made by Norma, one of the elementary principals. Second, Gavin acknowledged his strong role in the development of first and second-year teachers. He indicated that the portfolio was a tool he used to assist in their development. Neither of the other two high school principals made any Phase III comments relative to their role.

Time Investment. Keith was the only principal to comment about time. He simply indicated that he only gets approximately 40 minutes with teachers during their planning

period. Keith indicated that, due to a lack of time, he relied heavily on the teacher's reflection after one of his observation visits. He said that, "My biggest thing when I'm looking at that observation is I want that reflection to come from the teachers. I want them to think about what they're doing and how they might change; then to follow through on that. If I can see that taking place and being supported in the portfolio, it would have more meaning to me." Keith used his own observation, the teacher reflection, and the portfolio as a way to triangulate evidence efficiently. The critical pieces of observation and reflection were illuminated in his comment.

In summary, the high school principals continued to make coaching comments in Phase III. Specifically the principals indicated that they took seriously their role in helping the teacher develop his/her portfolio. Additional information continued to be a suggested coaching alternative. Only one comment relative to the time it takes to review portfolios and meet with teachers was made. The comment was essentially a problem-solving technique that allowed the principal to use multiple sources of data effectively and efficiently.

High School Principals and Critical Pieces in Phase III

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two "pieces" carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. The category of critical pieces

consisted of two sub-categories; critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO).

The high school principals made only three comments relative to critical pieces. Two of them were pertinent to the role of observation. Both were made by Gavin. He indicated that, "I would want to see this person in action; within his environment. That's another concept to the whole thing too. I can't get a true indication on what kind of a teacher they are until I see them work in the environment that they're working in." The importance of teacher reflection was reported under judging activities in conjunction with Keith's comments about time.

Summary of High School Principals' Thinking in Phase III

The verbal reporting summaries in Phase III clearly illustrated that the principals focused their comments mainly on judgment of the portfolio. The high school principals rated the portfolio as proficient; one was a low-proficient while the other two bordered on exemplar. In all cases, the principals indicated that knowing the teacher personally and/or observing him teach would have likely increased their ratings. Because the principals did not know the teacher, they made suggested alternatives asking the teacher to provide additional explanation throughout the portfolio; thus, providing insight into the thinking of the teacher relative to the evidence in the portfolio.

Phase III: Reflection via Guided Interview per Geographic/Demographic Region

In this section, the findings relative to research question one for Phase III (reflection via guided interview) thinking will be reported per geographic/demographic region. First, a brief overview of the data collected for geographic/demographic region is

provided. Total comment counts across geographic/demographic region for Phase III will be reported and illustrated. Then, findings for each geographic/demographic region for Phase III will be reported. Because the trends were so similar for each of the geographic/demographic regions, a single summary of the geographic/demographic data for Phase III will conclude the section.

It is important to be reminded that while principals in each academic level (elementary, middle school, and high school) reviewed the same portfolio, the geographic/demographic region (urban, suburban, and rural) verbal comment counts were representative of comments made by the *same* principals *across* academic level, i.e., the principals in geographic/demographic region did not review the same portfolio because they represented differing academic levels. Therefore, data from geographic/demographic region is representative of comparison between urban, suburban and rural factors only; academic level data will not be revisited. Data will not be reported using the four broad categories as was true with the academic level reporting because it would be repetitive. Rather, data comparing Phase I and Phase II results per each of the ten coded categories will be reported.

Geographic/Demographic Overview

Each geographic/demographic region represented one principal from each academic level (elementary, middle school, high school). The rural principals that participated in the study were Brenda, Ivan, and Keith. The suburban principals that participated in the study were Norma, Mike, and Gavin. The urban principals that participated in the study were Leo, Rob, and Kathy. The total number of verbal comment

counts made in the study for each geographic/demographic region was remarkably similar (see Figure 2). Suburban principals had the highest overall percentage of coded verbal comments with 34%. The rural and urban principals each accounted for 33% of the total verbal comment counts for the study.

Table 44 illustrates comment counts for Phase III by geographic/demographic region per the ten coded categories. As was true in Phase I and Phase II, all ten of the coded comment categories were represented in Phase III. Suburban principals made the most Phase III comments accounting for 37% of the total coded comments per geographic/demographic region. Comments made by urban principals accounted for 33% of the total. Rural principals made the fewest Phase III comments accounting for 30% of the total coded comments. As noted in academic results, Phase III comments accounted for 14% of the total coded comments for the study.

Table 44

Summary of Phase III Comment Counts per Geographic/Demographic Region

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Rural	41	7	1	6	6	3	5	1	4	5	3
Suburban	50	8	5	7	11	1	9	2	2	3	2
Urban	45	8	4	2	18	1	10	0	0	2	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS = Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

The number of comments declined significantly in Phase III. There was not a marked difference between the numbers of total comment counts among the principals per their geographic/demographic region. As expected, the coded categories that were significant in Phase III were judgment (J), coaching (C), principal's opinion (PO), and critical pieces (CP).

Rural Principals' Thinking in Phase III

Table 45 illustrates the breakdowns of Phase III comments made by the rural principals. All ten of the coded comment categories were present in Phase III.

Table 45

Summary of Phase III Comment Counts made by Rural Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Brenda	16	4	0	1	2	2	1	1	2	3	0
Ivan	11	1	0	4	1	0	2	0	1	0	2
Keith	14	2	1	1	3	1	2	0	1	2	1

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

The rural principals did not deviate from the norm for distribution of comments for geographic/demographic region. The categories that were significant across the regions were significant for the rural principals as well, i.e., judgment, coaching, principal's

opinion, and critical pieces. This trend coincided with the guided interview question that asked the principals to rate the portfolio. All three of the portfolios received a proficient rating.

Suburban Principals' Thinking in Phase III

Table 46 illustrates the breakdowns of Phase III comments made by the suburban principals. The suburban principals made the most comments of any group across geographic/demographic region. Phase III was the only Phase in the study where all ten of the coded comment categories were present for the suburban principals. There were no abnormal trends in terms of comment distribution among the principals or across the coded comment categories. Judgment, coaching, critical pieces, and principal's opinion were the most significant numerical categories for the suburban principals; a trend that existed across academic level and geographic/demographic region for Phase III. The

Table 46

Summary of Phase III Comment Counts made by Suburban Principals

	Total Comments	C	CL	CP	J	PFS	PO	PR	PS	T	TI
Norma	22	5	3	3	4	0	3	0	1	2	1
Mike	12	2	1	2	2	0	3	0	0	1	1
Gavin	16	1	1	2	5	1	3	2	1	0	0

C = Coaching, CL = Comfort Level, CP = Critical Pieces, J = Judgment, PFS – Portfolio Structure, PO = Principal's Opinion, PR = Principal's Role, PS = Process Steps, T = Tools, TI = Time

trend was largely due to the fact that the principals were asked to rate the portfolio. All three of the portfolios were rated proficient.

Urban Principals' Thinking in Phase III

Table 47 illustrates the Phase III comments made by urban principals. Urban principals attended to only seven of the ten coded categories. Absent in Phase III were principal's role, process steps, and time. Comments regarding principal's role and time were not highly attended to by urban principals in any of the three phases in the study. The process step category was present in Phase I and Phase II at a fairly significant rate.

Table 47

Summary of Phase III Comment Counts made by Urban Principals

	Total Comments	C	CL	CP	J	PFS	PO	T
Leo	17	1	1	1	9	1	4	0
Rob	12	1	2	1	4	0	4	0
Kathy	16	6	1	0	5	0	2	2

C = Coaching, CL = Comfort Level, CP = Critical Pieces, PFS – Portfolio Structure, PO = Principal's Opinion, PS = Process Steps, TI = Time

The urban principals exhibited a slight deviation from the normal comment distribution across the categories. The category of critical pieces (CP) did not garner the significant number of comments for the urban principals that it did for the rural and urban

principals. However, the categories of coaching (C), judgment (J), and principal's opinion (PO) continued to be numerically significant for the urban principals just as they were for the rural and suburban principals.

Judgment comments accounted the most comments in Phase III for the urban principals largely due to the guided interview question that asked the principal to rate the portfolio. The comment counts reflect the principals' comments regarding judgment of the portfolio. All three of the portfolios were rated as proficient.

In summary, there was little deviation between the geographic/demographic region comment counts. The verbal reporting data confirmed that, because principals were asked to rate the portfolio, the principals' comments naturally focused on judgment. As judgment comments increased, so, too, did comments relative to coaching, principal's opinion, and critical pieces. The only region to deviate was the urban group. The number of comment counts relative to critical pieces was lower for the urban principals than for the rural and suburban principals. However, their comments relative to judgment were laced with reference to the critical piece of observation.

Summary: Findings for Phase III Thinking

The findings in Phase III established that the principals who participated in the study (a) rated their respective portfolios as proficient and (b) placed great importance on observation in teacher evaluation. The findings in Phase III were, in part, a direct result of the guided interview questions posed by the researcher (Appendix K). In other words, the verbal reporting in Phase III was not a result of the less-guided think-aloud (TA).

The verbal reporting data for Phase III revealed that the nine principals attended most to the categories of judging, coaching, principal's opinion, and critical pieces. This was not surprising as the principals were asked in Phase III to rate the sample portfolio as unsatisfactory, proficient, or exemplar. All nine principals indicated that they would rate the portfolio as proficient; however, there were some degrees of proficiency voiced. For instance, one high school principal rated the high school portfolio as a "low" proficient. Another principal in the high school group rated the same portfolio as "nearly exemplar."

As the principals judged the portfolio and made a rating decision, they also made coaching comments. In general, the principals indicated that they felt a responsibility to partner with the teacher during portfolio preparation. In addition, principals made four suggestions for improvement of the sample portfolio. They suggested that the teacher include (a) comprehensive cover sheets including more detail in the area of reflection, (b) student work throughout, (c) digital photos of students in action, and (d) more evidence of how writing was being taught within the unit.

Perhaps the most consistent thread throughout Phase III was the emphasis placed on observation. Principals voiced opinions about the portfolio as evidence of good teaching saying that while they valued the portfolio, observation was deemed more important than the portfolio.

There were no dramatic differences in numbers of comment counts in Phase III across academic level or across geographic/demographic region. The principals focused on rating the portfolio and their responses were somewhat guided by the interview questions.

Summary: Research Question 1

Findings for research question one indicated that (a) the participating principals operated within a similar “thinking framework” as they evaluated the portfolio provided to them by the researcher, (b) the participating principals were able to successfully evaluate a foreign/sample portfolio, (c) the participating principals attended most to judging and coaching activities as they reviewed the portfolio, (d) the participating principals established a similar four-step rhythm when judging individual artifacts, (e) the participating principals varied across academic level in terms of judgment pattern across artifacts, (f) the participating principals placed a higher value on observation than on the portfolio, and (g) the participating principals rated their respective sample portfolio as proficient.

The verbal reporting data gathered via the TA/interview sessions revealed that the participating principals operated within a similar “thinking framework” during the review of the sample portfolio. The broad framework included three phases of thinking that, in this study, were identified as pre-assessment (Phase I), judgment (Phase II), and reflection via guided questioning (Phase III). Within the framework, principals’ verbal reports centered on processing, judging, and coaching activities. In addition, observation and teacher reflection were identified by the participating principals as being critical pieces in teacher evaluation.

Portfolio structure was cause for concern during the pilot study. The principal who participated in the pilot study was not able to completely move beyond thinking about the structure of the portfolio (pre-assessment/Phase I thinking) into Phase II

(judgment). Even during the judgment phase, he heavily critiqued the structure of the portfolio and was distracted as he attempted to assess the artifacts included in the portfolio. As a result of the pilot participant's experience, the think-aloud instructions provided by the researcher to the nine participants prior to their review were altered slightly to include a very brief description of the portfolio structure (Appendix F) the principal was about to review.

This slight alteration proved to be effective. The nine participants all completed the review of the portfolio without the structure distraction experienced by the pilot participant. Statements regarding portfolio structure were made during all phases of the respective reviews; however, the percentage of structure statements dropped off significantly as the principals moved from pre-assessment (Phase I) into the judgment stage (Phase II) of their reviews. In addition, verbal reporting confirmed that comment counts regarding comfort level were low during pre-assessment. In fact, comfort level comments throughout all phases of the review were minimal; further evidence that the participants were comfortable with the responsibility of evaluating a "foreign" portfolio.

Judging activities permeated every phase of the thinking framework; however, the percentage of judgment activities increased significantly as principals moved from Phase I to Phase II and then decreased again in Phase III. As the principals in the current study moved into the Phase II and began to judge the artifacts, a four-step rhythm became apparent. The principals would (a) identify and verbally describe an artifact aloud, (b) read aloud from the teacher reflection and/or artifact description, (c) make a judgment

statement regarding the artifact, and (d) sometimes make a coaching statement about the artifact. This type of rhythm was consistently demonstrated by all nine principals.

Differences did exist in the amount of time principals spent engaged in judgment. The bulk of judgment activity was evident in Phase II for all academic levels (see Table 12). During Phase II, middle school principals spent twice as much time judging artifacts as did elementary principals. The high school principals also spent significant time in judging but were still lower than the middle school principals. The high school principals spent one and one-half as much time engaged in judgment as did their elementary counterparts.

In general, the principals in the study judged the artifacts as positive, i.e., the artifacts sufficiently met the established ITS/criteria. In addition, the principals agreed, in most cases, with the teacher as to which ITS/criteria specific artifacts evidenced. In some cases, the principals in the study verbally indicated if they thought the teacher had met additional ITS/criteria not listed on the cover sheet or the cross-reference guide. This was especially true of the middle school principals. The middle school principals used cross-referencing at a much higher level; hence, their judgment counts were much higher in Phase II. The artifacts most often identified by principals as providing effective evidence were (a) observation summaries completed by the teacher's principal or a peer, (b) samples of student work, (c) pictures of activities, (d) copies of two-way email communication with parents and colleagues, (e) rubrics, (f) lesson plans, and (g) minutes from professional development sessions and committee assignments.

A discrepancy in judging pattern was revealed via the verbal reporting. It was discovered that, during artifact judgment, the elementary and middle school principals closely scrutinized the first 30% of the artifacts in their respective portfolios and then their judgment comments significantly decreased. It was anticipated that the high school principals would exhibit the same sort of judgment pattern, i.e., early scrutiny of artifacts and then a decrease in judgment comments. However, this expectation did not come to fruition. The high school principals closely scrutinized each of the seven artifacts in the portfolio (see Table 31). Overall, the artifacts garnered more positive than negative comments but the percentage of negative comments was higher for the high school principals than it was for either the elementary or middle school principals. In addition, 50% of the negative artifact judgment comments made by the high school principals occurred during judgment of the first three artifacts; thus, leading to closer scrutiny of the remaining artifacts. This finding suggested that principals used the first several artifacts to develop a sense of trust in the teacher which, in turn, led to decreased scrutiny for the remaining artifacts. For the high school principals, that sense of trust took longer to establish.

Coaching activity increased as the principals began their judgment activities in Phase II. Coaching comments relative to specific artifacts were common in the four-step judgment rhythm that was established by the principals. Principals most often made “suggested alternative” coaching comments specific to artifacts. Suggested alternatives were associated with artifacts that were judged negatively and positively. The suggested alternatives frequently reflected the principals need for clarification. Additional suggested

alternatives were the use of (a) more frequent reflection (teacher and student), (b) additional proof of outcomes, (c) more variety, (d) evidence of student achievement in each artifact, (e) grammar and spell check, (f) more detailed artifact description, (g) authentic student work, and (h) data from other sources, i.e., student and/or parent surveys.

On average, across academic level, roughly 28% of the combined judging/coaching comments represented coaching. The percentage was lowest for elementary principals (25%) and highest for high school principals (33%). The middle school principals coaching comments represented 28% of their combined coaching/judging comments. The somewhat higher percentage of coaching comments from the high school principals was expected given the higher level of negative comments surrounding the artifacts. The conclusion could be drawn that coaching comments increase as negative artifact judgment increases.

The comment counts also revealed a high number of coaching comments relative to meetings between principals and the teacher, and questions posed by principals as they thought aloud about said meetings. This trend was indicative of the emphasis principals placed on coaching/partnering with their teachers during portfolio development. For instance, if principals felt a teacher had a weak area, i.e., missing or inappropriate evidence, there was a willingness to assist the teacher in identifying appropriate artifacts; ones that the teacher may already be incorporating but not illustrating in the collection of artifacts. Principals in the current study indicated that the meetings they held with teachers were for purposes of clarification, dialogue, and coaching.

Consistent across all three phases was the emphasis placed on observation as a critical piece of teacher evaluation. As the principals moved between each phase of the portfolio evaluation and attended to processing, judging, and coaching activities, they were very clear that the portfolio was “just a piece” of teacher evaluation and that they placed higher value on observation.

When asked to rate the portfolio, all nine principals in the study rated their respective portfolio as proficient. This was not surprising based on the mostly positive artifact judgment comments. However, varying degrees of proficient were evident. The most variance was seen with the high school portfolio. One principal rated it as “low proficient” while the other two rated it as “nearly exemplar.” Via verbal reporting, it was evident that each of the nine principals would have considered a higher rating if they (a) knew the teacher and (b) had personally observed the teacher.

Research Question 2

What tools have principals developed to assist them in evaluating second year teacher portfolios?

Via verbal reporting data and follow-up questioning, the following were identified as tools principals used during portfolio evaluation:

1. A copy of the eight ITS/criteria as a reference guide as they move through the portfolio (Appendix A).
2. Sticky notes that they attached to pages in the portfolio. The sticky notes may have questions and/or comments for the teacher.

3. A cross-reference guide that lists the ITS/criteria. In parentheses, beside each standard/criteria, other standards/criteria that might be met are also listed (Appendix R).

4. Artifact cover-sheets listing the ITS/criteria next to which the teachers indicate which of their artifacts are being used to meet each of the ITS/criteria. The cover sheet is a tool for the teacher and the principal.

5. A “log” that supplies teachers with descriptors and a way to communicate to the principal in written form, at a glance, which artifact is meeting which ITS/criteria.

6. The DE summative evaluation form (Appendix B)

Each of the six tools identified were not used by every principal. Some of the tools were self-generated, some were generated by the district, and others were obtained from the DE website. In this section, further description of the tools and supporting comments will be reported.

Comment counts relative to tools accounted for approximately 4% of the total comments made in the study. The most comments relative to tools were made during Phase II as the principals focused on judging each artifact. Principals were not formally asked to bring any documents they used (self-generated or otherwise) as they evaluated portfolios. However, as the TA/interview progressed and the principal made reference to tools he/she used for evaluation purposes, the researcher made notes about the tool and asked the principal for copies following the TA/interview session. Asking the principal to bring documents to the interview might have inferred to the principal that they *should* be using some sort of document. Consequently, they may have created something

specifically for the TA/interview session that they did not otherwise use, thus contaminating the naturalistic atmosphere of the environment.

The tool most often referred to was a copy of the ITS/criteria (Appendix A). The principals indicated that they reviewed the copy prior to portfolio review and then used it as a guide during portfolio review. Leo, a middle school principal, was the only principal that did not reference the use of a copy of the ITS/criteria or any other tool for that matter. He indicated that, "I have simply done this off the cuff in a verbal kind of conversation with the teacher as opposed to having any kind of a paradigm or any kind of grid form."

Verbal reporting also indicated that sticky notes were a tool. Brenda, an elementary principal, and Kim, a high school principal, both made reference to the use of sticky notes. Each principal indicated that they attached sticky notes directly to artifacts in the portfolio as a way to remind themselves of what they needed to communicate to the teacher. Kathy described how she used sticky notes when she said, "Sometimes I just tacky sticky notes to pages and say, 'Get some of this'."

The principals in the study made use of cross-referencing, i.e., using one artifact to meet multiple ITS/criteria; none more than the middle school principals. Ivan, one of the middle school principals provided the researcher with a copy of the cross-referencing guide he used (Appendix R). Ivan used the cross-reference sheet extensively, relying on it to guide him and as an aide to make better use of his time. Ivan described the tool when he said, "What I have is a copy of the Iowa Teaching Standards and Criteria. We shrunk it down so it can be just on the face front side of the sheet of paper. It has the 8 standards

and underneath it, it has the 42 criteria listed just like they are. Behind each one what we have done is, in parentheses, cross referenced.” Then, Ivan described how he used the tool when he said, “For instance, 1a—Provides Evidence of Student Learning to Students, Families and Staff – we have in parentheses 1g, 5b, 5e and 8e. What we’re saying is that it is more than likely that if they did 1a they probably also evidenced 1g, 5b, 5e, 8e. Not necessarily, but it’s a quick check. I can look at those.” Ivan was the only principal to bring any kind of a cross-reference guide with him. However, principals made use of the teacher-included cross-referencing information provided in the sample portfolios (Appendixes I, J, and K)

Brenda and Norma, both elementary principals, attempted to describe artifact cover-sheets that they had developed. Norma indicated that the cover sheet she developed was placed in front of each artifact. The teacher indicated on the cover sheet which of the ITS/criteria the artifact met. During her review, Nancy would indicate on the cover sheet whether she agreed with the teacher or not. Brenda’s description sounded similar to Norma’s but neither principal had a copy of the cover-sheet available for the researcher.

Mike, a middle school principal, described a “log” that sounded similar to the cover-sheet used by Norma and Brenda. Mike provided a description of the log when he said, “Just so you know – our teachers have a log that I would lay out in front of me and it shows if they are applying the artifact in many standards and criteria. They’ll check mark that and they’ll tell me whether or not, as I’m looking at it, that it’s also gonna be found again later.” Mike did not provide the log to the researcher but he did provide a document titled “Artifact Identification Guide.” Appendix S contains a sample specific to

ITS/criteria number one. Mike indicated that the guide was to assist teachers as they identified appropriate artifacts for their portfolio.

Gavin, a high school principal, provided the researcher with a district-generated evaluation guide. A sample specific to ITS/criteria number two is contained in Appendix U. He indicated that teachers in his district had access to the guide as they prepared their portfolios. In addition, Gavin said that he used it as he evaluated portfolios. The evaluation guide provided by Gavin was similar to the “Artifact Identification Guide” provided by Mike. Both tools served as prompts for teachers and principals.

Three of the principals made direct reference to the DE summative evaluation form (Appendix B). Rob, a middle school principal, commented about his formal write-up during each phase of the study. He used the form to take informal notes as he reviewed the portfolio, observed the teacher, and then met with the teacher. The informal notes then became the framework for his formal summative evaluation. Brenda and Norma also made reference to the DE summative evaluation but only that they kept it in mind as they reviewed the portfolio.

Summary: Research Question 2

Via verbal reporting and follow-up questioning, the researcher identified six tools that principals used during portfolio evaluation. The tools identified were (a) a copy of the eight ITS/criteria, (b) sticky notes, (c) a cross-reference guide, (d) artifact cover sheets, (e) a log listing descriptors for the ITS/criteria, (f) the DE summative evaluation form. The varied functions of the tools were to (a) provide visual guidance to both principals and teachers, (b) provide a means for principals to quickly cross-reference

multiple ITS/criteria, (c) provide principals a means to informally and quickly communicate with teachers, (d) provide cues to the teacher and the principal relative to appropriate evidence, and (e) provide an informal framework for the summative evaluation. The tools were generated by varied entities including the principal, the district, and the DE. Several of the tools exhibited similarity in appearance and function; however, with the exception of the DE summative evaluation, there were no two tools that were identical. Three of the principals provided the researcher with samples of the tools they used (see Appendix R, S, and U).

Research Question 3

How much bearing does the portfolio evaluation have on the judgment the administrator makes regarding licensure?

Findings for research question three indicated that the portfolio would account for roughly 30% of a licensure decision made by the participating principals relative to their respective sample portfolio. Findings also indicated that, per academic level, (a) elementary principals put the most bearing on the portfolio (38%), (b) high school principals were second (30%), and (c) middle school principals put the least bearing on the portfolio (25%). Findings regarding geographic/demographic region showed that (a) suburban principals put the most bearing on the portfolio (33%), (b) rural principals were second (28%), and (c) urban principals did not commit to a percentage-based response.

Research question three was answered via a guided interview question that asked, "How much bearing would you assign to this portfolio when making a licensure decision?" The responses were varied. One of the principals could not commit to an

answer. Two others said, "Not much." One principal said, "Less than 10%." Three principals said, "Less than 25%" and two principals said, "Less than 50%." Because of the wide range of responses it was not possible to calculate an "absolute average bearing" for the nine principals. However, if the stated percentages alone were considered, the average bearing would be roughly 30%; accounting for one-third of the principal's decision.

It is important to note that the principals in the study rated all three portfolios as proficient. Unsolicited, four of the nine principals suggested that the sample portfolio would positively contribute to their licensure decision. For instance, Mike said, "We are not talking a third year." The other five principals did not make similar unsolicited comments; however, there was no indication in the verbal reporting that the sample portfolio would hinder any of the teachers from achieving licensure following their second year. In addition, every principal, when asked, indicated that the bearing they placed on the sample portfolio was exactly the same bearing they placed on portfolios they reviewed from their own buildings.

There was some very slight variance as to portfolio bearing between academic levels and geographic/demographic regions. In this section, the responses of the principals will be reported first by academic level and then by geographic/demographic region.

Academic Level Responses

Elementary principals. The elementary principals made varied responses to the question regarding the bearing of the portfolio on a licensure decision. Brenda said,

“About half.” Leo simply replied, “Not much” and Norma indicated, “Less than 25%.”

Based on the two definite percentages provided, the elementary principal average bearing was roughly 38%; slightly higher than the calculated average for the entire group. For the elementary principals, the recurring theme of the portfolio as evidence of good teaching was present. Brenda commented that, “I can’t just look at a book and say, ‘You’re [teacher] ready to go’. I can have a very good portfolio in front of me that would be exemplary and know that the teacher isn’t doing this stuff.” The critical piece of observation was apparent in Brenda’s statement and was consistent with the verbal reports of Leo and Norma.

Middle school principals. The middle school principals’ responses to the question regarding portfolio bearing were somewhat less varied than those of the elementary principals. Ivan and Mike both indicated that the portfolio would account for less than 25% of their licensure decisions. Rob simply replied, “Not much.” Based on the two definite percentages provided, the middle school principal average bearing was 25%; slightly lower than the calculated average for the entire group (30%) and considerably lower than the calculated average for the elementary principals (38%).

Ivan left no question that observation was the determining factor in terms of a licensure decision when he said, “Let’s just get down to it. It’s [licensure] is based on what I see and hear.” He even reneged on his earlier bearing percentage and stated, “A better way of saying it is that you observe 90% of it.” Mike and Rob were not quite as adamant as Ivan. Rob simply stated that, “Portfolios are not that important to me.” Mike

reiterated that the portfolio was proficient but that it was only one piece. Mike and Rob both alluded to observation as a significant piece of the licensure decision.

High school principals. The high school principals' responses to the question regarding the bearing of the portfolio on a licensure decision were very disparate. Kathy indicated that she could not commit to an answer. Gavin indicated that the portfolio would account for less than 50% of his licensure decision. The portfolio accounted for less than 10% of Keith's licensure decision. Based on the two definite percentages provided, the high school principal average bearing was 30%; the same as the calculated average for the entire group. The calculated average bearing of the high school principals was higher than that of the middle school principals (25%) and well below that of the elementary principals (38%).

The fact that Kathy rated the portfolio on the lower end of proficient influenced her response to the question asking her about the bearing of the portfolio on a licensure decision. She explained that she could not commit to an answer because she considered the sample portfolio a "work in progress." She added, "If it's a young teacher who is working towards things and I've seen him in the classroom and he's doing good work but just doesn't understand the collection and documentation process, then if he's a good teacher I'd work with him long and hard to make sure he ended up getting his license." Indirectly, Kathy indicated that observation was critical in teacher evaluation. In addition, her commitment to coaching teachers was very apparent; consistent with her comments throughout the study.

Despite the disparity between Gavin (50%) and Keith's (10%) responses, the two principals made similar comments relative to observation. Gavin said, "The reason [the bearing is less than 50%] is because I think there is so much more to teaching than what's on paper. This stuff really doesn't mean a whole lot to me if I've been in there to see the teacher." Gavin's response was somewhat surprising given his high percentage. Keith commented that, "The portfolio-10%. To me, that's not where the rubber meets the road. It just supports and is a place for the teacher to put things. But they take those items and if it's working in the classroom, that's where I want it to work."

Geographic/Demographic Region Responses

Grouping the principals per geographic/demographic region yielded a very slight skew to the calculated bearing averages. The overall calculated bearing for the entire group of nine principals was roughly 30%. The calculated bearing average for rural principals (Brenda, Ivan, and Keith) was 28%. The average for suburban principals (Norma, Mike, and Gavin) was slightly higher at 33%. None of the urban principals (Leo, Rob, and Kathy) provided a firm percentage. One of the urban principals could not commit to a percentage and the other two urban principals replied, "Not much." The two geographic/demographic groups that provided percentages were not out of range with each other nor were they out of range with the overall percentage calculated for the entire group.

Summary: Research Question 3

Findings for research question three indicated that the portfolio would account for roughly 30% of licensure decisions made by the participating principals relative to their

respective sample portfolio. Findings also indicated that, per academic level, (a) elementary principals put the most bearing on the portfolio (38%) relative to licensure decisions, (b) high school principals were second (30%), and (c) middle school principals put the least bearing on the portfolio (25%). Findings regarding geographic/demographic region showed that (a) suburban principals put the most bearing on the portfolio (33%) relative to licensure decisions, (b) rural principals were second (28%), and (c) urban principals did not commit to a percentage-based response.

The findings relative to research question three were not surprising given the emphasis that principals in the study placed on observation. As each principal described their decision on the bearing they would assign to the portfolio, a comment regarding observation was always included. The fact that the calculated bearings were somewhat similar suggested that consistency existed across academic lines and across geographic/demographic region. The consistency was further supported when the principals indicated that the bearing they placed on the sample portfolio was the same bearing they would place on a portfolio from one of their own teachers.

CHAPTER 5

DISCUSSION AND IMPLICATIONS

The purpose of this study was to examine the thinking of administrators as they reviewed and made evaluative judgment of a second year teacher portfolio. Qualitative data was collected using a think-aloud (TA) process and four open-ended questions.

Three research questions guided the study:

1. What do principals verbally report they are thinking as they review second-year teacher portfolios for purposes of evaluative judgment?
2. What tools have principals developed to assist them in evaluating beginning teacher portfolios?
3. How much bearing does portfolio evaluation have on the judgment the administrator makes regarding licensure?

Study findings indicated that (a) participants operated within a similar “thinking framework” as they evaluated the portfolio provided to them by the researcher, (b) the thinking of principals across academic level and geographic/demographic region closely linked to the Iowa Evaluator Training Model, (c) principals have developed unique tools for use during portfolio evaluation, and (d) the portfolio was not a significant consideration in licensure decisions.

In this chapter, results from the study will be used to discuss the connection between the verbal reporting (thinking) of the participants and (a) already-established frameworks for evaluating portfolios, (b) the Iowa Evaluator Approval Training Program (IEATP), (c) portfolio evaluation tools described in current literature, and (d) the

significance of the portfolio regarding licensure decisions. The findings provide insight into the effects of local control on a state-wide system. The findings also offer insight relative to licensure decisions in terms of judgment and the bearing of the portfolio on said licensure decisions. Each research question will be individually discussed.

Research Question 1: What do principals think during portfolio review?

With respect to the first research question regarding the thought processes of principals during portfolio review, the verbal reporting data gathered via the think aloud sessions established that participants operated within a similar “thinking framework” as they evaluated the portfolio provided to them by the researcher. The broad framework included three phases of thinking that, in this study, were identified as pre-assessment (Phase I), judgment (Phase II), and reflection via guided questioning (Phase III). Within the framework, principals’ verbal reports centered on processing, judging, and coaching activities. In addition, observation and teacher reflection were identified by the participating principals as being critical pieces in teacher evaluation. In the following discussion, each of the three broad phases will be addressed via the context of the common thinking activities of processing, judging, coaching, and critical pieces.

Processing

Processing activities included comments relative to portfolio structure, process steps, and comfort level. Processing occurred during all phases of the portfolio review; however, processing was most prevalent in Phase I thinking. Pre-assessment thinking (Phase I) referred to that period of time when principals prepared to judge the artifacts and ultimately, the portfolio. In short, principals *framed* their work of evaluating the

portfolio in the pre-assessment thinking phase. This kind of pre-assessment framing aligned with the work of Glatthorn (1996) and Wolf, Lichtenstein, and Stevenson (1997). These authors proposed that evaluators get a sense of the entire portfolio before commencing analytical judgment of the contents relative to standards and criteria. During pre-assessment, the principals in this study focused their thinking on overall structure of the portfolio with some attention given to individually developed process steps. While the amount of time spent in pre-assessment thinking varied somewhat across academic level and geographic/demographic region, the activities of pre-assessment thinking were consistently present with the exception of one middle school principal.

Portfolio structure was cause for concern during the pilot study. The principal who participated in the pilot study was not able to completely move beyond thinking about the structure of the portfolio (pre-assessment thinking) into the judgment phase. Even during the judgment phase, he heavily critiqued the structure of the portfolio and was distracted as he attempted to assess the artifacts included in the portfolio. As a result of the pilot participant's experience, the think-aloud instructions provided by the researcher to the nine participants prior to their review were altered slightly to include a very brief description of the portfolio structure (Appendix F) the principal was about to review.

This slight alteration proved to be effective. The nine participants all completed the review of the portfolio without the structure distraction experienced by the pilot participant. Statements regarding portfolio structure were made during the other two phases of the respective reviews; however, the percentage of structure statements dropped

off significantly as the principals moved from pre-assessment (Phase I) into the judgment stage (Phase II) of their reviews. In addition, verbal reporting confirmed that comment counts regarding comfort level were low during pre-assessment. In fact, comfort level comments throughout all phases of the review were minimal; further evidence that the participants were comfortable with the responsibility of evaluating a “foreign” portfolio.

The eight principals who participated in pre-assessment thinking compared the portfolio provided to them by the researcher to second-year teacher portfolios produced in their own districts/buildings by teachers with whom they were familiar. Each school district in Iowa, while mandated to engage in teacher assessment via multiple measures, including the review of “collected artifacts” (known as a portfolio for purposes of this study), is given the latitude to develop their own means of defining, collecting, and displaying evidence of teaching. In short, districts exercise local control.

The element of local control was a concern for the researcher for fear that a foreign way of structuring a portfolio might be a barrier to a review or that reviews would be radically inconsistent within academic level. However, the common denominator during pre-assessment proved to be the Eight Iowa Teaching Standards (ITS). The principals were consistent across academic level and geographic/demographic region in that they focused their pre-assessment efforts on discovering *how* the teacher who produced the portfolio chose to illustrate his/her teaching for each ITS in the assembled artifacts. If the portfolio the principal was reviewing was not structured in a way with which the reviewing principal was accustomed, it did not prove to be a barrier or cause

the principal to negatively judge the artifacts or the teacher. The element of local control appeared to be a non-factor for the principals in this study during processing activities.

Verbal reporting relative to what principals thought during Phase I (the phase in which the highest percentage of processing activities existed) was consistent in content, however, differences did exist relative to the *number* of pre-assessment comments across academic level (see Table 4) and geographic/demographic region (see Table 8). In regard to academic level, middle school principals had the highest comment count, i.e., spent the most time in pre-assessment activities. Middle school principals made twice the comments of elementary principals and one-and-a-half times as many as the high school principals. One explanation as to the difference in the amount of time committed to pre-assessment thinking across academic level could be attributed to the difference in the portfolio structure. Both the elementary teacher and the high school teacher structured their portfolios by artifact; the two levels with the fewest comment counts. The middle school teacher structured her portfolio by teaching standard. In short, while structure did not prove to be a barrier, differing structures may require differing time commitment on the part of the reviewing principal.

In regard to geographic/demographic region, rural principals spent very little time in pre-assessment thinking. The time urban and suburban principals spent in pre-assessment thinking was comparable and was nearly twice that of their rural counterparts. It could be speculated that enrollment might be a factor. Principals in schools with lower enrollment, i.e., rural, may not have assistant principals who share in the responsibility of teacher evaluation; thus, spend less time in processing and pre-assessment activities. The

purpose of the current study was to discover what principals thought; not to explain why they thought as they did. Further study targeting portfolio structure and/or portfolio review as it relates to enrollment could provide additional insight.

Judging

Judging activities included verbal comments relative to judgment (artifact, teacher, portfolio), principals' opinions, and tools. Judging activities permeated every phase of the thinking framework; however, the percentage of judgment activities increased significantly as principals moved from Phase I to Phase II. In Phase I, judgment comments were very general and did not suggest a positive or negative determination. Again, this trend aligned with the suggestions of Glatthorn (1996) and Wolf, et al. (1997) that judgment activities not commence until principals had some sense of the portfolio, i.e., how it was laid out. Judgment activities in Phase III (reflection) decreased significantly and verbal reporting in that phase was mainly in response to the four guided interview questions. Verbal comments relative to artifact judgment will be discussed in this section. Judgment activities relative to the teacher, the portfolio, principal opinion, and tools will be discussed while answering Research Questions Two and Three.

The artifact judgment activities displayed by principals paralleled training they received in the IEATP. As the principals in the current study moved into the Phase II and began to judge the artifacts, a four-step rhythm became apparent. The principals would (a) identify and verbally describe an artifact aloud, (b) read aloud from the teacher reflection and/or artifact description, (c) make a judgment statement regarding the artifact, and (d) sometimes make a coaching statement about the artifact. This type of

rhythm was consistently demonstrated by principals across academic level and geographic/demographic region.

The four-step rhythm aligned closely with training processes outlined in the IEATP. The IEATP includes a section in which principals are coached to look for clues when examining sample artifacts for evidence of the eight ITS/criteria. During training, principals are asked to first become familiar with a sample artifact and then decide which ITSs and/or criterion is supported by the artifact. Principals are cautioned during the exercise to NOT judge the quality of the individual artifact before correctly matching the artifact with the ITS/criteria.

As principals in the current study judged an artifact, they were careful to fully understand the artifact and make a connection between the artifact and one or more of the ITS/criteria. If they could not make sense of the particular artifact, they would probe the artifact more deeply and reread any teacher reflection statements that might have been included before making a decision as to whether or not the artifact evidenced a particular ITS/criteria. The teachers who provided the portfolios for the current study indicated which artifact evidenced which ITS/criteria via artifact cover sheets or cross-referencing guides; however, the principals in the study still reviewed each artifact to some degree and made their own determination regarding which ITS/criteria the artifact evidenced.

Generally, the principals agreed with the teacher as to which ITS/criteria the artifact (s) evidenced. In some cases, the principals in the study verbally indicated if they thought the teacher had met additional ITS/criteria not listed on the cover sheet or the cross-reference guide. This was especially true of the middle school principals. The

middle school principals used cross-referencing at a much higher level; however, consistent with the IEATP, every principal in the study considered the use of one artifact to meet multiple ITS/criteria. The judgment statements, positive or negative, were sometimes followed by a coaching statement.

The fact that all nine principals in the study established a similar four-step rhythm provided a clear connection between training and actual practice; a result sure to be welcomed by the Iowa DE. Additional evidence emerged that supported a correlation between the IEATP and actual practice. Principals in the current study were consistent in the particular types of artifacts they believed to effectively evidence the ITS/criteria. The artifacts most often identified by principals as providing effective evidence were (a) observation summaries completed by the teacher's principal or a peer, (b) samples of student work, (c) pictures of activities, (d) copies of two-way email communication with parents and colleagues, (e) rubrics, (f) lesson plans, and (g) minutes from professional development sessions and committee assignments. During IEATP training, principals are exposed to and asked to make judgment about samples similar to the above-identified artifacts. Results of the current study provided evidence that the IEATP caused consistency across academic level and geographic/demographic region in terms of types of evidence principals identify as effectively meeting ITS/criteria.

While consistency across academic level was evident in terms of the four-step rhythm, artifact identification, and the influence of IEATP, differences did exist in the amount of time principals spent engaged in judgment. The bulk of judgment activity was evident in Phase II for all academic levels (see Table 12). During Phase II, middle school

principals spent twice as much time judging artifacts as did elementary principals. The high school principals also spent significant time in judging but were still lower than the middle school principals. The high school principals spent one and one-half as much time engaged in judgment as did their elementary counterparts.

The verbal reports point to the judging pattern of the elementary school principals as a possible explanation for the lower judgment activity they exhibited. The elementary principals very closely scrutinized the first three of the ten artifacts (see Table 16). Their verbal comments regarding judgment of the first three artifacts accounted for nearly half of their artifact judgment comments in Phase II. In addition, their judgment comments for the first three artifacts were primarily positive. It may be that once the principals became comfortable with the artifacts and convinced that the teacher was correctly evidencing the indicated ITS/criteria, they did not deem it necessary to closely scrutinize the last seven artifacts; thus, accounting for lower judgment activity overall.

Middle school principals exhibited a judging pattern similar to the elementary principals, i.e., they closely scrutinized artifacts evidencing the first three ITS/criteria and then decreased their judgment comments for the remaining five (see Table 23). However, the similar judging pattern did not yield the same lower result in overall judging comments as seen with the elementary principals. In fact, the judgment comments made by the middle school principals were nearly twice that of the elementary principals. It could be speculated that middle school principals used the same logic as the elementary school principals in their judgment pattern in that they became comfortable with the artifacts evidencing the first three ITS/criteria and convinced that the teacher knew what

she was doing so decreased their scrutiny for the remaining five ITS. The significant difference in judgment comments between the elementary and middle school principals could be attributed to the considerable amount of cross-referencing done by the middle school principals; hence, more judgment comments. In short, perhaps middle school principals attended to their judgment work early in the artifact judgment phase by using cross-referencing to make sure they could account for each ITS/criteria. Once they did account for each ITS/criteria, they did not as closely scrutinize the remaining artifacts.

It was anticipated that the high school principals would exhibit the same sort of judgment pattern demonstrated by the elementary and high school principals, i.e., early scrutiny of artifacts and then a decrease in judgment comments. However, this expectation did not bear out. The high school principals closely scrutinized each of the seven artifacts in the portfolio (see Table 31). Overall, the artifacts garnered more positive than negative comments but the percentage of negative comments was higher for the high school principals than it was for either the elementary or middle school principals.

In an effort to explain the high negative comment counts relative to artifact judgment, the comment counts of the high school principals for comfort level were revisited. There appeared to be no correlation between comfort level and negative comments. The comfort level comments that occurred during the time when judgment of the artifacts was highest (Phase II) did not increase in number or level of concern. It could be speculated that higher across-the-board scrutiny relative to individual artifacts may have occurred due to the fact that 50% of the negative comments occurred during

judgment of the first three artifacts; thus, leading to closer scrutiny of the remaining artifacts.

In a sense, the increased scrutiny supported the judgment trend exhibited by the elementary and middle school principals; however, the elementary and middle school principals gained a positive “trust” for their respective teachers more quickly than did the high school principals for their teacher. Had positive comments been more plentiful for the first three artifacts in the high school portfolio, the high school principals may have decreased their judgment comments for the remaining artifacts as did their elementary and middle school counterparts. It might also be speculated that closer scrutiny occurs at academic levels where teachers are responsible for teaching one discipline.

In regard to geographic/demographic region (see Figure 2), the differences in Phase II comment counts relative to judgment were much less marked than they were for Phase I. As was true in processing, the rural principals spent the least amount of time in judging but they were within 20 comments of their suburban and urban counterparts with regard to judging. Suburban and urban principals were within two comments of each other in judgment. These results illustrated that judgment was the emphasis of the review *regardless* of enrollment and resources as was posited for processing activities.

Coaching

As might be expected, coaching activities emerged as principals began their judgment activities in Phase II. The verbal reporting data relative to coaching illuminated the formative purpose of the portfolio and the role of the principal in the professional development of teachers. Danielson and McGreal (2000) described the formative nature

of a professional development portfolio as a framework in which teachers can initiate, plan, and facilitate their growth while they build connections between their own interests and goals and those of the schools (p.110). The high number of comments relative to meetings between principals and the teacher, and questions posed by principals as they thought aloud about said meetings, was indicative of the formative process being emphasized in the nine districts involved in the study. Verbal reporting suggested that the principals in the current study perceived their role, in part, as assisting the teacher in the portfolio process. For instance, if principals felt a teacher had a weak area, i.e., missing or inappropriate evidence, there was a willingness to assist the teacher in identifying appropriate artifacts; ones that the teacher may already be incorporating but not illustrating in the collection of artifacts.

Principals in the current study indicated that the meetings they held with teachers were for purposes of clarification, dialogue, and coaching. Peterson (2004) and Zimmerman and Deckert-Pelton (2003) described the role of the principal in teacher evaluation as critical and key. Shinkfield and Stufflebeam (1995) said that staff evaluation is one of the most critical responsibilities of the principal. They continued by saying that the review/feedback loop affects individual teachers and ultimately the school itself. Davis, Ellett, and Annunziata (2002) contended that the principal's ability to make teacher evaluation meaningful has the potential to enhance quality teaching. The coaching activities exhibited by the principals in the current study aligned with the current literature in that the participating principals were cognizant of their role in the formative process of evaluating the portfolio and the artifacts within.

Coaching comments relative to specific artifacts were common in the four-step judgment rhythm that was established by the principals; further evidence that principals were cognizant of the importance of feedback in the evaluation process. Principals most often made “suggested alternative” coaching comments specific to artifacts. Suggested alternatives were associated with artifacts that were judged negatively and positively. The suggested alternatives frequently reflected the principals need for clarification. Additional suggested alternatives were the use of (a) more frequent reflection (teacher and student), (b) additional proof of outcomes, (c) more variety, (d) evidence of student achievement in each artifact, (e) grammar and spell check, (f) more detailed artifact description, (g) authentic student work, and (h) data from other sources, i.e., student and/or parent surveys.

The principals in the current study indicated that meetings with the teacher were an important part of the evaluation process. They further indicated that discussion of the artifacts was a good vehicle for professional dialogue with their teachers. The verbal reports of the principals also suggested that they valued the use of questioning and equal engagement. The verbal comments relative to principal’s role aligned with the description of good leadership relative to teacher evaluation as described by Davis, et al (2002). Davis and his co-authors described a “small jazz combo (SJC)” style of leadership that emphasized the opportunities for collaboration and focus on good teaching and learning via well-designed/orchestrated teacher evaluation. The authors also described a “knight in shining armor” style of leadership in which the principal believed that he/she was protecting his teachers from and unfair judgment system. Verbal comments in the current

study did not suggest that the participating principals engaged in the “knight in shining armor” leadership style.

As opposed to punitive or “gotcha” evaluation and leadership, verbal reporting by the principals in the current study indicated the principals’ desire to contribute to the formative process. This type of leadership is emphasized in the IEATP via a training module dedicated to conferencing, coaching, and feedback. The module includes sample vignettes about which the principal creates questions using a framework. The framework suggests that principals design the following types of questions as he/she conferences:

1. Objective questions (“What?”). These questions are easy to answer, get at the facts, relieve stress, and invite/initiate active participation. An example of a suggested objective question is, “Where does this lesson fit into the curriculum?”

2. Reflective questions (“Then What?”). These questions elicit more emotional response and personal reaction. They invite a deepened level of participation; think, feel, gauge. An example of a reflective question is, “As you look at these artifacts, what concerns/pleases you?”

3. Interpretive questions (“So What?”). These questions invite sharing, and they build consciousness. In addition, they are designed to generate options and possibilities. An example of an interpretive question is, “What do these results mean to you in terms of future planning?”

4. Decisional questions (“Now What?”). These questions develop opinions that lead to future actions. They clarify expectations for improvement. An example of a

decisional question is, “What supports will you need to continue to work on those areas of concern to you?” (IEATP Training Manual, 2005, p. 257)

Once again, the verbal reporting in the current study indicated clear alignment with training that principals receive via the IEATP. The questions that the principals in the study asked aloud during verbal reporting reflected the types of questions suggested in the IEATP.

Critical Pieces

The category of critical pieces (CP) included comments relative to the critical nature of both teacher reflection and the role of observation. Because of the weight these two “pieces” carried throughout the study, the category emerged as one of the four cluster categories; not so much based on the number of comments but on the importance placed on each of the sub-categories via a few comments. The category of critical pieces consisted of two sub-categories; critical pieces/teacher reflection (CP/TR) and critical pieces/role of observation (CP/RO).

The principals in the study relied on written teacher reflection to provide insight into teachers’ thinking. The written reflection was in lieu of a face-to-face meeting in which the principal could ascertain why a teacher included a specific artifact or the impact the artifact had on teaching practice. Further, the principals in the study indicated that teacher reflection was critical to improving teaching practice. Principals, in research conducted by Attinello, Lare, and Waters (2006), also felt that portfolios encouraged teacher self-reflection and ultimately improved teaching practice. Xu (2004) emphasized that the portfolio process was a reflective process and a means of increasing conversation

with those outside the classroom. The written reflections were the only means of communicating with the teacher available to the principals in the current study. Consequently the principals were very attentive to the written reflections that the teachers included and commented when there was not a sufficient amount of reflection.

The role of observation was very pervasive throughout the study. As the principals moved between each phase of the portfolio evaluation and attended to processing, judging, and coaching activities, they were very clear that the portfolio was “just a piece” of teacher evaluation and that they placed higher value on observation. However, they contended that the portfolio provided a means for teachers to document evidence not readily seen. In essence, the principals were interested in multiple data sources. Peterson (2000) indicated that multiple sources were essential to an effective teacher evaluation system. The principals in Attinello’s (2004) study agreed that portfolios were a more comprehensive measure and they supported an evaluation process that included multiple data sources.

The fact that the principals in the current study considered the importance of multiple sources was supported in the IEATP via training modules dedicated to portfolio evaluation and classroom observation. The comments of the principals repeatedly referred to the portfolio as “only one piece.” During training, principals are reminded that observation is “only one source.” The training clarifies that classroom observation is appropriate when data is needed on teacher behaviors related to (a) student interaction, (b) classroom management, (c) classroom climate, (d) instructional strategies, and (e) student learning progress. Furthermore, the training advises principals of the limitations

of classroom management. Per the IEATP, limitations of classroom observation include (a) student work samples, (b) written feedback and assessment, (c) long-range planning, and communication with parents and school community. Principals in the current study were very cognizant of artifacts referred to in the limitations; perhaps heightened by the fact that they did not have the opportunity to observe the teacher.

In summary, the most significant finding for research question one was the consistency that existed in the portfolio judgment process. Based on the findings for research question one, Iowa has successfully implemented two of the three essential elements of an effective teacher evaluation system to which Danielson and McGreal (2000) refer. The ITS/criteria have provided a sound framework for defining effective teaching. And, results from the current study indicate that the IEATP has produced trained evaluators who can make consistent judgments.

Research Question 2: What tools do principals use?

Via verbal reporting and follow-up questioning, the researcher identified six tools that principals used during portfolio evaluation. The tools identified were (a) a copy of the eight ITS/criteria, (b) sticky notes, (c) a cross-reference guide, (d) artifact cover sheets, (e) a log listing descriptors for the ITS/criteria, (f) the DE summative evaluation form. The varied functions of the tools were to (a) provide visual guidance to both principals and teachers, (b) provide a means for principals to quickly cross-reference multiple ITS/criteria, (c) provide principals a means to informally and quickly communicate with teachers, (d) provide cues to the teacher and the principal relative to appropriate evidence, and (e) provide an informal framework for the summative

evaluation. The tools were generated by varied entities including the principal, the district, and the DE. Several of the tools exhibited similarity in appearance and function; however, with the exception of the DE summative evaluation, there were no two tools that were identical. Three of the principals provided the researcher with samples of the tools they used (see Appendix S, W, and X).

Surprisingly absent from the tools used by the principals was a rubric. Green and Smyser (1996) indicated that evaluators should develop and validate rubrics. The National Board for Professional Teaching Standards (NBPTS), recognized by many as the hallmark in the use of portfolios, makes extensive use of rubrics when they evaluate the portfolios of their candidates so that assessments are accurate and fair. In addition, Kimball (2002) indicated that rubrics served an important function in teacher evaluation systems.

The Iowa DE does not require principals to use a rubric. However, during the IEATP, administrators are trained on the use of a rubric and an example rubric is included in the training manual. The IEATP training manual indicates that a rubric and associated descriptors are intended to increase consistency across administrators and settings. Information in the manual also reminds principals that use of the rubric is optional and that the only requirement by Iowa law is to determine if a teacher has “met” the ITS as defined through the criteria and descriptors. The suggested rubric in the IEATP includes four rating levels: (a) exceeds expectation, (b) meets expectation, (c) needs improvement, and (d) unsatisfactory.

The participating principals may have had a “mental rubric” but did not bring with them a rubric that described the differentiated levels of performance. The tools used by the principals in the current study that supplied the most consistency were the ITS/criteria sheets and the descriptors that listed appropriate evidence.

In summary, despite local control and the lack of a state-wide rubric, the coded comment counts indicated that the nine principals in the current study were fairly consistent as they judged the artifacts in the portfolio; suggesting that the tools that the principals used relative to the ITS/criteria and model descriptors provided the necessary consistency.

Research Question 3: Bearing of the portfolio on licensure decision.

Research question three was answered via a guided interview question that asked, “How much bearing would you assign to this portfolio when making a licensure decision?” Responses ranged from “not much” to as much as 50%. Based on reported numerical percentages, the average bearing was roughly 30%; accounting for one-third of the principal’s decision.

Findings indicated that, per academic level, (a) elementary principals put the most bearing on the portfolio (38%), (b) high school principals were second (30%), and (c) middle school principals put the least bearing on the portfolio (25%). Findings regarding geographic/demographic region showed that (a) suburban principals put the most bearing on the portfolio (33%), (b) rural principals were second (28%), and (c) urban principals did not commit to a percentage-based response.

Results indicating that principals based only 30% of their decision on the portfolio were mildly surprising. Given the amount of time that principals said they spent reviewing portfolios and writing the summative evaluation of the teacher that included the portfolio evaluation, it was anticipated that more bearing would be placed on portfolios. However, based on the principals' comments relative to the critical nature of observation in teacher evaluation, the low bearing was not a shock.

The principals in the current study were clear that a good portfolio was not always an indicator of good teaching. Peterson (2000) and Green and Smyser (1996) believe that good portfolios can make bad teachers look good and vice versa. The verbal reports of the principals in the current study indicated that they had strong opinions regarding the use of portfolios in teacher assessment. Consistent with literature (e.g., Danielson, 2001; Danielson & McGreal, 2000; Green & Smyser, 1996; Stronge & Tucker, 2003) the principals' comments suggested that they believed portfolios served as only one piece of teacher evaluation and should be used in conjunction with observation. In fact, comments made by principals in the current study indicated that observation was a critical piece of teacher evaluation, if not the most critical.

Implications for Further Research

The results of the current study clearly illustrated that consistency exists relative to portfolio evaluation across academic level and across geographic/demographic region. The ITS/criteria and the IEATP seem to have been fully implemented. Both are considered by Danielson (2001) as two of the three critical pieces of an effective teacher evaluation system. However, cause for concern may be the absence of the third piece of

an effective teacher evaluation system; a rubric. Rubrics provide clear data relative to quality. Evaluation will become even more of a high stakes proposition as policymakers move towards merit-pay programs (Jacob & Lefgren, 2006). As a result, clear and defensible data will be essential. The sample for the current study was small. It is suggested that additional research be conducted, using a much larger sample, relative to consistency in judgment and to what extent rubrics are being incorporated into the teacher evaluation system in Iowa.

In the current study, the participating principals were able to conduct a complete review of a portfolio that was created by a teacher other than someone working in their building. In short, the structure of the portfolios did not prove to be a barrier to a complete review. Perhaps it is conceivable that an outside person or team could perform scheduled periodic reviews and rate the portfolios for schools/districts as a way to further validate the judgment consistency of reviewing principals. It would be crucial to pilot such a system and track consistency of judgment; similar to how judgment was tracked in the current study.

The principals in the current study indicated that they would most likely recommend the teacher who created the sample portfolio for professional licensure. Collecting state-wide data regarding the number of teachers (a) who are not recommended for licensure by their principal after the second year, (b) who are recommended for a third year of mentoring, (c) who are counseled out of education in year one or two by mentors or principals, (d) who self-select i.e., choose to leave of their own accord during year one or two, and (e) the rating of the portfolio per each associated

decision made by the principal, mentor, or teacher could provide additional insight into the impact of the teacher evaluation system currently in place in Iowa. It would be important to be sensitive to confidentiality issues in the data collection process.

The current study provided qualitative data from the perspective of the principal. Information from the perspective of teachers is also critical. It is suggested that additional research be conducted to ascertain what teachers think about the portfolio process, the tools they use, and the bearing they believe should be placed on the portfolio for licensure decisions. Further, teacher perceptions would provide clear insight into the value they place on the portfolio as a tool for reflection and professional development. As a result, a much fuller picture of the practical implications of teacher assessment in Iowa would be created.

REFERENCES

- Attinello, J. (2004). Teachers' and administrators' perceptions of the value of portfolio-based teacher appraisals for evaluation and professional growth. ProQuest Information and Learning Company. (UMI No. 3120835)
- Attinello, J., Lare, D., & Waters, F. (2006). The value of teacher portfolios for evaluation and personal growth. *NASSP Bulletin*, 79(2), 132-152.
- Beerens, D. R. (2000). *Evaluating teachers for professional growth: Creating a culture of motivation and learning*. Thousand Oaks, CA: Corwin Press, Inc.
- Beyer, L. E. (2002). The politics of standards and the education of teachers. *Teaching Education*, 13(3), 305-316.
- Blake, J., Bachman, J., Frys, M. K., Holbert, P., Tamara, I., & Sellitto, P. (1995). A portfolio-based assessment model for teachers: Encouraging professional growth. *NASSP Bulletin*, 79(October), 37-46.
- Boeije, H. (2002). A purposeful approach to the constant comparative method in the analysis of qualitative interviews. *Quality and Quantity*, 36, 391-409.
- Bogdan, R. C., & Biklen, S. K. (2003). *Qualitative research for education: An introduction to theories and methods*. (4th ed.). Boston: Pearson Education Group.
- Brock, B. L., & Grady, M. L. (1998). Beginning teacher induction programs: The role of the principal. *The Clearing House*, 71(3), 179-183.
- Burke, K. (1997). *Designing Professional Portfolios for Change*. Arlington Heights, IL: Skylight Professional Development.
- Campbell, C., Cignetti, P., Melenyzer, B., Nettles, D., & Wyman, R. (2001). *How to develop a professional portfolio: A manual for teachers* (2nd ed.). Boston: Allyn and Bacon.
- Costantino, P. M., & De Lorenzo, M. N. (2006). *Developing a professional teaching portfolio: A guide to success* (2nd ed.). Boston: Pearson.
- Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks: Sage.

- Danielson, C. (2001). New trends in teacher evaluation. *Educational Leadership*, 58(5), 12-15.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Princeton: Educational Testing Services.
- Davis, R., Ellett, C., and Annunziata, J. (2002). Teacher evaluation, leadership and learning organizations. *Journal of Personnel Evaluation in Education*, 16(4), 287-301.
- Ellis, A. K. (2001). *Research on education innovations* (3rd ed.). Larchmont, NY: Eye on Education, Inc.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge: MIT Press.
- Gall, G. D., Gall, J. P., & Borg, W. R. (2003). *Educational research: An introduction* (7th ed.). Boston: Allyn and Bacon.
- Gallini, J. K., & Barron, D. (2001-2002). Participants' perceptions of Web-infused environments: A survey of teaching beliefs, learning approaches, and communication. *Journal of Research on Technology in Education*, 34(2), 139-156.
- Gay, L. R., & Airasian, P. (2003). *Educational research: Competencies for analysis and application*. (7th ed.). Upper Saddle River: Merrill/Prentice Hall.
- Gelfer, J., Xu, Y., and Perkins, P. (2004). Developing portfolios to evaluate teacher performance in early childhood education. *Early childhood education journal*, 32(2), 127-132.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine DeGuyter.
- Glatthorn, A. (1996). *The teacher's portfolio: Fostering and documenting professional development*. Rockport: Pro Active Publications.
- Good, T. L., & Mulryan, C. (1990). Teacher ratings: A call for teacher control and self-evaluation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers* (pp. 191-215). Newbury Park, CA: Sage Publications.
- Green, J., & Smyser, S. (1996). *The teacher portfolio: A strategy for professional development and evaluation*. Lancaster, PA: Technomic.

- Henneman, G. H., & Milanowski, A. T. (2003). Continuing assessment of teacher reaction to a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, 17(2), 173-195.
- Herring, M. (1997). *Design and training for implementation of constructivist-based distance learning environments*. Unpublished Doctoral Dissertation, Iowa State University, Ames.
- Iowa Code, Chapter 284. (2001). *Teacher performance, compensation and career development*. Retrieved October 7, 2007, from <http://search.legis.state.ia.us/NXT/gateway.dll/IowaState/ISLRoot/code.htm?f=templates&fn=default.htm>
- Iowa Department of Education. (n.d.1). *Iowa professional development model: Case studies*. Retrieved March 1, 2009, from http://www.iowa.gov/educate/index.php?option=com_content&view=article&id=232&Itemid=1286
- Iowa Department of Education. (n.d.2). Iowa teaching standards and model criteria. Retrieved March 1, 2008, from http://www.iowa.gov/educate/index.php?option=com_content&task=view&id=1684&Itemid=1
- Iowa Evaluator Approval Training Program Participants' Manual. (2005). Iowa Department of Education.
- Jacob, B., & Lefgren, L. (2006). When principals rate teachers. *Education Next*, 6(2), 59-64.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Jonassen, D. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory (Vol. II)*, pp. 215-239. Mahwah: Lawrence Erlbaum Associates.
- Kimball, S. (2002). Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems. *Journal of Personnel Evaluation in Education*, 16(4), 241-268.
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills: Sage Publications.

- Lane, S., & Stone, C. A. (2002). Strategies for examining the consequences of assessment and accountability programs. *Educational Measurement*, 21(1), 23-30.
- LeCompte, M. D., & Preissle, J. (1993). *Ethnography and qualitative design in educational research* (2nd ed.). San Diego: Academic Press, Inc.
- Marshall, M. (2004). School-based accountability systems: Data sources for teacher development and school improvement. In National Evaluation Systems, Inc, What is a qualified, capable teacher (p. 11-33). Retrieved January 11, 2006 from http://www.nesinc.com/PDFs/2004_03Marshall.pdf
- McNelly, T. A. (2002). Evaluations that ensure growth: Teacher portfolios. *Principal Leadership*, 3(4), 55-60.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.
- Muth, K. (1993). The thinking-aloud procedure: A diagnostic tool for middle school mathematics teachers. *Middle School Journal*, 24(March), 5-9.
- National Commission on Teaching and America's Future (1996). *What matters most: Teaching and America's future*. Retrieved October 23, 2003 from <http://www.nctaf.org/article/index.php?c=4&sc=42&>
- Nolan, J., & Hoover, L. A. (2004). *Teacher supervision and evaluation: Theory into practice*. Hoboken, NJ: John Wiley & Sons, Inc.
- Pedersen, S., & Liu, M. (2003). Teacher's beliefs about issues in the implementation of a student-centered learning environment. *Educational Technology Research and Development*, 51(2), 57-76.
- Peterson, K. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices* (2nd ed.). Thousand Oaks, CA: Corwin Press, Inc.
- Peterson, K. (2004). Research on school based evaluation. *NAASP Bulletin*. 88(639), 60-79.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale: Lawrence Erlbaum Associates.
- Protheroe, N. (2002). Improving instruction through teacher observation. *Principal*, 81(1), 48-51.

- Reese, S. (2004). Teacher Portfolios: Displaying the art of teaching. *Techniques*, 79(5), 18-21.
- Reeves, D. B. (2004). *Assessing educational leaders: Evaluating performance for improved individual and organizational results*. Thousand Oaks: Corwin Press.
- Ryan, G. W., & Bernard, H. R. (2003). Data management and analysis methods. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (pp. 259-309). Thousand Oaks, CA: Sage Publications.
- Schlechty, P. C. (1997). *Inventing better schools*. San Francisco: Jossey-Bass.
- Shavelson, R., Webb, N., & Burstein, L. (1986). Measurement of teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 50-91). New York: Macmillan.
- Shinkfield, A. J., & Stufflebeam, D. L. (1995). *Teacher evaluation: Guide to effective practice*. Boston: Kluwer Academic Publishers.
- St. Maurice, H., & Shaw, P. (2004). Teacher portfolios come of age: A preliminary study. *NASSP Bulletin*, 88(639), 15-25.
- Stronge, J. H. (2002). *Qualities of effective teachers*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Stronge, H. J., & Tucker, P. D. (2003). *Handbook on teacher evaluation: Assessing and Improving performance*. Larchmont, NY: Eye on Education, Inc.
- Tellez, K. (2003). In response to increasing state and national control over the teacher education profession. *Teacher Education Quarterly*, 30(1), 9-72.
- Tigelaar, D., Dolmans, D., Wolfhagen, I., & van der Vleuten, C. (2005). Quality issues in judging portfolios: Implications for organizing teaching portfolio assessment procedures. *Studies in Higher Education*, 30(5), 595-610.
- Tucker, P., Stronge, J., & Gareis, C. (2002). *Handbook on teacher portfolios for evaluation and professional development*. Larchmont, NY: Eye on Education, Inc.
- United States Department of Agriculture Economic Research Service. (n.d.). *Measuring rurality: What is rural?* Retrieved January 1, 2006, from <http://www.ers.usda.gov/Briefing/Rurality/WhatIsRural>
- Urban Education Network of Iowa by-laws (1984). *Section 2: Membership in the Urban Education Network of Iowa*.

- Wasley, P.A., & McDiarmid, G. W. (2003). *Tying the assessment of new teachers to student learning and to teacher preparation*. Unpublished manuscript, Denver, CO.
- Wolf, K., Lichtenstein, G., & Stevenson, C. (1997). Portfolios in teacher evaluation. In H. J. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 193-214). Thousand Oaks, Ca: Sage Publications.
- Xu, J. (2003). Promoting school-centered professional development through teaching portfolios: A case study. *Journal of Teacher Education*, 54(4), 347-361.
- Xu, Y. (2004). Teacher portfolios: An effective way to assess teacher performance and enhance learning. *Childhood Education*, 80(4), 198-201.
- Zimmerman, S., & Deckert-Pelton, M. (2003). Evaluating the evaluators: Teachers' perceptions of the principal's role in professional evaluation. *NASSP Bulletin*. 87(636), 28-37.

APPENDIX A

IOWA TEACHING STANDARDS AND MODEL CRITERIA

IOWA TEACHING STANDARDS AND MODEL CRITERIA

Standard 1

Demonstrates ability to enhance academic performance and support for implementation of the school district's student achievement goals.

Model Criteria

The teacher:

- a. Provides evidence of student learning to students, families, and staff.
- b. Implements strategies supporting student, building, and district goals.
- c. Uses student performance data as a guide for decision making.
- d. Accepts and demonstrates responsibility for creating a classroom culture that supports the learning of every student.
- e. Creates an environment of mutual respect, rapport, and fairness.
- f. Participates in and contributes to a school culture that focuses on improved student learning.
- g. Communicates with students, families, colleagues, and communities effectively and accurately.

Standard 2

Demonstrates competence in content knowledge appropriate to the teaching position.

Model Criteria

The teacher:

- a. Understands and uses key concepts, underlying themes, relationships, and different perspectives related to the content area.
- b. Uses knowledge of student development to make learning experiences in the content area meaningful and accessible for every student.
- c. Relates ideas and information within and across content areas.
- d. Understands and uses instructional strategies that are appropriate to the content area.

Standard 3

Demonstrates competence in planning and preparing for instruction.

Model Criteria

The teacher:

- a. Uses student achievement data, local standards, and the district curriculum in planning for instruction.
- b. Sets and communicates high expectations for social, behavioral, and academic success of all students.
- c. Uses student's developmental needs, backgrounds, and interests in planning for instruction.
- d. Selects strategies to engage all students in learning.
- e. Uses available resources, including technologies, in the development and sequencing of instruction.

Standard 4

Uses strategies to deliver instruction that meets the multiple learning needs of students.

Model Criteria

The teacher:

- a. Aligns classroom instruction with local standards and district curriculum.
- b. Uses research-based instructional strategies that address the full range of cognitive levels.
- c. Demonstrates flexibility and responsiveness in adjusting instruction to meet student needs.
- d. Engages students in varied experiences that meet diverse needs and promote social, emotional, and academic growth.
- e. Connects students' prior knowledge, life experiences, and interests in the instructional process.
- f. Uses available resources, including technologies, in the delivery of instruction.

Standard 5

Uses a variety of methods to monitor student learning

Model Criteria

The teacher:

- a. Aligns classroom assessment with instruction.
- b. Communicates assessment criteria and standards to all students and parents.
- c. Understands and uses the results of multiple assessments to guide planning and instruction.
- d. Guides students in goal setting and assessing their own learning.
- e. Provides substantive, timely, and constructive feedback to students and parents.
- f. Works with other staff and building and district leadership in analysis of student progress.

Standard 6

Demonstrate competence in classroom management.

Model Criteria

The teacher:

- a. Creates a learning community that encourages positive social interaction, active engagement, and self-regulation for every student.
- b. Establishes, communicates, models, and maintains standards of responsible student behavior.
- c. Develops and implements classroom procedures and routines that support high expectations for student learning.
- d. Uses instructional time effectively to maximize student achievement.
- e. Creates a safe and purposeful learning environment.

Standard 7

Engages in professional growth.

Model Criteria

The teacher:

- a. Demonstrates habits and skills of continuous inquiry and learning.
- b. Works collaboratively to improve professional practice and student learning.
- c. Applies research, knowledge, and skills from professional development opportunities to improve practice.
- d. Establishes and implements professional development plans based upon the teacher's needs aligned to the Iowa teaching standards and district/building student achievement goals.

Standard 8

Fulfills professional responsibilities established by the school district.

Model Criteria

The teacher:

- a. Adheres to board policies, district procedures, and contractual obligations.
- b. Demonstrates professional and ethical conduct as defined by state law and district policy.
- c. Contributes to efforts to achieve district and building goals.
- d. Demonstrates an understanding of and respect for all learners and staff.
- e. Collaborates with students, families, colleagues, and communities to enhance student learning.

Source: Iowa Department of Education (n.d. 2) Iowa teaching standards and model criteria. Educator Quality Link.

APPENDIX B

IOWA DEPARTMENT OF EDUCATION SAMPLE COMPREHENSIVE
EVALUATION

Iowa Department of Education
Sample Comprehensive Evaluation

Comprehensive Evaluation Summative Evaluation Form

Teacher: Bob Smith _____ Folder #: 0000 _____
 Evaluator: Bee A. Betterprincipal _____ Folder #: 1111 _____
 School Name: High Standards CSD _____
 Grade Level: 2 _____ Subjects: Elementary (all) _____ Year: 1 3

Bob Smith and I met on August 20, 2002 to go over our plan for the school year.

-We first discussed his participation in the mentor induction program and coordinated all dates with his mentor.
 -Next, we went over the comprehensive evaluation form that is being used for his final evaluation in March and set up dates for the three classroom observations that I would be conducting. The schedule is below:

Observation (1): Pre-observation conference: October 22

Observation (1): October 23

Post-observation conference 25

Observation (1) would be from 8:30 a.m. – 9:45 a.m. Subjects covered: reading

Observation (2): Pre-observation conference: January 20, 2003

Observation: January 23, 2003

Post-observation conference: January 24, 2003

Observation (2) would be from 8:25 a.m. – 11:50 a.m. Subjects covered: reading, spelling, language and math

Observation (3): Pre-observation conference: February 18, 2003

Observation (3): February 19, 2003

Post-observation conference: February 20, 2003

Observation (3) would be from 1:30 – 2:30 p.m. Subjects covered: science, writing

Final Evaluation: March 6th, 8:00 a.m.

-Bob and I both understood that circumstances could arise that would cause us to have to change this schedule, however, we will try to reserve these dates with any changes that are made adequate and timely notification given. Fortunately, we were able to maintain the conferencing and observation schedule this year without any changes.

-It was discussed at this time that informal observations would be conducted at random during the school year. He was encouraged to ask for my assistance at any time.

-A rubric that includes the Iowa Teaching Standards and Criteria plus our district's descriptors and what artifacts that he would need to include as evidence for his final evaluation was given to Bob on 8-20-02. We went over this rubric in great lengths, going through each standard/criteria and talking about district expectations. After the August 20 meeting we scheduled meetings monthly so we could develop a deeper meaning of the standards and examine artifacts and evidence collected. We developed a timeline in which we would need to have artifacts collected and determine the evidence still needed to collect. We discussed at this time some good ideas for filing, record keeping and putting-organizing any evidence of progress.

-Bob was given copies (8-20-02) of all evaluation guidelines and forms and is aware that he must show evidence of all eight Iowa Teaching Standards and Criteria in order to obtain licensure in Iowa.

-Copies of the district and building student achievement goals for 2002-03 were also given to Bob at this time (8-20-02) along with 2001-02 student achievement data.

Iowa Department of Education
Sample Comprehensive Evaluation

Directions:

In the narrative under each standard, the evaluator should incorporate and address each criterion.

1. DEMONSTRATES ABILITY TO ENHANCE ACADEMIC PERFORMANCE AND SUPPORT FOR AND IMPLEMENTATION OF THE SCHOOL DISTRICT'S STUDENT ACHIEVEMENT GOALS.

The teacher:

- Provides evidence of student learning to students, families, and staff.
- Implements strategies supporting student, building, and district goals.
- Uses student performance data as a guide for decision making.
- Accepts and demonstrates responsibility for creating a classroom culture that supports the learning of every student.
- Creates an environment of mutual respect, rapport, and fairness.
- Participates in and contributes to a school culture that focuses on improved student learning.
- Communicates with students, families, colleagues, and communities effectively and accurately.

Evidence to support attainment of or failure to meet standards:

Bob has shown that he documents students learning with meaningful measures using data that is understandable. He has shared individual and classroom goals, and results, with students families, and staffs this year. Bob has provided evidence such as the Achievement Level testing data, which is a standardized, criterion-referenced test that is used in our district to determine his flexible skill grouping in math.

He also shared with me the reporting that he does weekly to parents in his Friday folders. A copy of a student report card shows that he is aligning student's achievement goals with our district's goals. Bob also knows the short and long-range building and district goals for student learning, and does implement these goals in his classroom. He has posted in his room the building's goals for reading and has communicated what the students, parents and teacher's responsibilities for reaching these goals with each group. He also has submitted as evidence his unit plans, which align learning with our district's standards and benchmarks.

Included in Bob's collection of artifacts is the performance data such as our district's Achievement Level tests, the reading program probes, math timed test results and evidence of how he has used this to make decisions regarding the student's progress and planning for instruction.

He motivates students to make positive choices to enhance their learning. I have received many notes from parents supporting Mr. Smith and his ability to run his classroom in a safe and respectful manner. He has included one of these notes from a parent as evidence. He has also submitted as evidence his classroom rules, which are posted, along with the building level behavioral expectations. He has developed, and supported our district's character education goals, by incorporating the building level behavioral expectations into his unit plans for teaching. He has submitted evidence of this through a unit plan and has included as evidence a student artifact – a student-developed book, which is now available for check out in our school library. I have observed Mr. Smith on each formal classroom observation encouraging students to work cooperatively and independently and including students in decision-making when appropriate. He has a very responsible classroom culture.

Bob has submitted as evidence the strategies developed with colleagues to improve student learning by including team meeting planning notes and documentation of discussions he has made on the behalf of students with the special education teacher, AEA personnel and parents. The ability to communicate well is one of Bob's strong attributes. Bob has included as artifacts under Standard 1 a weekly Friday folder log to parents, e-mail communications, and articles written for the district and building newsletters that demonstrates communications with families that is effective and accurate.

All criteria for Standard 1 have been addressed using multiple sources and multiple data points.

Circle one:

Meets Standard

Does Not Meet Standard

Additional documentation/artifacts applicable to this standard are attached as Appendix A-1.

APPENDIX C

UNIVERSITY OF NORTHERN IOWA HUMAN PARTICIPANTS REVIEW
INFORMED CONSENT—PRINCIPAL

**UNIVERSITY OF NORTHERN IOWA
HUMAN PARTICIPANTS REVIEW
INFORMED CONSENT—PRINCIPAL**

Project Title: The Thought Processes of Administrators as They Review and Make Evaluative Judgment of a Second-Year Teacher Portfolio: A Qualitative Study

Name of Investigator: Terri Anne Lasswell

Invitation to Participate: You are invited to participate in a research project conducted through the University of Northern Iowa. The University requires that you give your signed agreement to participate in this project. The following information is provided to help you make an informed decision whether or not to participate.

Nature and Purpose of the Project: The purpose of this research is to ascertain how administrators evaluate beginning teacher portfolios.

Explanation of Procedures: As a participant, you will be asked to 1) complete a demographic questionnaire, 2) review a sample portfolio while verbalizing your thoughts, and 3) answer several framed questions after you have completed the review. You will be asked to complete the demographic questionnaire and mail it back to the investigator prior to the interview. It should take approximately 15 minutes to complete the questionnaire. Envelope and stamp will be provided.

The anticipated length of the interview is ninety minutes to two hours. The interview will take place at an agreed upon location between you and the investigator. The interviews will be audio taped and later transcribed (see below for confidentiality information). The investigator will act as the interviewer. Transcriptions will be provided to you by the investigator for accuracy approval.

Once the interviews are completed, then approved by you, emerging and consistent themes concerning the evaluation of portfolios will then be analyzed and interpreted for purposes of my dissertation.

Discomfort and Risks: No more than minimal risks (discomfort, burden, and inconvenience) are anticipated.

Benefits: You will receive no direct benefits from your participation in this study.

Confidentiality: Information obtained during this study, which could identify you, will be kept confidential. The summarized findings with no identifying information will be included in my dissertation and may be published in an academic journal or presented at a scholarly conference. The audiotapes and the transcriptions will be coded. Your identity will be kept separate from the coded data. Only the investigator will have access to the

identity of the audiotapes for clarification purposes should questions arise during the course of the study.

Right to Refuse or Withdraw: Your participation is completely voluntary. You are free to withdraw from participation at any time or choose not to participate at all.

Questions: If you have any questions about the study or desire more information, you may contact Terri A. Lasswell (investigator) at 319-236-3541 or my faculty advisors, Dr. John Henning in the Department of Educational Psychology and Foundations (319-273-7488) and Dr. Mary Herring, Department of Curriculum and Instruction (319-273-2368), University of Northern Iowa. You can also contact the Office of Human Participants Coordinator, University of Northern Iowa, at 319-273-2748, for answers to questions about the rights of research participants and the participant review process.

I am fully aware of the nature and extent of my participation in this project as stated above and the possible risks arising from it. I hereby agree to participate in this project. I acknowledge that I have received a copy of this consent statement. I am 18 years of age or older.

(Signature of participant)

(Date)

(Printed name of participant)

(Signature of investigator)

(Date)

(Signature of instructor/advisor)

(Date)

APPENDIX D
DEMOGRAPHIC INFORMATION

DEMOGRAPHIC INFORMATION

1. How many years have you been in education as a Teacher _____ Principal _____
2. Years in current position: _____
3. Gender: _____ M _____ F
4. Age: _____
5. Degrees Earned _____BA/BS _____ MS _____ Ed.D./Ph.D. _____ Other
(specify)_____
6. Year you completed Evaluator Approval Training/DDL: _____
7. Total Student Population in Your District: _____
8. How many 1st year teachers did you evaluate in the '04-'05 school year? _____
9. How many 2nd year teachers did you evaluate in the '04-'05 school year? _____

APPENDIX E

UNIVERSITY OF NORTHERN IOWA HUMAN PARTICIPANTS REVIEW
INFORMED CONSENT—TEACHER

**UNIVERSITY OF NORTHERN IOWA
HUMAN PARTICIPANTS REVIEW
INFORMED CONSENT—TEACHER**

Project Title: The Thought Processes of Administrators as They Review and Make Evaluative Judgment of a Second-Year Teacher Portfolio: A Qualitative Study

Name of Investigator: Terri Anne Lasswell

Invitation to Participate: You are invited to participate in a research project conducted through the University of Northern Iowa. The University requires that you give your signed agreement to participate in this project. The following information is provided to help you make an informed decision whether or not to participate.

Nature and Purpose of the Project: The purpose of this research is to ascertain how administrators evaluate beginning teacher portfolios.

Explanation of Procedures: Your role as a participant in this study is to provide your portfolio as a sample portfolio for review by principals (not in your district) who are also voluntary participants in this study. Any reference to you, your school, your community, or specific students will be blacked out prior to use of your portfolio in the study. I will personally team with you to mark up the portfolios for confidentiality purposes. To further insure confidentiality, media items, whether produced by you or by your students, (e.g., such as PowerPoint, streaming video, taped audio, and pictures) will not be included in the portfolio. A table of contents will be included indicating to the reviewing principals that you did indeed use this type of artifact.

The anticipated length time it will take to mark up your portfolio is one hour. The mark up session will take place at an agreed upon location between you and the investigator.

Discomfort and Risks: No more than minimal risks (discomfort, burden, and inconvenience) are anticipated.

Benefits: You will receive no direct benefits from your participation in this study.

Confidentiality: Information obtained during this study, which could identify you, will be kept confidential. The summarized findings with no identifying information will be included in my dissertation and may be published in an academic journal or presented at a scholarly conference.

Right to Refuse or Withdraw: Your participation is completely voluntary. You are free to withdraw from participation at any time or choose not to participate at all.

Questions: If you have any questions about the study or desire more information, you may contact Terri A. Lasswell (investigator) at 319-236-3541 or my faculty advisors, Dr. John Henning in the Department of Educational Psychology and Foundations (319-273-7488) and Dr. Mary Herring, Department of Curriculum and Instruction (319-273-2368), University of Northern Iowa. You can also contact the Office of Human Participants Coordinator, University of Northern Iowa, at 319-273-2748, for answers to questions about the rights of research participants and the participant review process.

I am fully aware of the nature and extent of my participation in this project as stated above and the possible risks arising from it. I hereby agree to participate in this project. I acknowledge that I have received a copy of this consent statement. I am 18 years of age or older.

(Signature of participant)

(Date)

(Printed name of participant)

(Signature of investigator)

(Date)

(Signature of instructor/advisor)

(Date)

APPENDIX F

THINK ALOUD INSTRUCTIONS TO PARTICIPANTS

THINK ALOUD INSTRUCTIONS TO PARTICIPANTS

“In this research, I am interested in what you think about as you review and make judgments concerning the portfolio of a second-year teacher. In order to do this, I am going to ask you to THINK ALOUD as you review the sample portfolio. What I mean by think aloud is that I want you to tell me EVERYTHING you are thinking from the time you first see the portfolio until the time YOU render the review complete. I don't want you to try to plan out what you say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself. It is most important that you keep talking. If you are silent for any long period of time, I will ask you to talk. Do you understand what I want you to do?” (Adapted from Ericsson and Simon (1993), p. 378)

The portfolio you are about to evaluate belongs to a _____ (grade level and content where appropriate). The portfolio is structured with a table of contents to guide you in terms of structure. This particular portfolio is structured:

Elementary Portfolio: by artifact. In other words, the teacher has divided the portfolio into ten artifact sections where the artifact represents one or more of the ITS. For instance, the teacher includes an observation artifact concerning writer's workshop. The artifact is cross referenced with each of the standards on a cover sheet and then the evidence is placed in the artifact section. (At this point, I opened the portfolio to show the principals the cross reference sheet for clarification).

Middle School Portfolio: by ITS. In other words, the teacher has organized the portfolio by ITS standards. Artifacts are included under each standard that provide

evidence that the standard was met. There is a divider identifying the ITS prior to the evidentiary material.

High School Portfolio: by artifact. This teacher has divided the portfolio into seven artifact sections where the artifact represents one or more of the ITS. For instance, the teacher includes a United States history lesson. The artifact is cross referenced with each of the standards on a cover sheet and then the evidence is placed in the artifact section. (At this point, I opened the portfolio to show the principals the cross reference sheet for clarification).

Once again, please let me remind you that I want you to keep talking. Are you ready to proceed?

APPENDIX G
EXPERT PANEL RUBRIC

EXPERT PANEL RUBRIC

Portfolio Identification Code _____ Evaluator _____

Criteria	Level of Performance		
	Unsatisfactory	Proficient	Exemplar
Artifact Inclusion	Portfolio Does Not Include artifacts relative to all 8 ITSs	All 8 ITSs are represented	All 8 ITSs are represented
Artifact Quality	Included artifacts are not representative of each corresponding ITS	Included artifacts are somewhat representative of each corresponding ITS	Included artifacts are representative of each corresponding ITS
Knowledge of ITS Concepts	Artifacts reflect recall and comprehension	Artifacts reflect analysis and synthesis	Artifacts reflect evaluation and application

Overall Rating (circle one): Unsatisfactory Proficient Exemplar

Note. Language in last row adapted from *Designing Professional Portfolios for Change* (p. 129), by K. Burke, 1997, Arlington Heights, IL: Skylight Professional Development

APPENDIX H

ELEMENTARY TEACHER PORTFOLIO CROSS-REFERENCE INFORMATION

APPENDIX I

HIGH SCHOOL TEACHER PORTFOLIO TABLE OF CONTENTS AND
CROSS-REFERENCE INFORMATION

Table of Contents

Artifact # 1: United States History lesson.

Standards met: 1-d, 2-a, 2-b, 2-c, 2-d, 3-a, 3-c, 3-d, 4-a, 4-e, 4-f, 5-c

Artifact # 2: Mr. Sullivan's web page

Standards met: 1-e, 1-g, 3-b, 3-e, 5-b, 7-c, 8-c, 8-e

Artifact # 3: Communication with parents

Standards met: 1-a, 1-b, 1-e, 1-f, 1-g, 3-b, 5-b, 5-e, 6-c, 8-a, 8-b, 8-c, 8-e

Artifact # 4: American government sampling activity

Standards met: 1-a, 1-c, 3-a, 3-b, 3-c, 3-d, 4-b, 4-c, 4-e, 5-d, 5-e, 7-c

Artifact # 5: Extracurricular/ off contract work

Standards met: 1-b, 1-f, 1-g, 5-f, 7-a, 7-b, 7-c, 7-d, 8-a, 8-b, 8-c, 8-d, 8-e

Artifact # 6: United States history portfolio assessment

Standards met: 2-b, 2-c, 2-d, 3-c, 3-d, 4-b, 4-d, 4-e, 5-a, 5-b, 5-c, 5-d, 5-e

Artifact # 7: Classroom expectations

Standards met: 1-a, 1-d, 1-e, 1-f, 3-b, 5-b, 5-d, 6-a, 6-b, 6-c, 6-d, 6-e, 8-d

Iowa Teaching Standards Cross Reference List

Standard 1: Demonstrates ability to enhance academic performance and support for implementation of the school district's student achievement goals.

Evidencing Artifacts:

- Artifact # 1: United States History lesson.
- Artifact # 2: Mr. Sullivan's web page.
- Artifact # 3: Communication with parents.
- Artifact # 4: American government sampling activity.
- Artifact # 5: Extracurricular/ off contract work
- Artifact # 7: Classroom expectations

Standard 2: Demonstrates competence in content knowledge appropriate to the teaching position.

Evidencing Artifacts:

- Artifact # 1: United States History lesson.
- Artifact # 6: United States history portfolio assessment

Standard 3: Demonstrates competence in planning and preparing for instruction.

Evidencing Artifacts:

- Artifact # 1: United States History lesson.
- Artifact # 2: Mr. Sullivan's web page.
- Artifact # 3: Communication with parents.
- Artifact # 4: American government sampling activity.
- Artifact # 6: United States history portfolio assessment
- Artifact # 7: Classroom expectations

Standard 4: Uses strategies to deliver instruction that meets the multiple learning needs of students.

Evidencing Artifacts:

- Artifact # 1: United States History lesson.
- Artifact # 4: American government sampling activity.
- Artifact # 6: United States history portfolio assessment

Standard 5: Uses a variety of methods to monitor student learning.

Evidencing Artifacts:

- Artifact # 1: United States History lesson.
- Artifact # 2: Mr. Sullivan's web page.
- Artifact # 3: Communication with parents.
- Artifact # 4: American government sampling activity.
- Artifact # 5: Extracurricular/ off contract work
- Artifact # 6: United States history portfolio assessment
- Artifact # 7: Classroom expectations

Standard 6: Demonstrates competence in classroom management.

Evidencing Artifacts:

- Artifact # 3: Communication with parents.
- Artifact # 7: Classroom expectations

Standard 7: Engages in professional growth.

Evidencing Artifacts:

- Artifact # 2: Mr. Sullivan's web page.
- Artifact # 4: American government sampling activity.
- Artifact # 5: Extracurricular/ off contract work

Standard 8: Fulfills professional responsibilities established by the school district.


Evidencing Artifacts:

- Artifact # 2: Mr. Sullivan's web page.
- Artifact # 3: Communication with parents.
- Artifact # 5: Extracurricular/ off contract work
- Artifact # 7: Classroom expectations

APPENDIX J

MIDDLE SCHOOL TEACHER PORTFOLIO ARTIFACT COVER PAGE


Description of Artifact:

Each mid-term and end of quarter I print off student summary reports for each of the students. Each student takes the progress report home and them signed by a parent or guardian and returns it the next day. 

Date Created:

First semester of  school year (on-going)

Alignment of Artifact to Iowa Teaching Standards and Criteria:

1a, 5b

Teacher Reflection on Student Learning, Teaching Performance, and Rationale for Selection:

This helps the parents stay current as to how their child is performing in school. It also keep the student up to date with their grades. It also helps students realize how important each of the scores they earn are in the overall grade. Students are allowed to make-up missing work for half the credit.

APPENDIX K
GUIDED INTERVIEW QUESTIONS

GUIDED INTERVIEW QUESTIONS

1. Given the rating possibilities of unsatisfactory, proficient, or exemplar, what rating would you assign this portfolio? Why?
2. How much bearing would you place on this portfolio when making a licensure decision? Explain.
3. Is this the same value that you place on the portfolios you review from your own second-year teachers? What is the same/different?
4. Was today's process similar to/different than how you typically evaluate your own teachers? How?

APPENDIX L
SAMPLE OF CODED TRANSCRIPT

Well, I do like the initial little diagram. Just looking at it tells me that this person is open to new ideas and new directions and isn't necessarily at this point of observing, focused on one particular thing and that's I think important as you look at where teachers, especially veteran teachers, tend to find themselves in the same mold and they do things the same way year in and year out. As an administrator, I continue to try to push for my staff to try new things and to make their teaching more relevant to the students. The table of contents is basically kind of meaningless as I look at it. It just talks about what the individual feels they met. It doesn't give ...to me it would look better if it would just talk about the artifact. This is what the artifact is and not saying exactly what it has met, because it's probably more my determination whether the standard has been met, versus the instructor.

I What are you seeing there?

S I see somebody who's trying ...it looks like they're trying to prove where they've met the standard. This is unique. I've not seen this done in the other portfolios that I've observed or as people are putting them together. That part of it I guess I'm good with, at least it's where the teacher is seeing that they felt like they've met the teaching standards and through what artifact. I find that there's a lot of apprehension when we're putting together the portfolio because people aren't exactly sure what would meet standard one, standard two, and so on. In education I think there's a great deal of cross over. As a building administrator, I need to tie the portfolio to the job that I'm seeing in the classroom and not just the portfolio for itself. So this is just a working tool to help support what's taking place in the classroom.

Comment [TAL1]: PFS

Comment [TAL2]: JT

Comment [TAL3]: PR

Comment [TAL4]: PFS

Comment [TAL5]: PFS

Comment [TAL6]: PO/PE

APPENDIX M
KWIC WORKSHEET

APPENDIX N
INITIAL CODING INSTRUCTIONS

INITIAL CODING INSTRUCTIONS

1. Please take as long as you like to review the coding categories, definitions, key words, and sample comments.

2. The comments that you will be coding are pre-empted by a number in red ink (1-25). The comment itself is bolded.

3. The yellow highlighted areas represent material not being coded, i.e., researcher instructions or the participant merely reading a portion of the material in the portfolio.

Code each comment with only **one** of the category definitions provided to you. If a comment appears to be representative of more than one category, select the category that best encompasses the **gist of the entire** comment.

4. Let's practice with the first five comments to get a rhythm.

APPENDIX O
CATEGORY DEFINITIONS

CATEGORY DEFINITIONS

1. **Portfolio Structure (PFS)** Comments in this category made reference to how the teacher physically structured and organized his/her portfolio. Typically, the principal referred to specific pages that provided structure within the portfolio they were viewing as well as the overall structure/organization of the portfolio. KWIC used: structure, organize, lay or laid out, figure out.
2. **Process steps (PS)** – Comments in this category refer to how the principal progresses through the portfolio; the steps he/she takes. KWIC used: first, next, second, last, finally, always, and usually.
3. **Principal's role (PR)** – Comments in this category refer to how the principal perceives his/her role in the mentoring and induction of the teacher with emphasis on portfolio preparation. KWIC used: role, job.
4. **Coaching:** Coaching comments refer to ways in which the principal assists the teacher as he/she prepares their portfolio. Comments in this category divided nicely into **four subcategories**.

Coaching/Portfolio Preparation (C/PP): This subcategory is specific to the overall process of preparation of the portfolio. Coaching comments in this category specifically refer to how the principal prepares and leads their own teachers through the “general” portfolio process. KWIC used: coach, instruct, direct, support, mentor, team.

Coaching/Meetings (C/M): In this category the principal expresses that a meeting will help clarify some issues in the portfolio. KWIC used: meet, meeting, discuss, conversation, sit down, conference.

Coaching/Questioning (C/Q): Comments in this category are very distinctive. The principal asks a series of questions that are designed to help the teacher reflect upon and clarify the function of a particular artifact. KWIC used: how, explain, what, tell, question, talk, ask.

Coaching/Suggested Alternatives for Artifacts (C/SA): Typically comments in this category follow a judgment statement about a particular artifact. The principal makes very direct, specific suggestions to the teacher that another artifact or another way of presenting the artifact might be more representative of the standard the teacher is attempting to evidence. KWIC used: Rather than, instead of, I would like to see more of, suggestion.

5. **Tools (T)** – Comments in this category refer to any sort of instrument or equipment a principal might use as he/she evaluates the portfolio. KWIC used: tools, check, checklists, sticky notes, cross-referencing sheets, notes in margin, grease pen.

6. **Judgment:** Comments in this category indicated that the principal was coming to a conclusion (judgment) concerning one of **three entities**:

Judgment/Artifact (J/A): The principal is making a judgment about a specific artifact. KWIC used: #s 1-8, artifact, evidence, support, judge, judgment.

Judgment/Portfolio (J/P): The principal is making a judgment about the portfolio as an entire entity. KWIC used: portfolio, artifacts, overall, in general, cookbook, proficient, satisfactory, strong, and weak.

Judgment/Teacher (J/T): The principal is drawing conclusions about the teacher's [the one who produced the sample portfolio] ability. KWIC used: teacher, judge, ability, strong.

7. **Critical Pieces:** Comments in this category center on those items that a principal considers to be critical when assessing a portfolio and hence, a teacher. **Two subcategories** surfaced in the transcripts.

Critical Pieces/Teacher Reflection (CP/TR): Principals' comments show interest in the amount and level of reflection offered by the teacher in the portfolio. KWIC used: reflection, critical, important, necessary, reflective piece.

Critical Pieces/Role of Observation (CP/RO): Comments in this category represent principals' comments concerning the role of observation in teacher evaluation. KWIC used: observation, see in action, principal's, summary.

8. **Time Investment (TI)** – Comments in this category refer to the time it takes principals to evaluate teacher portfolios in their own building and the one for the current study. KWIC used: time, hour, go through.

9. **Principals' Opinions:** This category includes comments made by the principal **outside** the judgment/evaluation comments about the particular portfolio they were viewing. These comments are more general in nature. **Two subcategories** emerged:

Principals' opinions/portfolio as evidence of good teaching (or not) PO/PE: Comments in this category are clearly comments made by the principal concerning the value of the portfolio in the assessment process. KWIC used: portfolio does not show greatness of teacher, portfolio can conceal deficiencies, fluff, scrapbook, filler.

Principals' opinions/State Department of Education (DE) Teacher Assessment Requirements (PO/DE): Comments in this category reflect opinions of the principals

concerning the DOE requirements for teacher assessment. KWIC used: new process, new standards, evaluator training, state requirements, department of education.

10. **Comfort level (CL):** Comments in this category indicate the principal's comfort or discomfort with evaluating a portfolio other than one produced by an instructor in his/her own building. KWIC: comfortable, uncomfortable, foreign, accustomed, used to, my teacher, discern.

APPENDIX P
REVISED CODING INSTRUCTIONS

REVISED CODING INSTRUCTIONS

1. Please take as long as you like to review the coding categories, definitions, key words, and sample comments.

2. The comments that you will be coding are pre-empted by a number in red ink (1-25). The comment itself is bolded.

3. The yellow highlighted areas represent material not being coded, i.e., researcher instructions or the participant merely reading a portion of the material in the portfolio.

4. Code each comment with only **one** of the category definitions provided to you. If a comment appears to be representative of more than one category, select the category that best encompasses the **gist of the entire** comment. For example, the following comment by Leo is somewhat ambiguous:

This pretty much is simply a pretty cut and dried look at the project. She has a tendency to do that on all of her artifacts, is to lean on or to look pretty strongly at the lesson plan as opposed to looking at it more from a child perspective.

Leo addresses *a* project but he focuses on *all* the artifacts and the spirit of the *entire* portfolio as he sees it. This comment would represent **portfolio judgment**.

5. Let's practice with the first five comments to get a rhythm.

APPENDIX Q

SAMPLE OF CATEGORY HOLDING PAGE

SAMPLE OF CATEGORY HOLDING PAGE

Category: Coaching/Suggested Alternatives (C/SA)

BMDES

Now I'm looking at a writing rubric that she's going to be using. Okay. I think I would like to have a little description of how you use this rubric. It's a little bit different than your typical rubric as far as a scale. There's a total of 20 points. We have all the things you can do under each of the areas that she's checking for, in writing. Evidently I'm gonna need a little explanation of the writing rubric. Something that just kind of tells me how it's used.

She might look at more 'let's kind of review some things. (198-199)

The only thing I'd ask that she'd do is like what is her goal for this project. I'm thinking maybe it is part of the previous one that she was working on and extending on that. (238-240).

Evidently, this was a very good day. One of her better assessments probably because 97%. A good grade work on that. It might be helpful if she would just kind of indicate on these the type of student. Like if we have a top student, we have a student that's resource, and seeing ...okay (262-265) So I'm going to have to ask her how Tyler's written project and rubric is good and I might give some examples. We have some examples of rubrics. I like to see rubrics that don't just give numbers, but are more or a 'what can I do if I wanted to' ...I don't know why he got 'excellent' is my question right here. I'm sure that that can be explained to me, but I'm not seeing it as I'm following with this one. If this is submitted and it has different criteria, I'd like to just see something covering it saying that this is the requirement and a different tool for evaluation on this one. Again, I have another one Native Americans. The student has the same assignment, they got 40 out of 40. No feedback on the writing. I kind of like to see comments along the side. If she thought it was good, a little more specific. You want the student to continue to do terrific things that they're doing, so sometimes just 'the comment that you made about this really gave me a good picture of what was happening' – those kind of things that help them know what they're doing right, is good. 'You know your stuff' is good too. How might you get this student to grow in their writing? I would like to see that in there. Get them into the habit of doing that. I would like her to do that. (277-291)

APPENDIX R

IVAN'S ITS/CRITERIA CROSS-REFERENCE TOOL

IOWA TEACHING STANDARDS & CRITERIA (indexed by Davidson and related words and concepts) v2.0

1. STANDARD: Demonstrates ability to enhance academic performance and support for implementation of the school district's student achievement goals.

The Teacher:

- ~~XX~~a. Provides evidence of student learning to students, families and staff. (1g, 5b, 5e, 8e)
- ~~X~~b. Implements strategies supporting student, building, and district goals. (3b, 4d, 4f, 5f, 8e, 8c, 4b, 7d)
- ~~X~~c. Uses student performance data as a guide for decision making. (3a, 5a, 5c, 5d, 4a)
- ~~XX~~d. Accepts and demonstrates responsibility for creating a classroom culture that supports the learning of every student. (3b, 3d, 5d, 6c, 6e)
- ~~XX~~e. Creates an environment of mutual respect, rapport, and fairness. (6e, 6a, 8d, 1e, 1f, 3b, 4d)
- ~~X~~f. Participates in and contributes to a school culture that focuses on improved student learning. (5f, 7b, 8d, 1d, 1e, 3b, 4d, 6a, 6c)
- ~~XX~~g. Communicates with students, families, colleagues, and communities effectively and accurately. (1a, 3a, 5b, 5c, 8c)

2. STANDARD: Demonstrates competence in content knowledge appropriate to the teaching position.

The Teacher:

- ~~XX~~a. Understands and uses key concepts, underlying themes, relationships, and different perspectives related to the content area. (2c, 7c)
- ~~XX~~b. Uses knowledge of student development to make learning experiences in the content area meaningful and accessible for every student. (3c, 4e, 4c)
- ~~X~~c. Relates ideas and information within and across content areas. (2a, 2d, 7c)
- ~~XX~~d. Understands and uses instructional strategies that are appropriate to the content area. (2a, 2c, 3d, 4b, 7c, 1b,)

3. STANDARD: Demonstrates competence in planning and preparing for instruction.

The Teacher:

- ~~XX~~a. Uses student achievement data, local standards, and the district curriculum in planning for instruction. (1c, 4b, 5a, 5c, 5d, 7a, 4a)
- ~~X~~b. Sets and communicates high expectations for social, behavioral, and academic success of all students. (1b, 1d, 3d, 5b, 5d, 6b, 6c, 6e, 1e, 1f, 6a)
- ~~X~~c. Uses student's developmental needs, backgrounds, and interest in planning for instruction. (2b, 4e, 4c)
- ~~X~~d. Selects strategies to engage all students in learning. (1d, 2d, 3b, 4b, 4d, 5d, 6c, 6e, 4c)
- ~~XX~~e. Uses available resources including technologies, in the development and sequencing of instruction. (4f, 8c)

4. STANDARD: Uses strategies to deliver instruction that meets the multiple learning needs of students.

The Teacher:

- ~~X~~a. Aligns classroom instruction with local standards and district curriculum. (1d, 4c, 1c, 3a, 5c, 8c)
- ~~X~~b. Uses research-based instructional strategies that address the full range of cognitive levels. (1b, 2d, 3a, 3d, 5a)
- ~~XX~~c. Demonstrates flexibility and responsiveness in adjusting instruction to meet student needs. (4a, 5c, 2b, 3c, 3d, 4c)
- ~~XX~~d. Engages students in varied experiences that meet diverse needs, and promote social, emotional, and academic growth. (1b, 3d, 4f, 1d, 1e, 1f, 3b, 6a, 6c)

- ~~X~~e. Connects students' prior knowledge, life experiences, and interest in the instructional process. (2a, 2b, 3d, 3c)
- ~~XX~~f. Uses available resources, including technologies, in the delivery of instruction. (1b, 3e, 4d)

5. STANDARD: Uses a variety of methods to monitor student learning.

The Teacher:

- ~~XX~~a. Aligns classroom assessment with instruction. (1e, 3a, 4b, 5b, 5c, 5d)
- ~~XX~~b. Communicates assessment criteria and standards to all students and parents. (1a, 1g, 3b, 5a, 5d, 5c, 8e)
- ~~XX~~c. Understands and uses the results of multiple assessments to guide planning and instruction. (1c, 3a, 4c, 5a, 5d, 7a, 4a)
- ~~X~~d. Guides students in goal setting and assessing their own learning. (1c, 1d, 2b, 2d, 3a, 5a, 5b, 5c, 6c, 6e)
- ~~X~~e. Provides substantive, timely and constructive feedback to students and parents. (1g, 1a, 5b, 8c)
- ~~XX~~f. Works with other staff and building and district leadership in analysis of student progress. (1b, 1f, 7b, 8c, 8d, 8e)

6. STANDARD: Demonstrates competence in classroom management.

The Teacher:

- ~~XX~~a. Creates a learning community that encourages positive social interaction, active engagement, and self-regulation for every student. (1e, 6c, 8d, 1d, 1f, 3b, 4d, 6c)
- ~~XX~~b. Establishes, communicates, models, and maintains standards of responsible student behavior. (3b, 8d)
- ~~XX~~c. Develops and implements classroom procedures and routines that support high expectations for student learning. (1d, 3b, 3d, 5d, 6d, 6e, 6a)
- ~~X~~d. Uses instructional time effectively to maximize student achievement. (6c)
- ~~X~~e. Creates a safe and purposeful learning environment. (1d, 1e, 3b, 3d, 5d, 6a, 6c, 8d, 1f, 4d, 6a)

7. STANDARD: Engages in professional growth.

The Teacher:

- ~~XX~~a. Demonstrates habits and skills of continuous inquiry and learning. (3a, 5c, 7b, 7c, 7d, 8c)
- ~~XX~~b. Works collaboratively to improve professional practice and student learning. (1f, 5f, 7a, 8c, 8d, 8e)
- ~~X~~c. Applies research, knowledge, and skills from professional development opportunities to improve practice. (2a, 2c, 2d, 7a, 7d)
- ~~X~~d. Establishes and implements professional development plans based upon the teacher's needs aligned to the Iowa teaching standards and district/building student achievement goals. (7a, 7c, 1b, 8c)

8. STANDARD: Fulfills professional responsibilities established by the school district.

The Teacher:

- ~~XX~~a. Adheres to board policies, district procedures, and contractual obligations.
- ~~XXXX~~b. Demonstrates professional and ethical conduct as defined by state law and district policy.
- ~~X~~c. Contributes to efforts to achieve district and building goals. (1b, 5f, 7a, 7b, 8c)
- ~~XX~~d. Demonstrates an understanding of and respect for all learners and staff. (1e, 1f, 5f, 6a, 6c, 7b, 8c, 1b, 4a, 7d)
- ~~XXXX~~e. Collaborates with students, families, colleagues, and communities to enhance student learning. (1a, 1b, 3e, 5f, 7f, 8c, 8d, 1g, 5c, 5f)

APPENDIX S

MIKE'S MODEL DESCRIPTORS SAMPLE

Artifact Identification Guide

IOWA TEACHING STANDARDS, CRITERIA, DESCRIPTORS AND SAMPLE DATA POINTS

Standard 1

Demonstrates ability to enhance academic performance and support for implementation of the school district's student achievement goals.

1 a. Provides evidence of student learning to students, family and staff.

Teacher documents student learning with meaningful measures using data that is understandable and shares individual and classroom goals and results with students, families and staff members.

The teacher uses multiple artifacts, including achievement trends for local standards and benchmarks, to document to document and provide evidence of student learning to students, families and staff members.

The teacher plans parent teacher conferences so his/her teacher and learning objectives have the greatest likelihood of serving the student's best interests.

Sample Data Points

- Calling log to parents
- E-mail/written communication to parent
- Grade updates/progress reports achieved benchmarks/report cards
- Copy of student progress report
- Classroom observations
- Log of staff or parent contacts
- Student achievement test data reports
- Assessment results shared with other staff
- Newsletters
- Student/ teacher conference to discuss progress/Parent-teacher conference log

Data Source

Teacher
Teacher

Teacher
Student
Administrator
Teacher
Teacher
Teacher/Administrator

Teacher/Parent

1 b. Implements strategies to support student, building and district goals.

The teacher knows the short and long-range building and district goals for student learning and implements classroom instructional strategies that clearly align with these established goals.

The teacher effectively communicates these goals and accomplishments to various constituents including students, parents, and colleagues.

Sample Data Points

- Lesson plans incorporate instructional strategies and assessments that address content benchmarks
- Shares standards/benchmarks for content area and shares progress on the benchmarks with colleagues
- Shares standards/benchmarks for content area and shares progress on the benchmarks with students
- Course syllabus/outline
- Classroom observation
- Classroom rules/assignment posters

Data Source

Teacher

Teacher

Teacher/Student
Teacher
Administrator
Teacher

APPENDIX T

MIDDLE SCHOOL TEACHER'S ARTIFACT: FOUR-QUADRANT VOCABULARY
FOR ITS 4B

Vocabulary Quadrant

Word name the property	Synonym / Descriptors give an example
Sentence say it in words	Illustration draw a picture

APPENDIX U

GAVIN'S EVALUATION GUIDE: ITS 2

Evaluation Guide to the State of Iowa Teaching Standards

Standard 1 Standard 2 Standard 3 Standard 4 Standard 5 Standard 6 Standard 7 Standard 8

Standard 2

Demonstrates competence in content knowledge appropriate to the teaching position.

Criteria: <i>The Teacher</i>	Teacher Behavior	Optional Documentation
a. Understands and uses key concepts, underlying themes, relationships, and different perspectives related to the content area.	<ul style="list-style-type: none"> • Weekly objectives follow district objectives. • Demonstrate different points of view • Uses district curriculum 	<ul style="list-style-type: none"> • Copies of lesson plan • Student work/portfolio
b. Uses knowledge of student development to make learning experiences in the content area meaningful and accessible for every student	<ul style="list-style-type: none"> • Uses adaptations to make all students successful • Makes curriculum accommodations 	<ul style="list-style-type: none"> • Copies of hand-outs • Video used, CD's • Notes on lesson plans • Log of GWAEA involvement
c. Relates ideas and information within and across content areas.	<ul style="list-style-type: none"> • Interdisciplinary units/ works with other teachers including specials 	<ul style="list-style-type: none"> • Pictures of students • Pictures of students and work • Lesson plans • Notes planning meetings
d. Understands and uses instructional strategies that are appropriate to the content area	<ul style="list-style-type: none"> • IEPlan lessons to meet district objectives • Teach to the students' needs • Use of multiple intelligences considered 	<ul style="list-style-type: none"> • Assessments, copies of study guides • Lesson Plans, differentiated student work