

# Florida State University Libraries

---

Electronic Theses, Treatises and Dissertations

The Graduate School

---

2005

## The Effect of Goal Setting, Self-Evaluation and Selfreflection on Student Art Performance in Selected 4th and 5th Grade Visual Art Classes

Marcia Schilling Meale



THE FLORIDA STATE UNIVERSITY  
COLLEGE OF VISUAL ARTS AND DANCE

THE EFFECT OF GOAL SETTING, SELF-EVALUATION AND SELF-  
REFLECTION ON STUDENT ART PERFORMANCE IN SELECTED 4<sup>th</sup> AND  
5<sup>th</sup> GRADE VISUAL ART CLASSES

By

MARCIA SCHILLING MEALE

A Dissertation submitted to the  
Department of Art Education.  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

Degree Awarded:  
Spring Semester, 2005

The members of the Committee approve the Dissertation of Marcia Schilling Meale defended on March 14, 2005.

Charles M. Dorn  
Professor Directing Dissertation

Gary W. Peterson  
Outside Committee Member

Tom Anderson  
Committee Member

Marcia L. Rosal  
Committee Member

Approved:

Marcia L. Rosal, Chair, Department of Art Education

Sallie McRorie, Dean, College of Visual Arts and Dance

The Office of Graduate Studies has verified and approved the above named committee members.

## ACKNOWLEDGEMENTS

Professors Charles Dorn, Marcia Rosal, Tom Anderson and Gary Peterson sustained this inquiry with their insight and encouragement. The task would have been far more difficult without their support and Charles Dorn's expertise in portfolio assessment.

The art teachers in Leon County, Florida are to be commended for their dedication to art education, their willingness to undertake this study in addition to their teaching responsibilities and their commitment to continuous improvement.

Leon County Schools, provided support for this research project. The Florida Division of Cultural Affairs, Department of State, contributed financial support. Salih Binici and Roumen Vesselinov provided statistical expertise.

## TABLE OF CONTENTS

List of Tables .....	vii
List of Figures .....	ix
Abstract .....	x
1. Introduction .....	1
Statement of the Problem .....	5
Research and Study Hypotheses .....	5
Procedures .....	6
Rational .....	6
Scope and Limitations .....	9
Research Design .....	11
Organization of the Chapters .....	15
Definition of Terms .....	16
2. Review of the Literature .....	18
Philosophy of Mind .....	18
Views of Cognition .....	22
Theories of Metacognitive/Self-Reflective Learning .....	37
Standards Movement .....	49
Beliefs and Attitudes Related to Assessment .....	52
Assessment .....	53
Standardized Tests .....	55
Alternative Assessment .....	58
Digital Portfolios .....	64
Primary Source Articles .....	65
3. Methods .....	68
Statement of the Problem.....	68
Hypotheses .....	68
Research Design .....	69
Description of the Sample .....	69
Description of the Instruments Used .....	71
Explanation of the Procedures .....	73
Reliability .....	76

Score Spread .....	77
Validity .....	77
External Validity .....	78
Variables to be Tested .....	78
4. Reporting the Data .....	82
Research Hypothesis and H1 .....	82
Descriptive Statistics Inclusive of All Five Schools .....	82
Inter-Rater Reliability .....	92
Validity .....	93
Inferential Statistics Inclusive of All Five Schools .....	94
Descriptive Statistics for Each School .....	97
Inferential Statistics for Each School .....	99
H2 – H4, Treatment of Extraneous Variables .....	102
Descriptive Statistics .....	102
Inferential Statistics .....	103
Additional Findings .....	104
Journal Scores .....	104
Additional Assessments .....	105
5. Analysis of the Data .....	106
Research Hypothesis and H1 .....	106
H2 – H4, Extraneous Variables .....	107
Equality of Groups .....	109
Inter-Rater Reliability .....	109
Score Spread .....	110
Validity .....	111
6. Results, Implications, Recommendations, Conclusions .....	112
Results .....	112
Recommendations for Further Study .....	118
Implications for the Field .....	122
Conclusion .....	123
APPENDICES .....	126
A Holistic Rubric for Standards-Based Assessment in Visual Arts .....	126
B Journal Questions .....	128
C Holistic Journal Rubric .....	132

D Human Subjects Documents .....	134
E Leon County Documents .....	138
REFERENCES .....	143
BIOGRAPHICAL SKETCH .....	150

## LIST OF TABLES

Table 1: Number of Students by School and Group .....	82
Table 2: Descriptive Statistics of Extraneous Variables .....	83
Table 3: Raw Data for School A .....	85
Table 4: Raw Data for School B .....	86
Table 5: Raw Data for School C .....	88
Table 6: Raw Data for School D .....	89
Table 7: Raw Data for School E .....	90
Table 8: Inter-Rater Reliability Coefficients .....	93
Table 9: Percent of Agreement .....	93
Table 10: Pretest & Posttest Mean Differences for All Schools .....	95
Table 11: Pretest & Posttest Paired T-Test .....	95
Table 12: Mean Posttest Comparison for All Schools .....	96
Table 13: Independent Samples T-Test for All Schools .....	97
Table 14: Descriptive Statistics Mean Comparison by School.....	97
Table 15: Descriptive Statistics Mean Posttest Comparison by School .....	98
Table 16: Posttest Independent Samples T-Test by School .....	100
Table 17: ANOVA Comparison, Pretest and Posttest Scores by School .....	101
Table 18: Effect Size by School .....	102



Table 19: Descriptive Statistics for the Extraneous by School .....	103
Table 20: Pearson r Correlations of Study Variables .....	103
Table 21: Experimental Groups' Journal Scores by School .....	105

## LIST OF FIGURES

Figure 1: Pretest & Posttest Score Distributions, Control Group .....	91
Figure 2: Pretest & Posttest Score Distributions, Experimental Group .....	91
Figure 3: Box Plot Distribution of Scores for All Schools .....	94
Figure 4: Box Plot Comparison, Control and Experimental for All Schools .....	96
Figure 5: Quartile and Median Posttest Distributions by School .....	99

## ABSTRACT

A self-regulatory intervention of metacognitive/self-reflective journal writing, specifically addressing goal setting, self-reflection and self-evaluation, was combined with an established digital portfolio assessment methodology established by Dorn, Sabol, Madeja, (2003) to measure student art performance in visual art classes. Two groups of students in seven schools were studied in this quasi-experimental research project, involving digitization of the students' artworks and an intervention of metacognitive/self-reflective journal writing. Beginning of the term portfolios and end of the term portfolios were collected and later scored using holistic rubrics. Statistical analyses were applied in order to judge and compare the two portfolios and the posttest results of the control group compared to the experimental group. The first portfolio, pretest, was a collection of the first three to four 2-D artworks produced by each student. The second portfolio, posttest, was a collection of the last three or four 2-D artworks produced near the end of the year or term. The metacognitive/self-reflective writing journals were collected and scored. The journal scores were then analyzed and used to determine the level of treatment implementation. The results suggest that metacognitive/self-reflective journal writing based on self-regulated learning (SRL) and metacognitive learning strategies may increase student art performance.

# CHAPTER 1

## INTRODUCTION

In the late 1980s and early 1990s proponents of the national standards movement predicted that within the next decade “national standards describing essential outcomes in various school subjects could become the glue holding together many pieces of the education system - curriculum frameworks and guides, textbook adoption, staff development, and perhaps most significantly, assessment - at the local, state and national level” (O’Neil, 1993, p. 4). Costa (1993) predicted the establishment of world-class standards. These higher standards would “shift learning *of* the content to learning *from* the content” (p. 50). Currently many states, including Florida, have taken an accountability through standardized testing approach which stresses assessment *of* content. The designers of the standards movement in contrast to the current climate of high stakes testing, predicted that the boundaries between subjects would dissolve, self-evaluation would be favored over external evaluation, content acquisition and knowledge retention would no longer be an end in itself, and uniformity would be dismissed and diversity embraced (Costa, 1993). Assessment in itself is not sufficient to bring about the results intended by legislators. Emphasis on what is being learned and how it is being learned must also be considered. In the case of the arts, the development of both the standards and the assessments on the national level were intertwined (Hausman, 1994).

In Florida, there is not currently a standardized state art assessment. Yet, increasingly, assessment is becoming an important aspect of K-12 school art programs. Several national accountability acts including a *Nation at Risk* (1983), the *National Assessment of Educational Progress* (NAEP) of 1997, the *America 2000: an Education Strategy* (1991), and the *Goals 2000 Educate America Act* (Public Law 103-227,1994) have recommended that schools provide evidence of accountability through assessment (Sabol, 1998).

Historically, however, assessment in visual art classes has not been emphasized and only occasionally included in school assessment programs. Traditionally, the school art experience has focused on the use of art as a tool for self-expression rather than art as cognitive learning (Gruber & Hobbs, 2002). In the 1960s, Manuel Barkan proposed the need for concepts of

structure and discipline in art education, which later, in the 1980s, became the basis for discipline based art education (DBAE) supported by The Getty Foundation (Gruber & Hobbs, 2002). As a result, the emphasis in assessment continues to remain primarily on the theoretical or on methodology rather than on the art product as a product of instruction (Eisner, 1996). In 1996 Broughton, Eisner, and Ligtovet edited a series of articles and rebuttals on assessment and evaluation's purposes and practices from an international perspective. Their assessment and evaluation focused on assessment in and evaluation of program goals not on methods for measuring student growth. Gardner (1996) in a contributing article recognized that in the United States, there has been little assessment beyond that undertaken casually and often informally by teachers.

At the 2004 National Art Education Convention, when the word assessment was mentioned, a buzz of "what are you doing?" quickly filled the room. Brandt (1987) stated, "Even people who are strong advocates of arts in schools are very much in conflict about assessment" (p. 31). The question, he noted, is no longer shall we assess but rather, how shall we assess. The states and districts are in fact demanding that assessments for accountability purposes be developed and implemented. Sabol's (1998) content analysis of state assessments and results of the 1997 NAEP (National Assessment of Educational Progress) leaves no doubt that current assessments efforts are inadequate. Art supervisors and art educators therefore are currently scrambling to find ways to measure performance outcomes in the visual arts that takes into consideration the unique character of learning in the arts.

Part of the assessment problem resides in how the arts are viewed in general education. Efland (2002) identified three problems. The first is the view that the arts are "modes of entertainment, frivolous occupations, and elective options - 'nice' cultural experiences to have if time and resources permit, but not major contributors to the mind or personality formation" (Efland 2002, p. 7). The second is the "lack of awareness of the substantive roles the arts can play in overall cognitive development" (Efland 2002, p. 7). The third problem addressed was the role of cognition in the arts. Efland (2002) stated, "educators are unsure of how to use the arts to develop cognitive abilities in children or of the means for assessing such attainments" (p. 7).

Further complicating the viability of the arts as cognitive, substantive, and assessable are historical beliefs about cognition in general and its role in education. Dating back to Plato and Socrates the belief prevailed that all matters of genuine knowledge were already implanted in the

human soul at birth. The task of instruction was to bring this innate knowledge to conscious awareness (Gardner, 1985). Until the early 1800's "people learned in real-life, on-the-job situations. At that time our industrial society required people to develop no more than a range of functional skills (such as reading, writing and calculation) that allowed them to fit into the dull routines of manufacturing" (Abbot, 1997, Create intelligence section, ¶ 1). In the 1920s and 1930s a "standardized test paradigm gained momentum....The establishment of the Educational Testing Service (ETS) in 1947 solidified the supremacy of tests as the ultimate assessment tool" (Cole, Ryan, Kick, Mathies, 2000). In the 1950s, psychological behaviorism dominated the educational scene and its influence on schooling persists (Efland, 2002). The behaviorists believed learning in all subjects was the same. They also believed that certain subjects, like art, were not cognitive in nature but, rather affective. Behaviorists, Clark L. Hull and B.F. Skinner "excluded all reference to mental processes and aspired to reduce all forms of learning to conditioning" (Efland, 2002, p. 14).

The objectivists and behaviorists today still have a firm hold on testing and assessment (Eisner, 1996). Standardized testing is primarily based on the assumption that the desired outcomes of educational activities can be known in advance, that there are correct answers, that knowledge can be segmented, and that knowledge is using or recalling products produced by others (Eisner, 1996). Sabol's (1998) content analysis of state art assessments indicated that only 1% of the test items are product-oriented while 99% reflect traditional paper and pencil testing methods. Cannatella (2001) expressed concern over this standardized approach to assessment in art education, "...the particular character and activity that goes into the making of art does not fit comfortably into any system of general assessment criteria. One major reason for this is that individuality of the art work and its making cannot be reduced into common properties independent of and detached from what is being creatively fashioned" (p. 319).

Today cognitive theories of learning have moved from a conditioning or recording mind to a mind that "builds mental constructions that order experience. The brain represents rather than records reality.... Like an artist, the brain selects, discounts most signal, and seeks constancies to make up our images of the world" (Spence, 2002, Power of learning section, ¶ 5). Current constructivist views on cognition have for the most part dissolved the distinction between cognitive and noncognitive subjects. Instead, all subjects have intellectual and affective components (Efland, 2002). Alternative forms of assessment support this constructivist pedagogy

including assessment in the arts. In the arts learning outcomes value surprise, idiosyncratic responses, works as whole and the work produced out of personal experiences (Eisner, 1996) features not in accordance with standardized testing.

Recently work done by Dorn, Sabol, Madeja (2003) suggested that authentic assessment techniques, specifically portfolio assessment, can be used to demonstrate growth in visual art classes as a means of satisfying accountability demands. Their work was based on the fact that most elementary and secondary art classes spend the majority of their time making art. They also found the primary assessment strategy used by art teachers included work samples. The holistic portfolio scoring methodology established by Dorn, Madeja, and Sabol (2003) resulted in a valid and reliable method for determining student progress or growth based on the actual art products students produce in their visual art classes. The process involves two portfolios being collected and scored using holistic rubrics. The first portfolio consists of the first three or four works of the term, while the second portfolio consists of the last three or four works of the term. The portfolio scores were then compared in order to determine growth. This same study found that there were certain ages at which students made little or no progress.

This study was based on the premise that art is a cognitive and metacognitive activity and as such, student art performance can be enhanced. Eisner (2002) described art as a cognitive endeavor that included a wide range of cognitive activities such as problem solving, refining the senses, enlarging the imagination and representing an idea or image in a material. He identified four cognitive functions of representation, inscription, editing, communication and discovery, all of which are important strategies used in visual art classes. Efland (2002) referred to Gardner's multiple intelligences theory and Eisner's "mind that develops multiple forms of representation through experience gained through the senses" (p.157) as examples of how the visual arts contribute uniquely to cognition as a whole. Both theories, Efland (2002) claimed, were built on the notion of nonredundancy, "the idea that the arts provide unique opportunities for the development of the mind that are not available in other modalities" (p. 157). Efland (2002) developed four arguments the arts provide that enhance cognitive development based on an integrated view of cognition that avoids the dualism of the past, the cognitive flexibility argument, the integration of knowledge argument, the imagination argument and the aesthetic argument.

Efland referred to Koroscik's work in general education as a possible model for examining learners' inadequacies, such as, inadequacies in the learner's knowledge base, weak or inappropriate knowledge-seeking strategies, or a poor disposition to learn. These difficulties fall into the realm of metacognitive theories of learning. This study was based on applying metacognitive strategies found in learning theories such as, Self-Regulated Learning (SRL) and the Dimensions of Learning (DOL) model, specifically goal setting, self-evaluation and self-reflection strategies. Self-regulated learning may provide a new cognitive strategy for improving students' art performance. Art teachers may find that paying attention to students' cognitive and metacognitive abilities may be more productive than the study of teaching methods and/or content emphasis, shifting their focus to learning *from* the content not learning *of* the content.

### **Statement of the Problem**

In light of current concerns focusing on the relationship of cognition and learning in art and on accountability, this research project will focus on the effect inclusion of cognitive and metacognitive/self-regulatory activities, specifically, student goal setting, self-evaluation and self-reflection, have on students' progress in two-dimensional visual art production. The digital portfolio assessment techniques and the holistic scoring methodology initiated by Dorn, Madeja, and Sabol (2003) will be used to determine the effectiveness of the intervention. These methods were designed specifically to support the expressive, "ill-structured" (Efland, 2002) nature of art education. In addition, research by Dorn, Madeja, and Sabol (2003) suggested that students showed little or no gain in art performance in grades 4 and 5. Their study took into account maturation and a teacher intervention focused on instruction in more advanced studio skills, a means of improving content. To see if this lack of gain on the part of the students can be counteracted, students in grades 4-5 are the primary focus of this study.

### **Research Hypothesis**

The hypothesis states that setting annual goals and answering self-reflection and self-evaluation questions after each art assignment will contribute positively to fourth and fifth grade student performance in visual art classes as measured by an authentic portfolio assessment process established by Dorn, Sabol, and Madeja (2003).



## **Hypotheses**

- H1 The alternative hypothesis states that students' mean art portfolio scores will increase as a result of the intervention, metacognitive/self-reflective journal writing.
- H2 The null hypothesis states that the extraneous variable of student FCAT-NRT reading scores will not affect student art performance.
- H3 The null hypothesis states that the extraneous variable of gender will not affect student art performance
- H4 The null hypothesis states that the extraneous variables of age, free/reduced lunch status, instructional time and budget will not affect student art performance.

## **Procedures**

The necessary steps and procedures to answer the research questions are to:

1. Establish a theoretical foundation to ground and frame the study through critically reviewing the salient literature.
2. Select and train the art teachers in the methods of the study.
3. Implement the strategies for the study in the field.
4. Synthesize, interpret and evaluate the findings, and draw implications for further research and practice.

## **Rationale**

Demonstrating student achievement and student growth in visual arts education has been mandated at the local, state and national levels. Assessment has been viewed as a tool for transforming instruction and learning, of revitalizing the education system (Resnick, 1994). National efforts to design assessment and set high standards in the visual arts at the national level were conducted simultaneously (Hausman, 1994). Later, the No Child Left Behind Act of 2001 required that not only certain levels of student performance be achieved but also that “adequately yearly progress” be demonstrated. As a result, testing and demonstrating student growth has become a high stakes endeavor. Annual teacher evaluations and teacher pay are dependent on demonstrating gains in student performance (Eisner, 2002).

Unlike classroom teachers, art teachers in addition to determining the instructional strategies, must also determine the content of the program and the criteria by which student performances will be assessed. The state and national standards and the district guidelines for education in the visual arts are broad and can be met in numerous ways. How the standards are met is usually left up to the individual art teacher. Efland (2002) categorized art as an “ill-structured” domain, which is based on case studies, versus a “well structured” domain like the physical sciences, which are organized around general or overarching principles. He noted also that, “Well structured representations of knowledge often are experienced as being easier to teach and learn, and thus textbooks tend to be biased in their favor. In effect, this misrepresents the inherent complexity (ill-structuredness) of learning in many domains, including the arts and humanities” (p. 11). Adding to the problem of art as an “ill-structured” domain, Bensure (2002) made four claims about the art standards and assessment, these are: 1) the standards are ambiguous in nature, 2) are difficult to define what all students should know and learn, 3) are for the teacher to determine what is essential for all students to learn, and that 4) one or two lessons could cover all the standards.

Art in most elementary and secondary school programs has focused primarily on the making of artworks or on the examination of art exemplars. Little attention is being paid to art as a cognitive and metacognitive activity. Current strategies have too often focused on procedures, processes and the supporting study of the elements and principles of design (Dorn, Madeja, Sabol, 2003). As a result, art as a cognitive process has not been considered as important as art making. Yet, when the National Standards and the 1997 NAEP assessment and standards were being developed, one of the expected learning outcomes was a shift from an emphasis on *information* to an emphasis on *meaning*, from *processing information* to the *construction of meaning* (Hausman, 1994).

One approach to introducing the study of the cognitive aspects of art making can be found in adding metacognitive instructional strategies to the art curriculum, especially those metacognitive strategies based on self-regulated learning theory (SRL) and the dimensions of learning (DOL) model. Both SRL and DOL take into consideration elements that effect students as learners beyond the acquisition of knowledge, such as, attitudes, work habits, and goals. In this study, goal setting, self-evaluation and self-reflective journal writing were the means for developing the students’ cognitive processes to enhance student performance in visual art

classes. Eisner (2002) defined cognition as “all those processes through which the organism becomes aware of the environment or its own consciousness” (p. 9). He identified three cognitive functions the arts perform, 1) to help one learn to notice the world, 2) to inscribe images and ideas in a material that stabilizes what would otherwise be evanescent, 3) as a means of exploring one’s interior landscape.

Self-regulated learning theory (Zimmerman, 2001) also attends to these cognitive functions. In self-regulated learning four processes are involved, 1) forethought and goal setting, 2)self-observation/self-monitoring, 3) self-judgments/self-regulation and 4)self-reaction/self-satisfaction. Self-regulated learning occurs when students 1) are metacognitively, motivationally, and behaviorally active in their own learning process, 2) generate thought, feelings, and actions in order to attain their learning goals, 3) are aware that self-regulations enhances their achievement, 4) they experience covert changes in self-perception, 5) they replace one learning strategy with another, and 6) self-instructional statements reinforce their behaviors.

In this study, metacognitive and self-regulatory learning strategies in the form of journal questions focus students’ attention on their cognitive processes and on what and how they are learning. In this case, through written responses to questions, assessment becomes a method for transforming instruction and learning. Currently, many students who attend school are unengaged. School is something that happens to them not something in which they engage actively. They do not, or perhaps have not been given opportunities to determine outcomes, develop goals, or set criteria related to their learning. Through goal setting the teacher can personalize instruction. The learning experience is no longer just a set of objectives established by the teacher, the school, the district, or the state. The learning experience is also established for the benefit and interest of the student. All too often gifted and talented students get by without using their time wisely or pushing their limits and still earn A’s. By incorporating self-regulated learning strategies, specific goals can be established for specific students. For example, for some students using time wisely and completing work may be the issue. For another student, not working too fast may be a concern. Planning multiple approaches to a problem may be beneficial for another student. All of these are exemplars of metacognitive strategies. By focusing on the students’ goals and needs in addition to the objectives, by taking into consideration metacognitive/self-regulatory strategies, the teacher shifts the responsibility for learning from the teacher to the student. Without the student’s involvement in the learning process, there are only

isolated bits of content reiterated for meaningless tests, at best the learning *of* content, not the learning *from* content. In the arts, the way knowledge is gained is a part of the knowledge itself and cannot be separated from what we find out (Hausman, 1994). Reflecting on the process and self-evaluating both the process and the product transform learning toward *constructing meaning* and away from simply reporting *information*.

In this study, students will set annual goals, answer self-assessment questions, and write reflections concerning their art making and thinking. The written responses will be kept in paper journals. In addition, teachers who implement the writing intervention fully will assist students in using self-regulated processes by setting criteria, reviewing the criteria for the lesson before the students' self-evaluate, question students as they write and assist in setting appropriate goals. The journals will be a record of these interactions.

### **Scope and Limitations**

This research study was designed to examine the role of metacognitive/self-regulated learning in relationship to students' art performance. The study focused specifically on goal setting, self-evaluation and self-reflection in the development of fourth and fifth grade students' progress or "growth" in visual art class.

### **Population**

This quasi-experimental research project was originally initiated in 12 schools in Tallahassee, Florida. The population was a volunteer convenience sample. The schools who participated reflect Tallahassee's racial and social-economic make-up and therefore lend credence to the generalizability of the study's findings. Throughout this study, processes and procedures were refined and noted in the teachers' logs.

Tallahassee, Florida, is a medium sized city located in the panhandle of Florida. Leon County does not have an art supervisor. Art teachers from the district volunteered to participate in this study. The primary motivation for their involvement was to meet accountability issues for either their principals or the district pay for performance bonus incentive program. For either purpose, teachers needed to demonstrate student growth overtime with x percent of the students meeting or exceeding the established criteria. The growth criteria stated that students would maintain a portfolio 3 or 4 score or show improvement in their portfolio 1 to portfolio 2 scores.

The percentage was 80% at the district level for the pay for performance bonus and established individually at the schools as part of the yearly teacher evaluation process.

### **Population Selection**

Initially, a convenience sample of volunteer art teachers from nine elementary schools, one middle school and two high schools (in Leon County, Florida) selected two classes of the same grade level and same course content to participate in this study. Seven schools, one middle and six elementary schools completed the study. Only five elementary schools collected enough data to undergo statistical analysis and were reported in this study. The middle school, and one elementary school were unable to collect the required 3 or 4 works of art per portfolio from the majority of their students. The final sample population included in the study, therefore, consisted of a convenience sample of volunteer art teachers from five elementary schools. They each choose two intact classes of either fourth or fifth grade students. One school chose two fourth grade classes. The other five schools chose two fifth grade classes. The visual art classes created primarily two-dimensional works of art, i.e. paintings, prints, drawings. English Speakers of Other Languages (ESOL) and Exception Student Education (ESE) students in inclusion classes were included in the study. To be included in the study, the students needed to be present from the start of the school year to the end, August to May.

To be considered for statistical analysis the schools needed to complete the study requirements and the majority of their school's portfolios had to include a minimum of three artworks for the pretest and three artworks for the posttest. It was recommended, by the researcher, that all artworks be digitally recorded at the completion of the lesson, finished or not. It was also recommended that sketches, drawings, paintings, prints, fibers, ceramics and preliminary drawings be digitally recorded and entered in the individual student portfolios and dropped before scoring if there was excess work. This was an attempt to ensure the required number of artworks were included in the portfolios for the scoring process. Five elementary schools met the study requirements.

The five elementary schools represented diverse populations and resources. Florida Comprehensive Achievement Test – Norm Referenced Test (FCAT-NRT) mean reading scores for the study group ranged from a low of 56.5 percent at one school to a high of 80.82 percent at another. Students receiving free/reduced lunch ranged from zero to 64 percent. Supply and instructional materials budgets ranged from a low of \$1 per student per year to a high of \$5.15

per student per year. Hours of art instruction ranged from a low of 12.5 hours per year to a high of 30.75 hours per year. Ethnicity varied as well from mostly minority students to mostly majority students. Students in the study were between the ages of 9.48 and 11 years of age and in the 4<sup>th</sup> or 5<sup>th</sup> grade. In chapters, 3 and 4 the individual schools' demographics are described.

### **Limitations**

First, with the exception of the art teacher/researcher, this study was the rest of the participating art teachers' initial training in both the digital portfolio process and the writing intervention methodology. Second, it was recognized that these teachers have varying levels of computer skills, access to computers, digital cameras and other technology and that training in the use of the hardware and software would be needed. Third, it was agreed that the participating teachers would implement the journal writing in a variety of ways. Four of the five art teachers felt "all they could do" this year was learn the digital portfolio process and let the students answer the journal questions. The art teachers kept notes and recorded changes they made to the journal questions. Most felt the information collected from the students' writing would provide a baseline of information regarding student self-evaluation and self-reflection. As a result, only one, the art teacher/researcher implemented the use of metacognitive/self-regulated learning theory with her students. Fourth, there were no curriculum requirements beyond those required by the state, district and school. As a result, the lessons and units of study varied considerably from school to school. Fifth, several of the art teachers wanted and did administer additional assessments to their experimental group. The nature of these assessments was not always consistent with best practices based on self-regulated learning theory. Several of the teachers felt that accountability issues at the parental and school administrative levels required them to assess students in the manner in which they were accustomed.

### **Research Design**

A quasi-experimental design was used for this study. The design of the study consisted of an experimental group and a control group. The experimental group underwent an intervention of metacognitive/self-reflective journal writing while the control group did not participate in any writing or assessment activities. The design was a static-group comparison design,  $X_1 0, X_2 0$ . Each art teacher included in the study, chose two intact classes of students from the same grade level, either two fourth grade classes or two fifth grade classes. One class acted as the control

group, only digitally collecting their artworks. The other class served as the experimental group, who in addition to collecting their artworks digitally, also set annual goals, answered self-reflection, and self-evaluation questions saved in paper journals. The treatment for the experimental group began with the completion of the first artwork of the year. At that time, the students filled out the first set of journal questions. The sample was a convenience sample. Information on the following extraneous variables was collected: FCAT (Florida Comprehensive Achievement Test) reading scores, grade, sex, free-reduced lunch status, ethnicity, hours in art class, and art budget to determine their impact if any on the study.

### **Collection of Data**

The dependent variable was student art performance gains. The independent variable was self-evaluation/reflection. Portfolios included at minimum, the students' first three artworks and the last three of the term. The term was defined as a full year for elementary schools, August to May.

### **Description of the Portfolios**

The art teachers each taught the units and lessons they designed based on state and school standards, and personal choice. As a result, each school's work was unique to that school. The teachers taught the same lessons to both their control and experimental groups. The art teachers and university students collected the students' artworks in both the control and experimental classes. The artworks were then photographed digitally and entered into individual student electronic PowerPoint portfolios. A template was provided which contained the format, a black background and ten slides. A minimum of the first three artworks of the year and last three artworks were collected with a dividing slides indicating portfolio 1 (pretest portfolio) and portfolio 2 (posttest portfolio). These digital portfolios were stored either on the art teacher's computer desktop or on the school server. Backup storage devices and CD's were also created. Before the scoring session, the art teachers created two CD's for each group, experimental and control, for a total of four CDs; portfolio 1 - control group; portfolio 2 - control group; portfolio 1 - experimental group; portfolio 2 - experimental group. Each CD contained all the portfolio 1 and portfolio 2 images of the students' artworks in the two groups. These portfolios contained identification codes for both the schools and students with dividing slides between students. For example, slide 1 contained the first student's identification code; slide 2 was an image of the

student's first artwork; slide 3 was the student's second artwork; slide 4 was the students' third artwork; slide 5 was the next student's identification slide, etc.

### **Treatment of the Data**

District art teachers blind scored the portfolios. The art teachers did not score their own students' portfolios. They used the 3-5 holistic rubric for standards-based assessment in the visual arts developed by Dorn, Madeja, and Sabol (Appendix A). The experimental groups' journals were scored using another holistic rubric (journal rubric, appendix C) to determine the students' ability to accurately and in depth self-evaluate and self-reflect, an indicator of the level of implementation of the journal writing intervention and students' use of self-regulated learning strategies.

### **Analysis of the Data**

To test inter-rater reliability two tests were conducted: 1) Hoyt's method (Alpha) and 2) the Intra-Class Correlation Coefficient (ICC). In addition, percentage of judge agreement was determined. Independent and Paired Samples T-tests were conducted to determine differences between groups and the significance of those differences at  $p < .05$ . The Pearson  $r$  was conducted to check for any effect the extraneous variables may have had on the results of the study. A significant difference in growth from the beginning of the term to the end of the term was expected in the experimental group. If the goal setting, self-evaluation and self-reflective journal writings positively influenced the development of art performance then the students creating the self-evaluation/reflection portfolios would show the most growth.

### **Description of Journal Questions**

The journal questions were developed based on the literature related to metacognition, self-regulated learning (SRL), Dimensions of Learning (DOL) and Project Zero (Brandt, 1987). Only the experimental group participated in the journal writing (appendix B). The students also rated their effort level after each assignment on a scale of one to five, one being low, five being high. The first question in the journal related to goal setting: "Name two things you want to learn to do better in art class this year? What skills will you need to learn?" Following the completion of each artwork four sets of questions were answered by each student: 1) "To what extent did you meet the requirements of this assignment?" 2) "What sections of this artwork or skills would you like to improve? How? Why?" 3) "What did you do to make this artwork better?" 4) "How has this project helped improve your art making or thinking? What did you learn? How will this



lesson help you make better works of art?” After completion of the last artwork of the year, a final set of questions related to the attainment of the students’ goals was asked: “Did you improve the two things you wanted to do better? Did your art making and thinking improve this year? Did your skills improve? How?”

Journal entry forms were provided. The teachers photocopied these forms and created paper journals for each student in the experimental group. There was some discussion over the wording of the questions and the questions were revised during the first training meeting with the teachers. One teacher, school D, changed the questions by adding a qualifying line after each question. Because this study took place in schools with varying needs, some of the teachers felt they needed to give assessments in addition to the journal writing. The primary reason for giving these additional assessments was for accountability purposes to parents and principals. The art teachers did not give any writing activities or assessments to the control group; they only made works of art.

### **Implementation Procedures**

First, a notice was sent to each art teacher in the district briefly describing the project and asking for those interested to respond. Second, a training and information session was scheduled. At that meeting, teachers were introduced to the project, its goals, and procedures. The journal questions’ language was also revised. The revised journal pages and questions were sent by email to each of the participating art teachers to be photocopied and made into individual student journals. Third, after the art teachers the initial training session, the art teachers selected two classes in the same grade level, 4<sup>th</sup> or 5<sup>th</sup> to participate in the study. One class was the control group. The other class was the experimental group.

Fourth, the art teachers collected all the artworks from the two classes throughout the year. With the help of students from the Florida State University art teacher education program the students first four artworks for the term (pretest) were photographed and entered into individual student digital PowerPoint portfolios. Later the participating art teachers photographed and entered the last four artworks (posttest) into the students’ digital portfolios. The artworks for students in both the experimental and the control groups were entered into individual student digital portfolios. Fifth, these individual portfolios were later combined into four scoring portfolios for each school, a pretest portfolio and a posttest portfolio for all the students in the control group, and a pretest portfolio and a posttest portfolio for all the students in

the experimental group. Sixth, three separate art teacher raters then blind scored these portfolios, with each individual student's pretest and posttest portfolios each receiving three separate scores.

Throughout the year, there were support sessions for the teachers. At the conclusion of the scoring session, the raw data was entered into excel spreadsheets and submitted for statistical analyses.

### **Teacher Training**

The art teachers in the study met almost monthly throughout the project for over 15 hours of training, scoring, and support sessions with the researcher. The training sessions covered: 1) the study process and procedures, 2) technology support, 3) suggestions for assisting students with the journal writing process, and 3) scoring the portfolios. The main aim of the training sessions was to support the teachers in implementing the study. The primary focus, in response to the teachers concerns and problems, revolved around creating the digital portfolios.

### **Scoring the Portfolios**

Prior to scoring the portfolios the art teachers created a pretest portfolio for each group, experimental and control which included each students first three or four artworks and a posttest portfolio for each group which included each students last three or four artworks of the term. Six district art teachers were trained in the scoring methodology and use of the holistic rubric. Each of the raters were art teachers but not all of them participated in the study. Each portfolio was then blind scored independently by three different art teacher raters, using the holistic rubrics for standards-based assessment in the visual arts (Appendix A) developed by Dorn, Madeja, and Sabol (2003) on a scale of 1 to 4, (4 being a high and 1 being a low). No raters scored their own students works. The scores were collected for later analysis.

## **Organization of the Chapters**

Chapter 1 presents an overview of the study, the rationale, aims of the study, population, research design, methods, limitations, and definition of terms. Chapter 2 contains a review of the pertinent literature including the philosophy of the mind, art as a cognitive activity, theories of metacognition, theories of self-regulated learning, the effect of high stakes testing and assessment of art learning. Chapter 3 covers the procedures for the study including the population, the training of the teachers, the variables to be tested, the instruments used, how the data will be collected and reported, the statistical analysis of the data and results of the study.

Chapter 4 reports the data and analyzes the data. Chapter 5 reports the results, makes recommendations for further study and implications for the field.

### **Definition of Terms**

Analytic rubric: breaks the portfolio into parts and rates each part on a scale. (Weidmer, 1998).

Authentic assessment: “An assessment is authentic when it involves students in tasks that are worthwhile, significant and meaningful. Such assessments will appear as learning activities, involve conceptual and higher order activities, and involve conceptual and higher order thinking skills, and make connections among several different forms of knowing” (Dorn, 2002, ¶ 15).

Cognition: “all those processes through which the organism becomes aware of the environment or its own consciousness” (Eisner, 2002, p. 9).

Digital portfolio: “A digital or electronic portfolio is a purposeful collection of work, captured by electronic means, that serves as an exhibit of individual efforts, progress and achievements in one or more areas” (Wiedmer, 1998, ¶ 3).

Holistic rubric: considers the whole portfolio and rates the overriding skills without doing a small scale analysis. (Weidmer, 1998).

Increased growth: for this project will be determined by finding a measurable improvement in the students’ electronic portfolio score, beginning of the term work compared to end of term work.

Metacognition: “the ability to reflect upon, understand, and control one’s learning” (Schraw & Dennison, 1994, p. 460).

Performance assessment: “in the arts involves testing what we generally do in the process of teaching art in schools, which is to make things and evaluate them in process” (Dorn, 2002, P17).

Posttest : throughout this study is used interchangeably with portfolio 2, referring to the last three or four artworks of the term.

Pretest : throughout this study is used interchangeably with portfolio 1 referring to the first three to four artworks collected of the term.

Primary trait rubric: assesses the demonstrated performance in one or more major areas of emphasis. (Weidmer, 1998).

Process Portfolio: “A Process Portfolio is the entire collection of work done on an assignment leading up to and including, the final solution to the project....it may include, but not be limited to writing, concept maps, thumbnail sketches, rough sketches computer generated rough and the final solution to a given problem” (Walker, 1998, p. 2).

Product: refers to the aesthetic quality of the student’s finished result or work of art (Clark, 2002).

Rubric: “A rubric...is a chart or matrix that describes varying levels of competency or success” (McCollister, 2002, P 2). The rubric may serve as a developmental guide or scoring device. They articulate qualities in the finished work or within a process. Rubrics are sets of criteria that describe levels of performance or understanding (Goldsby, 2001; Huffman, 1998).

Self-evaluation: when students determine and record the worth their work using scoring instruments.

Self-regulated learning (SRL): “refers to learning that results from students’ self-generated thoughts and behaviors that are systematically oriented toward the attainment of their learning goals” (Schunk, 2001).

Student reflections: will be journal type entries recording the students’ thoughts about their artwork, art making process and/or the effectiveness of their artwork.

Visual art classes: include courses where students study and make drawings, paintings, and print artworks.

Web-based portfolio: is a digital portfolio that incorporates web-based materials into teaching and learning. This purposeful collection of work, captured by electronic means, is an exhibit of individual efforts, progress and achievements in one or more areas that combines aspect of traditional and electronic or digital portfolios Goldsby (2001).

## CHAPTER II

### REVIEW OF THE LITERATURE

This review of literature examines: 1) the philosophy of mind, 2) art as a cognitive activity, 3) theories related to metacognition, 4) theories related to self-regulated learning, and 5) beliefs, attitudes and assessment strategies.

#### Philosophy of Mind

Dating back to Plato and Socrates the belief prevailed that all matters of genuine knowledge were already implanted in the human soul at birth. The task of instruction was to bring this innate knowledge to conscious awareness (Gardner, 1985). However, the development of the concept of mind as a separate entity dates back to Descartes.

“Rationalism, specifically, the ideas of Rene Descartes (1596-1650), is considered by some cognitive psychologists as the prototypical philosophical antecedent of cognitive science. In particular Descartes’ concept of his own mind laid the conceptual groundwork for a cognitive science, where the mind stands apart from and operates independently of the human body rejecting, as it were, the behaviorist notion of a rational mind and a mechanical body” (Dorn, 1999, p. 53).

This dualistic concept can be found in the work of Kant (Nelson, 1999; Copleston, 1994), Hegel (Baillie, 2003), Efland (2002), Eisner (2002) and Gardner (1985, 1999) whereas Piaget (Bruner, 1997), Dewey (Prawart, 2002) and Vygotsky (Prawart, 2002) rejected dualism. Kant divided human knowledge into two types, sensitive knowledge and intellectual knowledge or understanding. The distinction is understood in terms of objects, “the objects of sensitive knowledge being sensible things, *sensibilia*, capable of affecting the sensibility (*sensualitas*) of the subject, which is the latter’s [intellectual knowledge] receptivity or capacity for being affected by the presence of an object so as to produce a representation of it” (Copelston, 1994, p.197). Intellectual knowledge or rational knowledge was knowledge of objects that does not affect the senses, what Kant called *intelligibilia*. Copelston (1994) stated, “Sensitive knowledge is knowledge of objects as they *appear*... whereas intellectual knowledge is knowledge of things

*as they are (sicuti sunt)*. The empirical sciences come under the heading of sensitive knowledge, while metaphysics is the prime example of intellectual knowledge” (p. 199).

For most modern cognitive scientists:

“The troublesome part of Kantian logic for them is not so much with Kant’s duality (i.e., in being accepting of mind as a priori or that an external world shapes the mind [as seen above]) but rather with Kant’s notion of self, which is both a subject that has unity through awareness of its own existence and as an object in the process of discovery....This provides a conception of self that is both aware of what it is and what it is becoming, which is at the same time both the subject and the object (Dorn, 1999, p. 55).

The problem lies in that objects and events in the external world are not separated from the self that undergoes them in experience (Dorn, 1999).

Hegel, like Kant, thinks that “one’s capacity to be ‘conscious’ of some external object as something *distinct* from oneself requires the reflexivity of ‘self-consciousness’, that is, it requires one’s awareness of *oneself* as a subject *for whom* something distinct, the object, is presented *as known*” (Nelson, 1999; Stanford). Hegel goes beyond Kant, however, “in making this requirement dependent on one’s recognition (or acknowledgement --*Anerkennung*) as a subject by other self-consciousness whom one recognises in turn. ...It [self-consciousness] comes about only indirectly via one’s recognising other conscious subjects’ *recognition* of oneself!” (Stanford, *Phenomenology of spirit* section, ¶ 6). In the words of Hegel:

“For self-consciousness, then, otherness is a fact, it does not exist as a distinct moment; but the unity of itself with this difference is also a fact for self-consciousness occupies the position of consciousness, and the whole expanse of the world of sense is conserved as its object, but at the same time only as related to the second moment, the unity of self-consciousness with itself....Consciousness has, *qua* self-consciousness henceforth a twofold object -- the one immediate, the object of sense-certainty and of perception, which however, is here found to be marked by the character of negation; the second, viz. itself, which is the true essence, and is found in the first instance only in the opposition of the first object to it. Self-consciousness presents itself here as the process in which this opposition is removed, and oneness or identity with itself established” (Baillie, *The Truth which conscious* section ¶ 3).

Efland (2002) wrote, “By becoming mindful of our perceptual activity, we have made perception itself into an *object of thought*,...*In becoming an object for thought, perception itself has become a concept in our thinking*. Thinking arises in our consciousness and we are no longer perceiving. *We are thinking*” (p. 17).

Eisner claimed:

“Consciousness is the product of attention, and attention is guided by past experience and moderated by current need or purpose. Consciousness is also a form of awareness, and awareness is fed initially by sensibility. Thus, sensibility is the mother of consciousness and provides the content for reflection, analysis and the making of connections” (Eisner, 2002 p. 108).

Consciousness then according to Kant, Hegel, Efland and Eisner involved the recognition of representation and an awareness of self.

Kant established a philosophical foundation for the concept of experience that was later applied by Dewey to education. Experience was both active and reflective. For Kant, “The union of form and matter precedes all reflection. That is to say, because the human subject is what it is it necessarily perceives sensible objects in space and time. The act of distinguishing between form and matter is the work of philosophical reflection” (Copleston, 1994, p. 198). To clarify this concept Copleston (1994) stated:

“There is spatial and temporal co-ordination. We then have ‘appearances’. The mind then, through what Kant calls the logical use of the intellect, organizes the data of sense intuition, while leaving their fundamentally sensuous character intact. We then have the phenomenal world of ‘experience’. ‘From appearance to experience there is no way except by reflection according to the logical use of the intellect.’ In its logical use or function the mind simply organizes the data of sense intuition; and we then have the empirical concepts of experience” (p. 198).

To understand what is meant by matter and form in Kant’s writing Copleston (1994) explained, “The matter is what is given, namely sensations, that which is produced by the presence of sensible objects. The form is that which co-ordinates the matter; it is contributed, as it were, by the knowing subject and is the condition of sensitive knowledge. There are two such conditions space and time.... These ‘singular concepts’ [space and time] are described as ‘pure intuitions’ ...Their function is simply to co-ordinate the sensations which are received and thus to make sensitive knowledge possible”(Copleston, 1994, p. 197).

According to Kant space and time are not objective or real but they are *pure intuition* and *subjective*; subjective meaning empty of all empirical content. Given refers to the synthesis of *a priori* sense intuitions of space and time and sense-experience itself, form and matter. “Understanding then further synthesizes the data under its own pure (non-empirical) concepts or categories. Sensibility and understanding, therefore, co-operate in constituting experience and in determining objects as objects ...” (Copleston, 1994, p. 229).

In the case of the natural sciences and mathematics:

“where sense intuition supplies the data or material and where the intellect is employed only according to its logical use (that is, logically comparing and organizing the data but not supply concepts and axioms from its own inner nature), ‘use provides the method’. That is to say it is only after these sciences have already acquired a certain degree of development that we reflect on and analyse the method employed, considering how the method can be improved in detail” (Copelston, 1994, p. 201).

The mind of man for Kant then “is not passive wax upon which experience and sensation write their absolute and yet whimsical will; nor is it a mere abstract name for the series or group of mental states; it is an organ which transforms the chaotic multiplicity of experience into the ordered unity of thought” (Durant, 1961, p. 202)

Like Kant’s form that co-ordinates matter Dewey’s experience required a relationship between doing and undergoing, “The action and the consequence must be joined in perception. This relationship is what gives meaning: to grasp it is the objective of all intelligence. The scope and content of the relations measure the significant content of experience”(Dewey, 1934/58, p. 44). However, it was not enough just to have a relationship. Doing or trying had to involve a conscious connection. “Experience as trying [doing] involves change, but change is meaningless transition unless it is consciously connected with the return wave of consequences which flow from it. When an activity is continued into the undergoing of consequences, when the change made in us, the mere flux is loaded with significance. We learn something” (Dewey 1916, Nature of Experience section, ¶ 1). Thinking occurs. Knowledge Dewey believed was grounded in action, where each participant is a creator or user of knowledge (Fenstermacher & Sanger, 1998, The student’s nature section, ¶ 3). “Thinking, in other words, is the intentional endeavor to discover specific connections between something which we do and the consequences which result, so that the two become continuous” (Dewey, 1916, Reflection in experience section, ¶ 2). To “learn from experience” then is a circular or the backward and forward making of connections between what we do to things and what we enjoy or suffer from things in consequence (Dewey, 1916). Dewey stated, “An experience has pattern and structure, because it is not just doing and undergoing in alternation, but consists of them in relationship....The action and the consequence must be joined in perception. This relationship is what gives meaning: to grasp it is the objective of all intelligence. The scope and content of the relations measure the significant content of experience”(Dewey, 1934/58, p. 44).



Experience according to Dewey could be limited. He stated, “Experience is limited by all the causes which interfere with perception of the relations between undergoing and doing. There may be interference because of excess on the side of doing or of excess on the side of receptivity, or undergoing. Unbalance on either side blurs the perception of relations and leaves the experience partial and distorted, with scant or false meaning” (Dewey, 1934/58, p. 44). Dewey understood that there could be an impediment to one’s learning.

## **Views of Cognition**

### **Learning Theories**

The view of art as a “nice” cultural experience was founded in the objectivist and behaviorist beliefs about cognition and the role education has played historically (Eisner, 2002). Dating back to Plato and Socrates the belief prevailed that all matters of genuine knowledge were already implanted in the human soul at birth. The task of instruction was to bring this innate knowledge to conscious awareness (Gardner, 1985). Until the early 1800s, “people learned in real-life, on-the-job situations. Then our industrial society required people to develop no more than a range of functional skills (such as reading, writing and calculation) that allowed them to fit into the dull routines of manufacturing” (Abbot, 1997, How do we create intelligence section ¶ 1). In the 1950s psychological behaviorism dominated the educational scene and its influence on school still persists (Efland, 2002). The behaviorist believed learning in all cognitive subjects was the same. They also believed that certain subjects like art were not cognitive in nature. It was affective. Clark L. Hull and B.F. Skinner, “explicitly excluded all reference to mental processes and aspired to reduce all forms of learning to conditioning” (Efland, 2002, p. 14). Piaget’s work broke the “thrall of old-line associationisms and learning theories that dated back to Aristotle, and that had been regularly renewed in more recent times by empirical philosophers from Hobbes and Locke onward” (Bruner, 1997, p. 66). “Such theories were too subject to the contingencies of encounter to satisfy Piaget’s need; all failed to deal with the inherent *systematicity* of mental growth by putting the systematicity in the world rather than in the growth of the mind itself” (Bruner, 1997 p. 67).

The findings of cognitive scientist moved theories of learning from a conditioning or a recording mind to a mind that “builds mental constructions that order experience. The brain represents rather than records reality....Like an artist, the brain selects, discounts most signals,

and seeks constancies to make up our images of the world.” (Spence, 2002, The power of learning section, ¶ 5) Efland (2002) stated, “Current views on cognition have largely dissolved the distinction between cognitive and noncognitive subjects” (p. 10). Instead, all subjects have cognitive and affective components. Gardner (1985) defined cognitive science “as a contemporary, empirically based effort to answer long-standing epistemological questions - particularly those concerned with the nature of knowledge, its components, its sources, its development and its deployment”(p. 6).

Bruner’s five models of the learner, *tabula rasa*, hypothesis generator, nativism, constructivism, novice-to-expert reflect some of the historic views of learning. The theory of learning that best supports the view that learning in art is a cognitive activity is built on constructivist theories. The work of Dewey, Piaget, Bruner, Vygotsky, Gardner and Eisner helped inform the emerging view of art in a cognitive light.

Dewey’s notion of the student differed from many of his contemporaries by, 1) understanding the “individual mind as a function of social life,” 2) understanding the child from the “perspective of emotion and endeavor as well as knowledge and intellect,” 3) understanding that the mind is not a static entity but grows and is “characterized by distinctive phases of capacity and interest” (Simpson, 2001). These understandings lead to what has been termed progressive education. Students construct through experience and action rather than act as vessels of knowledge.

The early work of Dewey hung everything on action. “The purpose of action, including reflex action was to ensure that an organism maintained a viable relationship with the environment” (Prawat, 2002, p.18). Vygotsky favored action, which promotes self-regulation in individuals (Prawat, 2002). Prawat (2002) found similarities between Dewey’s and Vygotsky’s early work on the role of action as the mediator between organisms and environment and their latter work on the origin and role of concepts in human meaning making. “They both rejected nominalism in favor of a transactional approach that views meaning making as something that goes on in the world and not just in the head” (Prawat, 2002, p.19).

Vygotsky also established a new psychological theory that did not separate individuals from their socio-cultural setting. “His approach can be characterized by three themes: (a) the best way to understand mind is to look at how it changes; (b) higher mental functions have their

origins in social activity; and (c) higher mental functions are mediated by tools and signs.” (Hausfather, 1996, Vygotsky and education section ¶ 4).

Mind for Piaget can be described as an organized group of logical operations that mediate between the world and our knowledge of the world. Since the world cannot be known directly it is then a construction to be further tested against ongoing action in the world. “For Piaget, knowledge is made, not found.” (Bruner, 1997, p.66).

While for Vygotsky, as for Piaget, “mind mediates between the external world and individual experience, Vygotsky never conceived of mind as expressing a logical calculus. Mind, rather, comprised process for endowing experience with meaning. (Bruner, 1997, p. 68).

For Bruner:

“...the term *mind* means the *process* within which a person actively constructs knowledge by relating incoming information to a previously acquired psychological frame of reference or model of reality. The use of amplifiers of mind - skills, images, and conceptions, - requires a commonly shared human capacity, but each society fashions and perfects this capacity to its own needs. (Bigge & Shermis, 1999, p. 142).

Both Vygotsky and Bruner addressed the second theme, all higher mental functions originate in social activity. Vygotsky’s view required not only language but also a grasp of the cultural context in which language is used. Mental development consisted in mastering higher order, culturally embodied symbolic structures, each of which may have incorporated or even displaced what existed before (Bruner, 1997). Mental life first expressed itself in interaction with others. “The results of such interaction then became internalized and enter the stream of thought. Since social interaction is principally constituted and mediated by speech, what gets internalized into the child’s stream of thought are meanings and forms generated in verbal exchange in which themselves are products of the broader cultural-historical system (Bruner, 1997, p. 68). For Bruner, there “can never be the case that there is a ‘self’ independent of one’s cultural-historical existence” (Bigge & Shermis, 1999, p.137).

The third theme in Vygotsky’s approach was “the mediation of higher mental functions by tools and signs” (Hausfather, 1996, Vygotsky and education section ¶ 4). His system of signs restructured completely the psychological process, regulating attention and creating new forms of culturally-based psychological processes. According to Vygotsky cultural artifacts, “tools and signs create who we are and how we view the world, while we recreate and transform the cultural artifacts we have inherited”(Hausfather, 1996, Vygotsky and education section ¶ 4). For

Vygotsky the meaning of every word is a generalization or concept making the connection between the word and its meaning a “matter of ideological cognitive structure as contrasted with a simple association” (Bigge & Shermis, 1999, p. 128).

Vygotsky’s concern was to grasp how somebody interprets or understands something which:

“requires that we take into account their cultural and linguistic background and the context in which they find themselves both ‘in the small’, in the sense of a particular communicative situation, and ‘in the large’ of a patterned cultural system. Vygotsky’s emphasis, accordingly, was on situated meanings and on situated meaning-making, which inevitably generates a cultural-historical approach.” (Bruner, 1997, p. 72).

Bruner also referred to tools, his “cross-cultural studies of cognition lead him to say that intelligence is to a great extent the internalization of ‘tools’ provided by a given culture” (Bigge & Shermis, 1999, p.134).

The role of education for Vygotsky was to provide children with experiences that are within their respective Zones of Proximal Development (ZPD). These activities would challenge the child but with the help of sensitive adult guidance could be accomplished by the child. It is within this zone that the possibility for cognitive change existed (Hausfather, 1996). Scaffolding was the concept used to describe the process of the child actively constructing oneself and one’s environment. “The social environment is the necessary scaffolding and framing that permits a child to move forward and continue to build new competencies” (Bigge & Shermis, 1999, p. 130).

While Vygotsky reasoned that within human development, there were two main strands that intersect, individual and cultural development (Bigge & Shermis, 1999) Piaget reasoned, “that an organisms’ intelligence was embodied in a series of structures with latent tendencies for development, which could be brought out by appropriate interaction with the environment. But, again, the organism was not a passive reflector but rather possessed active potentialities which could unfold to a greater or lesser extent, depending upon the nature of the interaction with the environment”(Gardner, 1972, p. 59). Piaget considered mental growth as a process of the child moving from simpler to more complex systems of logical operations, “the process being effected by the transformation and internalization of action into thought” (Bruner, 1997, p. 66).

Piaget deduced three principal stages through which normal individuals in Western culture pass; the sensorimotor stage, the concrete operations stage and the formal operations

stage. What was important for Piaget was that the concrete-operational stage must always succeed the preoperational stage. "...what is not possible is for the child to skip a stage, or to vary the normal order of progression through successive stages" (Gardner, 1972, p.64).

In addition, three broad trends emerge. The first was the decline of egocentrism; not that the child is selfish, but rather that the child was incapable of differentiating himself from the world (Gardner, 1972). The second trend entailed the child gradually developing internalization of thought. And finally, "a growing child increasingly relies upon various kinds of symbols- words, pictures, mathematical or artistic concepts" (Gardner, 1972, p. 64).

Three other concepts were important to Piaget's theories, assimilation, accommodation and equilibrium. Accommodation is "a basic process of adaptation in which the child alters a behavioral pattern or scheme in accordance with the conditions he finds in the outer world" (Gardner, 1972, p. 75). Assimilation was the aspect of adaptation in "which an increased number of objects or events are subsumed under, and trigger, the exercise of a given behavioral scheme..." (Gardner, 1972, p. 75). By about 18 months, stage six of the sensorimotor stage, the child can carry out actions mentally rather than physically. Assimilation and accommodation could occur on a mental plane. (Gardner, 1972). Equilibrium, Piaget conceived as "a mechanism of change and continuity, a state of balance between competing actions; a system is in equilibrium when a perturbation which modifies the state of the system has its counterpart in a spontaneous action which compensates for its effect" (Gardner, 1972, p.104).

In summary, Piaget:

"has undermined 'common sense' notions of the child as either a passive reactor to the environment, a mere imitator, or one in whom 'innate ideas' will automatically unfold. He has replaced them with a more comprehensive and intricate concept of the child as an active constructor, one who acts upon the world and, in so doing, comes to increase his knowledge of the world as well as of his own thought and person. Piaget has made an impressive attempt to replace the behaviorist emphasis on the primacy of the environment and the Gestaltist emphasis upon the primacy of developed structures with a less elementary and more enriched picture of intelligence. Intelligence was a product of the interaction between constantly evolving structures in the child's mind and ever-varying aspect of reality, which the child becomes able to assimilate or accommodate. The world is not just 'out there', waiting to impress itself on a blank slate; it is a product of our actions upon it, of the relation between these actions, of the symbolic embodiments of those actions" (Gardner, 1972, p. 106).

According to Bruner (1997) "Piaget's genius was to recognize the fundamental role of logic-like operations in human mental activity. Vygotsky's was to recognize that the individual

human intellectual power depended upon our capacity to appropriate human culture and history as tools of mind” (p. 65). Piaget’s strove to “explain” while Vygotsky sought to interpret the human growth and the human condition (Bruner, 1997). Both, however, “dedicated their lives to the study of how human beings grow to construct and exchange theories about the world and each other (Bruner, 1997, p. 65). They both put forth theories that are essentially developmental in nature.

Bruner had two central unifying themes, 1) that the acquisition of knowledge was an active process and 2) that one actively constructs one’s knowledge by relating incoming information to a previously acquired frame of reference (Bigge & Shermis, 1999). Bruner identified three simultaneous processes for learning: 1) acquisition of new information, 2) transformation of knowledge, 3) checking the pertinence and adequacy of knowledge (Bigge & Shermis, 1999). He also identified three modes of representation, enactive, ionic and symbolic representation. First, the “*enactive* mode of representation was highly manipulative in character. It consisted of knowing some aspect of reality without the use of imagery or word.” (Bigge & Shermis, 1999, p. 139). In the enactive mode one represented past events through making appropriated motor responses. Second, the “*ionic* mode of representation is based upon internal imagery. Knowledge, in the ionic mode, is represented by a set of summary images or graphics that *stand* for a concept but do not fully *define* it” (Bigge & Shermis, 1999, p. 139). These representations depended upon visual or other sensory organization along with the use of summarizing images that are representative of greater ‘chunks’ of the environment.” (Bigge & Shermis, 1999, p. 139). Third, the “*symbolic representation*, [was] based upon an abstract, arbitrary, and more flexible system of thought”. (Bigge & Shermis, 1999, p. 140). Symbolic representation was the principal tool of reflective thinking. “The symbolic mode of representation is evidenced by a person’s ability to consider propositions rather than objects, to give concepts a hierarchical structure, and to consider alternative possibilities in a ‘combinatorial’ fashion” (Bigge & Shermis, 1999, p. 140). These modes appeared in order and depended on the previous one for its development but unlike Piaget, they were not linked to age, for Bruner, “some environments can slow the sequence down or bring it to a halt, others move it along faster” (Bigge & Shermis, 1999, p. 139).

Like Vygotsky, Bruner believed language and culture play an important role in the child’s learning. For Bruner *how* one *talks* eventually becomes how one *represents* what one

talked about (Bigge & Shermis, 1999). In addition, Bruner believed one grew up internalizing the ways of acting, imagining, and symbolizing that existed in the person's culture (Bigge & Shermis, 1999).

The task of schools, then, was to convert knowledge into structures that were within the grasp of learners of various ages and to arrange the structures in an optimal sequence of materials to be learned. Bruner believed we should not think in terms of *an* optimal sequence for presenting *a* body of knowledge but, rather, we should recognize that optimal sequences of learning cannot be specified independently of the criteria by which final learning is to be judged. For example, the criteria may include speed of learning, resistance to forgetting, transferability, and effective power of what has been learned in terms of its ability to generate new hypotheses and combinations (Bigge & Shermis, 1999).

Howard Gardner (1999) also defined intelligence within cultural terms. His revised definition of intelligence stated:

*“biopsychological potential to process information that can be activated in a cultural setting to solve problems or create products that are of value in a culture. This modest change in wording is important because it suggests that intelligences are not things that can be seen or counted. Instead, they are potentials -presumably, neural ones- that will or will not be activated, depending upon the values of a particular culture, the opportunities available in that culture, and the personal decisions made by individuals and/or families, schoolteachers, and others.”* (Gardner, 1999, p. 33.)

Gardner originally identified seven different kinds of minds, seven different ways of knowing the world, through language, logical-mathematical analysis, spatial representation, bodily-kinesthetic, interpersonal and intrapersonal intelligences. Recently he has added naturalistic and existential intelligences. Gardner (1991) believed that since students possessed different kinds of minds they therefore learned, remembered, performed and understood in different ways. Gardner's theory of multiple intelligences linked various regions of the brain to various intelligences (Efland, 2002).

Gardner (1991) also identified three kinds of learners, the intuitive learner, the traditional student, and the disciplinary expert. The intuitive learner also called the natural, naive, or universal learner was “the young child who is superbly equipped to learn language and other symbolic systems and who evolves serviceable theories of the physical world and of the world of other people during the opening years of life....”(p. 6). The traditional student or scholastic

learner was “the youngster from age seven to age twenty, roughly, who seeks to master the literacy’s, concepts and disciplinary forms of the school....” (p. 7). The disciplinary expert or skilled person was an individual of any age who has mastered the concepts and skills of a discipline or domain and can apply such knowledge appropriately in new situations....” ( p. 7).

While the intuitive learner’s understandings are often immature, misleading or fundamentally misconceived, the traditional student’s understandings are often rote, ritualistic or conventional in performance. Whereas disciplinary understandings “occur when students are able to take information and skills they have learned in school or other settings and apply them flexibly and appropriately in a new and at least somewhat unanticipated situation...Disciplinary understanding is always changing and never complete...” (Gardner, 1991, p. 9).

### **Brain Theory**

As cognitive scientists studied the brain and learning, another theory developed based on their findings. Brain theory took into consideration the physical networks of the brain itself. Valiant (1998) found that life begins with neural networks for seeing, hearing, speaking, moving and so on already in place and that learning consisted of the growth of neural connections. Neural growth was stimulated by the passage of electrical current along nerve cells and enhanced by chemicals (neurotransmitter) discharged into the gap (synapse) between neighboring cells. “In early childhood, these connections are made at a phenomenal rate and there are windows of opportunity for interventions to encourage growth of connections for particular attributes such as language development” (Valiant, 1998, What does brain research say about learning section ¶ 1). Abbot (1997) stated:

“All brain activity occurs spontaneously, automatically, in response to challenge. The brain does not have to be taught to learn. To thrive, the brain needs plenty of stimulation, and it needs suitable feedback systems. Effective learning depends on emotional energy. We are driven (the ancestral urges of long ago) as much by emotion as by logic.” (Going with the grain of the brain section ¶ 1).

Too much stimulation and the challenge was seen as a threat and the brain turned off. This turning off was documented using Magnetic Resonance Imagings (MRIs) (Abbot, 1997).

Caine and Caine (1990) discussed the implications on education of brain-based learning stating, “What we are beginning to discover about the role of emotions, stress, and threat in learning and about memory systems and motivation is challenging basic assumption about traditional education” (p. 66). The brain is a parallel processor where thoughts, emotions,



imagination, and predispositions operate concurrently. The human brain searches for meaning, simultaneously perceives parts and wholes, involves both conscious and unconscious processes, has two memory systems (spatial and rote), and each is unique (Caine and Caine, 1990).

Abbot (1997) described an archaeological and an anthropological approach to the development of intelligence over about a million years. He believed humans developed many discrete skills than have only recently, the past 30,000 years, been combined to create the broad intelligence that now gives us “amazing versatility” (Abbot, 1997, ¶ 14). These discrete skills included social intelligence, technological intelligence, natural history intelligence, language intelligence. As a result, archeology is starting to endorse Howard Gardner’s work on multiple intelligences (Abott, 1997).

### **Art as a Cognitive Activity**

Creating and perceiving art, two of the main functions of art education, require the use of higher order thinking skills. Within the concepts put forth by Dewey (1934), Vygotsky (Bruner,1997), Piaget (Bruner, 1997), Bruner (Bigge & Shermis, 1999), Gardner (1972), Arnheim (1969), Efland (2002), and Eisner (2002), the foundations of higher order thinking skills were conceived. Dewey and Vygotsky were concerned with the role of concepts in meaning making. Vygotsky believed tools and signs mediated higher mental functions. In Piaget’s formal-operations stage, the child was able to act upon not only real or imagined objects but upon propositions or a statement or possible statements about the object, which the child weighed against one another, and performed operations upon (Gardner, 1972). Bruner identified it as a stage called symbolic representation. Gardner’s disciplinary expert and Eisner’s inscribing, editing and communicating capacities all require the use of higher level thinking skills. Arnheim (1969) described a theory of visual thinking based on visual perception. Arnheim (1969) argued:

“cognitive operations called thinking are not the privilege of mental processes above and beyond perception but the essential ingredients of perception itself...such operations as active exploration, selection, grasping of essentials, simplification, abstraction, analysis and synthesis, completion, correction, comparison, problem solving, as well as combining, separating, putting into context (p. 13).

Arnheim (1969) defined cognitive as “all the mental operations involved in receiving, storing and processing of information: sensory, perception, memory, thinking, learning” (p. 13). Nelson Goodman argued that each of the arts is a symbol system and as such is cognitive in

nature (Efland, 2002). Wolf (1987) stated, “that the arts exercise not just hand and heart, but mind as well...like other demanding cognitive activities, the arts involve people in symbol-use, analysis, problem-solving and invention” (p. 26). Brandt (1987) also referred to artistic symbols as a distinctive way of using the mind commenting “one that is usually down played in school” (p. 30).

Three of Gardner’s intelligences, spatial intelligence, bodily-kinesthetic intelligence and existential intelligence, directly relate to art education. Central to spatial intelligence Gardner identified the following capabilities: “to perceive the visual world accurately, to perform transformation and modifications upon one’s initial perceptions, and to be able to re-create aspects of one’s visual experience, even in the absences of relevant physical stimuli” (Gardner, 1983, p. 176). These capacities included the “ability to recognize instances of the same element; the ability to transform or to recognize a mental transformation of one element into another; the capacity to conjure up mental imagery and then to transform the imagery; the capacity to produce likeness or spatial information and the like” (Gardner, 1983, p. 176).

The capabilities of spatial intelligence closely resemble a list of art teaching strategies or objectives. Gardner (1983) claimed the centrality of spatial thinking was self-evident in the visual arts, “The enterprises of painting and sculpting involve an exquisite sensitivity to the visual and spatial world as well as an ability to recreate it in fashioning a work or art” (p.196).

In addition, Gardner identified another capacity, which he claimed was more abstract and elusive. This one “involves sensitivity to the various lines of force that enter into a visual or spatial display. I refer here to the feelings of tension, balance, and composition that characterize a painting, a work of sculpture, and many natural elements (like a fire or a waterfall) as well. These facets, which contribute to the power of display, occupy the attention of artists and viewers of art” (Gardner, 1983, p.176). A final facet of spatial intelligence may contribute to the transfer of knowledge. This facet of spatial intelligence “grows out of the resemblances that may exist across two seemingly disparate forms or, for that matter across two seemingly remote domains of experience. In my [Gardner’s] view, that metaphoric ability to discern similarities across domains derives in many instance from a manifestation of spatial intelligence” (Gardner, 1983, p. 176). For example, when the gifted essayist Lewis Thomas drew analogies when he depicted the sky as a membrane, or described mankind as a heap of earth (Gardner, 1983).

Another intelligence closely related to art education was the bodily-kinesthetic intelligence which “entails the potential of using one’s whole body or parts of the body (like the hand or the mouth) to solve problems or fashion products” (Gardner, 1999, p. 43). Gardner claimed, dancers, athletes, actors, craftsperson, and surgeons are likely to have strong senses of this intelligence. Characteristic of this intelligence was the “ability to use ones body in highly differentiated and skilled ways, for expressive as well as goal-directed purposes...” (Gardner, 1983, p. 206).

Gardner related bodily-kinesthetic intelligence to the work of Piaget:

“Though Piaget did not himself view his work in relation to bodily intelligence ...his description of the unfolding of sensori-motor intelligence, in fact, illuminates its initial evolution. One can see in Piaget’s description how individuals progress from the simplest reflexes - such as those involved in sucking and looking - to behavioral acts that fall increasingly under the control of environmental variation and individual intentions” (Gardner, 1983, p. 220).

Two capacities were core to bodily intelligence, control of one’s bodily motions and the ability to handle objects skillfully (Gardner, 1983). Bodily-kinesthetic intelligence in combination with spatial intelligence, Gardner (1983) claimed, would contribute to one’s ability to understand how a mechanism worked and the ability to manipulate the mechanism. For example, the art student would not only know how to use a brayer for instance, but would also be able to determine which situations call for its use to create the desired effect.

The third multiple intelligence, which applied directly to art was existential intelligence. A core ability of existential intelligence was:

“the capacity to locate oneself with respect to the furthest reaches of the cosmos - the infinite and the infinitesimal - and the related capacity to locate oneself with respect to such existential features of the human condition as the significance of life, the meaning of death, the ultimate fate of the physical and psychological worlds, and such profound experiences as love of another person or total immersion in a work of art” (Gardner, 1999, p. 60).

These “existential features” often comprise the themes or enduring ideas artists pursue and attempt to express in their works of art. There is no denying that Picasso’s *Guernica* expressed the meaning of war, of death, and of the human condition.

Two additional intelligences, which apply to art education, Gardner would primarily categorize as emotionally pervaded experiences. These were the interpersonal and intrapersonal

intelligences. Although Gardner identified them as separate entities, he recognized that all intelligences have both a cognitive and affective aspect.

Eisner agreed with Gardner that emotion and imagination play a role in cognition. Eisner stated:

“that many of the most complex and subtle forms of thinking take place when students have an opportunity either to work meaningfully on the creation of images - whether visual, choreographic, musical, literary, or poetic - or to scrutinize them appreciatively. To be able to create a form of experience that can be regarded as aesthetic requires a mind that animates our imaginative capacities and that promotes our ability to undergo emotionally pervaded experience. Perception is, in the end a cognitive event” (Eisner, 2002, p. xii).

However, “for Eisner, the mind is biologically rooted in the senses, whereas Gardner’s ideas about the various intelligences are linked to regions in the brain rather than to the eye or ear” (Efland, 2002, p. 62). Learning for Eisner “moved from sensory perception to conception and then to representation in forms that could be shared publicly” (Efland, 2002, p. 63).

Representation, sensibility and imagination perform critically important cognitive functions according to Eisner (2002). Education according to Eisner (2002) was a process of the child learning to become the architect of his or her own experience, of creating his or herself. The arts make distinctive contributions to that end “through their emphasis on the expression of individuality and through the exercise and development of the imaginative capacities” (Eisner, 2002, p. 24). Imagination also played another role, “Imagination, that form of thinking that engenders images of the possible....provides a safety net for experiment and rehearsal” (Eisner, 2002, p. 5). Arnheim (1969) stated, “Thinking calls for images, and images contain thought...It [art] makes things visible that are invisible or inaccessible or born of fantasy” (p. 254). Efland (2002) stated, “Imagination is the act or power of forming mental images of what is not actually present to the senses or what has not actually been experienced. It is also the act or power of creating new ideas or images through the combination and reorganization of previous experiences” (p. 133).

Eisner claimed inscribing, editing, communicating, and discovery are the four cognitive functions of representation. These apply directly to the main feature of art classes, making art. Inscribing was “aimed at transforming the contents of consciousness within the constraints and affordances of material...Representation stabilizes the idea or image in a material and makes possible a dialogue with it.” (Eisner, 2002, p. 6). Second, editing, “is the process of working on

inscriptions so they achieve the quality, the precision, and the power their creator desires.....Editing is paying attention to relationships and attending to details; it is the process of making the work” (Eisner, 2002 p. 6). Third, communication, transformed:

“consciousness into a public form, which is what representation is designed to do....culture depends upon these communications because communication patterns provide opportunities for members of a culture to grow.... Thus the social contribution of the educational process it to make it possible for individuals to create symbiotic relationships with others through the development of their distinctive and complimentary abilities and in so doing to enrich one another’s lives” (Eisner, 2002, p. 7).

The fourth, discovery, was the discovery of ends in process, which in turn generates surprise. “Opportunities in the process of working are encountered that were not envisioned when the work began, but that speak so eloquently about the promise of emerging possibilities that new options are pursued” (Eisner, 2002, p. 7).

In addition, according to Eisner, “the arts are a means of exploring our own interior landscape. When the arts genuinely move us, we discover what it is that we are capable of experiencing” (Eisner, 2002, p.11). This could be accomplished through either the making of art or the looking at other artists’ works of art. Both, Eisner contended were important practices in a quality art education. One of the major aims of art education according to Eisner (2002) was to “promote the child’s ability to develop his or her mind through the experience that the creation or perception of expressive forms makes possible” (p. 24).

Eisner (2002) described three ways in which artists treat forms of representation so that they affect how meanings are conveyed. One mode of treatment was *mimetic*. By mimetic Eisner (2002) referred to forms that “look or sound like what they are intended to represent” (p. 15). For example, the use of perspective to create the illusion of a third dimensions. The second mode of treatment was the creation of *expressive form*, “The representation of feeling is achieved in many ways. Perhaps the most important is the way in which visual form - line, color, shape, value, texture, all aspects of form - is composed” (Eisner, 2002, p. 16). If all responses to formal relationships were alike then everyone’s responses would be the approximately the same. Since this is not the case “the meaning secured from a work depends not only on the features of the work but also on what the individual brings to it” (Eisner, 2002, p. 17). Here culture and personal experience interact (Eisner, 2002).

Eisner (2002) stressed, “the primary point should not be lost: the way forms are treated by the artist - or by the child - has a great deal to do with what the work expresses. And it is the possession of fertile imagination and an array of technical skills that enable the artists to shape forms that influence how we feel in their presence” (p. 17). Again, a major concern of art education is developing both the child’s expressive and technical skills.

The third mode of treatment occurred through the use of conventional signs, “Conventional signs are socially agreed-upon symbols that refer to ideas, objects, or events and the like” (Eisner, 2002, p. 18). In visual art classes, students and teachers often use and/or examine the use of signs in either the making of or understanding of the works of art.

Eisner (2002) summarized the cognitive nature of art succinctly, “The ability to create images in which mimesis, expressiveness, and conventional signs convey the creator’s aims is a substantial cognitive accomplishment. It requires a repertoire of technical skills, a sensitivity to relationships among the forms, and the ability to use appropriate conventional signs” (p. 18).

Efland (2002) put forth an integrated view of cognition that took into consideration three factors, “(1) that the mind is a computational function using symbols; (2) that cognition is a constructive process used to enable individuals to secure meaning; and (3) that learning includes the acquisition of social reality, the idea that learning becomes meaningful when it occurs in a sociocultural or situational context” (p. 156). Efland (2002) believed his integrated view of cognition avoided the “dualisms that have plagued educational practice throughout the modern era” (p. 159).

Four arguments were developed by Efland (2002) that support the cognitive justification for the arts, 1) the cognitive flexibility argument, 2) the integration of knowledge argument, 3) the imagination argument and 4) the aesthetic argument. Cognitive flexibility according to Efland (2002) was the ability to “change strategies as one becomes mindful of the structural demands of each domain, and the ability to activate the appropriate means to secure meaning or understanding” (p. 160).

The integration of knowledge argument was based on the belief that “the expressive content of a work of art reflects the larger situation surrounding the work, namely the social and cultural influences and source of motifs initially perceived and felt by the artist” (p. 164). Links between domains such as physical science, social science, humanities and/or history were established orienting learners toward knowledge and understanding. This argument also takes

into consideration the life world of the learner. The lifeworld was described as the “individual’s construction of reality, the commonsense, symbolic matrix into which school subjects are placed in the effort to make them meaningful” (Efland, 2002, p. 167).

Efland’s (2002) imagination argument was based on the role of metaphor and narrative as tools for understanding what cannot be fully comprehended such as our feelings, moral practice, aesthetic experiences and spiritual awareness. Efland (2002) believed metaphor should become the principal object of study within the arts. The fourth argument, the aesthetic argument, established the “point that perceptually vivid aesthetic encounters in the arts have educative value” (Efland, 2002, p. 159) by arousing and sustaining our interest and attention.

In creating and perceiving works of art, students and artists must go beyond recall. The elements and principles of design are molded to create meaning, personal, expressive, or cultural. Art students learn to inscribe, edit and communicate through a variety of media. Imagination, possibilities and discovery are key elements embodied in works of art. Making art requires the student to synthesize and bring together a variety of elements and concepts. Art works are constantly being evaluated as they are formed.

Bloom’s hierarchical Taxonomy (Karthwohl, 2002) in its revised form clearly illuminates the necessity of using higher level thinking skills in creating and examining works of art. Making and evaluating art involves the highest levels of the taxonomy. The cognitive dimensions of the revised taxonomy are, starting with the lowest, 1) Remember - retrieving relevant knowledge from long-term memory. This level includes recognizing, and recalling. 2) Understand - Determining the meaning of instructional messages, including oral, written, and graphic communication. This level includes interpreting, exemplifying, classifying, summarizing, inferring, comparing, and explaining. 3) Apply - carrying out or using a procedure in a given situation. This level includes executing and implementing. 4) Analyze - breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose. This includes level differentiating, organizing, and attributing. 5) Evaluate - making judgments based on criteria and standards. This level includes checking and critiquing. 6) Create- putting elements together to form a novel, coherent whole or make an original product. This level includes generating, planning, and producing. (Karthwohl, 2002).

Marzano’s (1992) Dimensions of Learning model resembles Bloom’s hierarchical Taxonomy. Dimension 2, Thinking Involved in Acquiring and Integrating Knowledge, deals

with declarative knowledge and procedural knowledge. At this level, constructing meaning, organizing, storing, shaping and internalizing are the main skills that students need to develop. Dimension 3, Thinking Involved in Extending and Refining Knowledge, has students compare, classify, use induction and deduction, error analysis, constructing support, abstracting and analyzing perspectives. Dimension 4, Thinking Involved in Using Knowledge Meaningfully, requires students to use the following processes; decision making, investigation, experimental inquiry, problem solving and invention. Dimensions 3 and 4 are considered higher level thinking skills. Again, making art necessitates the use of these higher level thinking skills. As Eisner (2002) stated, “Every task and each material with which we work both imposes constraints and provides opportunities for the development of mind” (p. 12).

Brown (1989) recognized the importance of an active mind stating, “The new literacy of thoughtfulness calls for a quite different technology of teaching and testing: it rests on quite different notions about the nature of knowledge and learning because it is about the making of meaning, not just the receiving of it. Thoughtfulness is a construction, not a passive undertaking” (p. 32).

### **Theories of Metacognition**

There are several variations concerning the definition of metacognition. Flavell (1979) referred to metacognition as “knowledge and cognition about cognitive phenomena...” (p. 906). Schwartz and Perfect (2002) stated, “Metacognitive knowledge refers to explicit knowledge about our own cognitive strengths and weaknesses” (p.5). Yussen (1985) defined metacognition in terms of reflection, “metacognition is that mental activity for which other mental states or processes become the object of reflection” (p. 253). Wellman (1985) stated, “*metacognition* has increasingly been employed to refer to a person’s cognition about cognition, that is, the person’s knowledge of cognitive processes and states such as memory, attention, knowledge, conjecture, illusion” (p. 1). Pintrich (2002) described metacognitive knowledge as “knowledge about cognition in general, as well as awareness of and knowledge about one’s own cognition” (¶ 2). Schraw and Dennison (1994) referred to metacognition as “the ability to reflect upon, understand, and control one’s learning” (p. 460). Bransford (1999) also identified prior knowledge, self-regulation and reflection as important aspects of metacognition in determining



performance. Three terms or phrases in these definitions, experiences, conscious understanding and reflect upon, can be traced back to the writings of Kant, Hegel and Dewey.

Modern metacognitive research is relatively new and went in two parallel directions, “one in the emerging cognitive psychology of the 1960s (e.g. Hart, 1965) and the other in the post-Piagetian developmental psychology of the 1970s (e.g. Flavell, 1979)” (Perfect & Schwartz, 2002, p. 2). Hart was interested in the “accuracy of judgments people make about memory” (Perfect & Schwartz, 2002, p. 2) while Flavell, a developmentalist was “interested in finding out if the improvement in children’s memory abilities was a function of greater conscious understanding of the rules that govern memory and cognition” (Perfect & Schwartz, 2002, p. 3). The developmentalists were “interested in the development of metacognitive thinking, that is, the ability to reflect on one’s cognitive processes” (Perfect & Schwartz, 2002, p 3).

There are many models of metecognition that divide metacognition into “two distinct classes (a) knowledge of cognition and (b) the process involving the monitoring, control and regulation of cognition” (Pintrich, 2002; Schraw, 1998). Flavell (1979) on the other hand identified four classes of phenomena: 1) metacognitive knowledge, 2) metacognitive experiences, 3) goals (or tasks), and 4) actions (or strategies). While Wellman identified five sets of knowledge, Perkins (1994) identified three categories of intellegences, and Gardner seven to nine separate intelligences.

First, consider the views of cognitive scientists whom identify two distinct classes, knowledge of cognition and regulation of cognition. Knowledge of cognition involved at least three kinds of metaconitive awareness: declarative, procedural and conditional knowledge (Schraw, 1988). “Declarative knowledge includes knowledge about oneself as a learner and about what factors influence one’s performance.... Procedural knowledge refers to knowledge about doing things. Much of this knowledge is heuristics and strategies.... Conditional knowledge refers to knowing when and why to use declarative and procedural knowledge” (Schraw, 1998, p. 114).

Monitoring or control of cognition referred to the activities that help students control their learning (Pintrich, 2002; Schraw & Dennison, 1988). Metacognitive regulation improves performance, including better use of resources, better use of existing strategies, and a greater awareness of comprehension breakdowns. Three essential regulatory skills were included planning, monitoring, and evaluation. Planning involved the selection of strategies and the

allocation of resources that affected performance (Schraw, 1988). “Monitoring refers to one’s on-line awareness of comprehension and task performance...Evaluating refers to appraising the products and efficiency of one’s learning” (Schraw & Dennison, 1988, p. 115).

Second, Flavell in his much-cited 1979 article *Metacognition and Cognitive Monitoring: A New Area of Cognitive - Development Inquiry*, as stated above delineated four classes of phenomena. According to Flavell (1979) “metacognitive knowledge consists primarily of knowledge or beliefs about what factors or variables act and interact in what ways to affect the course and outcome of cognitive enterprises” (Flavell, 1979, p. 907). Three categories or variables were identified, person, task and strategy. The person category encompassed everything one comes to believe about the nature of oneself and other people as cognitive processors. Concerning the task category, the “metacognitive knowledge in this subcategory is an understanding of what such variations imply for how the cognitive enterprise should best be managed and how successful you are likely to be in achieving its goal” (Flavell, 1979, p. 907). The strategy category dealt with knowledge about what strategies are likely to be affective in reaching goals. The third category goals or tasks referred to the objectives of a cognitive enterprise, while the fourth category, actions or strategies referred to the cognitions or other behaviors employed to achieve them (Flavell, 1979).

Perkins (1994) who proposed that there are three broad categories of intelligence presented a third view of metacognition. Valiant (1998) summarized these three views:

“Neural intelligence is that which is established by the neural networks we receive as a result of our genes. Experimental intelligence is the result of the new connections we are able to grow through our life experiences, including education. Finally, reflective intelligence results from reflecting on how we behaved in the past and what alternative routes we might take in the future when face with a similar situation” (What does brain research say about learning section ¶ 2).

Abbot (1997) stated this about the relationship between Perkins’ reflective intelligence and metacognition, “Through metacognition, we can develop skills that are genuinely transferable. These skills are linked to reflective intelligence or wits” (Abbot, 1997, How do we create intelligence section ¶ 3). Perkins (1994) defined reflective intelligence as, “the contribution of mindful self-management and strategic deployment of one’s intellectual resources to intelligent behavior” (p. 14). Efland (2002) also referred to Perkins’ view of the relationship between *experiential* and *reflective* intelligence emphasizing one’s ability to cultivate

awareness of one's own thinking as a means of steering our experiential intelligence in positive directions. This steering was directed by asking ourselves good questions and guiding ourselves with strategies.

Gardner's (1999) Interpersonal and Intrapersonal intelligences were metacognitive in nature and represent a fourth view. "*Interpersonal intelligence* denotes a person's capacity to understand the intentions, motivations, and desires of other people and, consequently, to work effectively with others....*intrapersonal intelligence* involves the capacity to understand oneself, to have an effective working model of oneself -including one's own desires, fears, and capacities - and to use such information effectively in regulating one's life" (p. 43). Gardner later added a possible ninth intelligence, existential intelligence which is both metacognitive and affective in nature. This intelligence encompassed:

"the capacity to locate oneself with respect to the furthest reaches of the cosmos - the infinite and the infinitesimal - and the related capacity to locate oneself with respect to such existential features as the human condition as the significance of life, the meaning of death, the ultimate fate of the physical and the psychological worlds, and such profound experiences as love of another person or total immersion in a work of art" (Gardner, 1999, p. 61).

Wellman (1985) presented a fifth view of metacognition that identified five different but overlapping sets of knowledge. The first, Existence, in this case the person must know that thoughts and internal mental states exist, but that they are not the same as external acts or events. The second, Distinct Processes, that humans engage in many cognitive processes. They can remember or forget, they can visualize images, and they can dream they can reason. The third, Integration, that while there are numerous possible distinctions among different mental acts, all mental processes are also similar and related. In some sense, there is only one mind. The fourth, Variables, that a number of other factors or variables influence any one mental performance. In remembering, for example, how much one can remember depends on how hard the task is, the nature of the items, and the memory strategies used. The fifth set of knowledge, Cognitive Monitoring, that humans are often able to 'read' their own mental states, or monitor their ongoing cognitive processes. Cognitive monitoring referred to one's abilities to accurately assess the state of information within one's own cognitive system.

Wellman's (1985) "five sets of knowledge" ties the theories of the philosophers together with those of the cognitive scientists. Wellman's identification of existence, as a set of

knowledge, seems to fit Descartes' separation of mind and body as do the concepts of consciousness found in the work of Kant and Hegel. The role of consciousness was also found in the writings of Flavell, Efland, Green and Perfect and Schwartz. Perfect and Schwartz (2002) stated that metacognitive control is, "the conscious and non-conscious *decisions* that we make based on the output of our monitoring processes. Control processes are revealed by the behaviors a person engages in as a function of monitoring" (p. 4). Flavell believed that "metacognitive experiences are especially likely to occur in situations that stimulate a lot of careful, highly conscious thinking" (Flavell, 1979, p. 908). Green (1999) writing on brain-based learning had this to say, "Learning involves conscious and unconscious processes, thus students need to reflect on how and what they learn for the development of personal meaning" (Brain-based learning section ¶ 2). Efland (2002) wrote, "By becoming mindful of our perceptual activity, we have made perception itself into an *object of thought*" (p. 17). These perspectives of consciousness link the current theories to the writings of Dewey and Kant and support what Wellman calls integration, for example, Dewey's notion of "doing and undergoing," and Kant's writings on the self as both object and subject and on matter and form.

In addition, metacognition was considered affective in nature as well as cognitive. The affective aspects of metacognition dealt with feelings and beliefs. Flavell (1979) stated the following about the affective nature of metacognition, "Many metacognitive experiences have to do with where you are in an enterprise and what sort of progress you are making or are likely to make: You believe/feel..." (p. 908). For example, how you feel you are doing on an assignment, do you believe you can memorize the instructions, or the decision concerning whether or not you should keep studying or get a good night's sleep (Flavell, 1979). Perfect and Schwartz (2002) identified what is called metacognitive awareness which, "refers to the feelings and experiences we have when we engage in cognitive processes" (p. 5). Caine and Caine (1990) discussed the affective aspects of brain-based learning, "What is learned is influenced by and organized by emotions and mind-sets involving expectancy, personal biases and prejudices, self-esteem, and the need for social interaction" (p. 67). Gardner's claim that the body was both subject and object reflected both the cognitive and affective nature of what he called bodily-kinesthetic intelligence. He stated, "We have seen how dancers and athletes use their bodies as 'mere' objects, and noted how inventors and other workers use parts of the body - particularly the hands - in order to

manipulate and transform objects in the world....It [the body] is also a vessel of the individual's sense of self, his most personal feelings and aspirations..." (Gardner, 1983, p. 235).

### **Theories of Metacognitive/Self-Reflective Learning**

Three models, Bloom's Revised Taxonomy (Krathwohl, 2002) and Dimensions of Learning (Marzano, 1988; 1992) and Self-Regulated Learning (Zimmerman, 2001) are based on theories that are both cognitive and affective in nature. In the following descriptions the metacognitive and affective aspects will be stressed.

**Bloom's Taxonomy.** Bloom's original Taxonomy developed definitions for six major categories in the cognitive domain, Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. These categories were ordered from simple to complex, from concrete to abstract and represented a cumulative hierarchy (Krathwohl, 2002). The revised Taxonomy was a two-dimensional framework: Knowledge and Cognitive Processes. The knowledge dimension includes: a) factual knowledge, b) conceptual knowledge, c) procedural knowledge and d) metacognitive knowledge. The metacognitive aspects were incorporated into the Taxonomy's three subcategories, 1) strategic knowledge, 2) knowledge about cognitive tasks, including appropriate contextual and conditional knowledge, and 3) self-knowledge (Krathwohl, 2002). The cognitive process dimension includes six dimensions, 1) remember, 2) understand, 3) apply, 4) analyze, 5) evaluate, and 6) create (Krathwohl, 2002). In the revised Taxonomy the fourth category metacognitive knowledge was based on Pintrich's definition of metacognition, "Metacognitive knowledge includes knowledge of general strategies that might be used for different tasks, knowledge of the conditions under which these strategies might be used, knowledge of the extent to which the strategies are effective, and the knowledge of self" (Pintrich, 2002, ¶ 4).

**Dimensions of Learning.** Another approach, Dimensions of Learning (DOL) was a direct descendant of the comprehensive research-based framework on cognition and learning described in the 1988 ASCD book *Dimensions of Thinking* (Marzano, et al. 1988). Dimensions of Learning translated research and theory into a practical model that K-12 teachers could use to improve the quality of teaching and learning in any content area (Marzano, 1992). The approach was based on the belief that "learning involves a complex system of interactive processes that includes five types of thinking - the five dimensions of learning" (Marzano, 1992, p. viii).

Dimension 1, Positive Attitudes and Perceptions About Learning and Dimension 5, Productive Habits of Mind, dealt with the psychological space or life space of the student, while dimensions two, three and four addressed the use and acquisition of knowledge. Learning according to this model was both cognitive and affective.

Dimension 1 contained two major areas; Positive Attitudes and Perceptions Related to Classroom Climate (acceptance, comfort and order); and Attitudes and Perceptions Related to Classroom Tasks (value, ability, and clarity). Strategies for developing positive attitudes and perceptions about classroom climate included establishing a relationship with each student, activities that involved physical movement, introduced the concept of "bracketing", established and communicated rules and procedures, stopped malicious teasing or threats. Attitudes and perceptions about classroom tasks included developing a sense of academic trust, linking classroom tasks to student's interests, providing positive feedback, teaching the students positive self-talk, breaking tasks into small steps.

Dimension 5, Productive Habits of Mind, focused on three main areas: critical thinking, creative thinking, and self-regulation.

Critical Thinking included: 1) being accurate and seeking accuracy, 2) being clear and seeking clarity, 3) being open-minded, 4) restraining impulsivity, 5) taking a position when the information warrants it, and 6) being sensitive to other's feelings and level of knowledge.

Creative Thinking included: 1) engaging intensely in tasks even when answers or solutions are not immediately apparent, 2) pushing the limits of your knowledge and abilities, 3) generating, trusting, and maintaining your own standards of evaluation, and 4) generating new ways of viewing a situation outside the boundaries of standard conventions.

Self-Regulation included: 1) being aware of your own thinking, 2) planning, 3) being aware of necessary resources and 4) evaluating the effectiveness of your actions.

Gardner's Personal Intelligences and the findings based on metacognitive research explain why Dimension 1's and 5's strategies for supporting student learning were so important to student success. These two dimensions form the backdrop or foundation upon which the traditional concept of learning takes place. According to Gardner, school-age children, between the start of school and the beginning of adolescence, "may acquire feelings of learned helplessness, as they become convinced that there are certain pursuits that they cannot carry out. (For instance, many young girls come to feel that they cannot solve mathematics problems, thus

launching a vicious cycle of diminishing expectations and diminishing achievements)" (Gardner, 1983, p. 250). The metacognitive dimensions of learning can have a significant impact on a student's learning ranging for example, from how a student regulates oneself to the affect of a student's beliefs about his/her ability to learn.

**Self-Regulated Learning.** Zimmerman (2001, 2002) reported that the theory and research related to self-regulated learning emerged in the mid-1980s. This research was an effort to discover how students became masters of their own learning processes. They found it was neither a mental ability nor an academic performance skill. Instead, self-regulation referred "to the self-directive *process* through which learners transform their mental abilities into task-related academic skills" (Zimmerman, 2001, p. 1). The key issue defining self-regulated learning (SRL) was whether the student displayed personal initiative, perseverance, and adaptive skill in his/her pursuit of learning (Zimmerman, 2001). Self-regulated learning shifted the focus of education, 1) to an activity students do for themselves rather than something that happens to them, 2) to students' personally initiating strategies to improve learning and environments rather than accepting fixed learning abilities and environments and 3) to students playing a significant role in choices related to the amount and form of instruction. .

Zimmerman (2001) found that theorists' definitions of self-regulated learning contained four common features although each theorist focused attention on different factors and motivations. The theorists included operant theorists, phenomenologists, Vygotskians, constructivist, cognitively oriented theorist, volitional theorists and social cognitive theorist. The first common feature found was, "Students are self-regulated to the degree that they are metacognitively, motivationally, and behaviorally active participants in their own learning process" (Zimmerman, 2001, p. 5). The second feature found, a feedback loop during learning that was self-oriented. The third common feature found was that all the definitions described "the how and why students choose to use a particular self-regulated process, strategy or response" (Zimmerman, 2001, p. 6). In addition, the final feature assumed that students' "efforts to self-regulated their academic learning often require additional preparation time, vigilance, and effort" (Zimmerman, 2001, p.7).

Self-evaluation was one aspect of self-regulated learning, which had different interpretations depending on the theorist (Zimmerman, 2001). According to operant theorists, individuals are required to compare some dimension of their behavior to a standard. Zimmerman

(2001) found that from the operant perspective self-evaluations fulfilled three roles, 1) to influence self-corrective responses, 2) to modify standards if they are insufficient or unnecessary, and 3) were the basis for self-administering rewards. These rewards were external and it was recommended that the teacher have a backup contingency when students do not meet the stringent reward criteria.

From a phenomenological perspective, the best way to enhance self-regulated learning was to understand one's belief in one's self as a self-regulator (McCombs, 2001). Self-evaluation was one of three processes identified by McCombs (2001) as key in developing self-regulated learners. In the self-evaluation process, the learner measured "task requirements against personal needs for competence and control and against self-system structures" (Zimmerman, 2001, p. 14). Self-structures were the learners' subjective perceptions of which self-worth and self-identity were key processes in the learner's psychological functioning. Motivation and persistence were dependent on one's self-concept. Other self-regulation processes, like planning and goal-setting were arrived at through the self-evaluation process, which then affected the self-system structures and processing in a reciprocal or looping fashion (McCombs, 2001).

Winne (2001), an informational processing theorist, also found a cycle of control and monitoring processes were key to self-regulated learning. During the monitoring process, students' compared current outcomes to the standards set by the student. These matches or mismatches became the driving force for learning. Two processes, control and monitoring, were used during four phases: perceiving the tasks, setting goals and plans, enacting studying tactics, and adapting tactics. Again, a feedback loop was established to enhance learning.

The social cognitive view of self-regulated learning, according to Zimmerman (2001) was based on the work of Bandura, Schunk and Zimmerman. The social cognitive approach found "the motivation to self-regulate involves two cognitive sources: self-efficacy and outcome expectations and goals" (Zimmerman, 2001, p. 20). Self-efficacy was defined as "the perceived ability to implement actions necessary to attain designated performance levels" (Zimmerman, 2001, p. 20). In this view, personal goals served as standards against which future performances would be evaluated and three key interactive subprocesses were identified, self-observation, self-judgment, and self-reaction (Zimmerman, 2001; 2002). Self-observations prompted learners to self-evaluate which lead to personal and behavioral self-reactions. Behaviors could be assessed on such dimensions as quantity, quality, and originality. Self-observation was aided through self-



recording with two important criteria, regularity and proximity (Schunk, 2001). Self-judgments were made by comparing existing performance levels with one's learning goals. This was an environmentally oriented approach where "particular attention is paid to contextual properties of students' goals" (Zimmerman, 2001, p. 21). Schunk (2001) listed factors that can affect self-judgments these included "the type of standard employed, the properties of the goal, the importance of goal attainment and the attributes made for outcomes" (p. 131). Self-reactions involved judging one's goal progress and affected students' motivation. The social cognitive view "postulates that the anticipation of consequences enhances motivation" (Schunk, 2001, p. 133) where as reinforcement theories contend that consequences alter behavior. Once again, the self-regulatory process was cyclical.

In addition, students' developmental level was considered with four self-regulatory competence levels identified: observational, emulative, self-control and self-regulation. Overtime the learner moved from a level of developing self-regulatory skills and strategies that were initiated from social sources to sources that were initiated by self-sources. Schunk (2001) explicitly stated, "self-regulation does not develop automatically nor is it acquired passively from the environment" (p. 142).

Three self-regulatory processes, forethought, performance and self-reflection interacted and lead to the development of self-efficacy and skill growth. Goal setting was a key element of the forethought phase in the self-regulatory process. Schunk (2001) found that self-set goals promoted greater self-efficacy, motivation and confidence. Proximal goals proved more effective than distant goals in increasing productivity, self-efficacy and achievement.

Regarding performance control, Schunk (2001) addressed two strategies, feedback and self-verbalization. He found feedback an effective means of relating prior success to current learning. Although "ability feedback for early success led to higher ability attributions, self-efficacy and achievement, compared with effort feedback for early successes" (Schunk, 2001, p.138) effort feedback was an effective tool in cases where prior experiences were unsuccessful. Self-verbalization of self-regulatory strategies also led to greater self-efficacy and achievement.

The self-reflection phase of the self-regulator process focused on three processes: 1) self-monitoring, 2) reward contingencies, and 3) progress feedback and self-evaluation. Schunk (2001) reported monitoring progress, self or external, enhanced achievement and students' self-efficacy for continued learning. Rewards also contributed positively to student achievement and

self-efficacy when they were based on performance contingencies rather than participation contingencies. The highest levels of self-efficacy and achievement were attained when both rewards and goals were combined. Concerning progress feedback and self-evaluation Schunk (2001) found in a writing study that a combination of process goals and feedback were the most effective in increasing student's self-efficacy, writing achievement, self-evaluated learning progress and self-regulated strategy usage. Learner's belief that they were making progress toward goal attainment was a necessary element of the feedback. Schunk (2001) determined that learners could self-evaluate progress on tasks that had clear criteria. When standards or criteria were unclear, learners could rely on feedback indicating progress to sustain self-efficacy and motivation.

Schunk (2001) also stressed the contextual nature of self-regulated learning and identified six areas in which one can use self-regulated processes: motives, methods, time, outcomes, physical environment and social environment.

Volitional views of self-regulated learning assumed the existence of "a covert psychological force or forces that control action" (Zimmerman, 2001, p. 23). According to this view, volition was distinguished as separate from motivation. Volitional approaches focused on strategies that affect learners' intentions rather than on learning itself and recommended training learners' in the use of those strategies. On the other hand, motivational strategies involved methods of increasing the learners' intent to learn by "imagining positive or negative consequences of success or failure" (Zimmerman, 2001, p. 25). The impact of the environment on emotions and motivation were recognized but were considered secondary to cognitive factors.

Those researchers interested in the work of Vygotsky and self-regulated learning focused their attention on two specific features emphasized in Vygotsky's theory: "inner speech as a source of knowledge and self-control, and interactive dialogue between adults and children as a vehicle for conveying and internalizing linguistic skill" (Zimmerman, 2001, p. 26). Vygotsky distinguished two types of inner speech that affect the learners' motivation, self-involved and task-involved inner speech. Self-involved inner speech dealt with affective statements that were used to improve student self-control while task-involved speech referred to problem solving strategy statements. As word meaning became internalized, evoked consciousness, the learner could increasingly "consciously guide, plan and monitor their own activities" (Zimmerman, 2001, p. 27). The key self-regulation process then was egocentric speech.

Schunk (1986) described a related theory based on the work of Vygotsky the cognitive developmental theory. The cognitive developmental theory established a link between private speech and self-regulation. Three stages in the development of verbal control of motor behavior were identified (Schunk, 1986). First, from the ages of 1.5 -2.5 speech of others was primarily responsible for directing the child's behavior. Second, from about 3-4 years of age, overt speech by the child could initiate behavior but not inhibit behavior. At about 4.5-5.5 years of age the child was capable of using private speech to initiate, direct, and inhibit motor behavior. From the age of 4 or 5 to age 8, the amount of private speech declined however, the proportion of private speech that was self-regulating increased with age (Schunk, 1986).

Schunk (1986) described several uses for verbalization in the instructional process. Instructional processes such as, attending, coding, associating, rehearsing and monitoring were amenable to verbal self-regulation by learners. Verbalization may have helped learners attend to important features of a task and disregard others. Rehearsal, a form of verbalization may assist students in coding and retaining information. Verbalization can also assist students monitor their behavior or learning. Research has shown that verbalization is an effective instructional process for students who are impulsive, learning disabled, or remedial learners (Schunk, 1986). Verbalization for students who can handle the task may actually limit performance.

The cognitive constructivist views were founded on the work of Bartlett and Piaget and their work related to a cognitive schema with conflict as the basis for learning and recall. "Cognitive constructionist assume that a human motive to construct meaning from experience is inherently compelling" (Zimmerman, 2001, p. 30). Recent cognitive constructivists Paris, Byrnes, and Paris (2001) saw self-regulated learning as a multifaceted set of processes where learners "construct theories to regulate four components of their learning: self-competence, agency and control, schooling and academic tasks, and strategies" (Zimmerman, 2001, p. 31). Instead of advocating cognitive conflict these theorist placed the learner in a situated cognitive perspective where the learners' self-concepts and use of self-regulatory methods were "adaptive to social and historical contexts, including the tools, values, and customs of local communities" (Zimmerman, 2001, p. 32). Constructivist emphasized the value of personal skill in developing learning strategies, deconstructing tasks, and organizing the parts into a hierarchical sequence.

Zimmerman (1986) summarized that:

“self-regulated learning theorists view students as metacognitively, motivationally, and behaviorally active participants in their own learning process. Metacognitively, self-regulated learners are persons who plan, organize, self-instruct, self-monitor, and self-evaluate at various stages during the learning process. Motivationally, self-regulated learners perceive themselves a competent, self-efficacious, and autonomous. Behaviorally, self-regulated learners select, structure and create environments that optimize learning” (p. 308).

Zimmerman (1986) reported that using self-regulated strategies was highly correlated to academic achievement for high school students. “These strategies included self-evaluation, organizing and transforming, subgoal setting and planning, seeking information, keeping records and self-monitoring, environmental structuring, self-consequences, rehearsing and memorizing, seeking peer, teacher or adult assistance, reviewing notes, tests, or text books” (p. 308). These self-regulated strategies correspond with a number of the strategies also found in the Dimensions of Learning Model.

### **Standards Movement, School Reform, High Stakes Testing**

Demonstrating achievement and student growth has been mandated at the state and national levels. The No Child Left Behind Act of 2001 (NCLB, HR1) requires that not only certain levels of student performance be demonstrated but also “adequately yearly progress” attained. As a result testing has become a high stakes endeavor. Assessment has been seen as a tool for transforming instruction and learning, of revitalizing the education system.

Cohen (1988) found that the states are on average paying about 50% of education costs and education was the largest single budget item of almost every state. As a result, states had a stake in student performance. At a national level, O’Neil (1993) found that a bipartisan task force created by Congress, the National Council on Education Standards and Testing, was charged with the task of examining the feasibility and desirability of a national system of educational standards and tests. Specifically, according to O’Neil (1993), the Council recommended that National standards should be developed that included 1) “*content standards* (what students should know and be able to do)” (p. 4), 2) “*student performance* (the level (s) of competence in the content)” (p. 4), and 3) a “*system performance standard* (to assess the success of schools, districts, states, and the nation as a whole in helping all student attain high performance standards” (p. 4). In addition, the Council recommended that states should develop “*school*

*delivery standards*” to judge the extent to which schools were providing students with opportunities to attain these high standards (O’Neil, 1993).

Over a decade ago Simmons and Resnick (1993) participated in the “New Standards Project”, a consortium of 17 state and district partners working with specialists, was charged with the task “to design and implement a system of performance standards, authentic assessments, and professional development intended to change the way the American school system works.” (p. 11). The goal was to “build a revitalized education system using assessment as a tool for transforming instruction and learning.” (p. 12) They determined there was a need to develop world-class content and performance standards. They envisioned an assessment system that embodied the standards and was performance-based with rubrics and procedures for scoring students’ work that was reliable and fair (Resnick, 1993). They predicted the NSP [New Standards Portfolios] would contain three kinds of work: 1) work chosen by the district, school, teacher, and/or student, 2) prescribed projects and extended learning activities and 3) responses generated by NSP matrix examination tasks. (Resnick, 1993).

Proponents of the national standards movement predicted that within the next decade “national standards describing essential outcomes in various school subjects *could* become the glue holding together many pieces of the education system - curriculum frameworks and guides, textbook adoption, staff development, and perhaps most significantly, assessment - at the local state and national level” (O’Neil, 1993, p. 4). The standards intended to spell out the “types of knowledge and skills that *all* students should attain as a result of schooling” (O’Neil, 1993, p. 6). Yet, by virtue of the testing and reporting processes, state assessments could focus the attention of educators on certain goals, process and outcomes (Cohen, 1988). Costa (1993) predicted the establishment of world-class standards. These higher standards would “shift learning *of* the content to learning *from* the content” (p. 50). Boundaries between subjects would dissolve, self-evaluation would be favored over external evaluation, content acquisition and knowledge retention would no longer be an end in itself, and uniformity would be dismissed and diversity embraced. Teaching would become a complex, collaborative, and reflective endeavor (Costa, 1993). Resnick (1993) believed standards could impact teaching and learning and provide the impetus for educational reform if the standards were “*internalized*”. By internalized she means the standards would be held first by teachers and then students as personal goals.

Critics warned that students may be the ones to pay the price if national standards and assessments were created before basic inequalities were addressed such as class size, outmoded facilities, and inequitable school financing (O’Neil, 1993). Another concern, revolved around the questions, 1) “Is it really possible to create standards that apply to all students?” and 2) “What kinds of tests will be used to measure whether students have attained the new standard?” (O’Neil, 1993, p. 5). Contrary to Resnick’s performance assessment model, there has been a proliferation of standardized tests, many of which assess low-level skills. As a result, many curriculums have focused on the basic skills (O’Neil, 1993). Cohen (1988) also noted that standards had an impact on the curricular and instructional decisions at the district, school and classroom level. In addition, Cohen (1988) cautioned that “by setting performance standards and attaching sanctions for students, teachers, or entire schools, state assessment systems can have powerful effects at the local level” (p. 585). Florida and South Carolina have school incentive programs that attach bonuses (financial) to assessment results at the school level (Cohen, 1988).

Eisner (1993) argued, “the current emphasis on standards will provide no panacea in education (p. 23). Instead, he stated we needed something more and that something more was criteria. Standards he stated quantify qualities but possessed none of the qualities of what was measured. Criteria, on the other hand, “facilitate the search for qualities we value within an essay, a scientific experiment, a painting, a work of history and the like” (Eisner, 1993, p. 22). Eisner (1993) contended that the ends of education “seek work that displays the student’s intelligent judgment” (p. 22), “work that displays ingenuity, complexity, and the student’s personal signature” (p. 22). Eisner (1993) also questioned the fairness of applying the same standards to all students and how the information was used after the standards were applied.

In 1993 Down and Mitchell wrote about the process of creating standards for the arts and the related assessment frameworks for the National Assessment of Education Progress (NAEP). Two important motivators, disappearing dollars for the arts and the absence of the arts from the national goals, provided impetus for establishing both the standards and the accountability systems for the arts - music, visual arts, theatre, and dance. Establishing standards and accountability systems were seen as a way for “those who believed the arts are essential in American education” (Down and Mitchell, 1993, p. 32) to get more. A common architecture for all the arts was “divided according to four aspects of learning about the arts: creation and performance, cultural and historical context, perception and analysis, and nature and value of the

arts” (Down and Mitchell, 1993, p. 34). These aspects were identified and accepted based on the belief that “not all students are talented, but all can appreciate and understand they are essential components of being human” (Down and Mitchell, 1993, p.33).

Almost a decade later, Bensure (2002) made four points related to the visual art standards: 1) the standards are ambiguous in nature, 2) it is difficult to define what all students should know and learn, 3) it is left to the teacher to determine what is essential for all students to learn, 4) one or two lessons could cover all the standards. Six content standards for visual art specified what students should know and be able to do: 1) “understanding and applying media, techniques, and processes, 2) using knowledge of structures and functions, 3) choosing and evaluating a range of subject matter, symbols, and ideas, 4) understanding the visual arts in relation to history and cultures, 5) reflecting upon and assessing the characteristics and merits of their work and the work of others, and 6) making connection between visual arts and other disciplines” (NAEA, 1994). The ambiguity of the standards, fact that the standards are not mandatory for school districts and states, combined with the standardized, objective nature of the assessment instrument contributed to a dismal result on the NAEP where half the students taking the NAEP art assessment failed (Dorn, 2002).

### **Beliefs and Attitudes Related to Assessment in Art Education**

Gruber and Hobbs (2000) in their article, *Historical Analysis of Assessment in Art Education* trace the history of assessment in art education, finding overall an aversion by art educators to embrace assessment, favoring process over product. Art experience as a tool for self-expression has dominated art instruction. In the 1920s, however, there was a brief period of respect for assessment with the primary purpose of identifying artistic aptitude or intelligence. Two examples are the McAcory Art Test of 1929 and Goodenough Measure of Intelligence by Drawing, 1926. During the post World War II era art education along with education in general entered a child-centered and self-expression movement based on the work of Freud and Dewey. Emphasis was on the art-making process, not the product. Lowenfeld’s first edition of *Creative and Mental Growth* contained no assessment. Brittain who updated Lowenfeld’s book believed grading in art had no function (Gruber & Hobbs, 2000). Manuel Barkan, who proposed the concept of structure and discipline in art education emphasized progress rather than the art product. Discipline Based Art Education (DBAE) and the Getty, which based their instructional

strategies on Barkan's work, dealt with the issue of evaluation on a theoretical level, if at all (Gruber & Hobbs, 2000).

Adding to this history, Clark (2002) captured a pervading belief about art, "You've either got talent or you don't...' Anybody can pass art.' Such attitudes may exemplify the familiar 'fuzzy' nature of assessment in the arts...In fact, the arts traditionally have been treated as nonacademic subjects; as such, grading and evaluation has been neglected" (§ 1). Schwartz (1996) added, "people unfamiliar with the arts often mistakenly believe that excellence and quality are mere matters of opinion ('I know what I like') and that one opinion is as good as another" (§ 7). These statements, while widely held by the public, are directly counter to the current emphasis on accountability. Added to art teachers overall aversion to assessment, the emphasis on self-expression and progress rather than the art product, the multiple goals for assessment combined with ambiguous standards and the very nature of art, there is no wonder teachers, districts, and states are searching for a way to meet the current accountability demands.

### **Assessment**

The focus of assessment for the past decade has been placed in the accountability arena. Armstrong (1994) defined assessment in such terms, "assessment provides the basis for schools to be accountable to their communities for student learning in all subjects (p. 1). Beattie (1997) defined assessment in broader, more commonly accepted terms as "the method or process used for gathering information about people, programs, or objects for the purpose of making an evaluation" (p. 2). Dorn, Madeja, and Sabol (2004) explained that "assessment, in the broadest sense, involves identification of goals and purposes, selection of procedures, methods, and measures, coordination of timing, analysis of data, interpretation of results, and formulation of responses to the results" (p. 22). Lazear (1999) reminded us that "assessment should be an opportunity to enhance, empower, and celebrate students' learning.

Attitudes about assessment held by art teachers are both positive and negative. In a recent study conducted by Dorn, Madeja, and Sabol (2004) teachers attitudes were collected and ranked. Responses were similar regardless of grade level.

"In order of priority drawback included:

1. 1. Too many students and not enough time to assess.



2. Lack of uniform performance standards, guidelines, procedures, inefficient assessment tools.
3. Changes the focus of art education from art learning to assessment results.
4. Involves too much subjectivity.
5. Inability of assessments to measure a broad range of learning
6. Increased student anxiety, lowered self-esteem, emotional upsets.
7. Inability to accurately and precisely assess personal expression.
8. Stifling of creativity, restrictive.
9. Increased teacher anxiety.
10. Lack of assessment knowledge and training.
11. Assessments drive curriculum.
12. Takeaway from studio time” (p. 21).

Art teachers also listed positive effects of assessment. Again, listed in order of priority:

1. “Makes students more aware of goals for the program and more accountable.
2. Provides feedback for students and teachers about learning, shows growth.
3. Helps students better understand assignments, improves work.
4. Improves student motivation, provides accountability for students.
5. Provides credibility for the art education program.
6. Indicates whether goals and objectives of the program are being met.
7. Improves student self-esteem.
8. Improves teaching and makes teachers more introspective.
9. Improves students’ understanding of their grades.
10. Makes parents aware of the program’s goals.
11. Increases respect from administrators.
12. Motivates student to work harder” (p. 21).

The fifth positive reason, credibility for the program, corresponds with the insistence of arts educators to be included in the standards movement. Lehman (1999) forthrightly stated, “perhaps, the most important feature of the arts assessment is its very existence. The inclusion of the arts within NAEP reinforces the opinion that the arts belong among the disciplines of the curriculum” (p. 13). Hanley (2003) found that the design of the NAEP assessments in the arts, 1971-72, 1978, and 1997, had moved from “an interest in student achievement to the implementation of a curriculum that is in many ways a political document” (p. 34). Zerull (1990) noted the political influence almost a decade earlier stating that the role of testing had begun “to move toward the determination of teacher effectiveness and status of local school districts” (p. 19). He predicted a profound effect on the arts curriculum similar to the effect found on the core curriculum subjects of mathematics, science, and language arts.

Further complicating assessment in arts education, Wilson (1992) found a lack of understanding among arts educators concerning the role and functions of assessment and evaluation. He also found that “when teachers develop tests, they tend to focus their efforts on easily identified, trivial bits and pieces of knowledge and skills rather than on the holistic and integrated dimensions of interpretation, performance, and creation” (p. 35). Teachers saw the principle purpose of assessment and evaluation as a way to assist in grading (Wilson, 1992). Zelman, Daniels, and Hyde (1998) stated:

“under pressure to ‘justify’ grades with copious scores and marks in their gradebooks, teachers expend enormous energy feeding the grading machine – finding ways to quantify, measure, score, compute, and record assorted aspects of kid’s behaviors....This futile expenditure of time should remind us that the main legitimate purpose of evaluation in education is to guide instruction. Anything we do to gather and interpret information about kid’s learning should provide accurate, helpful input for nurturing children’s further growth” (p. 247).

There are two broad categories of assessment, formative and summative. Formative assessment is a strategy used during the course of instruction. It is informal and diagnostic, with little or no emphasis on recording scores (Beattie, 1997; Zelman, Daniels & Hyde, 1998). Summative evaluation, on the other hand, doesn’t aim to nurture, it is formal (Beattie, 1997) and quantifies what has been learned up to a given point translated into a score that can then be used to rank or compare students to one another. Summative assessment is a way to report to outsiders about what has been studied and learned (Zelman, Daniels & Hyde, 1998). Zerull (1990) stated “summative evaluation suggests that there is a terminal point of arrival” (p. 20), as well as a specific set of goals, facts and outcomes that can be measured. He continued indicating that summative evaluation was “not an appropriate assessment in the arts because it can reasonably examine only limited memory for facts and technical judgments” (Zerull, 1990, p. 20). Formative assessment is the style of assessment Zerull (1990) recommended for the arts.

### **Standardized Tests**

Traditional standardized tests are summative assessments that require students to show their knowledge in a predetermined manner (Scholastic,1999), create artificial situations that “look like no other situations in life and are used to determine if a student can perform in a particular way” (Eisner, 2002, p. 181). Typically, academic knowledge was assessed with

arbitrary problems that a student had little intrinsic interest in or motivation to answer (Gardner, 1991; Resnick, 1993). The ends were well defined and used to formulate the means (Eisner, 2002) and the answers were “correct answer compromises” (Gardner, 1999, p. 208). In addition, achievement tests provide very limited information even in the domains tested (Eisner, 2002). Brown (1989) expressed a similar perspective, “The concept that testing is initiated externally from the student, separate from the learning process, and primarily aimed at determining whether inert knowledge is in students’ short-term memories exercises far too much influence over school people today” (p. 33). Wolf (1989) was more emphatic stating:

“...much school-based assessment actually *prevents* students from becoming thoughtful respondents to, and judges of, their own work. The ‘surprise’ nature of many test items, the emphasis on objective knowledge, the once-over and one-time nature of most exams - all offer students lessons that are destructive to their capacity to thoughtfully judge their own work: (1) assessment comes from without, it is not a personal responsibility; (2) what matters is not the full range of your intuitions and knowledge but your performance on a slice of skills that appear on tests; (3) first-draft work is good enough; and (4) achievement matter to the exclusion of development.” (p. 35).

Gardner agreed. He felt what American students needed was to get “more deeply interested in things, more involved in them, more engaged in wanting to know; to have projects they can get excited about and work on over longer periods of time; to be stimulated to find things out on their own” (Brandt, 1987, p. 33).

Shepard (1989) made two important points concerning large-scale testing, first, “large-scale assessments must be formal, objective, time-efficient, cost-efficient, widely applicable, and centrally processed” (p. 7) Second, the results of these assessments must be in a “form useful to policymakers, which usually means reducing the complexity to a single score.” (p.7). Mehrens (1992) also found that any assessment used for accountability purposes had to be “administratively feasible, professionally credible, publicly acceptable, legally defensible, and economically affordable”(p. 3). Mehrens (1992) also pointed out that paper and pencil objective tests are biased, measure irrelevant content in a format that “demands only the ability to recognize an answer - not to actually work problems” (p. 4) and that there is neither a breath of content inclusiveness nor a perfect match with the curriculum. Yet, as Gardner (1991) stated:

“In the United States, so-called ‘standardized tests’ have been developed to the greatest extent. The tests typically feature multiple-choice items or very short answers; in so far as possible the test is scored by machine, and the results locate students within various target groups with statistical precision” (p. 133).

Brown (1989) found that almost everyone they talked to was; 1) determining educational success or progress based on scores on commercial standardized norm-referenced tests and 2) believed these tests are valid indicators of student and school achievement. However, Brown (1989) cautioned standardized tests: 1) have serious limitations, 2) can be subject to misuse, and 3) do not address higher-order thinking skills. Mehrens (1992) countered Brown's last point and argued that the notion that multiple-choice items cannot measure higher order thinking skills is incorrect. He further supported the use of multiple-choice objective tests stating, "many different educational tasks require simple recall, particularly tasks in the lower grades and in introductory courses" (p. 5). O'Neil (1993) found that many critics were "concerned that it's premature to construct an examination system designed for high-stakes uses ultimately planned for that system, such as judgment about the success or failure of individual students or schools" (p. 6).

Many states have responded to the accountability movement of the last decade and the increased public access to student and school achievement on standardized tests by developing art assessments. Art educators have acknowledged several concerns. Sabols's (1998) content analysis of state art assessments discovered that only 1% of the items were product oriented while 99% reflected traditional paper and pencil testing methods. Brandt (1987) recognized the need for assessment due to the expense of art programs but cautioned art educators not to make the error of assessing the arts in the same way other kinds of learning are assessed, i.e. multiple-choice tests. Dorn (2002) pointed out another concern related to the content of tests:

"...test, which is designed as a standardized test, requires every student in every school to answer the same question and thereby assumes all U.S. art teachers teach - or ought to teach - the same art activities in the same way and that all students are likely to receive the same amount and quality of art instruction in all U.S. Schools" (§ 12).

Cannatella also (2001) expressed concern over this standardized approach to assessment in art education "...the particular character and activity that goes into the making of art does not fit comfortably into any system of general assessment criteria. One major reason for this is that the individuality of the art work and its making cannot be reduced into common properties independent of and detached from what is being creatively fashioned." One of Eisner's (2002) concerns focused on a related dichotomy. The dichotomy between the facts that art educators have been encouraged to think about curriculum planning in terms of intended outcomes while

“at the same time, one of the features of the arts is the encouragement of improvisation and the cultivation of a personal rendering of one’s ideas” (p. 160).

Another concern Cannatella (2001) had, regarding standardized testing in art education, was whether 50 minutes spent on multiple choice and essay questions related to aesthetics, criticism and art understanding and only 30 - 40 minutes spent on performance tasks was an accurate measure of student achievement given that perhaps up to 90% of all school art programs are devoted to creative studio activity. He stated, “It is often unconvincing to assess art work only on a conceptual-perception framework where aesthetic experience is involved...Indeed, the perception of the art form is inextricably linked to an understanding of particular mediums” (p. 324).

In addition, one must keep in mind that each assessment seeks a different goal: “the NAEP to test student knowledge to compare kids and schools, the Getty to measure the knowledge, skills, and attitude of discipline-based art education (DBAE), and the states Goals 2000 assessment whatever goals are sought by the state departments of education “ (Dorn, 1998, ¶ 6).

Wolf (1989) stated, regarding assessment in schools, “Never do we stop to ask how we could make our evaluative gate keeping model the kind of self-observation and informed critique that separates ball tossers from fine pitcher, doodlers from artists, or instrumentalists from musicians” (p. 35).

### **Alternative Assessment**

Art teachers overall aversion to assessment, their emphasis on self-expression and progress rather than the art product, the multiple goals for assessment combined with the ambiguous standards and the very nature of art have left educators confused and questioning how and what to assess. The conceptual-intellectual framework of many standardized tests has created concerns about what we are testing and the validity of those tests. Are we testing what we want students to know and be able to do in the arts? Do current assessments measure what students are learning in art classrooms and are these assessments related to art itself? Dorn (1998) stated, “For us to know that the student knows, then requires the adoption of an evaluation process that is concerned with assessing the processes of thinking and making, manifested in the products and methods of inquiry required for learning in the arts” (¶ 19).

Resnick (1994) believed, “New forms of assessment, assessments based on complex task performances, scored by trained and thoughtful judges, can [could] release educators from the grip of testing programs that drive instructional attention away from thoughtfulness and complex application of knowledge” (p. 410). Efland (2002) suggested that, “evaluation of instruction in complexly structured domains [i.e. art] should show evidence of knowledge assembled from many cases or examples and require looking for knowledge application in new situations, that is, transfer”(p. 105). Alternative assessments can measure the thoughtfulness and complexity of art education.

Armstrong (1994) stated, “Alternative assessment fulfills a need for valid assessment in art where actual outcomes of learning are difficult to observe by traditional measures. Direct observation of student performance is consonant with the curriculum-instruction-assessment relationship in art” (p. 132). Beattie (1997) defined alternative assessment as “assessment that deviates from traditional pencil-and-paper item formats” (p. 3). Armstrong (1994) stated the definition in terms of “alternative to traditional (objective tests and essays)” (p. 110). Authentic and performance assessment strategies represented two main categories of alternative assessments according to Beattie (1997). Authentic assessments are describe as realistic, meaningful, utilize open-ended problems, involve conceptual and higher order activities, are true to the discipline, and involve students in tasks that are worthwhile (Beattie, 1997; Dorn, 2002). Cole, Ryan, Kick, and Mathies (2000) described the components of authentic assessment and claimed authentic assessment: “assists in learning, encourages good instruction, relates to curriculum outcomes, fosters higher order learning, follows developmental perspective, uses testing sparingly, supports time efficiency, reports meaningful information, promotes partnering of parents, educators, and students, fosters student metacognition and reflection and is individualized” (p. 7). They also claimed, “Authentic assessment correlates with what students are learning and helps them gain confidence in their ability to master the subject” (p. 7).

Performance assessments refer to the “task, method, or activity used to assess the performance in which students demonstrate their achievement and learning” (Beattie, 1997). Students can develop a process or a product or both. Performance assessments can include a wide range of strategies including: 1) portfolios, 2) journals, diaries or logs, 3) integrated performances, 4) group discussions, 5) exhibitions, 6) audio tapes and video tapes and 7) computers (Beattie, 1997).

Assessment Beattie (1997) believed had a larger role to play in education than simply reporting to parents and the community. She recognized the power of effective assessment techniques to “improve classroom instruction, empower students, heighten student interest and motivation, and provide teachers with ongoing feedback on student progress” (p. 2).

Art educators, particularly at the secondary level, have historically used the performance assessment strategy of portfolio assessment. Often these portfolios were collected with the goal of seeking entry into college or credit for advanced work. Gardner (1996) stated, “In the familiar version of portfolio, a student assembles the best of his or her work, usually in the effort to gain admission to a selected program, to win some kind of prize, or to secure showing of works” (p. 144). Product portfolios, such as these, represent only one type of portfolio.

Clarke (2002) identified two types of portfolios that teachers have historically used for grading, either the process portfolio or the product portfolio. Clarke (2002) stated, “Process, in this case, refers to the entire range of activities required to produce a work of visual or performance art from concept to creation...Product, refers to the aesthetic quality of the student’s finished result or work of art” (¶ 3). Walker (1998) elaborated on the contents of a process portfolio. According to Walker (1998) process portfolios included “the entire collection of work done on an assignment leading up to and including, the final solution to the project... it may include, but not be limited to writing, concept maps, thumbnail sketches, rough sketches, computer generated roughs and the final solution to a given problem” (p. 2).

Beattie (1997) defined four types of portfolios the best-works portfolio, the expanded art portfolio, the mini-portfolio and the process folio. First, the best-works portfolio “includes only the students’ best work, accumulated over a specified time period” (p. 15). Second, the expanded art portfolio is a “thick base of information representing rich qualitative and ample quantitative evidence pertaining to a wide variety of thinking processes and art products” (p. 15). Third, the mini-portfolio is “an abbreviated portfolio that focuses on a collection of art products based on a single theme, project or artistic activity; a collection of the four disciplines; or a single unit of study” (p. 15). Fourth, the process folio is “a portfolio of evidence about student growth and learning, with an emphasis on process” (p. 15).

Schwartz (1996) recognized a change or growth in the portfolio process from a strategy that consisted of “mainly finished art products, to include samples of conceptual ideas, art works in progress, written reports and essays, test samples, photographs and videos, audiotapes,

and journals and self-critiques thus presenting a comprehensive picture of a student's progress and achievement" (§ 2). Gardner (1996) termed process portfolios, processfolios. These processfolios were a way of furthering the student's own learning by recording a student's progress on a project including: "an initial idea, early sketches, false starts, pivotal pieces (where an idea gels), journal entries (in whatever medium seems appropriate), interim critique and self-critiques, the final product, critiques of that product, and plans for further revisions of the project or for new projects that in some way build upon the works in the processfolio" (p. 144). Walker (1998) stated, "by using Process Portfolios, an instructor is able to evaluate a student's understanding of and effort in the design process separately from the final product."

Beattie (1994) expressed a concern that a portfolio may result in a fragmented assortment of unrelated tasks if attention was not paid to the idea and process as well as the product. She defined the terminology: 1) Idea - "the conceptual developmental process", 2) Process - "carrying into the effect the outcomes of the conceptual process," and 3) Product - "the final synthesis of Idea and Process, or the final work pieces" (p. 14). Dorn (1998) also believed the assessment plan should center on more than the final product and should focus on the student's development of "artistic intent, expression and skill that made the vision possible" (p. 3).

Schwartz (1996) and Huffman (1998) found creating portfolios empowered students. They believed students should be involved in selecting the criteria for the portfolio's assessment. Huffman (1998) believed student involvement in the criteria development process built confidence in the student's work and ability to make decisions, established clarity about the assessment criteria and created a commitment on the part of the student to the assessment process. McCollister (2002) found, "Student articulation is enhanced when brainstorming the content of rubrics...The student's ability to use language that represents varying cognitive and manipulative processes and results is practiced and refined, and; as expectations are clarified and strengthened, student work improves" (§ 26).

Eisner (2002) identified three features or criteria for assessing students' artworks: 1) the technical quality of the work produced, 2) the extent to which it displays and inventive use of an idea or process and 3) the expressive power or aesthetic quality the work displays. For Eisner (2002) criteria are "features one can look for in a work; they are not fixed descriptions that obey some formulaic recipe" (p. 184).



Portfolio and authentic assessment strategies often rely on the use of rubrics as the main evaluation tool. However, not all rubrics are alike. Weidmer (1998) identified three kinds of rubrics, the holistic, the primary trait rubric, and the analytic rubric. Teachers primarily make use of the analytic rubric when either grading or identifying the criteria for a project. McCollister (2002) stated, “Rubrics can be particularly serviceable to inform students of ways to improve their learning and to inform teachers of ways to improve their teaching”(¶ 2). McCollister (2002) found that, “The teacher can use the rubric as a tool in intervention to inform a conversation with the student concerning work-in-progress. This informed dialogue can assist the learner with 1) increased personal responsibility, 2) independence, 3) accurate self-assessment, 4) application of concepts, 5) selective attention to what is important” (¶ 15). Walker (1998) expressed similar formative uses of the analytic rubric. Goldsby (2001) advised one to decide on the number of levels of performance needed, describe the standards of performance to use as benchmarks, use demonstrative verbs that describe observable behavior. Huffman (1998) summarized that, “The content for rubric assessment and evaluation can include knowledge and understanding about the arts, including the student’s personal, historical, cultural, and social contexts for art, and his or her perceptions, technical, expressive and intellectual skills” (p. 64).

However, McCollister (2002) also warned, “Extensive use of criteria rubrics can hamper personal responsibility, creativity and independence” (¶ 23). Like many educators (Schwartz, 1996; Cannatella, 2002; Dorn, 2002; Clarke, 2002), McCollister (2002) supports the use of a variety assessments to evaluate student cognition, affect and psychomotor skills. Critiques, tests, performance samples, and observations are some of the assessment strategies suggested. Clarke (2002) emphasized the role of student performance stating, “In a standards based curriculum, student performance generally is expected to be the most important indicator that learning has taken place” (¶ 2).

While, Beattie (1994) and McCollister (2002) found verbalizing, student articulation and journal entries as playing a role in student assessment and learning Cannatella (2001) and Eisner (2002) emphasized assessment based on performance. Canatella (2001) recognized non-linguistic practices stating, “A carpenter, a potter, a musician, a dancer, and a furniture maker can all be ‘masters’ of their respective crafts without developing substantially the skills of linguistics” (p. 232). Eisner(2002) believed student artworks were important indicators of what the student had learned.

He stated:

“An important data source for determining what students have learned is the artwork that students create. These works are potent sources of evidence regarding the outcomes of effective arts education when competence in qualitative reasoning is an educational aim. The works that students create can serve as proxies for their ability to think within the affordances and constrains of a material, to employ their imaginative abilities, to apply technical skills and, indeed, to use the various forms of thinking I have already described. The work is an expression, a representative of their ability to think intelligently about the perception and creation of the visual arts. If you want to know what student have learned in visual arts, one way to find out is to look at their artwork and to compare it with their earlier efforts” (Eisner, 2002, p. 92)

Gardner (1996) in the ARTS PROPEL project identified four main areas for assessing processfolios; production, reflection, perception, and work. First, production involves thinking in the domain. “The evidence for assessing work on the dimension of production lies in the work itself” (p. 152). Four areas were identified as sources of the evidence: 1) Craft, 2). Pursuit , 3) Invention, and 4) Expression. Second, reflection involves thinking about the domain. “The evidence for assessing reflections comes from the student’s journals and sketchbooks, and from observations of the kinds of comments the student makes in class” (p. 152). Five areas were examined: 1) the ability and proclivity to assess own work, 2) the ability and proclivity to take on the role of critic, 3) the ability and proclivity to use criticisms and suggestions and 4) the ability to learn from other works of art within the domain. 5) the ability to articulate artistic goals. Third, perception involves perceiving in the domain. “The evidence for assessing a student’s perceptual skills comes from the student’s journal entries and from observations of the student’s comments made in critique sessions” (p. 152). Three areas contributed to the perception domain: 1) the capacity to make fine discriminations about works in the domain, 2) the awareness of sensuous aspects of experience, 3) the awareness of physical properties and qualities of materials. The fourth domain is the approach to work. “The evidence for assessing a student’s approach to work lies in observations of the student in classroom interactions and in the student’s journal entries.” (p. 152). Four areas are examined: 1) engagement, 2) the ability to work independently, 3) the ability to work collaboratively, and 4) the ability to use cultural resources.

Processfolios contain evidence of students’ metacognitive processes as well as their cognitive processes. Gardner’s (1996) list of dimensions that illuminate the unique potential of processfolios follows:

- Awareness of their own strengths and weaknesses [metacognitive]

- Capacity to reflect accurately [metacognitive]
- Ability to build upon self-critique and to make use of critiques of others [metacognitive]
- Sensitivity to their own developmental milestones [metacognitive]
- Ability to use lessons from domain projects productively [cognitive]
- Capacity to find and solve new problems [cognitive]
- Ability to relate current projects to those undertaken at earlier times and those that they hope to undertake in the future. [cognitive and metacognitive]
- Capacity to move comfortably and appropriately from one aesthetic stance or role to another and back again.” [metacognitive] (Gardner, 1996 p. 144).

Efland (2002), examined the work of Koroscik, Perkins and Prawat identifying relationships between the metacognitive and the cognitive. Three groups of difficulties were identified that can affect students’ intellectual development: 1) inadequacies in the learner’s knowledge base, 2) weak or inappropriate knowledge-seeking strategies and 3) a poor disposition to learn. Perkins (1994) stressed the linkage between students’ state of willingness or unwillingness to learn and their selection of appropriate knowledge-seeking strategies while Prawat (1989) distinguished student dispositions as either a performance or a mastery disposition. “A performance orientation may direct the student to get the job done as quickly as possible, whereas a mastery orientation directs the student to increase knowledge because he or she is intrinsically interested in the material.” (Efland, 2002, p. 118). Processfolios provide evidence revealing not only the students’ end results but their thinking, amount of effort, attitudes, and problem-solving strategies.

### **Digital Portfolios**

As the paper portfolio systems of the 1990’s began to overflow folders and file cabinets and posters, artwork, video and audio tapes were added, teachers felt there had to be a better way to deal with these collections of student performances. Niguidula (1997) reported on a team working with the Annenberg Institute for School Reform and the Coalition of Essential Schools with the support of IBM. They investigated one technology, digital portfolios from 1993-1996 with six Coalition Schools in multiple states. Five areas were considered in the planning and implementation: 1) vision (school) - what should students know, 2) assessment - how to demonstrate the vision, what to collect, how to identify quality, 3) technology - hardware, software, networking, 4) logistics - when to digitize, who, who selects the work, 5) culture - Is work used to discuss student work? Is the school open to tuning standards? Two important

points were made in the article. First, “Without time for reflection, the digital portfolio might be no different from a paper portfolio filed away in a locked cabinet. Time for reflection is critical” (Nicuidula, 1997, p. 28). Second, “ The key elements of a school’s culture that make a digital portfolio system work are the relationships within the school, regular discussions of student work, and an openness to examine the school’s vision with others outside the school” (Nicuidula, 1997, p. 28).

Weidmer (1998) added to Nicuidula’s five areas stressing the importance of not focusing evaluation too strongly on the appearance of the final portfolio product. The performance is more important than the style or look of the portfolio. Weidmer (1998) also encouraged the establishment of a list of standards, competencies or proficiencies against which all products were judged and recommended feedback from both the formative and summative perspectives.

“It is essential for the developer and designer of CD-ROM [digital] portfolios to take the time to ensure that stakeholders draw accurate conclusions about 1) what the portfolios show and how they are representative of performance, 2) how the portfolios employ clear criteria to represent what is valued in model performance, and 3) when portfolios provide more authentic views of what individuals know and are capable of performing” (Weidmer, 1998, ¶ 29).

Weidmer recommended the digital portfolio be saved in a CD-ROM format, listing the following advantages: CD’s store up to 650 megabytes, are inexpensive, have little weight, data cannot be erased, are in an incorruptible format, the data encoding process is standardized and any computer can read them. One disadvantage was noted: if files were updated frequently, the cost and time for burning new CD’s increased.

### **Primary Source Articles**

Sherry, Billig, and Jesse (2001) described a portfolio WEB project that was part of a Technology Innovation Challenge Grant, conducted in Vermont K-12 schools and completed in September of 2002. The purpose of the project was “to infuse standards-based instruction in multi-media, digital art, music composition and online discourse into the general arts and humanities curricula of Vermont K-12 schools” (¶ 1). Methods included students who shared and received feedback via a WEB server from teachers, students, digital artists, traditional artists, musicians, composers, web designers and other experts. The students used and filtered the feedback for use in making their final artistic projects. They extended the Sternberg Developing

Expertise model to define “expertise” as student achievement measured by teacher created rubrics. The Sternberg model’s premise is motivation drives metacognition which in turn stimulates the development of thinking and learning skills. A rubric for measuring student learning processes was also selected from Marzano, et al’s (1993) Dimensions of Learning model. A survey was administered to 165 students in nine schools in January of 2000 and again in May 2000. Seventy-five percent of the students surveyed were high school students, 25 % were middle school students.

Sherry, Billig, and Jesse (2001) found there was a significant correlation between motivation and metacognition and all the students met the standards for both the teacher-created student product assessment and learning process assessment. Their conclusion stated, “These preliminary findings suggest that teachers should emphasize the use of metacognition skills, application of skills and inquiry learning as they infuse technology into their respective academic content areas” (Sherry, Billig and Jesse, 2001, ¶ 17). They suggested this study be replicated with several cautions: 1) articulate specific knowledge and skills, 2) carefully observe student behavior, and 3) create and benchmark rubrics that are used to evaluate student work.

Dorn, Madeja and Sabol’s (2002, 2003) “Models for Assessing Art Performance Project” (MAAP) 2000-2001 portfolio assessment project involved three Universities, Florida State University, Purdue University and Northern Illinois University, 2000 portfolios, 16,000 student artworks, 3 states, and 15 school districts. The focus was on meeting three needs: 1) how to administer an authentic assessment model for evaluation of the teachers own students, 2) to help teachers develop an assessment plan for their classrooms and schools, 3) to devise a data collection system that meets the needs of art students, schools, and the state and national art assessment standards.

In the MAAP study it was found that art teachers can with training “govern themselves and set their own standards for providing valid and reliable estimates of their own students’ performances” (Dorn, 2002, ¶ 24). The project rubrics in the authentic assessment setting could effectively measure students’ expressive outcomes and “student art performance and progress will vary among different classrooms at different grade levels and in different school districts” (Dorn, 2002, ¶ 24). A significant level of agreement, ( $p > .01$ ) between the art teacher and the other judges was found at all grade levels. There was 85% to 90% score agreement among the raters, establishing inter-rater reliability.

Study recommendations included teachers in each state developing their own authentic assessment instruments, development of school and district assessment plans which use a peer review process, and development of ways to document student progress and establish sensible and appropriate record-keeping systems that will meet agreed upon goals of the district and state. In addition, early results from scoring digital portfolios using the same methodology suggest the use of digital portfolios could be reliably used as an alternative to the physical portfolios (Dorn, 2003).

## **CHAPTER III**

### **METHODS**

#### **Statement of the Problem**

In light of current concerns focusing on the relationship of cognition and learning in art and on accountability, this research project will focus on the effect inclusion of cognitive and metacognitive/self-regulatory activities, specifically, student goal setting, self-evaluation and self-reflection, have on students' progress in two-dimensional visual art production. The digital portfolio assessment techniques and the holistic scoring methodology initiated by Dorn, Madeja, and Sabol (2003) will be used to determine the effectiveness of the intervention. These methods were designed specifically to support the expressive, "ill-structured" (Efland, 2002) nature of art education. In addition, research by Dorn, Madeja, and Sabol (2003) suggested that students showed little or no gain in art performance in grades 4 and 5. Their study took into account maturation and a teacher intervention focused on instruction in more advanced studio skills, a means of improving content. To see if this lack of gain on the part of the students can be counteracted, students in grades 4-5 are the primary focus of this study.

#### **Hypotheses**

The research hypothesis states that setting annual goals and answering self-reflection and self-evaluation questions after each art assignment will contribute positively to fourth and fifth grade student performance in visual art classes as measured by an authentic portfolio assessment process established by Dorn, Sabol, and Madeja (2003).

H1 The alternative hypothesis states that students' mean art portfolio scores will increase as a result of the intervention, metacognitive/self-reflective journal writing.

H2 The null hypothesis states that the extraneous variable of student FCAT-NRT reading scores will not affect student art performance.

H3 The null hypothesis states that the extraneous variable of gender will not affect student art performance

H4 The null hypothesis states that the extraneous variables of age, free/reduced lunch status, instructional time and budget will not affect student art performance.

## **Research Design**

A quasi-experimental design was used for this study. The design of the study consisted of an experimental group and a control group. The experimental group underwent an intervention of metacognitive/self-reflective journal writing while the control group did not participate in any writing or assessment activities. The design was a static-group comparison design,  $X_1 0, X_2 0$ . Each art teacher included in the study, chose two intact classes of students from the same grade level, either two fourth grade classes or two fifth grade classes. One class acted as the control group which, only had their artworks digitally collected. The other class served as the experimental group, who in addition to having their artworks digitally collected, also set annual goals, answered self-reflection, and self-evaluation questions saved in paper journals. The treatment for the experimental group began with the completion of the first artwork of the year. At that time, the students filled out the first set of journal questions. The sample was a convenience sample. Information on the following extraneous variables was collected: FCAT (Florida Comprehensive Achievement Test) reading scores, grade, gender, free-reduced lunch status, ethnicity, hours in art class, and art budget to determine their impact if any on the study.

## **Description of the Sample**

A convenience sample of volunteer elementary art teachers in Leon County Florida selected two classes of the same grade level, either two fourth grade classes or two fifth grade classes were to participate in this study. Initially, nine elementary schools, one middle school and two high schools chose to participate in the study. Seven schools, one middle school and six elementary schools completed the portfolio process. Five elementary schools collected the required number of images for inclusion in the study. The middle school and one elementary school did not collect the required data for scoring and analysis. Four schools chose two fifth grade classes; one school chose two fourth grade classes. English Speakers of Other Languages (ESOL) and Exception Student Education (ESE) students in inclusion classes were included in the study.

To be included in the study, the students needed to be present from the start of the school year to the end, August to May. The five schools included in the study for statistical analysis provided student portfolios where the portfolios consisted of a minimum of three artworks for the pretest and three artworks for the posttest.



The schools included in the study represented diverse backgrounds. The control group at school A consisted of students with: 1) an average age of 11, 2) was 45 percent male, 3) was 61 percent minority, 4) had a free or reduced lunch percentage of 64 percent, 5) had an average FCAT reading NRT of 56.5 percent, 6) had 22.5 hours of 2- D art instruction per year, and 7) had an annual budget of \$2 per student per year. The experimental group at school A consisted of students with: 1) an average age of 11, 2) was 56 percent male, 3) was 78 percent minority, 4) had a free or reduced lunch percentage of 61 percent, 5) had an average FCAT read NRT of 63.12 percent, 6) received 25.5 hours of 2-D art instruction per year, and 7) had an annual budget of \$2 per student per year. Both groups were made up of fifth grade classes.

School B's control group consisted of students with: 1) an average age of 10.3, 2) was 47 percent male, 3) was 47 percent minority, 4) had a free or reduced lunch percentage of 59 percent, 5) had an average FCAT reading NRT of 70.1 percent, 6) had 29 hours of art instruction per year, and 7) had an annual budget of \$3 per student per year. The experimental group consisted of student with: 1) an average age of 10.6, 2) was 37 percent male, 3) was 50 percent minority, 4) had a free or reduced lunch percentage of 56 percent, 5) had an average FCAT reading NRT of 73.3 percent, 6) received 29 hours of art instruction per year, and 7) had an annual budget of \$3 per student per year. Both groups were made up of fifth grade classes.

School C's control group consisted of students with: 1) an average age of 10.6, 2) was 52 percent male, 3) was 29 percent minority, 4) had a free or reduced lunch percentage of 12 percent, 5) had an average FCAT reading NRT of 72.33 percent, 6) had 23.5 hours of art instruction per year, and 7) had an annual budget of \$5.15 per student per year. The experimental group consisted of student with: 1) an average age of 10.7, 2) was 75 percent male, 3) was 20 percent minority, 4) had a free or reduced lunch percentage of 0 percent, 5) had an average FCAT reading NRT of 80.82 percent, 6) received 29.5 hours of art instruction per year, and 7) had an annual budget of \$5.15 per student per year. Both groups were made up of fifth grade classes.

School D's control group consisted of students with: 1) an average age of 10.6, 2) was 31 percent male, 3) was 0 percent minority, 4) had a free or reduced lunch percentage of 15 percent, 5) had an average FCAT reading NRT of 59.8 percent, 6) had 13.5 hours of 2-D instruction per year, and 7) had an annual budget of \$1 per student per year. The experimental group consisted of student with: 1) an average age of 10.4, 2) 0 percent male, 3) was 8 percent minority, 4) had a

free or reduced lunch percentage of 58 percent, 5) had an average FCAT reading NRT of 72.18 percent, 6) received 12.5 hours of 2-D instruction per year, and 7) had an annual budget of \$1 per student per year. Both groups were made up of fifth grade classes.

School E's control group consisted of students with: 1) an average age of 9.48, 2) was 45 percent male, 3) was 25 percent minority, 4) had a free or reduced lunch percentage of 24 percent, 5) had an average FCAT reading NRT of 76.33 percent, 6) had 30.75 hours of art instruction per year, and 7) had an annual budget of \$3 per student per year. The experimental group consisted of student with: 1) an average age of 9.53, 2) was 48 percent male, 3) was 26 percent minority, 4) had a free or reduced lunch percentage of 26 percent, 5) had an average FCAT reading NRT of 74.3 percent, 6) received 28.5 hours of art instruction per year, and 7) had an annual budget of \$3 per student per year. Both groups were made up of fourth grade classes.

### **Description of the Instruments Used**

After completing an artwork the art teacher or university student assistant digitally recorded the students' artwork and inserted the image into the students' individual PowerPoint portfolio. A template was provided. At minimum, the digital portfolios were to include the first three (pretest) and last three (posttest) artworks of the year, August to May. In the case of the schools involved in this study, digital cameras were used to digitally record the students' works of art. A PowerPoint presentation (portfolio) was created for each student in both groups, control and experimental with a black background for each slide. These portfolios were kept in class folders on either the teacher's computer desktop or on the school server. A coding system was used to assure anonymity. Numbers identified the students. Letters identified the schools. Before scoring the digital portfolios, the teachers made four distinct portfolios, two for each class. Portfolio 1/Control contained all the students in the control groups' pretest artworks. Portfolio 1/Experimental contained all the students in the experimental groups' pretest artworks. Portfolio 2/Control contained all the students in the control groups' posttest artworks. Portfolio 2/Experimental contained all the students in the control groups' posttest artworks. These combined portfolios contained each student's first 3 to 4 artworks (pretest) or each student's last 3 to 4 artworks (posttest) and the student's identification number. For example for the pretest the slides ran: student 1, artwork 1, artwork 2, artwork 3, student 2, artwork 1, artwork 2, etc. Using the same format, combined posttest portfolios were created. Each combined portfolio was burned

on to a separate CD-ROM. The researcher coded and covered all identifying markings with removable CD labels and attached a score sheet which contained the hidden code.

The digital portfolios were scored in early June using the holistic rubrics for standards-based assessment in the visual arts (appendix A). Three independent art teachers (raters) blind scored each portfolio. The raters were trained in the use of the holistic rubrics immediately prior to the judging session to help benchmark the use of the rubrics. After the training session, the six raters moved to the computer lab to begin blind scoring each portfolio.

The computer lab contained 25 Dell GX260 Pentium IV computers with 15" monitors. The computers were equipped with CD-ROM drives, floppy drives, and network connections. Windows XP was the operating system and Microsoft PowerPoint was the program used for viewing the digital portfolios. The raters moved from computer to computer scoring one portfolio at each station. When the raters completed scoring a school's portfolios they covered the scores with post-it-notes before the next rater moved to that station. Each portfolio was scored three times. Three moderators assisted in the scoring process. Their primary functions were to: 1) set up each portfolio for scoring, 2) ascertain that three different raters independently blind scored each portfolio, 3) check the raters' scores within each class for a sufficient score spread and 4) make sure no art teacher scored her own students' work. The researcher collected the CD-ROM's, experimental groups' journals, score sheets and teachers' anecdotal notes and samples of any other assessments administered.

The portfolios were scored using the holistic rubrics (appendix A) developed by Dorn, Madeja, and Sabol (2004) which matched the grade level of the students, 3-5. The scoring process was based on other portfolio or performance based assessments such as National Board of Teaching Standards (NBPTS), Advanced Placement (AP), and International Baccalaureate (IB), which require agreement of two judges. In the AP portfolio process used in this study a 1-point difference in scores between judges was considered acceptable, but a wider spread would indicate that the judges were looking at different features of the work or suffering from "reader fatigue". The 1-point difference standard has been used by the Educational Testing Standards in scoring the studio art program and similar measures are employed by the IB program (Dorn, Madeja and Sabol, 2004). The scoring for this study followed the AP format. When there was a disagreement of more than one point, the judges with scores covered were asked to look at the portfolio again. Then asked what score they thought they had given the portfolio. In most cases,

one judge or the other changed their score, which brought the scores to within a point of agreement. However, some judges chose to leave their original score and a greater than 1-point variation was accepted.

The six art teacher raters were all elementary art teachers from Leon County with ten or more years of teaching experience at the elementary level. Five of the six art teachers were nationally board certified. One of the nationally board certified art teachers also had experience benchmarking art portfolios for the National Board of Professional Teaching Standards. There were two moderators. One was the designer of the portfolio assessment methodology and a university professor. The other moderator works at the county office writing and modifying textbooks and assessments. She participated in all of the art teacher training sessions and had a firm grasp of the processes and procedures for both the collection of the data and the scoring process.

## **Explanation of the Procedures**

### **Theoretical Perspective**

There were two focuses for this study. First, could student art performance be enhanced by developing students' metacognitive awareness through answering self-evaluation and self-reflection questions upon the completion of each artwork? Second, could student accountability issues be addressed through an authentic assessment process designed to enhance student performance?

The study was originally designed to include instruction in both the methodologies of the digital portfolio process and in the use of metacognitive/self-regulated learning strategies for both the teachers and the students in the experimental group. However, the participating teachers felt all they could handle the year of the study was learning the digital portfolio assessment process. Only one teacher and school had prior training in self-regulated learning theories, school A. After meeting with the teachers, it was agreed the students would write in their journals without formal training in metcognitive/self-regulated learning instruction. The journal writing therefore, would provide baseline information on the effect answering the journal questions had on students' art performance as well as indicate what instruction in metacognitive/self-regulated learning would be needed in the future. The teachers' primary motivation for participating in the study was to meet accountability demands as a part of the

teachers' annual evaluation or to qualify for the school district pay for performance bonus. In both cases, the teachers would take the results of each student's pretest and compare it to their posttest results to determine individual student growth. For example a teacher's annual student performance measure might read, "eighty percent of the students in fifth grade will maintain a level 3 or 4, or show improvement." Results for this purpose were not the focus of this study and were not reported in this study. Teachers used the raw data provided by the study to determine whether or not the criteria were met for their principal or the district.

Adding another layer of complexity to the intervention, all but one participating teacher felt the need to do assessments in addition to the journal writing with their experimental groups. Accountability to parents and administrators seemed to be the primary motivation. The teachers all agreed to refrain from giving assessments or writing activities to the control groups. The general assumption made by the teachers was that assessment has a positive effect on student art performance. As a result, the findings of this study provided insight into the effectiveness of current testing, assessment, and evaluation practices in relationship to student art performance and the literature on self-regulated learning.

In addition, computer experience, familiarity with the portfolio software, and the attitude of the teacher toward the project may have affected the implementation. The researchers main motivation to include schools in addition to her own in the study was an attempt to increase the size of the sample and as a service to the teachers in the district who were attempting to demonstrate student growth in their visual art classes to meet accountability demands.

### **Training of the Teachers**

During the 2003-2004 school year, the art teachers met six times at monthly or bi-monthly intervals with the researcher for over 15 hours of training, scoring and support sessions. The teachers and researcher decided that for the year of the implementation, the art teacher could choose to have the students write in their journals without formal training in metacognitive/self-reflective learning instruction. However, the art teachers were encouraged to model possible answers, review objectives of the lesson, and ask probing questions. In addition, the teachers would continue to use writing, checklists, and tests as they had been doing in prior years. The control group would only make art. The students in the control groups did no writing, assessments, or tests for art class.

The first training session focused on the goals, methods for the study, setting criteria for district individual improvement plans (accountability) and refining the journal questions. The October, November, and February sessions focused on how to create the digital portfolios, how to take better photos, solve technological issues, and how to work effectively with the university students.

The June session was a scoring and training session. During the scoring/training session six art teachers, raters, were introduced to the holistic rubrics, looked at a series of portfolios that represented sample portfolios rated 1 - 4 and then scored three sets of sample portfolios. The raters were trained in the use of the holistic rubrics immediately before the judging session to help benchmark the use of the rubrics. This training session used traditional paper portfolios to facilitate whole group instruction in the scoring process. Prior work by Dorn, Madeja, and Sabol (in press) indicated that there were no significant differences between the scores actual portfolios received compared to their digital representations. Following the instruction on the application of the holistic rubrics, the raters looked at four traditional paper portfolios and were asked to identify the level 1, 2, 3, and 4 portfolios. This identification process was repeated with three separate sets of portfolios to assure consistent scoring. The teachers then began the scoring session.

### **Teacher Implementation**

In the fall, the five art teachers selected two classes for inclusion in the study. Four schools chose two fifth grade classes and one school chose two fourth grade classes. Parental permission and child assent was obtained. Each teacher then designated one class as the control group and one as the experimental group. The teachers began instruction without any interventions or alterations to their lesson or unit plans. After completion of the first lesson, the art teachers had their experimental groups answer the questions on the first page of their journals (appendix B). At this time, the students also set goals for the year. The art teachers could and did explain the meaning of “goals”. The art teachers were encouraged to review the lesson objectives and the evaluation criteria before the students began writing in their journals. The children then wrote their answers to the study’s questions (appendix B) in their paper journals. The students in the control groups did not do any evaluation or assessment activities. When they finished a lesson they went on to the next art lesson. After the experimental group finished writing, they began the next lesson. After the final art lesson of the year, the experimental group answered the

last page of the journal, which included a question requiring students to reflect on their goals in relationship to the progress they made throughout the year (appendix B).

Both groups' artworks were digitally photographed and entered into PowerPoint portfolios. For ease of identification, 2 inch by 8 inch identification cards were made for each student. Their first name and student identification number were written on the cards and photographed next to the students' artworks. One PowerPoint portfolio was created for each student with two parts, pretest and posttest. These portfolios were saved in class folders (control group, experimental group) on either the teachers' computer desktop or the school server. Later, before the scoring process, the art teachers made class portfolios as follows: 1) a control group pretest portfolio, 2) a control group posttest portfolio, 3) an experimental group pretest portfolio, and 4) an experimental group posttest portfolios. These four portfolios contained everyone in the groups' artworks. The background graphics and student names were removed from the portfolios leaving a plain black background, the school identification letters and the student identification numbers to assure anonymity. The portfolios were then scored as already discussed in the description of the implementation.

### **Other Data Collected**

In addition to having their artworks digitally photographed or scanned, the experimental group set annual goals, and answered self-reflection and self-evaluation journal questions (appendix B) after each art lesson was completed. The journals were written and saved in traditional paper or file folders in hard copy. In addition, most of the teachers used writing, checklists, and tests with the experimental group similar to those used in prior years. In order to score the journals a holistic journal rubric (appendix C) was developed and applied to determine the students' level of involvement and understanding of the metacognitive/self-reflective journal writing strategy. The researcher scored the journals.

The art teachers also gathered and reported data related to the extraneous variables. These variables included: FCAT (Florida Comprehensive Achievement Test) reading scores, grade, gender free-reduced lunch status, ethnicity, hours in art class, and art budget to determine their

### **Reliability**

To ensure reliability of the results obtained two tests were used to determine the inter-rater reliability, Hoyt's Alpha and an Intra-Class Correlation Coefficient (ICC). In addition, the

researcher did additional calculations to determine in more detail the percent of agreement among the judges, exact match, two out of three, and no match.

In previous studies, (Dorn, Madeja, Sabol, 2002, 2003) the reliability using the same scoring methodology had an 85%-90% match between judges.

### **Score Spread**

Since the portfolios were scored within the context of each school's set of artwork, it was necessary to determine two factors. First, was there a sufficient score spread, 1-4? And second, did the instrument (the rubrics) discriminate among the portfolios.

### **Validity**

The dependent variable in this study was art performance. The independent variable was goal setting, self-reflection and self-evaluation. Several threats to the validity were addressed; these included subject characteristics, maturation, mortality, location and history. To control for subject characteristics, the classes selected came from a variety of schools, which represented different social and economic groups. To control for maturation, the students in the control groups were the same age as the students in the experimental groups. There is little control on mortality except students usually enter and leave schools randomly. To control for location, each school had its own control group. The researcher also gathered information on art class contact time and supply budget. To control for history, the use of control groups should help account for any unexpected events. Descriptive statistical analysis was used to examine equivalency of groups. Independent and Paired Samples T-tests was conducted to determine if there was a difference between control and experimental groups and to determine whether or not the difference was significant at  $p < .05$ . The Pearson r was conducted to check for any effect the extraneous variables may have had on the results of the study. Data was collected on seven extraneous variables: (1) age, (2) gender, (3) grade level, (4) socioeconomic level, and (5) achievement level (FCAT/NRT reading score) (6) art class contact time and (7) supply budget.

Beattie (1997) identified twelve validation criteria for performance-based assessments; relevance, content fidelity and integrity, exhaustiveness, cognitive complexity, equity, meaningfulness, straightforwardness, cohesiveness, consequences, directness, cost and efficiency



and generalizability. She also noted that not all the criteria needed to be addressed for every assessment.

### **External Validity**

The results of this study should be applicable for all students in K-12 two dimensional visual art classes. There was a wide range of schools participating with very diverse populations, which increased the external validity of the research project. If the expected findings hold, visual art teachers will have a reliable, valid assessment instrument that measures student growth and an intervention that increases student performance.

### **Variables to be Tested**

The primary variable tested in this study was the effect metacognitive/self-reflective journal writing (independent variable) had on student art performance (dependent variable). To determine the effectiveness of the intervention several tests were conducted.

### **Research and H1**

The research hypothesis states that setting annual goals and answering self-reflection and self-evaluation questions after each art assignment will contribute positively to fourth and fifth grade student performance in visual art classes as measured by an authentic portfolio assessment process established by Dorn, Sabol, and Madeja (2003).

H1 The alternative hypothesis states that students' mean art portfolio scores will increase as a result of the intervention, metacognitive/self-reflective journal writing

The research design called for a comparison of the posttest results. To confirm the appropriateness of this approach, descriptive statistics were used to also compare the extraneous variables and determine the equivalency of the two groups, control and experimental inclusive of all five schools. Score spread and distribution data was also collected to illustrate graphically changes in art performance levels.

Next, two tests were chosen to determine the inter-rater reliability, Hoyt's method (Alpha) and 2) the Intra-Class Correlation Coefficient (ICC) base on Analysis of Variance. Inter-rater reliability refers to the consistency among raters or to the probability that different raters would come to the same conclusion. These two instruments were needed to determine the level of agreement among the three raters. Instruments that require judgment on the part of the scorers

(raters) are vulnerable to observer differences (Fraenkel and Wallen, 2003). The methods used for scoring the students' portfolios in this study required rater judgment. Fraenkel and Wallen (2003) stated, "a desired correlation of at least .90 among the scorers or agreement of at least 80 percent" (p.169). While Armstrong (1994) stated that a .80 level of agreement was appropriate for portfolio assessment methods in the visual arts. Data on the percentage of agreement was also collected. A combination of these different methods adds a level of certainty that the results, if found to be significant, were not a product of chance.

A series of Paired Sampled and Levene's Independent Samples T-tests were then conducted. "The t-test is a parametric statistical test used to see whether a difference between the means of two samples is significant" (Fraenkel and Wallen, p. 241). Parametric statistic techniques are generally more powerful than nonparametric techniques and are more likely to reveal a true difference in relationships if one exists. A t-test of independent means is used to compare the mean scores of two different or independent groups. Several such comparisons were applied. First, a paired sample t-test was conducted comparing the control group's and experimental group's pretest and posttest art performances inclusive of all five schools. These mean comparisons would indicate whether there was a significant difference in student art performance or growth from the pretest to the posttest. The box plot or quartile distributions also provided information on which group made the greatest gain and whether there was a sufficient score spread. A posttest to posttest comparison between the control and experimental group was then administered to determine whether the difference in the means was significant at  $p < .05$ . In other words, did the intervention have a positive effect on student performance?

In addition, comparisons were made at the individual school level. Descriptive statistics were used to compare mean scores. Levene's Test for Equality of Variances was used to compare the posttest results, experimental to control, for each school. An analysis of variance (ANOVA) was then conducted to determine the level of significance based on the degree of mean change, pretest to posttest, by individual school. An ANOVA is a more general form of the t-test appropriate for use when there are three or more groups. "Like the t-test the ANOVA helps determine whether the sample means differ enough to conclude that they do not differ just by chance" (Cohen and Lea,2004). In this instance, there were five schools. The posttest scores of the individual schools were then compared. In schools where the pretest to posttest mean score

differences were significant at  $p < .05$ , one measure for determining the effectiveness of the intervention, metacognitive/self-reflective journal writing would have been examined.

Effect size is another indicator of the effectiveness of the intervention. Effect size is an indicator of the strength of the relationship (Cohen and Lea, 2004). The effect size quantifies the size of the difference between two groups, and is therefore said to be a true measure of the significance of the difference. Effect size reported with the confidence interval provides the same information as a test of significance, but with an emphasis on the significance of the effect, rather than the sample size. In terms of percentiles, “An effect size is exactly equivalent to a Z-score of a standard Normal distribution. For example, an effect size of 0.8 means that the score of the average person in the experimental group exceeds the scores of 79% of the control group” (Coe, R., electronic, 2005). Cohen’s  $d$  and the effect-size correlation were used to determine the effect size in this study. A small effect size ranges from 0.0 to 0.2, a medium effect size ranges from 0.3 to 0.5, and a large effect size ranges from 0.6 to 0.8 or above.

## **H2 – H4, Extraneous Variables**

H2 The null hypothesis for all schools stated that the extraneous variable of FCAT-NRT reading scores did not affect student performance.

H3 The null hypothesis for all schools stated that the extraneous variable of gender did not affect student performance.

H4 The null hypothesis for all schools stated that the extraneous variables of age, free/reduced lunch status, instructional time and budget did not affect student performance.

The secondary variables tested were for the extraneous variables. The purpose of testing these variables was to determine if any of the extraneous factors may have affected the results of the study. If none of the extraneous variables affected the outcome of the findings and the experimental group’s posttest scores were higher than the control group’s then the results would suggest the intervention was causal. The art teachers recorded extraneous variable information regarding students’ gender, academic achievement level based on their FCAT reading scores, free or reduced lunch status, ethnicity, age, amount of time each class met and annual art budget information.

First, descriptive statistics, related to each of the schools, provided information relevant to the nature of the schools and groups. These descriptive statistics added a context within which to interpret the findings. Second, a Pearson  $r$  was conducted to determine if any of the extraneous

variables had a significant correlation with the posttest scores. The Pearson r is “an index of correlation appropriate when the data represents either interval or ratio scales; it takes into account each pair of scores and produces a coefficient between 0.00 and plus or minus 1.00” (Fraenkel and Wallen, 2003 p.G-6). The raw data for the extraneous variables was coded using an interval scale and then a Pearson r was conducted. This test provided important information related to the effectiveness of the intervention. If a positive correlation was found that affected the results of the study then further methods and analyses would be needed to determine the effectiveness of the intervention. If there was not a positive correlation the effected the study and the findings were significant then the intervention could be interpreted as causal.

### **Reporting the Data**

The results were reported in a series of charts and graphs. Each set of hypotheses are first described in terms of descriptive statistics. The raw data for the five elementary schools follows the descriptive statistics related to the research hypothesis and H1. In the raw data tables each of the raters’ portfolio scores was recorded as well as the extraneous variables for each student in the study. Next, the inter-rater reliability was reported followed by the inferential statistics. H2 through H4 hypotheses relate to the effect of the extraneous variables on the study results if any. Again, descriptive statistics are reported first followed by the inferential statistics.

**CHAPTER IV**  
**REPORTING THE DATA**

**Research Hypothesis and H1**

The research hypothesis states that setting annual goals and answering self-reflection and self-evaluation questions after each art assignment will contribute positively to fourth and fifth grade student performance in visual art classes as measured by an authentic portfolio assessment process established by Dorn, Sabol, and Madeja (2003).

H1 The alternative hypothesis states that students’ mean art portfolio scores will increase as a result of the intervention, metacognitive/self-reflective journal writing.

**Descriptive Statistics**

**Number of Students in the Control and Experimental Groups.** A total of 171 cases had both pretest and posttest portfolio scores. The control group had 82 cases and the experimental group had 89 cases. All analysis was conducted using matched cases (all cases that met the study criteria were included).

There were 171 students in the study. All of the 171 students have both Pretest (Portfolio 1) and Posttest (Portfolio 2) scores. The sample size of the study and the corresponding percentages are given on the following table listed by school and group, control or experimental.

**Table 1.** Number of Students Disaggregated by Schools and Group.

School ID	Number	Groups		Total
		Control Group	Experimental Group	
<b>A</b>	N	14	18	32
	%	43.8%	56.3%	100%
<b>B</b>	N	17	16	33
	%	51.5%	48.5%	100%
<b>C</b>	N	17	20	37
	%	45.9%	54.1%	100%

**Table 1 – Continued.**

School ID	Number	Groups		Total
		Control Group	Experimental Group	
<b>D</b>	N	13	12	25
	%	52.0%	48.0%	100%
<b>E</b>	N	21	23	44
	%	47.7%	52.3%	100%
<b>Total</b>	N	82	89	171
	%	48.0%	52.0%	100.0%

**Composition of the Control and Experimental Groups - Group Equivalency.** To determine the equality of the groups, control and experimental, descriptive statistics were used to compare the extraneous variables. This comparison dictated which method of statistical analysis to apply to the data.

**Table 2.** Descriptive Statistics for Extraneous Variables by Implementation Group Inclusive of All Five Schools

		N	Minimum	Maximum	Mean	Std. Deviation
<b>Control Group</b>	Age	81	9.0	12.0	10.3	.9
	Gender	82	1.0	2.0	1.5	.5
	Pretest Mean	82	1.0	4.0	2.54	.7
	Posttest Mean	82	1.3	4.0	2.71	.7
	Ethnicity	82	1.0	3.0	1.9	.6
	FCAT NRT 02-03	78	14.0	99.0	68.4	22.0
	Free\Reduced Lunch	82	.0	1.0	.3	.5
	Hours Per Year	82	13.5	30.8	24.7	5.9
	Supply Money	82	1.0	4.0	2.7	1.0

**Table 2 - Continued.**

	N	Minimum	Maximum	Mean	Std. Deviation	
<b>Experimental Group</b>	Age	88	9.0	13.0	10.4	.9
	Gender	89	1.0	2.0	1.5	.5
	Pretest Mean	89	1.0	4.0	2.60	.8
	Posttest Mean	89	1.0	4.0	2.65	.7
	Ethnicity	89	1.0	3.0	1.9	.6
	FCAT NRT 02-03	80	1.0	99.0	72.8	25.0
	Free\Reduced Lunch	89	.0	1.0	.4	.5
	Hours Per Year	89	12.5	29.5	26.0	5.6
	Supply Money	89	1.0	4.0	2.8	1.0

The control and experimental groups had equivalent gender results with mean scores of 1.5 each indicating an equal number of boys and girls. The mean age for the control group was 10.3 with a slightly higher mean age of 10.4 for the experimental group. Both control and experimental groups had mean scores of 1.9 indicating a larger number of majority students than minority. The FCAT (Florida Comprehensive Assessment Test) mean score for the control group was a little lower than for the experimental group, 68.4 and 72.8 respectively. The control group had a free/reduced lunch mean score of 30 %. The experimental group had a higher mean score of 40% of its students receiving free or reduced lunch. The control group received a mean of 24.7 hours of instruction per year while the experimental group received a slightly higher mean with 26 hours per year. The control group had a mean score of \$2.70 per child per year for materials and supplies. The experimental group had a slightly higher mean score of \$2.80 per child per year. The portfolio 1 mean score for the control group was 2.54 while the experimental group had a slightly higher mean score of 2.60. There were no significant differences between the control and experimental groups for the pretest period. Therefore, equivalent statistical methods were used to analyze the data.

**Raw Data from the Five Schools.** Student scores on the 13 variables for the experimental and control groups included age, gender, three portfolio 1 scores (one from each rater), three portfolio 2 (one from each rater) scores, ethnicity, FCAT-NRT reading scores,

free/reduced lunch status, hours of instruction per year and annual materials and supply budget per student are reported in the following tables.

Key for Extraneous Variables for Raw Data Tables

Group	1 = Control	2 = Experimental		
Gender	1 = Female	2 = Male		
Ethnicity	1= Black	2 = White	3= Other	
Free/Reduced Lunch	0 = don't receive	1 = receive		
Supply/Materials Money	1 = \$1.00	2 = \$2.00	3 = \$3.00	4 = \$5.15

**Table 3.** Reporting the Raw Data on the 13 Variables Linked with the Study for School A.

School ID #	Control /Experimental	Student Number	Age	Gender	Journal Score	Portfolio 1Score 1	Portfolio 1Score 2	Portfolio 1 Score 3	Portfolio 2 Score 1	Portfolio 2 Score 2	Portfolio 2 Score 3	Ethnicity	FCAT/NRT02-03	Free/Reduced Lunch	Hours/per year	Supply Money
A	2	25	11	1	3	4	4	4	3	4	4	1	70	0	25.25	2
A	2	29	11	2	1	2	2	3	3	3	3	1	8	1	25.25	2
A	2	19	12	2	2	4	3	4	4	4	4	3		0	25.25	2
A	2	16	10	1	3	4	3	3	3	3	3	1	38	0	25.25	2
A	2	17	12	1	3	3	2	3	4	4	4	2	67	0	25.25	2
A	2	18	10	2	1	2	2	3	2	3	3	1	65	1	25.25	2
A	2	20	11	2	1	2	2	3	3	3	3	1	22	1	25.25	2
A	2	21	12	2	2	4	3	4	4	4	4	2	88	0	25.25	2
A	2	22	11	1	3	4	4	3	3	4	3	1	81	1	25.25	2
A	2	23	11	2	2	3	3	4	3	4	4	3	93	0	25.25	2
A	2	24	12	1	2	3	2	1	3	4	3	1	91	1	25.25	2
A	2	26	11	2	1	3	2	3	3	3	2	1	99	1	25.25	2
A	2	27	11	2	1	3	2	2	2	2	2	1	65	1	25.25	2
A	2	28	10	1	3	3	2	3	3	4	4	2	99	0	25.25	2
A	2	30	10	2	2	3	2	2	3	2	3	1	30	1	25.25	2
A	2	31	11	1	2	3	2	2	3	2	2	1	30	1	25.25	2
A	2	32	10	1	1	4	4	4	2	2	1	1	75	1	25.25	2
A	2	33	11	2	1	1	2	2	2	3	2	2	52	1	25.25	2



**Table 3 – Continued.**

A	1	2	11	1		3	3	3	2	3	3	1		1	22.25	2
A	1	3	11	1		3	3	3	3	4	4	2	93	0	22.25	2
A	1	4	11	1		3	4	4	3	4	3	2		1	22.25	2
A	1	5	10	2		3	4	4	3	3	3	1	79	1	22.25	2
A	1	6	12	2		2	3	2	2	2	2	1	67	1	22.25	2
A	1	7	10	1		3	3	3	3	4	3	2	79	0	22.25	2
A	1	8	12	2		2	2	1	1	2	2	1	14	1	22.25	2
A	1	9	11	1		2	2	2	2	3	3	1	29	1	22.25	2
A	1	10	11	1		2	2	3	3	3	3	1	79	1	22.25	2
A	1	11	12	1		2	2	3	2	2	2	1	52	1	22.25	2
A	1	12	11	2		2	3	3	2	3	2	1	32	1	22.25	2
A	1	13	12	2		2	2	3	1	2	2	1	88	0	22.25	2
A	1	14	11	1		3	4	4	4	4	3	1	17	0	22.25	2
School ID #	Control /Experimental	Student Number	Age	Gender	Journal Score	Portfolio 1Score 1	Portfolio 1Score 2	Portfolio 1 Score 3	Portfolio 2 Score 1	Portfolio 2 Score 2	Portfolio 2 Score 3	Ethnicity	FCAT/NRT02-03	Free/Reduced Lunch	Hours/per year	Supply Money

**Table 4.** Reporting the Raw Data on the 13 Variables Linked with the Study for School B.

School ID #	Control/ Experimental	Student Number	Age	Gender	Journal Score	Portfolio 1Score 1	Portfolio 1Score 2	Portfolio 1 Score 3	Portfolio 2 Score 1	Portfolio 2 Score 2	Portfolio 2 Score 3	Ethnicity	FCAT/NRT02-03	Free/Reduced Lunch	Hours/per year	Supply Money
B	2	108	11	2	2	1	2	2	4	3	3	2	85	1	29	3
B	2	109	11	2	1	3	4	3	2	2	2	1	73	1	29	3
B	2	110	10	1	1	2	2	2	3	2	2	2	34	0	29	3
B	2	111	10	1	3	3	3	3	4	3	3	2	97	1	29	3
B	2	112	10	2	2	4	4	3	4	3	3	3	97	0	29	3
B	2	113	10	2	1	3	4	4	3	3	2	3	93	0	29	3
B	2	114	11	1	1	2	4	2	3	3	3	1	36	1	29	3
B	2	115	10	1	2	3	4	3	4	3	3	2	95	1	29	3
B	2	116	11	1	2	2	2	2	2	2	2	2	61	1	29	3

**Table 4 – Continued.**

B	2	117	11	1	2	2	3	3	2	2	2	1	99	1	29	3
B	2	119	11	1	2	2	2	2	3	2	2	2	79	0	29	3
B	2	120	10	1	2	3	3	3	4	3	3	1	91	0	29	3
B	2	122	12	1	1	3	2	2	3	2	2	2	73	1	29	3
B	2	123	10	1	2	4	4	3	4	3	3	2	85	0	29	3
B	2	124	11	2	3	3	2	1	3	2	2	1	73	0	29	3
B	2	125	11	2	1	1	1	1	1	1	1	1	1	1	29	3
B	1	127	11	1		4	3	3	2	3	3	1	61	1	29	3
B	1	128	10	2		4	4	3	4	3	4	3	99	0	29	3
B	1	130	11	1		3	3	3	2	2	3	2	75	1	29	3
B	1	131	10	1		3	3	4	4	3	3	2	46	1	29	3
B	1	132	10	1		3	2	3	3	2	3	1	73	1	29	3
B	1	133	10	2		4	2	3	2	2	3	2	88	0	29	3
B	1	136		2		2	2	3	3	2	2	2	34	0	29	3
B	1	137	12	1		2	2	2	2	3	2	1	32	1	29	3
B	1	138	10	1		3	2	2	3	3	3	2	73	0	29	3
B	1	139	10	2		3	2	2	3	3	4	1	52	1	29	3
B	1	141	10	1		4	3	3	3	3	3	1	79	1	29	3
B	1	142	10	1		2	2	2	2	2	3	2	73	1	29	3
B	1	144	10	2		1	1	1	1	1	2	2	85	0	29	3
B	1	145	11	2		2	2	2	2	2	3	2	73	1	29	3
B	1	146	10	2		3	3	4	3	2	3	1	56	1	29	3
B	1	147	9	1		3	2	2	3	3	4	1	93	0	29	3
B	1	148	11	2		2	2	3	4	4	4	3	99	0	29	3
School ID #	Control / Experimental	Student Number	Age	Gender	Journal Score	Portfolio 1 Score 1	Portfolio 1 Score 2	Portfolio 1 Score 3	Portfolio 2 Score 1	Portfolio 2 Score 2	Portfolio 2 Score 3	Ethnicity	FCAT/NRT02-03	Free/Reduced Lunch	Hours/per year	Supply Money

Key for Extraneous Variables for Raw Data Tables

Group	1 = Control	2 = Experimental		
Gender	1 = Female	2 = Male		
Ethnicity	1= Black	2 = White	3= Other	
Free/Reduced Lunch	0 = don't receive	1 = receive		
Supply/Materials Money	1 = \$1.00	2 = \$2.00	3 = \$3.00	4 = \$5.15

**Table 5.** Reporting the Raw Data on the 13 Variables Linked with the Study for School C.

School ID #	Control/Experimental	Student Number	Age	Gender	Journal Score	Portfolio 1 Score 1	Portfolio 1 Score 2	Portfolio 1 Score 3	Portfolio 2 Score 1	Portfolio 2 Score 2	Portfolio 2 Score 3	Ethnicity	FCAT/NRT02-03	Free/Reduced Lunch	Hours/per year	Supply Money
C	2	60	10	1	2	2	2	2	3	2	3	3	49	0	29.5	4
C	2	61	11	2	2	3	3	2	4	4	4	2		0	29.5	4
C	2	62	10	2	1	2	2	1	1	1	1	2		0	29.5	4
C	2	63	12	2	1	3	3	2	3	2	2	2	15	0	29.5	4
C	2	64	11	2	2	1	1	2	2	2	2	2	93	0	29.5	4
C	2	65	13	2	1	3	4	4	2	2	2	2	79	0	29.5	4
C	2	67	10	2	2	2	3	2	2	2	1	2		0	29.5	4
C	2	68	11	2	1	2	3	3	2	2	2	1	70	0	29.5	4
C	2	69	10	1	2	3	3	3	3	2	2	2	88	0	29.5	4
C	2	70	10	2	2	3	2	2	3	2	2	2	97	0	29.5	4
C	2	71	12	2	1	2	3	2	3	4	4	2	75	0	29.5	4
C	2	72	11	2	2	4	4	3	3	3	3	1	95	0	29.5	4
C	2	76	10	1	2	3	2	2	2	3	3	1	85	0	29.5	4
C	2	77	11	2	1	1	2	1	2	2	1	2	79	0	29.5	4
C	2	78	11	2	2	2	2	2	2	2	1	2	88	0	29.5	4
C	2	79	10	1	1	3	2	3	2	3	2	2	93	0	29.5	4
C	2	80	10	2	2	2	2	2	4	3	3	2	95	0	29.5	4
C	2	82	10	1	2	4	4	4	2	3	4	2	97	0	29.5	4
C	2	83	10	2	1	3	3	3	2	3	3	2	85	0	29.5	4
C	2	85	11	2	1	2	3	3	2	3	3	2	91	0	29.5	4
C	1	86	10	2		2	3	2	3	2	2	2	71	0	23.5	4
C	1	87	9	1		3	4	4	3	4	3	2		0	23.5	4
C	1	88	10	1		4	3	4	4	4	4	3	99	0	23.5	4
C	1	89	10	1		3	3	3	2	2	3	2	61	0	23.5	4
C	1	90	11	2		1	2	2	2	2	3	2	81	0	23.5	4
C	1	91	11	1		2	2	2	2	3	3	2	90	0	23.5	4
C	1	92	10	2		3	2	3	4	3	3	1	71	1	23.5	4
C	1	93	11	1		1	2	1	2	2	2	1	54	1	23.5	4
C	1	96	12	2		3	4	3	4	4	4	1	61	0	23.5	4
C	1	97	12	1		3	3	3	4	3	3	2	46	0	23.5	4
C	1	99	10	2		1	1	1	3	3	4	2	68	0	23.5	4
C	1	100	11	2		2	3	2	3	2	3	2	78	0	23.5	4
C	1	102	11	2		1	1	2	2	3	2	2	89	0	23.5	4
C	1	103	11	2		2	1	2	1	2	2	2		0	23.5	4
C	1	104	10	1		3	4	3	4	3	4	2	81	0	23.5	4
C	1	105	10	2		1	1	3	2	2	3	3	54	0	23.5	4
C	1	106	11	1		2	3	4	3	3	3	2	81	0	23.5	4

**Table 6.** Reporting the Raw Data on the 13 Variables Linked with the Study for School D.

School ID #	Control/Experimental	Student Number	Age	Gender	Journal Scores	Portfolio 1 Score 1	Portfolio 1 Score 2	Portfolio 1 Score 3	Portfolio 2 Score 1	Portfolio 2 Score 2	Portfolio 2 Score 3	Ethnicity	FCAT/NRT02-03	Free/Reduced Lunch	Hours/per year	Supply Money
D	2	35	10	1	1	4	3	3	2	2	3	2	93	0	12.5	1
D	2	36	10	1	1	3	2	2	2	3	2	2	70	1	12.5	1
D	2	37	10	1	1	4	3	4	2	3	3	2	90	1	12.5	1
D	2	38	11	1	2	2	2	1	3	2	2	2		0	12.5	1
D	2	39	10	1	1	3	4	4	3	3	2	2	87	0	12.5	1
D	2	40	10	1	1	2	1	2	2	2	2	2	64	0	12.5	1
D	2	41	12	1	1	3	3	2	3	3	2	1	73	1	12.5	1
D	2	42	11	1	2	4	3	3	2	3	2	2	42	1	12.5	1
D	2	43	10	1	1	4	4	4	3	4	3	2	87	1	12.5	1
D	2	44	11	1	1	2	3	2	3	4	4	2	76	1	12.5	1
D	2	45	10	1	2	2	3	2	2	2	3	2	67	1	12.5	1
D	2	46	10	1	1	3	2	3	2	2	2	2	45	0	12.5	1
D	1	47	12	2		2	3	2	2	2	2	2	78	0	13.5	1
D	1	48	10	1		3	4	3	3	4	3	2	74	0	13.5	1
D	1	49	10	1		2	2	2	3	3	3	2	71	0	13.5	1
D	1	50	10	1		2	3	2	2	3	2	2	53	0	13.5	1
D	1	51	11	1		2	2	2	2	3	2	2	32	1	13.5	1
D	1	52	11	1		3	3	3	2	3	2	2	81	0	13.5	1
D	1	53	11	1		4	4	4	4	4	4	2	45	0	13.5	1
D	1	54	12	1		1	1	1	1	2	2	2	28	1	13.5	1
D	1	55	10	1		2	2	2	2	3	3	2	38	0	13.5	1
D	1	56	10	2		3	3	3	3	3	3	2	62	0	13.5	1
D	1	57	11	2		4	4	3	4	4	4	2	78	0	13.5	1
D	1	58	10	1		3	2	2	2	3	3	2	67	0	13.5	1
D	1	59	10	2		3	2	2	3	4	3	2	71	0	13.5	1

Key for Extraneous Variables for Raw Data Tables

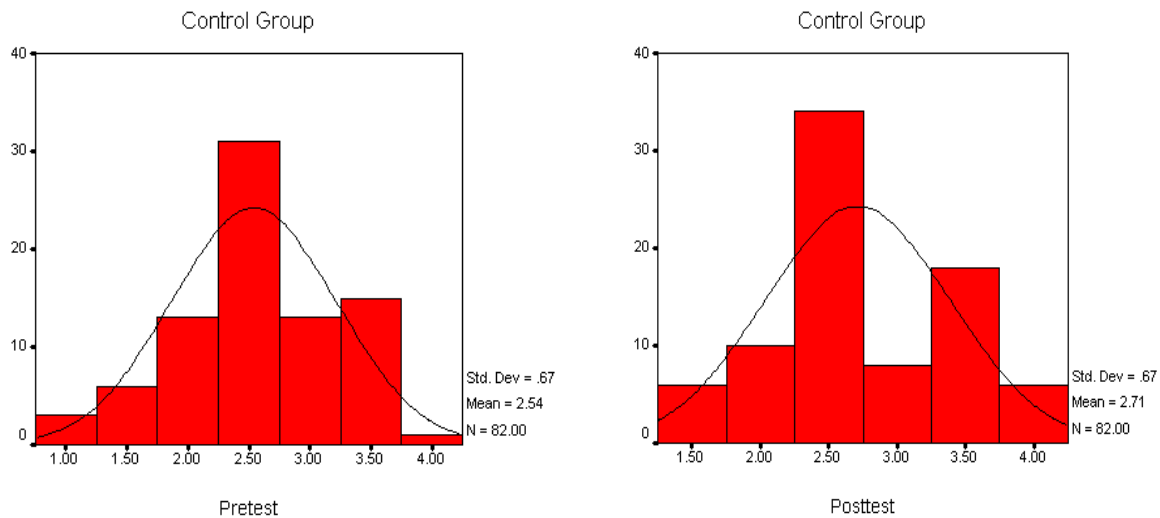
Group	1 = Control	2 = Experimental		
Gender	1 = Female	2 = Male		
Ethnicity	1= Black	2 = White	3= Other	
Free/Reduced Lunch	0 = don't receive	1 = receive		
Supply/Materials Money	1 = \$1.00	2 = \$2.00	3 = \$3.00	4 = \$5.15

**Table 7.** Reporting the Raw Data on the 13 Variables Linked with the Study for School E.

School ID #	Control/ Experimental	Student Number	Age	Gender	Journal Score	Portfolio 1 Score 1	Portfolio 1 Score 2	Portfolio 1 Score 3	Portfolio 2 Score 1	Portfolio 2 Score 2	Portfolio 2 Score 3	Ethnicity	FCAT/NRT02-03	Free/Reduced Lunch	Hours/per year	Supply Money
E	2	210	11	2	1	2	4	2	3	3	2	2	32	1	28.5	3
E	2	211	9	1	1	2	3	2	3	2	2	2		0	28.5	3
E	2	213	9	2	1	1	1	2	2	1	2	2	64	0	28.5	3
E	2	214		1	1	4	4	3	2	1	2	3	23	0	28.5	3
E	2	215	9	1	2	4	4	4	4	3	3	2	70	0	28.5	3
E	2	216	9	1	2	3	3	2	3	3	2	2	93	0	28.5	3
E	2	217	10	2	2	2	1	2	2	2	2	2	84	1	28.5	3
E	2	218	9	1	2	2	2	1	4	3	3	2	42	0	28.5	3
E	2	219	9	1	2	3	3	3	3	3	3	3	72	1	28.5	3
E	2	220	10	2	3	3	2	2	3	3	3	3	93	0	28.5	3
E	2	221	10	2	2	2	1	3	2	2	1	2	90	0	28.5	3
E	2	222	9	1	2	3	2	2	2	3	3	2	52	0	28.5	3
E	2	223	10	1	2	2	3	3	3	4	4	2	95	0	28.5	3
E	2	224	9	1	2	3	4	3	4	4	3	2		0	28.5	3
E	2	225	10	1	2	2	2	3	3	4	3	3		1	28.5	3
E	2	226	9	1	2	3	3	2	2	3	2	2	98	0	28.5	3
E	2	227	9	1	1	3	2	2	2	3	3	2	93	0	28.5	3
E	2	228	10	2	2	3	2	3	3	4	2	3	96	0	28.5	3
E	2	229	10	2	1	3	2	1	2	3	2	2	97	0	28.5	3
E	2	230	10	2	1	3	2	2	2	3	2	2	93	0	28.5	3
E	2	231	9	2	2	2	2	2	2	3	2	2	93	0	28.5	3
E	2	232	11	2	1	1	1	1	1	2	1	3	32	1	28.5	3
E	2	233	9	2	2	1	1	1	2	2	2	2		1	28.5	3
E	1	234	9	1		3	2	2	1	2	1	2	64	0	30.75	3
E	1	235	9	1		3	3	3	3	3	3	2	95	0	30.75	3
E	1	237	9	2		3	2	2	3	2	2	2	90	0	30.75	3
E	1	238	10	1		2	3	3	4	3	4	3	57	0	30.75	3
E	1	239	9	1		2	2	2	2	2	3	2	62	1	30.75	3
E	1	240	9	2		2	3	2	3	2	2	2	99	0	30.75	3
E	1	241	10	2		3	3	2	2	2	2	2	60	0	30.75	3
E	1	243	9	1		2	3	2	4	3	3	2	93	0	30.75	3
E	1	244	10	2		3	3	3	2	2	2	2	99	0	30.75	3
E	1	245	10	2		3	2	2	2	2	2	2	93	0	30.75	3
E	1	246	10	2		4	3	4	2	2	3	2	88	1	30.75	3
E	1	247	9	1		3	3	3	3	3	2	3	60	0	30.75	3
E	1	248	10	1		2	2	2	3	2	2	3	23	1	30.75	3
E	1	249	9	1		3	2	2	2	2	2	2	62	0	30.75	3
E	1	250	9	1		3	3	2	3	2	2	2	62	0	30.75	3
E	1	251	9	2		3	2	2	4	4	4	2	97	0	30.75	3
E	1	252	9	2		2	2	2	2	2	3	2	95	0	30.75	3
E	1	253	11	2		2	2	2	2	2	2	3	34	1	30.75	3
E	1	254	9	1		2	3	3	4	3	3	2	97	0	30.75	3
E	1	255	10	1		2	2	2	3	2	3	3	95	1	30.75	3
E	1	257	10	2		3	2	2	2	2	2	2	78	0	30.75	3

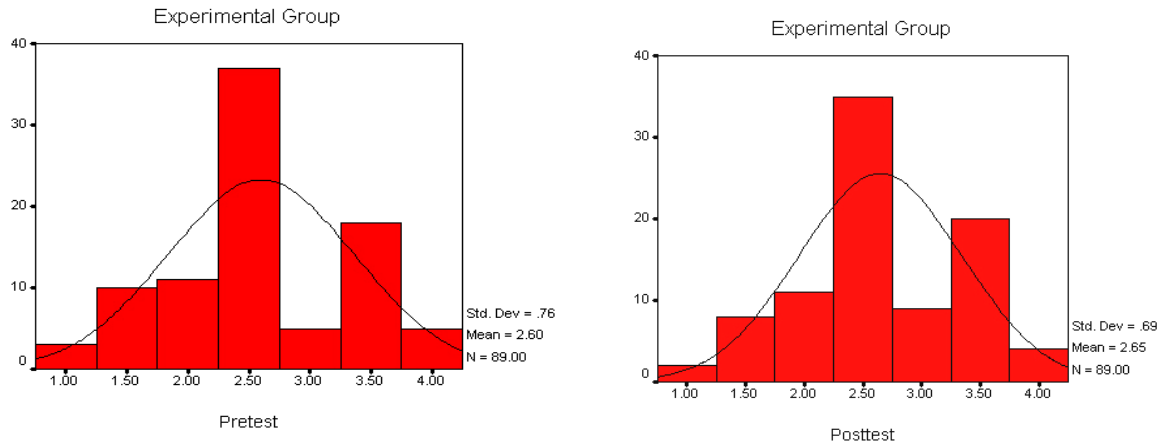
## Score Spread and Distribution for All School, Control and Experimental Groups.

Descriptive figures were chosen to illustrate trends in the students' art portfolio scores.



**Figure 1.** Pretest and Posttest Art Performance Score Distributions with Normal Curve for the Control Group Inclusive of all Five Schools.

Figure 1 shows the score distribution with a normal curve for the control group. The mean pretest score was 2.54. The mean posttest score was 2.71. Overall the posttest scores are higher than the pretest scores. More students received scores of 4 and 3.5.



**Figure 2.** Pretest and Posttest Art Performance Score Distributions with Normal Curve for the Experimental Group Inclusive of all Five Schools.

. Figure 4 shows the score distribution with a normal curve. The mean pretest score for the experimental group was 2.60 and the mean posttest score was 2.65. The posttest results indicate fewer students received scores of 1.0 and 1.5 than on the pretest. There was also a rise in the number of students receiving scores of 3.0 and 3.5.

### Inter-Rater Reliability

Inter-rater reliability refers to the consistency among raters or to the probability that different raters would come to the same conclusion. To examine whether the inter-rater reliability was acceptable for both the control and experimental groups in the pretest and posttest, the following analyses were conducted: 1) The internal consistency among multi-rater using Hoyt's method (Alpha) and 2) The Intra-Class Correlation Coefficient (ICC) based on Analysis of Variance.

**Table 8.** Inter-rater Reliability Coefficients for Pretest and Posttest Portfolio Scores for the Control and experimental groups Inclusive of all five schools.

Groups	N	Pretest		Posttest	
		<i>ICC</i>	<i>Alpha</i>	<i>ICC</i>	<i>Alpha</i>
Control Group	82	.6086	.8217	.6137	.8287
Experimental Group	89	.6053	.8237	.6064	.8284
Total	171	.6051	.8215	.6073	.8228

Table 8 shows the reliability levels among the raters is acceptable for portfolio assessment across the pretest and posttest sessions with the reliability slightly higher for the posttest scores.

**Table 9.** Percent of Agreement Among the Judges

Level of Agreement	Pretest	Posttest
3 out of 3	28 %	29 %
2 out of 3	68 %	70 %
1 out of 3	4 %	1 %

Table 9 shows the percentage of judges who gave the portfolios an identical score. For example 28% of the time all three judges gave a portfolio the same score on the pretest portfolios. 68% of the two out of three judges gave the pretest portfolio the same score.

**Validity**

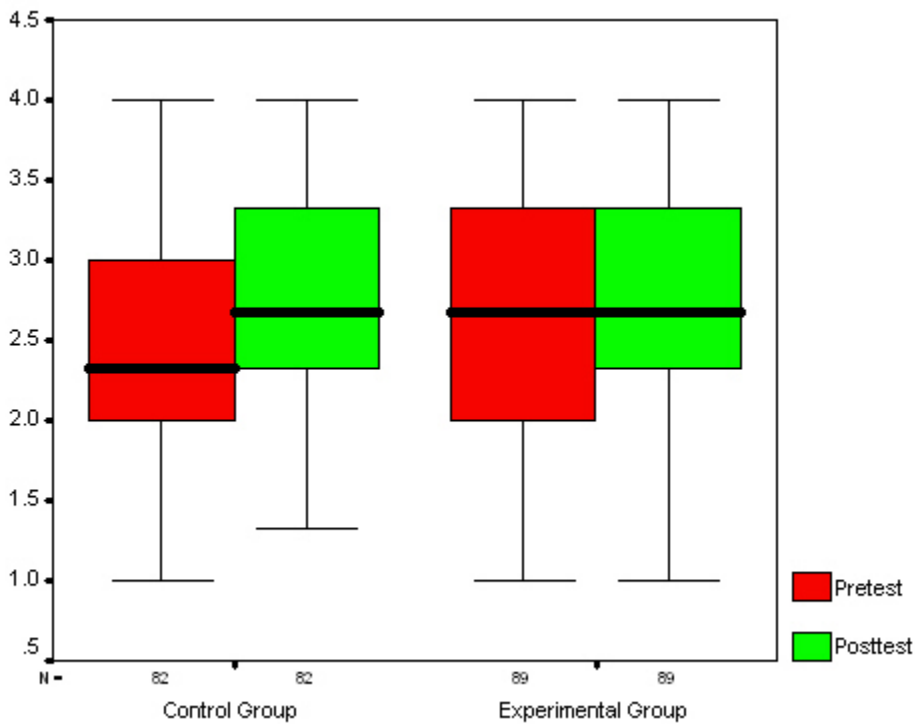
The artworks created by the students throughout the year as part of their regular art instruction was the source of the data for both the pretest and the posttest and provided the evidence of learning. The assessment method was straightforward and based on the assignments the students completed in art class where the expectations for the lessons were openly clarified and discussed. By using student artworks created in class as a part of their regular instruction



content fidelity was high incorporating major content and process in the field of art education, art making.

### Inferential Statistics

**Paired Sample T-Test comparing Pretest and Posttest Performance.** Paired Sample T-Tests were conducted to examine differences between the control and experimental group's pretest and posttest scores on art performance for all schools.



**Figure 3.** Quartile (Box Plot) Distribution of Art performance Pretest, Posttest and Median Scores by Control and Experimental Group Inclusive of All Five Schools.

Figure 3 shows the quartile distribution, the median scores and the score spread for both control and experimental groups for both the pretest and posttest. The box plots indicate adequate score spreads for both the pretests and the posttests covering the full range of possible scores, 1- 4.

**Table 10.** Art Performance Mean Pretest and Posttest Scores by Control and Experimental Groups and their Mean Differences Inclusive of All Five Schools.

Groups	Measure	N	Mean	Std. Dev.
Control	Pretest	82	2.54	.67
	Posttest		2.71	.67
	Difference		-0.17*	.62
Experimental	Pretest	89	2.60	.76
	Posttest		2.65	.69
	Difference		-0.05	.76
All	Pretest	171	2.57	.72
	Posttest		2.68	.68
	Difference		-0.11*	.70

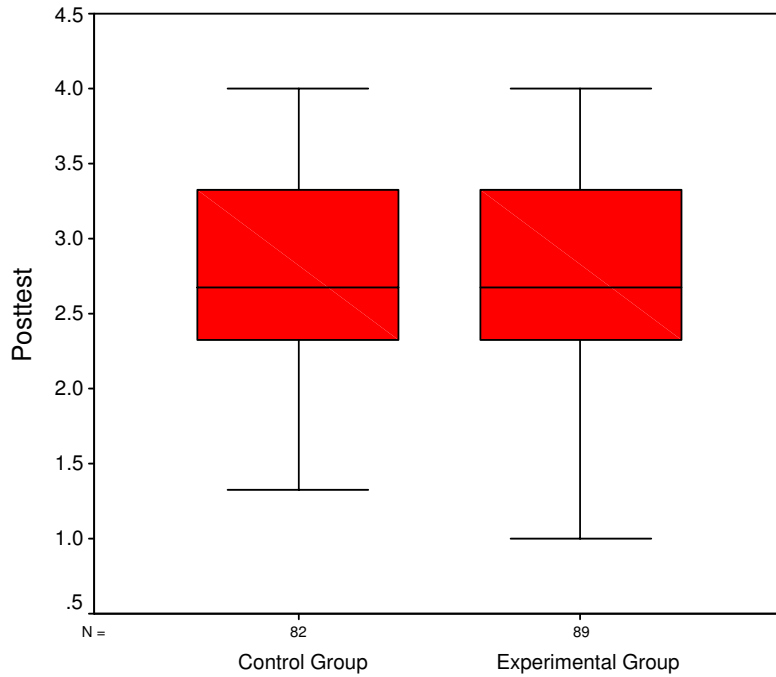
\* The difference is significant at .05 level.

**Table 11.** Pretest & Posttest scores by Control and Experimental groups and their mean differences tested by Paired Sample T-Test

Paired Samples Test										
			Paired Differences					t	df	Sig. (2-tailed)
			Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
						Lower	Upper			
Control Group	Pair 1	Pretest - Posttest	-.17	.62	.07	-.31	-.03	-2.49	81.00	.015
Experimental Group	Pair 1	Pretest - Posttest	-.05	.76	.08	-.21	.11	-.60	88.00	.551

Tables 10 and 11 reveal that in general the median and mean posttest scores for both the control and experimental groups were higher than their pretest scores. However, the results varied considerably between the groups. The control group showed a gain of .17 and was significant. The experimental group showed a slight gain of .05 and was not significant.

### Independent Samples T-Test Comparing Posttest Scores.



**Figure 4.** Box Plot (Quartile) Comparison of Posttest Art Performance Scores by Control and Experimental Groups Inclusive of All Five Schools.

Figure 4 indicates that the first, second and third posttest quartile scores for both the control group and the experimental groups inclusive of all five schools were nearly identical.

**Table 12.** Mean Posttest Art Performance Scores by Control and Experimental Groups Inclusive of All Five Schools.

		Group Statistics			
		N	Mean	Std. Deviation	Std. Error Mean
Posttest	Control Group	82	2.71	.67	.07
	Experimental Group	89	2.65	.69	.07

**Table 13.** Results of Independent Samples T-test for the Comparison of Art performance Posttest Scores between Control and Experimental Groups Inclusive of All Five Schools.

		Independent Samples Test								
		Levene's Test for Equality of Variances			t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Posttest	Equal variances assumed	.001	.974	.602	169	.548	.06	.10	-.14	.27
	Equal variances not assumed			.603	168.544	.547	.06	.10	-.14	.27

Tables 12 and 13, indicate that the mean of the posttest score for control group, inclusive of all five schools, at 2.71 was higher than the experimental group's posttest score mean, which was 2.65. However, the difference of 0.6 was not statistically significant (0.547).

### Descriptive Statistics for Each of the Five Schools

**Table 14.** Descriptive Statistics of Pretest and Posttest Scores for the Control and Experimental Groups by School Showing Mean Art Portfolio Score Comparisons.

School ID	Groups	Group	N	Mean	Standard Deviation
School A	Pretest	Control Group	14	2.78	.65
		Experimental Group	18	2.85	.73
	Posttest	Control Group		2.69	.68
		Experimental Group		3.06	.71
School B	Pretest	Control Group	17	2.60	.69
		Experimental Group	16	2.62	.81
	Posttest	Control Group		2.76	.64
		Experimental Group		2.62	.70

**Table 14 – Continued.**

School ID	Groups	Group	N	Mean	Standard Deviation
School C					
	Pretest	Control Group	17	2.41	.87
	Posttest			2.86	.70
	Pretest	Experimental Group	20	2.51	.78
	Posttest			2.45	.73
School D					
	Pretest	Control Group	13	2.56	.81
	Posttest			2.82	.71
	Pretest	Experimental Group	12	2.80	.78
	Posttest			2.55	.50
School E					
	Pretest	Control Group	21	2.47	.43
	Posttest			2.49	.65

Table 14 indicates that there were variations among the schools. In four schools, B, C, D and E the control groups' pretest mean scores were lower than their posttest mean scores indicating a positive change in the mean art portfolio scores. School A's control group had a higher pretest mean score than a posttest mean score indicating a decline in the mean art portfolio score. In two schools, A and E, the experimental groups' posttest mean scores were higher than their pretest scores indicating a positive change in their mean art portfolio scores. In one school, B, the experimental group's pretest and posttest mean art portfolio scores remained the same. In two schools, C and D the experimental group's posttest mean scores were lower than their pretest mean scores indicating a decline in art portfolio scores.

**Table 15.** Descriptive Statistics of Posttest Scores for the Control and Experimental Groups by School Showing Mean Score Comparisons.

School ID	NEW_GRPS	N	Mean	Std. Deviation	Std. Error Mean	
A	Posttest	Control Group	14	2.69	.68	.18
		Experimental Group	18	3.06	.71	.17
B	Posttest	Control Group	17	2.76	.64	.16
		Experimental Group	16	2.62	.70	.17

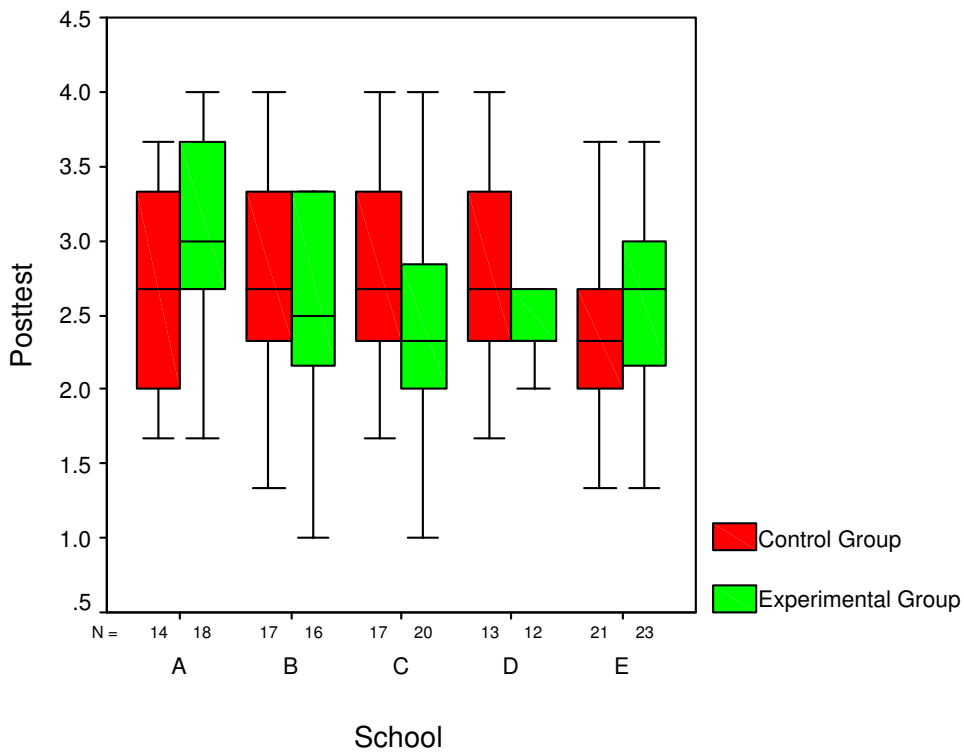
**Table 15 – Continued.**

School ID	NEW_GRPS	N	Mean	Std. Deviation	Std. Error Mean
C	Posttest Control Group	17	2.86	.70	.17
	Posttest Experimental Group	20	2.45	.73	.16
D	Posttest Control Group	13	2.82	.71	.20
	Posttest Experimental Group	12	2.55	.50	.14
E	Posttest Control Group	21	2.49	.65	.14
	Posttest Experimental Group	23	2.57	.66	.14

Table 15 shows variations among the schools. There was a mean posttest score range of 2.49 to 2.86 for the control group. The experimental group’s posttest means ranged from 2.45 to 3.06. The mean differences between the pretest and posttest scores ranged from minus 0.41 to plus 0.37 among control and experimental groups posttest scores. There are no significant differences in the posttest art performances of students among the schools.

**Inferential Statistics**

**The range, median and quartile distributions by school.**



**Figure 5.** Quartile (Box Plot) and Median Posttest Distributions for Control and Experimental groups on Art Performance by School

Figure 5 shows variations among the schools. The median scores for schools A and E were higher for the experimental group, stayed the same for school D and were lower for schools B and C. The third quartile scores for schools A and E were higher for the experimental groups, were the same for school B and were lower for schools C and D. However, the differences are not significant.

### T-Test Comparing Posttests by School.

**Table 16.** Levene’s Test for Equality of Variances (T- Test) Comparing Posttest Art Performance Scores for the Control and Experimental Groups by School

Independent Samples Test											
		Levene's Test for Equality of Variances					t-test for Equality of Means				
School ID			F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Differenc e	95% Confidence Interval of the Difference	
										Lower	Upper
A	Posttest	Equal variances assumed	.05	.83	-1.47	30	.15	-.36	.25	-.87	.14
		Equal variances not assumed			-1.47	29	.15	-.36	.25	-.87	.14
B	Posttest	Equal variances assumed	.35	.56	.60	31	.55	.14	.23	-.34	.62
		Equal variances not assumed			.60	30	.55	.14	.23	-.34	.62
C	Posttest	Equal variances assumed	.08	.78	1.75	35	.09	.41	.24	-.07	.89
		Equal variances not assumed			1.75	34	.09	.41	.23	-.07	.89
D	Posttest	Equal variances assumed	2.03	.17	1.07	23	.30	.27	.25	-.25	.78
		Equal variances not assumed			1.08	22	.29	.27	.25	-.24	.77
E	Posttest	Equal variances assumed	.14	.71	-.38	42	.71	-.07	.20	-.47	.32
		Equal variances not assumed			-.38	42	.71	-.07	.20	-.47	.32

Table 16 indicates there are no significance differences in the posttest art performances of students among the schools.

## Analysis of Variance.

**Table 17.** Analysis of Variance (ANOVA) of Pretest and Posttest Mean Score Comparisons by School Indicating Level of Significance for Art Performance.

School A			Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	(Combined)	0.034	1	0.034	0.072	0.791
	Within Groups		14.407	30	0.480		
	Total		14.441	31			
Posttest	Between Groups	(Combined)	1.05	1	1.05	2.156	0.152
	Within Groups		14.603	30	0.487		
	Total		15.653	31			
School B			Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	(Combined)	0.002	1	0.002	0.004	0.948
	Within Groups		17.358	31	0.560		
	Total		17.36	32			
Posttest	Between Groups	(Combined)	0.161	1	0.161	0.385	0.554
	Within Groups		13.92	31	0.449		
	Total		14.081	32			
School C			Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	(Combined)	0.101	1	0.101	0.161	0.691
	Within Groups		22.001	35	0.629		
	Total		22.102	36			
Posttest	Between Groups		1.565	1	1.565	3.069	0.089
	Within Groups		17.852	35	0.51		
	Total		19.417	36			
School D			Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	(Combined)	0.364	1	0.364	0.572	0.457
	Within Groups		14.632	23	0.636		
	Total		14.996	23			
Posttest	Between Groups		0.438	1	0.438	1.135	0.298
	Within Groups		8.877	23	0.386		
	Total		9.316	24			
School E			Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	(Combined)	0.181	1	0.181	0.474	0.495
	Within Groups		16.011	42	0.381		
	Total		16.192	43			
Posttest	Between Groups		0.059	1	0.059	0.137	0.713
	Within Groups		.18.012	42	0.429		
	Total		.18.071	43			

a. With fewer than three groups, linearity measures for pretest and posttest scores cannot be computed.



Table 17 Indicates that there are no significant differences in the pretest to posttest means for any of the schools.

### Effect Size

**Table 18.** Effect Size and Cohen's *d* by Individual School

School	Cohen's <i>d</i>	Effect-size <i>r</i>
A	0.532	0.257
B	- 0.208	- 0.103
C	- 0.573	- 0.275
D	- 0.439	- 0.214
E	0.122	0.060

Two schools A and E had positive effect sizes ranging from a small effect size of 0.06 to a medium positive effect size of 0.257. Three schools, B, C and D had negative effect sizes ranging from a small effect size of 0.103 to a medium negative effect size of 0.275.

### H2 – H4, Treatment of Extraneous Variables

H2 The null hypothesis for all schools stated that the extraneous variable of FCAT-NRT reading scores did not affect student performance.

H3 The null hypothesis for all schools stated that the extraneous variable of gender did not affect student performance.

H4 The null hypothesis for all schools stated that the extraneous variables of age, free/reduced lunch status, instructional time and budget did not affect student performance.

### Descriptive Statistics

**Table 19.** Descriptive Statistics for the Extraneous Variables by school

	Age Years	Gender Percent Male	Ethnicity Percent Minority	FCAT-NRT Average Reading Percent	Free/ Reduced Lunch, Percent Receive	Hours of Art per year	Supply Money Dollars per Student per Year
School A control	11	45	61	56.5	64	22.5	2.00
School A experimental	11	56	78	63.12	61	25.25	2.00
School B control	10.3	47	53	70.1	59	29	3.00
School B experimental	10.6	37	50	73.3	56	29	3.00
School C control	10.6	52	29	72.33	12	23.5	5.15
School C experimental	10.7	75	20	80.82	0	29.5	5.15
School D control	10.6	31	0	59.8	15	*13.5	1.00
School D experimental	10.4	0	8	72.18	58	*12.5	1.00
School E control	9.48	45	25	76.33	23	30.75	3.00
School E experimental	9.53	48	26	74.3	26	28.5	3.00

\* Excluding free drawing on Fridays and 6 hours of clay projects

## Inferential Statistics

**Table 20.** Correlations (Pearson r) Between Posttest Art Performance Scores and Other Study Variables

		Posttest
Age	Pearson Correlation	-.02
	Sig. (2-tailed)	.81
	N	169
Gender	Pearson Correlation	-.19*
	Sig. (2-tailed)	.01
	N	171
Ethnicity	Pearson Correlation	.10
	Sig. (2-tailed)	.21
	N	171
FCAT NRT 02-03	Pearson Correlation	.23**
	Sig. (2-tailed)	.00
	N	158
Free or Reduced Lunch	Pearson Correlation	-.14
	Sig. (2-tailed)	.07
	N	171

**Table 20 – Continued.**

		Posttest
Hours Per Year	Pearson Correlation	-.02
	Sig. (2-Tailed)	.07
	N	171
Supply Money	Pearson Correlation	-.08
	Sig. (2-tailed)	.29
	N	171

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level.

Tables 20 indicates that two extraneous variables had a significant correlation with the posttest scores. They were gender and F-CAT NRT scores. Although there is a significant effect of gender and FCAT scores on the posttest scores, statistical significance of the pair wise comparisons did not change after taking into account of the confounding effect of gender and FCAT from the study results.

### **Additional Findings**

#### **Journal Scores**

Journals were scored by the researcher using a holistic rubric (appendix C).

**Table 21.** Experimental Groups' Journal Scores by School.

School	1	2	3
A	38%	34%	28%
B	38%	49%	13%
C	45%	55%	0%
D	75%	25%	0%
E	35%	61%	4%

Three schools, A, B and E had over 60 % of their students with a journal score of either 2 or 3; A and B with 62%, E 65%. School A had the largest percentage of students with journal scores of 3 at 28%, followed by school B with 13% and school E with 4%. Two schools, C and D, had no journals receiving a score of 3.

### **Additional Assessments**

Schools A, C, D and E, all used assessments in addition to the journal writing. School B used only the journal writing. School A administered only three assessments spread throughout the year. They included on skill oriented painting checklist, one textbook vocabulary worksheet and one teacher made content test. School C's art teacher included one or two assessments for each unit. Each unit included several lessons.

The art teacher at School D added clarifying sentences following the journal questions. The phrases add to each question corresponds with the number of the original journal questions. After the first artwork, the additions included: 1) "to be a better artist?" 2) "How well did you follow directions?" 3) "What parts of this artwork would you like to do better?" and 4) "How does this lesson help me be a better artist - to reach my goal?" Additions to the middle journal questions following completion of all but the first and last artwork of the year included: 1) "How well did you follow directions?" 2) "What parts of this artwork would you like to do better?" and 4) "What did you learn from this art lesson?" Changes to the journal questions to be answered after the last artwork included: 1) "How will did you follow directions?" 2) "What parts of this artwork would you like to do better", and 3) "Did your art skills get better?" School D also used two additional assessments that required the students to rate themselves 0-3 on a number of factors. School E had students do the largest number of additional assessments in addition to the journal writing, often three per art project. These assessments included, self and peer checklists and questions, vocabulary definitions, short answer content questions, and extended journal questions.

## CHAPTER V

### ANALYSIS OF THE DATA

#### Research Hypothesis and H1

**Research Hypothesis:** The research hypothesis states that setting annual goals and answering self-reflection and self-evaluation questions after each art assignment will contribute positively to fourth and fifth grade student performance in visual art classes as measured by an authentic portfolio assessment process established by Dorn, Sabol, and Madeja (2003).

**H1** The alternative hypothesis states that students' mean art portfolio scores will increase as a result of the intervention, metacognitive/self-reflective journal writing.

The data does not reveal significant gains in the mean art performances of the students in the experimental group inclusive of all five schools. The experimental group made a small gain of .05, which did not represent a significant gain.

The mean art performance posttest score gains however, varied substantially from school to school. The data supports meaningful gains at school A, some gain at school E, little or no gain at school B, and some reduction in gains at schools C and D. School A had a meaningful gain while school E had a small gain in their portfolio 1 to portfolio 2 mean scores. In addition, the first, second and third quartile scores of their experimental groups were higher than the respective scores for their control groups. School A's and E's experimental groups showed positive mean score difference of .37 and .08 respectively compared to their control groups. In school A's case there was a clear and meaningful difference in student performance. In school A's case, there was a positive effect size of .257. With only 14 and 18 respectively in the control and experimental groups proving statistical significance is difficult and the gain was therefore not statistically significant. However, an effect size of .257 indicates that the intervention was mildly to moderately effective. Perhaps most striking in this case was the first quartile of the experimental group equaled the median of the control group indicating a dramatic gain in the majority of the lowest performing students' growth in visual art class. School E also showed higher first, second and third quartile scores indicating a meaningful positive change.

Schools B's, C's, and D's experimental groups all had lower median posttest scores than their control groups. Examining the box plots in Figure 5 revealed further differences between

the control and experimental group's posttest scores. School D's posttest score range for the experimental group was lower than the control group's with a posttest range of approximately 1.7 and a high of 2.6. School B's posttest scores for the experimental group also indicated a lower range than the control group's posttest scores. The fourth quartile collapsed into the third with a posttest score spread of 1 to approximately 3.4. School C's posttest score range for the experimental group covered the full range from 1 to 4. However, the first, second and third quartile scores of the experimental group were lower than the control group's.

Schools C and D journal scores were lower than schools A's, B's and E's journal scores. Schools C and D had no journals receiving a score of 3. School C had 55 % of its journals receiving a score of 2 while school D only had 25%. In both cases the control groups had an increase in their mean scores pretest to posttest while their experimental groups had a decrease in mean scores pretest to posttest. In school C's case the control group had an increase in mean score of .45 while the experimental group had a decrease of .10. School D's control group had an increase in mean score of .26 while the experimental group had a decrease in mean score of .25. In both cases, the considerable swing in mean scores was counter to the research hypothesis.

School A with the highest percentage of journals scoring a 3, 28%, showed the most gain in overall student performance. School E with 4% of the journals receiving a 3 also showed an overall gain in student performance. However, school B showed an overall decrease even though 13% of their journals received a score of 3 and the mean score remained the same pretest to posttest. One difference between schools A and E and school B is that schools A and E both gave assessments in addition to the journal writing. School B only had the students write in the journals. Schools C and D also gave additional assessments.

## **H2 through H4, Extraneous Variables**

**H2** The null hypothesis stated that the extraneous variable of FCAT-NRT reading scores did not affect student performance.

The statistical analysis indicated that higher FCAT-NRT (norm referenced test) reading scores positively correlated with students' posttest scores but did not affect the overall study results. This held true when the individual school results were examined. With the exception of school E, all the control groups had lower FCAT-NRT scores than their experimental groups. One therefore would have expected all but one school to have higher posttest scores. This was

not the case, only two experimental groups had higher posttest mean scores. School E's experimental group had lower FCAT-NRT reading scores than the control group however, the experimental group out performed the control group.

**H3** The null hypothesis stated that the extraneous variable of gender did not affect student performance.

Being female positively correlated with the posttest scores. However, gender did not affect the overall results of the study. Gender findings were mixed, schools A and C had more girls in the control group than the experimental group, schools B and D had more boys in the control group, and school E had the same number of boys and girls in the control and experimental groups. Only schools A's and E's experimental groups had higher posttest scores. Results for School A's experimental group's went against the gender trend with a higher percentage of boys and higher scores. Whereas, schools B and D both had more girls in the experimental group than the control group but their control groups had higher posttest scores. As a result, gender did not affect the overall results of the study.

**H4** The null hypothesis stated that the extraneous variables of age, free/reduced lunch status, instructional time and budget did not affect student performance.

The demographics of the two schools with mean posttest score gains were different. School A's experimental group had the lowest FCAT-NRT score with 63.12; the highest percentage of students receiving free/reduced lunch of 61%; the most minority students; the second smallest supply budget, \$1.00 per student; the oldest students, average age 11; and the highest pretest and posttest scores, 2.85 and 3.06 respectively. School E's experimental group in contrast had the youngest students with an average age of 9.53, the second highest FCAT-NRT score with a mean of 74.3, the second lowest free/reduced lunch percentage at 26 percent, and the highest ethnicity mean of 2.26 indicating primarily majority and other students in this case, Asian students, a supply/materials budget of \$3.00 per student, and the lowest pretest and second lowest posttest score, 2.35 and 2.57 respectively. Since both schools, A and E, contrast demographically and their experimental groups showed more gain than their control groups the extraneous variables did not account for the findings in this study.

At first glance, School D's number of hours of instruction seems to explain the differences in the experimental group's posttest mean score drop. School D's hours of instruction for 2-D lessons were considerably lower than any other school's with 12.5 for the experimental

group for the year. However, School D's control group with 13.5 hours of instruction showed a gain in mean score of .45 while the experimental group had a mean score drop of .25 which indicated that factors other than instructional time were responsible for the difference in posttest scores. School A also recorded only the hours spent doing 2-D work for the study. School A's experimental group spent 25.25 hours on instruction for the study while the control group was able to complete the 2-D work for the study in 22.5 hours. School A's instructional time records indicated that the intervention did add to the instructional time needed to complete assignments for the experimental group.

School C had the highest FCAT-NRT scores, the fewest students receiving free/reduced lunch, the lowest number of minority students and the largest supply/materials budget with \$5.15 per student per year. The pretest and posttest results were unexpected and counterintuitive. The control group had a gain of .45 while the experimental group had a drop of .07. These results indicated that the extraneous variables, FCAT-NRT reading scores, percentage of students receiving free/reduced lunch, the size of the budget and the percentage of minority students did not determine growth in student performance.

### **Equality of Groups**

Since this study was quasi-experimental in nature it was necessary to determine the kind of statistical analysis necessary for examining the data by determining whether the control and experimental groups were equivalent for the pretest period. Analysis determined that there were no significant differences between the groups and therefore equivalent statistical methods were used to analyze the data in this study.

### **Inter-Rater Reliability**

Inter-rater reliability reports the level of agreement among raters. Two statistical instruments were used to determine the inter-rater reliability, Hoyt's Alpha and the Intra-Class Correlation Coefficient (ICC). The ICC for both the experimental and the control groups at .61. The Alpha for both groups was .82. The scores for the ICC and the Alpha are consistent with prior studies which were considered acceptable (Dorn and Sabol, FSU).



Since many portfolio or performance based assessments for example, National Board of Professional Teaching Standards (NBPTS), Advanced Placement (AP), and International Baccalaureate (IB), require the agreement of two judges. In the AP portfolio process used in this study a 1- point difference in scores between judges was considered acceptable, but a wider spread would indicate that the judges were looking at different features of the work or suffering from “reader fatigue”. The 1- point difference standard has been used by the Educational Testing Standards in scoring the studio art program and similar measures are employed by the IB program (Dorn, Madeja and Sabol, 2004). The scoring for this study followed the AP format. When there was a disagreement of more than one point, the judges were asked to look at the portfolio again and asked what score they thought they had given the portfolio. In most cases, one judge or the other changed their score, which brought the scores to within a point of agreement. However, some judges chose to leave their original score. With this in mind the researcher did additional calculations to determine the percent of agreement among the judges. The researcher found that on the pretest 28% of the time all three judges had an exact match, 68% of the time two out of three judges scores matched and only 4% of the time did all three judges rate the portfolio differently. Two or more judges were in agreement 96% of the time with no more than a 1-point difference for the third judge. For the posttest 29% of the time all three judges rated the portfolios the same, 70% of the time two out of three judges scored the portfolios the same and less than 1% of the time the judges disagreed. For the posttest 99% of the time, two or more judges were in agreement with no more than a 1- point difference for the third judge.

### **Score Spread**

Since the portfolios were scored within the context of each school’s set of artwork, it was necessary to determine two factors. First, was there a sufficient score spread, 1- 4? And second, did the instrument (the rubrics) discriminate among portfolios.

Figure 3 indicated that the judges discriminated giving scores covering the full range from 1 to 4 on the pretest and the posttest. Figure 3 discriminated further indicating that there were definite differences among schools. The judges’ concern arose from the fact that they had been directed to use the whole score spread and score the portfolios within the context of that school’s set of lessons. Some schools’ posttest portfolios were definitely stronger than other schools’ posttest portfolios. School A’s experimental group had the highest mean and second and

third quartile scores while school D's scores collapsed into the middle. There were definite and distinct differences and the schools posttest could be ranked high, medium, or low. School A had the highest mean posttest score and school C had the lowest mean posttest score while school D had the lowest posttest score range overall.

### **Validity**

Beattie (1997) identified twelve validation criteria for performance-based assessments; relevance, content fidelity and integrity, exhaustiveness, cognitive complexity, equity, meaningfulness, straightforwardness, cohesiveness, consequences, directness, cost and efficiency and generalizability. She also noted that not all the criteria need to be addressed for every assessment. In this study, there was a solid match between the purpose of the assessment and the assessment format. The purpose of the assessment was to measure students' art performance overtime both with and without metacognitive/self-reflective journal writing. The students' actual artworks created throughout the year as a part of their regular class instruction provide the data for the assessment. The students' actual artworks were the evidence of learning. The assessment method was straightforward based on the assignments the students complete in art class where the expectations for the lessons were openly clarified and discussed. By using student artworks as the evidence content fidelity was also high incorporating major content and processes in the field of art education, making art. Equity was one of the main strengths of this assessment method. All students have an opportunity to succeed. Success was achieved through either maintaining high portfolio scores or increasing portfolio scores. Students are compared to themselves not others. The assessment method was also very direct, artworks were made and artworks were assessed. Meaningfulness and cognitive complexity were addressed with greater depth by the experimental group where processes, skills, metacognitions and attitudes were recorded in their journals and students were empowered to make independent, meaningful and new connections. The cost and efficiency issue primarily affected the teachers' out of class time. Although the experimental group did need class time to write in their journals. The artwork and journal data were both easily collected. Having met many of the criteria identified by Beattie (1997) this study had a high degree of validity.

## CHAPTER VI

### RESULTS, IMPLICATIONS, RECOMMENDATIONS AND CONCLUSION

#### Results

Prior research by Dorn, Madeja, and Sabol (2003) indicated that students at some grade levels, particularly fourth and fifth grades, experienced a leveling off in their art performance. This study was designed to address the leveling off issue and to see whether student art performance could be increased while at the same time addressing accountability issues using an authentic assessment methodology appropriate for visual art classes. An intervention, metacognitive/self-reflective journal writing, based on the literature related to metacognitive learning theory, self-regulated learning theory (SRL) and the Dimensions of Learning (DOL) model was designed and implemented. In addition, descriptions of projects such as Project Zero and PROPEL (Brandt, 1987; Wolf, 1989) contributed to the development of the journal questions. The implementations of the intervention was limited in this study by two factors, the teachers' desire to first gather baseline writing data, to see what students would write on their own, and second the teachers' lack of experience in the digital portfolio process. In the case of the art teacher/researcher, who was familiar with the literature, had experience in using both the Dimensions of Learning program and the digital portfolio process, the data indicated student portfolio scores improved meaningfully and the implementation was effective. Conversely, in cases where the implementation of the metacognitive/self-reflective journal writing was not consistent with the intervention theory the results were inconclusive and in two cases counter to the research hypothesis, although not statistically significantly.

The findings of schools A and E suggest that a self-regulatory intervention of metacognitive/self-reflective journal writing, may positively affect student art performance. However, the findings also suggest that some writing strategies are stronger than others. Strategies that incorporated the journal writing as an instructional tool, as a way of developing students' awareness of themselves as learners and artists resulted in improved performance scores. In cases where the value and purpose of the journal writing was unclear, the results were not conclusive.

Examining the literature (Schunk, 2001) on self-regulated learning (SRL) in relationship to the implementation overall, provides insight into the inconsistencies found in the study's results. The social cognitive view of self-regulated learning theory suggested that students' willingness to self-regulate was based on two cognitive sources, 1) goals, self-efficacy and outcome expectations and 2) involved three self-regulatory processes. The three processes were forethought, performance control and self-reflection. According to SRL theory, personal goals (forethought) served as standards against which future progress would be evaluated. Unfortunately, there were instances when the students only addressed the art material they wanted to use, i.e. clay, not what they wanted to be able to know or do better. In such cases, two problems arose. First, the students may not have understood how to set a learning goal and second, the teachers were unclear on how to include student goals in their instructional strategies. It is interesting to note that a number of students in the study stated goals related to drawing more realistically, a developmentally appropriate goal. Schunk (2001) found that self-set goals promoted greater self-efficacy, motivation and confidence. He also indicated that proximal goals proved more effective than distant goals in increasing productivity, self-efficacy and achievement. In this study, setting both long-term and short-term learning goals may have resulted in more consistent student performances, school to school.

Examining the data from the journal writing, differences in the schools' journal scores (only the experimental groups wrote in journals) were observed. In three cases, schools A, B and E, over 60% of their students scored a 2 or 3 on their journals. Two of these schools, A and E, also had an increase in their pretest to posttest mean portfolio scores. The increase for both schools, A's and E's, experimental groups was greater than the increase for their control groups who did not participate in any assessment or metacognitive journal writing. School B's experimental group's pretest to posttest mean score remained the same. School B only used the metacognitive journal writing intervention while schools A and E both did additional assessments. School A only incorporated three additional assessments during the entire year and relied primarily on the metacognitive/self-reflective writing strategy as a learning/assessment intervention.

Schools C's and D's journal scores were lower than schools A's and B and E's journal scores. Schools C and D had no journals receiving a score of 3. School C had 55% of its journals receiving a score of 2 while school D only had 25%. In both cases, the control groups had an

increase in their mean scores pretest to posttest while the experimental groups had a decrease in mean scores pretest to posttest. In both cases, the considerable swing in mean scores was counter to the research hypothesis.

The journal scores reflect, in part, the level of implementation of the intervention, metacognitive/self-reflective journal writing. Schools A, B and E according to the journal scores implemented the process more fully than schools C and D. Moreover, in the cases of schools A and E the experimental groups showed a definite increase in student performance. School A with the highest percentage of journals scoring a 3, 28%, showed the most gain in overall student performance. School E with 4% of the journals receiving a 3 also showed an overall gain in student performance.

The art teacher's logs also revealed implementation differences and challenges. The art teachers at schools C and D expressed frustration with the journal questions. The art teacher at school D's restatement of the questions not only demonstrated her concern with the meaning of the questions but in some cases changed the question from an open ended format to a close ended format, i.e. "How well did you follow directions?" A self-reflective or self-evaluative answer would have explained how the artwork was made as well as choices the student made in the process of making the artwork not just whether the directions were followed. The art teacher at school C stated in her teacher log, "the questions were not child friendly and difficult to understand. I had to continually remind them [the students] what each question meant and was often asked, 'Can I just answer this question the same way I did the last assignment.'" The students' desire to answer the question the same as the last indicated that the student did not understand the metacognitive context specific nature of the journal questions. The students' answers needed to be made in the context of each assignment; not a generic, "I followed all the directions." These students and possibly their teachers treated the implementation as an added assignment not an integral instructional strategy designed to improve their performance. In other cases, especially schools A, B and E, students demonstrated their understanding of the metacognitive nature of the questions by writing for example, "I took my time and planned". These students were able to express an awareness of their learning.

Time was another factor many of the teachers reported as a challenge that affected the implementation of the journal writing. Even the art teacher/researcher admitted to having the students write about the first four projects at one time, calling students back during another class

to finish the journal entries and having the students write in another class outside of the artroom. One teacher stated in her teacher log, "...we ran low on time and could not continue taking an entire period to fill out our journals." The art teachers felt a constant pull between making enough artwork to qualify for inclusion in the project and fully implementing the journal writing. All of the art teachers were surprised when they totaled their total minutes of instruction for the year. All believed and did have more instructional time per year than they were actually able to count. Holidays, fieldtrips, and testing were the main sources of missed instructional time. One elementary school was disqualified from the study due to a lack of work for the majority of the students. Pull out programs like strings, chorus, and gifted removed students from the art class to participate in those programs. All the schools included in the study lost instructional time for some of their students due to pull out programs.

In the social cognitive approach to SRL performance control, the second self-regulated learning process, included two strategies, feedback and self-verbalization. Both were of particular importance to this study. Schunk (2001) found that ability feedback to early successes as opposed to participation feedback led to higher ability attributions, self-efficacy and achievement. Effort feedback, on the other hand, was appropriate only where prior experiences were unsuccessful. The other teacher constructed assessments some of the art teachers created may have been counter productive. In the cases of schools C and D where the control groups showed greater gains than the experimental groups the other assessments also tended to focus on what Schunk would classify as effort feedback. While appropriate for the unsuccessful student, very few students in the study fell into this category.

The self-reflection phase, according to Schunk (2001) focused on three subprocesses, self-monitoring, reward contingencies and progress feedback and self-evaluation. In this study, we implemented the self-evaluation process but not the reward contingencies. The only self-monitoring process used and recorded by the students in the study was level of effort.

The data from this study was consistent with Schunk's (2001) theories of self-regulated learning. Schunk (2001) found that student achievement and self-efficacy were highest when rewards were based on performance contingencies rather than participation contingencies, when learners could determine they were making progress toward goal attainment, and when learners could self-evaluate based on clear criteria. Should the study be replicated, guidelines for setting clear criteria need to be developed and implemented according to Schunk's (2001) research.

Three interactive subprocesses were identified in the social cognitive view that leads to self-efficacy and skill growth in self-regulated learning process. The three interactive subprocesses identified were self-observation, self-judgment and self-reaction. Self-observations prompted learners to self-evaluate which lead to personal and behavioral self-reactions. Regularity and proximity were identified as two important aspects of recording self-observations. Implementing the metacognitive/self-reflective writing strategy with both regularity and proximity were constant challenges the art teachers faced during the study. In many cases, the artwork was completed a week or more before the journal writing took place. In other cases, teachers admitted to having the students write about several artworks all at one time or had another teacher supervise the writing process in another class. The art teachers in the study also reported the pressure to produce six to eight 2- D works of art made it difficult to schedule time for the journal writing.

Self-judgments according to Schunk (2001) are made by comparing existing performance levels with one's learning goals. Factors that affect self-judgments include types of standards employed, properties of the goal, and importance of attaining the goal. Throughout this study, the standards, the properties, the performance levels, were set by the five teachers or in a few cases with the teacher. In the schools, A and E, where the standards and properties were clear and focused on art processes and art content, even though the teacher set them, students tended to show increased art performance. In the two cases where the teachers used checklists that focused on completion and inclusion, students' art performance was lower than students who did no writing or assessments.

Self-reaction according to Schunk (2001) involved students judging goal progress as acceptable and anticipating goal accomplishment. Students who believed they could perform better persisted longer and worked harder. This belief in the ability to improve or accomplish the goal was a motivating factor. Occasionally in this study, students' answers addressed how they felt they were doing. This is an issue that needs to be addressed more directly in future studies.

In this study, the implementation of the self-regulated learning theory by the five teachers was inconsistent. Developing teaching strategies that address the forethought, performance and self-reflection loop of self-regulated learning along with consistent implementation of the metacognitive/self-reflective writing may be a key factors in achieving positive results in students' visual art performance. As Perkins (1996) stated, "...experiential intelligence needs

reflective intelligence to manage its powers for a fuller perception of art – and more generally for better thinking about anything” (p. 16).

**Extraneous Variables.** While it is tempting to explain the findings in terms of the extraneous variables, statistical analysis confirmed that they did not affect the overall findings in this study. Two extraneous variables, gender and FCAT-NRT reading scores, had a significant correlation with the posttest scores. Being female and having higher FCAT-NRT scores had a positive correlation with higher portfolio scores. However, in both cases, there were counter examples to the two trends that support the statistical findings. The experimental group at school E had a lower FCAT-NRT reading score mean than school E’s control group and school A’s experimental group had more boys. The other extraneous variables, age, free/reduced lunch status, instructional time and budget did not correlate significantly with the posttest scores or the findings in the study. In fact, the school, A, with the lowest FCAT-NRT reading scores, the most boys, the largest number of minority students, the largest percentage of students receiving free/reduced lunch and the second lowest budget had the highest experiment group mean score, which was also the highest posttest score. School A had an advantage in that both the art teacher and the students’ classroom teacher had prior experience with the Dimensions of Learning model that fosters the use of metacognitive/self-reflective strategies. For example, the students already knew how to set goals, and stated “we just did that for Ms. W.” The art teacher did not have to explain what a goal was, only direct the students to set goals that would positively affect their art performance. In addition, the art teacher had prior experience with the digital portfolio process and was the researcher for this project. Because the technological aspects of the study were not an issue for the art/teacher researcher, she could focus on implementing the SRL intervention of metacognitive/self-reflective writing.

### **Summary**

A number of implementation issues became apparent at the beginning of the project and were known compromises. While other, unforeseen issues arose later. The most significant implementation compromise at the beginning of the study was the art teachers’ lack of willingness to incorporate metacognitive/self-regulated learning strategies into their teaching methods. As a result, only school A gathered data in a manner somewhat consistent with the literature. What became clear is that the teachers needed to be introduced to new methods of assessment and strategies for improving student performance gradually, over a period of a couple



of years. First, mastering the digital portfolio process and then utilizing new instructional strategies. In a couple of cases, had the teachers been able to examine their students' digital portfolio scores in terms of individual and over all progress prior to the study, motivation to try the metacognitive/self-regulated learning strategies may have been found in the data. Not having any prior performance information the teachers went into the study expecting the majority of their students to do well and show improvement. The teachers also, believed the assessment methods they historically used supported student performance. It never occurred to the teachers that some methods of assessment could hinder student performance. On the positive side, having the teachers vary from the study's ideal methodology provided insight and raised important questions regarding effective assessment strategies for visual art classes.

Time was an unforeseen complication in the study. In mid-November, when a large number of the participating schools only had one or almost two artworks completed it became apparent that a number of the schools were going to have a very difficult time meeting the portfolio requirements, 3 to 4 completed artworks for each student by the end of the semester. This proved to be true, only half of the original sample finished the study. The remaining teachers all expressed concern during the support meetings about the difficulty of finding time to write and also finish the required number of artworks for the study. This was an on going battle and an issue that needs to be examined, possibly at an administrative or scheduling level.

Art teachers need both a reason and time to change practice. The art teachers, at least in Leon County, had no prior measure of student art performance. As a result, there was no motivation to change or examine current practices making it difficult for the researcher to impose not only on their time but also on their teaching strategies.

### **Recommendation for Further Study**

The researcher recommends that a follow up study be replicated in other school settings using other methods and other forms of analysis for evaluation. In addition, the varying strategies for self-regulated learning should be tested in other school environments. These studies should include the four common features of self-regulated learning (SRL) identified by Zimmerman (2001) and should serve as basic premises for the implementation of the intervention, metacognitive/self-reflective journal writing. First, students must be active participants in their own learning process, metacognitively, motivationally, and behaviorally. Second, students must use a self-oriented feedback loop. Third, students must be aware of how and why they chose a

particular self-regulated process, strategy or response. Fourth, students and teachers must understand that self-regulation takes more time preparation, vigilance and effort on the part of the student.

In this approach, forethought and goal setting, addressed in both the self-regulated learning literature and the Dimensions of Learning model, needs to be implemented thoughtfully and directly instructed. Students need to be taught how to set goals. Long term goals as well as short term goals should be established. Teachers should find ways to address students' goals within the context of the lesson. This may be as simple as directly instructing a particular technique or as complicated as finding ways to include personal relevance into a skill practice session.

Feedback and self-verbalization, two performance control strategies, should be ability linked to early success or failing early success to effort. Questioning students using one or more of Dimension 5's (DOL) critical, creative or self-regulation skills could lead to a deeper understanding of one's performance.

During the self-reflective phase of self-regulated learning three subprocesses play a role, self-monitoring, reward contingencies and progress feedback and self-evaluation. Monitoring should be undertaken internally and externally. Reward contingencies should be based on performance criteria not completion criteria unless a student is not producing work. Feedback and self-evaluation should again be based on performance contingencies or observations that one is making progress toward one's goal. Brandt (1987), Wolf (1989), Arter and Spandel (1992) all discussed the use of self-reflection as a means of promoting student learning. Arter and Spandel (1992) stated, "self-reflection is one thing that makes a portfolio instructional" (p.37). They asked questions to prompt student self-reflection. Examples include:

"Describe the process you went through to complete this assignment. Include where you got ideas, how you explored the subject, what problems you encountered, and what revision strategies you used. ... What makes your most effective piece different from your least effective piece? How does this activity relate to what you have learned before? What are the strengths of your work? What still makes you uneasy?" (p. 40).

Project Zero, Brandt (1987) reported used reflection as a means to step back from both ones perception and production and ask, "What am I doing? Why am I doing it? What am I learning? What am I trying to achieve? Am I being successful? How can I revise my performance in a desirable way?"(p. 32). The foregoing questions exemplify the self-reflective phase of self-

regulated learning and represent some of questions art teachers should ask students as they circulate and observe the students art making and journal writing.

Three subprocesses of self-regulated learning identified by Schunk (2001), self-observation, self-judgment, and self-reaction also need to be considered. Quality, quantity and originality are factors that can be observed during self-observations. Attention should be paid to two key characteristics, proximity and regularity. The journal writing should be a regular practice and take place immediately following the completion of an artwork. Gaps in time make it difficult for the students to remember what they were doing, thinking and feeling. Concerning originality teachers should refer to dimension 5 of the DOL model as well as the literature on creative thinking. SCAMPER, an acronym for divergent thinking (Starko, 2001), is concrete and offers a variety of strategies to push one's limits.

Self-judgment takes place when the students compare their present performance to goals. The goals may be absolute standards, normative standards, social standards or personal standards. Regardless, three factors must be taken into consideration when selecting the goals, specificity, proximity, and difficulty level. First, the more specific the standard the easier it is for the student to compare their work to that standard. Second, comparing their performance to short term goals is more beneficial than comparing to long term goals. Finally, the level of difficulty should be challenging but attainable. The researcher's recommendations for practice suggests that goals should included short term as well as long term goals, include specific standards and should be challenging but attainable.

Self-reaction, the third subprocess of self-regulated learning is concerned with students' self-reactions. Students' motivation is increased when the student perceives the level or rate of progress toward goal attainment as satisfactory. Negative evaluation does not necessarily decrease motivation if the student believes one can improve. Tangible consequences can enhance student self-efficacy when tied to the student's actual accomplishments. Rewards and grades need to be clearly tied to the student's accomplishment. The researcher recommends that students include self-reaction statements. These statements should be focused on what has been achieved, progress that has been made and ways to improve performance when the student feels progress toward the goals is deficient.

The role of the art teachers must not be forgotten. In order for students' to implement the metacognitive/self-reflective writing process training, direct instruction in the processes and

monitoring on the part of the art teacher are necessary. Students must be active participants in their art learning, making and evaluation. They must make use of a self-oriented feedback loop comparing their performance to goals. Time must be included in the schedule for the journal writing. The writing must be valued as a tool that enhances student performance, not as an add on or afterthought.

- Further research is needed based on a longitudinal study over two or more years. Teachers stated they found it difficult to complete the minimum of six 2-D artworks required for the pretest and posttest portfolios in a single school year. They also had students writing outside of the art room or writing several entries on the same day. It is recommended that the metacognitive/self-reflective journal writing continue even when artworks for scoring are not being collected. Since gathering enough works has been such a challenge, a two year process is recommended as one approach. Fall of one year for the pretest portfolio and fall of the following year for the posttest portfolio. Accountability issues could still be met while giving teachers and students time to engage in the process fully.
- Another testing scenario that would meet accountability demands would be to collect student portfolios at certain grade levels, i.e. 1<sup>st</sup>, 3<sup>rd</sup>, and 5<sup>th</sup>. These portfolios would be scored then compared to demonstrate growth over time.
- In order to more accurately assess the impact of the metacognitive/self-reflective writing and learning strategies on student performance, future studies should either eliminate the use of other assessments or give specific parameters for those assessments. Unexpected findings occurred in this study. In the cases of two schools, the students' scores for the experimental groups' posttests were lower than the control groups' posttest scores. Two factors seem to have contributed to these lower scores. First, none of the students received journal scores of three indicating the level of implementation of the self-regulated writing process was limited. Second, the nature of the other assessments often involved checklist or criteria that were based on inclusion and completion. The researcher as well as the literature recommends that completion and inclusion criteria only be used for students who are either not working or are not meeting other criteria successfully.
- A study comparing students who participate in other assessments which focus on specific criteria that is not completion or inclusion oriented in addition to the journal writing and students who only do the journal writing may be very useful in determining the role other

assessments play in developing students' art performance. In this study, the two schools' whose experimental groups scored higher than the control groups' both used other assessments that focused on art processes and elements and principles. In the case of the school that only did the journal writing students' performance for both the control and experimental groups was almost identical. The question remains, what role do other assessments play in developing student art performance, if any? Examining the role of other assessments in developing student art performance is definitely a topic for further research.

- If at all possible, conduct a study in the future that has a more controlled environment, i.e. common course content when comparing different classes on the effects of SRL on student performance, or selecting a small group of students from the same class with level 1 or 2 portfolio one scores and providing these selected students with a SRL intervention. In this study, each art teacher developed their own curriculum and course content, there was no consistency from school to school.

### **Implications for the Field**

While not all of the findings can be empirically confirmed, some of the study's findings do suggest:

1. That the encouragement of metacognitive behaviors in art learning may have a positive effect on the art performances of K-12 students in art.
2. That student self-regulation may encourage the cognitive growth of students in the art classroom.
3. That metacognitive strategies encourage self-efficacy and independent self directed learning in the art classroom through goal setting, self-assessment and self-reflection.
4. That the self-regulation art learning process encourages student involvement through becoming a stake holder in the learning process and in the setting of standards for their own behavior.
5. That self-regulated learning itself confirms the true nature of expressive behaviors which find its origins within the conscious thinking of the student rather than in the external forms imposed on the student by tradition and technique.

## Conclusion

The findings in this study suggest that metacognitive/self-reflective journal writing may improve student art performance when the implementation is consistent with metacognitive and self-regulated learning theories. One of the strengths of this study was the fact that the portfolio process provided a method for determining student growth over time to meet accountability concerns while the metacognitive/self-reflective journal writing provided a method of self-regulated learning designed to enhance student art performance through increasing students awareness of art making and art thinking processes. Assessment was more than an accountability tool it became a means of improving student performance. The metacognitive/self-reflective journal writing process involved students as active participants in their learning. The students set goals, they observed their own practices, they self-evaluated and they reflected on what they learned, and what they still needed to learn; skills and thinking processes used by artists in the field. Students were engaged in learning that came *from* the content of the art lessons. However, the findings indicate that simply asking the questions without instructing students in the use of strategies and developing the skills necessary to accurately self-evaluate and reflect do not in themselves lead to thinking, self-aware, self-regulated learners.

The advantage of the portfolio assessment process used in this study is that accountability was addressed through student performance, what students' actually do and make in visual art classes rather than some extraneous, disembodied test or assessment. Work was scored within the context of that specific school, that teacher's assignments and those students' performances. State and national standards could be met while at the same time local and individual goals were achieved. Student performance was quantified using the holistic rubrics and student progress was measured by comparing portfolio one scores to portfolio two scores. The digital assessment processes used in this study supports the ill-structured nature and multiple perspectives written into the arts standards. The art lessons varied considerably from school to school. Each teacher developed a curriculum unique to their school's interests based on the state and national standards. Hausman (1994) stated, "In writing standards and designing assessments strategies, we need always to keep in mind the spirit of art - an openness to ideas, a generosity of spirit, multiple perspectives, and an awareness of great traditions in creating and responding."

The results of the individual schools data also indicated differences that were measurable. One school clearly scored higher overall than the others, one school lower and three schools fell

in the middle. This study was not designed to make school to school comparisons but it was a concern stated during the scoring of the portfolios by the art teacher doing the scoring. The art teachers felt there were a couple of instances where the overall quality of a school's portfolios was lower than the other portfolios the teachers had scored. Their concern arose from the scoring process where they were requested to score the portfolios within the context of that school's body of work. Giving scores ranging from 1 to 4. The findings of this study supported what they observed. Even with the portfolios scored within the context of that school's body of work, one school clearly scored lower overall, another clearly higher. The implication of this finding may be that both outstanding and weak art programs could be at least initially identified through this portfolio process and then further evaluated to determine if the program is meeting school, district, state and national goals. This would add another layer of assessment for the purpose of accountability, program evaluation avoiding the use of standardized tests and forgoing the inherent limitations of those tests.

The digital portfolio assessment and metacognitive/self-reflective journal writing processes used in this study addressed three purposes of assessment or evaluation. First the digital portfolio process measured students' progress by comparing students' portfolio 1 scores to their portfolio 2 scores for accountability purposes. Second, the journal writing was a metacognitive/self-regulated instructional strategy designed to increase students' awareness of themselves as learners and as a result enhance student art performance. The effectiveness of the intervention strategy was measured by comparing the posttest scores of the control group to the posttest scores of the experimental group. Third by making school to school posttest comparisons individual school performance trends were evident indicating that this process may be valuable as a tool for evaluating art program effectiveness.

Art teachers in their search to develop methods of meeting the demands of accountability need to examine carefully their practices. While issues of completion, turning work in on time, and having a positive attitude may be factors in determining a student's grade for a term, these issues may not be appropriate individual goals that contribute to the students' art performance. Art may also need to redefine, in addition to meeting state and national standards, what we really want students to be able to know and do. Are the goals set by the state, the district, and the art teacher the same goals students would establish for themselves? Do the art teachers and students

have goals that have common ground? Should instruction be adjusted to meet students' learning and art performance goals?

While there was a need to address accountability, the strength of this study was in the use of self-regulated learning, metacognitive, and Dimensions of Learning strategies to enhance student art performance. These strategies' foundations were embodied in the work of Kant, Hegel and Dewey. Learning is a process of "doing and undergoing". All too often in art programs, educators concentrate on the "doing" and "how to do" and disregard the "undergoing". The methods used in this study required students in the experimental groups to at least make an attempt at undergoing, of reflecting on their work, on the process, on the product and how they felt about it. While not all of the findings were empirically confirmed, there was evidence that metacognitive/self-reflective journal writing as a self-regulated strategy enhanced student performance in 4<sup>th</sup> and 5<sup>th</sup> grade art classes. This is an outcome worthy of further study and investigation.



**APPENDIX A**

**HOLISTIC RUBRIC FOR STANDARDS-BASED  
ASSESSMENT IN THE VISUAL ARTS**

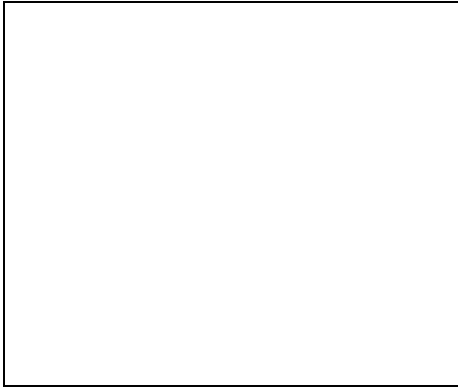
**HOLISTIC RUBRIC FOR STANDARDS-BASED  
ASSESSMENT IN THE VISUAL ARTS**

Grades 3-5

<p><i>EXCELLENT</i></p> <p><b>LEVEL 4</b></p>	<p>Shows objects from different viewpoints Consciously creates symbols Generalizes things observed Uses recombined shapes Notes subtle relationships between objects Makes invented shapes and objects Shows interest in future events Pairs and distributes similar forms Produces fantasy pictures Recognizes patterns Uses geometric shapes</p>
<p><i>VERY GOOD</i></p> <p><b>LEVEL 3</b></p>	<p>Makes color vary in value Shows objects in different environments Makes objects from memory Reveals actual or past events Orders similar shapes Represents events in a literal way Shows overlapping forms Varies positions of objects Sees and arranges similar shapes Uses texture</p>
<p><i>SATISFACTORY</i></p> <p><b>LEVEL 2</b></p>	<p>Places object in relation to where work began Makes shapes correspond with appearances Makes shapes in isolation Shows evidence of color</p>
<p><i>INADEQUATE</i></p> <p><b>LEVEL 1</b></p>	<p>Places objects randomly Makes object from one point of view Uses stereotypes rather than seeking likeness Places shapes or objects in unrelated spaces</p>

**APPENDIX B**  
**JOURNAL QUESTIONS**

## JOURNAL QUESTIONS



Artwork # or Title \_\_\_\_\_

Date \_\_\_\_\_

Student ID # \_\_\_\_\_

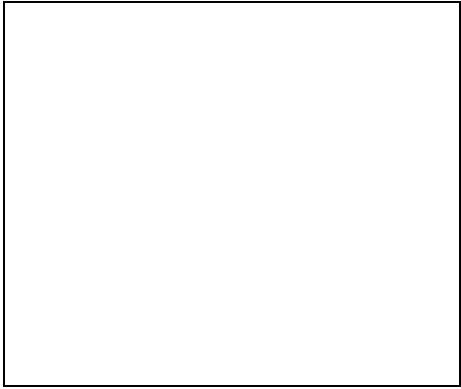
Effort Level \_\_\_\_\_

1-5

### **After first artwork:**

1. Name two things you want to learn to do better in art class this year? What skills will you need to learn?
  
  
  
  
  
  
  
  
  
  
2. To what extent did you meet the requirements of this assignment?
  
  
  
  
  
  
  
  
  
  
3. What sections of this artwork or skills would you like to improve? How? Why?
  
  
  
  
  
  
  
  
  
  
4. How has this project helped improve your art making or thinking? What did you learn? How will this lesson help you make better works of art?

## JOURNAL QUESTIONS



Artwork # or Title \_\_\_\_\_

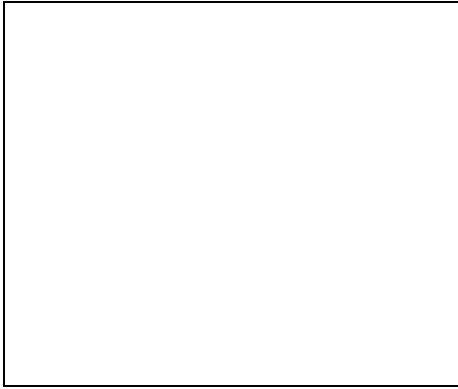
Date \_\_\_\_\_

Student ID # \_\_\_\_\_

Effort Level \_\_\_\_\_  
1-5

1. To what extent did you meet the requirements of this assignment?
2. What sections of this artwork or skills would you like to improve? How? Why?
3. What did you do that made this artwork better?
4. How has this project helped improve your art making or thinking? What did you learn? How will this lesson help you make better works of art?

## JOURNAL QUESTIONS



Artwork # or Title \_\_\_\_\_

Date \_\_\_\_\_

Student ID # \_\_\_\_\_

Effort Level \_\_\_\_\_

1-5

### **After Last Work**

4. To what extent did you meet the requirements of this assignment?

5. What sections of this artwork or skills would you like to improve? How? Why?

6. What did you do that made this artwork better?

4. Did you improve the two things you wanted to do better? Did your art making and thinking improve this year? Did your skills improve? How?

**APPENDIX C**  
**HOLISTIC JOURNAL RUBRIC**

## HOLISTIC JOURNAL RUBRIC

<b>1 Inadequate</b>	<b>Seldom refer to specific qualities in their artworks</b>	<b>Answered with, i.e. “I followed directions”</b>	<b>Skipped the question on self-reflection or answered with short one or two words that did not relate to project</b>
<b>2 Developing</b>	<b>Sometimes refer to specific qualities in their artworks</b>	<b>Occasionally talks about their art making processes and decisions</b>	<b>Occasionally answered the self-reflection question, talking about their art making or learning</b>
<b>3 Shows Some Insight</b>	<b>Frequently refers to specific qualities evident in their artworks</b>	<b>Specific references to their art making process and or thinking, i.e. decisions or their way of working</b>	<b>Regularly made self-reflective statements about their art making or learning</b>



**APPENDIX D**  
**HUMAN SUBJECTS DOCUMENTS**



Office of the Vice President For Research  
Human Subjects Committee  
Tallahassee, Florida 32305-2763  
(850) 644-8633 FAX (850) 644-4392

## REAPPROVAL MEMORANDUM *(duplicate)*

Date: 3/26/2004

To:  
Charles M. Dorn  
MC:4480

Dept.: Art Education

From: John Tomkowiak, Chair

A handwritten signature in black ink that reads "John Tomkowiak".

Re: Reapproval of Use of Human subjects in Research:  
Field Test of 20 K-12 Arts Rubrics in Two Public School Settings: Dade and Pinellas  
Counties, Florida Mododis For Assessing Art Performance

Your request to continue the research project listed above involving human subjects has been approved by the Human Subjects Committee. If your project has not been completed by 3/9/2005 please request renewed approval.

You are reminded that a change in protocol in this project must be approved by resubmission of the project to the Committee for approval. Also, the principal investigator must report to the Chair promptly, and in writing, any unanticipated problems involving risks to subjects or others.

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

Cc: Barbara Edwards  
HSC No. 2004.122-R

Dr. Dorn: Attached please find all stamped/approved consent forms for use in your study. Good luck w/your research.

A handwritten signature in black ink that reads "Heidi Hodupe".

Institutional and Federal regulations require that all protocols involving human studies be reviewed at least yearly. Please complete and submit the following.  
**RESEARCHERS: PLEASE COMPLETE ALL SHADED AREAS OF THIS FORM.**

To: OFFICE OF RESEARCH  
 HUMAN SUBJECTS COMMITTEE  
 MAIL CODE 2763, FAX No. 644-4392

FROM:	RESEARCHER'S NAME (PLEASE PRINT): Charles M. Dorn
	MAILING ADDRESS: 126 Carothers Hall MC 4480
DATE (THIS FORM WAS COMPLETED):	2/18/04
PROJECT TITLE:	Field Test of 20 K-12 Arts Rubrics in Five Public School Settings: Duval, Hillsborough, Pinellas & Leon Counties, FL and Americus, Georgia. Models for Assessing Art Performances (MAAP) Project
DATE PROJECT WAS PREVIOUSLY APPROVED:	4/17/03

I hereby certify that the below-listed responses accurately reflect my research protocol to date.

*Charles M. Dorn*  
 Researcher's Signature

2/18/04  
 Date Signed

DO YOU WANT TO CONTINUE THIS PROTOCOL IN AN ACTIVE STATUS? (IF SUBJECTS ARE STILL UNDERGOING INVESTIGATION, THIS PROTOCOL SHOULD NOT BE INACTIVATED.)	Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>
SINCE THE STUDY COMMENCED, HAS ANYTHING APPEARED IN THE PERTINENT MEDICAL LITERATURE THAT AFFECTS THE CONDUCT OF THIS STUDY, THE ANTICIPATED BENEFITS, OR THE POTENTIAL RISKS? (IF YES, PLEASE EXPLAIN ON A SEPARATE SHEET.)	Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>
DOES THIS PROTOCOL INVOLVE THE USE OF AN INVESTIGATIONAL DRUG ("IND") OR INVESTIGATIONAL DEVICE ("IDE")? (IF ANY CHANGE IN STATUS OF IND OR IDE, EXPLAIN ON A SEPARATE SHEET.)	Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>
DOES THIS PROTOCOL INVOLVE THE USE OF IONIZING RADIATION? (IF ANY CHANGE IN STATUS OF IONIZING RADIATION, EXPLAIN ON A SEPARATE SHEET.)	Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>
HAVE ANY ADVERSE OR UNDESIRABLE REACTIONS BEEN NOTED IN ANY SUBJECTS PARTICIPATING IN THE PROTOCOL? (IF YES, EXPLAIN ON A SEPARATE SHEET.)	Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>
HAVE ANY CHANGES IN THE SAMPLE, AIMS, METHODS OF PROCEDURE, OR CONSENT FORM OCCURRED THAT HAVE NOT BEEN APPROVED BY THE IRB SINCE THE PROTOCOL WAS LAST APPROVED (INITIAL APPROVAL OR LAST PROGRESS REPORT)? (IF YES, EXPLAIN ON A SEPARATE SHEET.)	Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>
HAVE ANY CHANGES OCCURRED IN THE PROFESSIONAL PERSONNEL PARTICIPATING IN THE STUDY? (IF YES, EXPLAIN ON A SEPARATE SHEET.)	Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>
HAS ANY CHANGE IN THE FUNDING FOR THIS PROTOCOL OCCURRED SINCE IT WAS LAST APPROVED (INITIAL APPROVAL OR LAST PROGRESS REPORT)? (IF YES, EXPLAIN ON A SEPARATE SHEET.)	Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>

**PLEASE ATTACH**

(YOUR RENEWAL REQUEST WILL NOT BE PROCESSED WITHOUT THE FOLLOWING ATTACHMENTS):

1. A current or final progress report (200 words or less).
2. The Consent Form now in use.

PLEASE ALSO NOTE THAT NO CHANGES IN RESEARCH PROJECT(S) AND/OR PROCEDURES THERE TO CAN BE MADE WITHOUT FIRST OBTAINING THE COMMITTEE'S APPROVAL. ALSO, IF CONDITIONAL APPROVAL REGARDING CHANGES IS GRANTED BY THE COMMITTEE, THE COMMITTEE RESERVES THE RIGHT TO TERMINATE ITS APPROVAL IF THOSE CHANGES ARE NOT TIMELY SUBMITTED.

**PARENTAL CONSENT LETTER**  
For Florida State University Arts Assessment Field Test

Dear Parent:

I am a professor of Art Education in the College of Visual Arts and Dance at Florida State University. I am conducting a research study to assess teaching in Art, Music, Dance and Theatre in the Leon County elementary and secondary schools. Your child's participation will involve arts teachers in Leon County evaluating the students' performance either through the review of a visual art portfolio or observation of a student's art performance normally done as part of the student's regularly assigned work. The assessment will require no more than a one time five or ten minute performance review or, as in visual art, no use of class time. In the case of Music, Theatre and Dance, the arts performance may be videotaped. These video tapes will be viewed only by the assessment team and the researcher. The tapes will be kept confidential to the extent allowed by the law, in a locked filing cabinet and will be destroyed on July 31, 2005. Your participation, as well as that of your child, in this study is voluntary. The results of the research study may be published, but your child's name will not be used.

Although there may be no direct benefit to your child, the possible benefit of your child's participation will help the Leon County schools to assess and improve its arts program.

If you have any questions concerning this research study or your child's participation in the study, please call me at 850-644-2158 or by fax at 850-644-3067.

Sincerely,

Charles M. Dom

\*\*\*\*\*

I give consent for my child \_\_\_\_\_ to participate in the above study.

Parent's Name: \_\_\_\_\_

Parent's Signature \_\_\_\_\_ Date \_\_\_\_\_

If you have any questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Committee, Institutional Review Board, through the Vice President for the Office of Research at 850-644-8633.



**APPENDIX E**  
**LEON COUNTY DOCUMENTS**

BOARD CHAIRMAN  
Dee Cumber

BOARD VICE-CHAIR  
Georgia "Joy" Brown

DIRECTOR, PROGRAM MONITORING  
AND EVALUATION SERVICES  
Margaret Soutard, Ph.D.



BOARD MEMBERS  
Sheila Costigan  
Maggie H. Lewis  
H. Fred Warr

SUPERINTENDENT  
William J. Montford, III

August 14, 2003

Marla Meale  
1452 Mitchell Avenue  
Tallahassee, Florida 32303

Topic: "Digital Portfolio Assessment and Self-Reflection/Assessment"

Dear Ms. Meale:

The Leon County Research Review Board has approved your request for research. Based on your proposal, your research request will be approved for the period of August 2003 through August 2004. Should you desire to continue your research efforts after this period of time, you must submit a progress report on the status of your research and request renewed approval for continuation of the project. Any significant changes or amendments to the procedures or design of this study must be approved by resubmitting the request for research to the Research Review Board.

You need to contact the principals of the schools in which you wish to conduct your study as soon as possible. The principal is responsible for making the decision relative to his or her school. It is your responsibility to return the enclosed "Principal's Consent for Research Participation," signed by the principal(s) of the school(s) to be involved, prior to the start of any research. Receipt of this consent form by this office will complete the approval process.

In the interest of continued research benefits and the coordination of research interests, please send this office one copy of your results and discussion. This information, and any other relevant information you may have, will be filed in our research library and added to the annotated listing of research projects. We look forward to your results and any suggestions they may offer toward improving the educational process in Leon County Schools.

Please feel free to call me if I can be of further assistance. I can be reached at 488-7007.

Sincerely,

A handwritten signature in black ink that reads "Margarita F. Soutard".

Margarita F. Soutard, Ph.D.  
Program Monitoring and Evaluation  
Chair, Research Review Board

MFS/dh

cc:	Hal Pitts/Buck Lake	Desoto Trail/Jarris Johnson	Paul Green/Ft. Braden
	Edna Owens/Gilchrist	Penny Brinson/Hawke Rise	Kathleen Rodgers/Kate Sullivan
	Peggy Youngblood/Roberts	Doug Cayler/Roadside	Tom Inarra/Nealey
	Jackie Pons/Dear Lake	Dirma Callaway/Ran	Alan Cox/Chiles
	Martha Burch/Lincoln		

2803 West Pensacola Street • Tallahassee, Florida 32304-2090 • Phone (904) 488-7007 • Fax (904) 922-8070 • Sunbeam (888) 278-7007

**Teaching, Learning, Caring for the Future**

Alternative Action/Equal Opportunity Employer • (904) 487-7100

## PARENTAL CONSENT LETTER

For Florida State University Art Portfolio Assessment Field Test

Dear Parent:

I am a graduate student and art teacher here in Leon County working with Professor Dorn. Dr. Dorn is a professor of Art Education in the College of Visual Arts and Dance at Florida State University. I am conducting a research study to determine the affect of creating digital art portfolios both with and without self-reflection/evaluation activities on student growth in visual art classes. Your child's participation will involve the visual art teachers in Leon County evaluating students' performance through review of their digital portfolios of their regularly assigned work. Some students will also write in a journal, (self-reflection), or do a self-assessment activity designed by his/her art teacher. The portfolio review will not affect your students' grade in visual art class and is designed to measure growth overtime, not performance for a grade.

Participating schools and students will be assigned numbers to assure anonymity. Only your student's art teacher will know you child's number. The only information collected from your child's records will be the FCAT or NRT reading level from last year. Your student's art teacher/or classroom teacher will collect this data. All portfolios and data collected for evaluation will assure student anonymity. The student name/number list will be destroyed at the end of the school year by your child's art teacher. Your participation, as well as that of your child, in this study is voluntary. If you or your child chooses not to participate or to withdraw from the study at any time, there will be no penalty. The results of the research study may be published, but your child's name, artwork, and identity will not be used.

Although there may be no direct benefit to your child, the possible benefit of your child's participation will help the Leon County Schools identify best practices and improve its art programs.

If you have any questions concerning this research study or your child's participation in the study, please call me at 850 321-9075, by email at: mealem@mail.ruediger.leon.k12.fl.us

Sincerely,

Marcia Meale

.....  
I give consent for my child \_\_\_\_\_ to participate in the above study.

I would prefer that my child not be included in this study. \_\_\_\_\_

Parent/Guardian's Name \_\_\_\_\_

Paernt/Guardian's Signature \_\_\_\_\_ Date \_\_\_\_\_

If you have any questions about your rights as a subject/participant in this research, or if you feel; you have been placed at risk, you can contact the Chair of the Human Subjects Committee, Institutional Review Board, through the Vice President for the Office of Research at 850 644-8633.



## **Child Assent**

For Florida State University Art Portfolio Assessment Field Test

Name: \_\_\_\_\_

We would like permission to view your work in art and if your class is chosen, review your self-reflection/assessment journals.

Your parents have given their permission for you to participate, but you have the right to choose whether you want to do it or not. If you decide that you do not want to participate, you will not be punished or penalized.

Child's signature \_\_\_\_\_ Date \_\_\_\_\_

## REFERENCES

- Abbot, J. (1997). To be intelligent. *Educational Leadership*, 54 (electronic 8/01/03), 6-10.
- Armstrong, C. L. (1994). *Designing assessment in art*. Reston, VA: National Art Education Association.
- Arnheim, R. (1969). *Visual thinking*. Berkeley, CA: University of California Press.
- Arter, J. A., & Spandel, V. (1992). Using portfolios of student work in instruction and assessment. *Educational Measurement: Issues and Practice*, 11, 36-44.
- Baillie, J. B.G.W.F. *Hegle: The phenomenology of mind*. Retrieved August 2, 2003, from <http://www.class.uidaho.edu/mickelsen/ToC/Hegel%20Phen%20ToC.htm>
- Beattie, D. K. (1997). *Assessment in art education*. Worcester, MA: Davis Publications, Inc.
- Bensur, B. J. (2002). Frustrated voices of art assessment. *Art Education*, 55, 18-23.
- Bigge, M. L., & Shermis, S. S. (1999). *Learning theories for teachers* (6th ed.). New York, NY: Addison Wesley Longman.
- Boughton, D. (1996). Evaluating art education in schools. In D. Boughton, E.W. Eisner, & J. Ligtvoet (Ed.), *Evaluating and assessing the visual arts in education* (pp. 1-16). New York, NY: Teachers College Press.
- Brandt, R. (1987/1988). On assessment in the arts: A conversation with Howard Gardner. *Educational Leadership*, 45, 30-34.
- Bransford, J. D., & Brown, A.L.; Cocking, R.R. (Ed.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Brown, R. (1989). Testing and thoughtfulness. *Educational Leadership*, 46, 31-33.
- Bruner, J. (1997). Celebrating divergence: Piaget and Vygotsky. *Human Development*, 40, 63-73.
- Caine, R. N., & Caine G. (1990). Understanding a brain-based approach to learning and teaching. *Educational Leadership*, 48, 66-70.
- Cannatella, H. (2001). Art assessment. *Assessment and Evaluation in Higher Education*, 26, 319-326.
- Clark, R. E. (2002). Performance assessment in the arts. *Kappa Delta Pi Record*, 39 (electronic), 29-32.

- Coe, R.(2005). *What is an 'effect size'*. Retrieved March 7, 2005, from <http://www.cemcentre.org/ebeuk/research/effectsize/ESguide.htm>
- Cohen, B.H., Lea, R.B. (2004). *Essentials of statistics for the social and behavioral sciences*. Hoboken, NJ: John Wiley & Sons, Inc.
- Cohen, M. (1988). Designing state assessment systems. *Phi Delta Kappan*, 69, 583-588.
- Cole, D. J., Ryan, Charles W., Kick, F, & Mathies, B. K. (2000). *Portfolios across the curriculum and beyond* (Second ed.). Thousand Oaks, CA: Corwin Press, Inc.
- Copelston, F. S. J. (1994). *Modern philosophy: From the French enlightenment to Kant* (Vol. VI). New York, NY: Doubleday.
- Costa, A. (1993). How world-class standards will change us. *Educational Leadership*, 50, 50-51.
- Dewey, J. (1916/1944). *Democracy and education*. Retrieved August 2, 2003, from <http://www.ilt.columbia.edu/publications/Projects/digitexts/dewey.html>
- Dewey, J. (1934/1958). *Art as experience*. New York, NY: Capricorn Books.
- Dorn, C. M. (1998). The aesthetically testable object. *Arts Education Policy Review*, 99, 3-10.
- Dorn, C. M. (1999). *Mind in art: Cognitive foundations in art education*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Dorn, C. M. (2002). The teacher as stakeholder in student art assessment and art program evaluation. *Art Education*, 55 (electronic), 40-45.
- Dorn, C. M. (2003). Models for assessing art performance (MAAP): A K-12 project. *Studies in Art Education*, 44, 348-369.
- Dorn, C. M., Madeja, S., & Sabol, F.R. (2004). *Assessing expressive learning*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Dorn, C. M., & Sabol, F.R. *Actual versus digital portfolio scores for pretest and posttest sessions across four counties*. Unpublished manuscript, Florida State University, Tallahassee, FL.
- Down, G. A.,& Mitchell, R. (1993). Shooting for the moon: Standards for the arts. *Educational Leadership*, 50, 32-35.
- Durant, W. (1961). *The story of philosophy: The lives and opinions of the great philosophers of the western world*. New York, NY: Simon and Schuster, Inc.

- Efland, A. D. (2002). *Art and cognition: Integrating the visual arts in the curriculum*. New York, NY: Teachers College Press.
- Eisner, E. W. (1993). Why standards may not improve schools. *Educational Leadership*, 50, 22-23.
- Eisner, E. W. (1996). Overview of evaluation and assessment: Conceptions in search of practice. In D. Boughton, E.W. Eisner, & J. Ligtvoet (Ed.), *Evaluating and assessing the visual arts in education* (pp. 1-16). New York, NY: Teachers College Press.
- Eisner, E. W. (2002). *The arts and the creation of mind*. New Haven, CT: Yale University Press.
- Fenstermacher, G. D., & Sanger, M. (1998). What is the significance of Dewey's approach to the problem of knowledge? *The Elementary School Journal*, 98 (electronic), 467-478.
- Flavell. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Fraenkel, J.R., & Wallen, N.E. (2003). *How to design and evaluate research in education* (Fifth ed.). New York, NY: McGraw-Hill Higher Education.
- Gardner, H. (1972). *The quest for mind*. New York, NY: Alfred A. Knopf.
- Gardner, H. (1983/1993). *Frames of mind: The theory of multiple intelligences* (Tenth anniversary ed.). New York, NY: Basic Books.
- Gardner, H. (1985). *The mind's new science: A history of the cognitive evolution*. New York, NY: BasicBooks.
- Gardner, H. (1991). *The unschooled mind: How children think & how schools should teach*. New York, NY: BasicBooks.
- Gardner, H. (1996). The assessing of student learning in the arts. In D. Boughton, E.W. Eisner, & J. Ligtvoet (Ed.), *Evaluating and assessing the visual arts in education*. New York, NY: Teachers College Press.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st Century*. New York, NY: BasicBooks.
- Goldsby, D., & Fazal, M. (2001). Now that your students have created web-based digital portfolios, how do you evaluate them? *Journal of Technology and Teacher Education*, 9, 607-616.
- Green, F. E. (1999). Brain learning and research: Implications for meeting the needs of diverse learners. *Education*, 119, electronic 08/01/03 (682-687).

- Gruber, D. D., & Hobbs, J.A. (2002). Historical analysis of assessment in art education. *Art Education*, 55, 12-17.
- Hanley. (2003). Policy issues in arts assessment in Canada: "let's get real". *Arts Education Policy Review*, 105, 33-37.
- Hausfather, S. J. (1996). Vygotsky and schooling: Creating a social context for learning. *Action in Teacher Education*, 18, 1-10.
- Hausman, J. (1994). Standards and assessment: New initiative and continuing dilemmas. *Art Education*, March, 9-13.
- Huffman, E. S. (1998). Authentic rubrics. *Art Education*, Jan., 64-68.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41, 212-218.
- Lazear, D. G. (1999). *Multiple intelligence approaches to assessment: Solving the assessment conundrum* (Revised ed.). Tucson, AZ: Zephyr Press.
- Lehman, P. R. (1999). Introduction to the symposium on the NAEP 1997 arts report card. *Arts Education Policy Review*, 100, 12-15.
- Marzano, R. J., Arredondo, D.E., Brandt, R.S., Pickering, D.J., Blackburn, & G.J, Moffet, C.A. (1988). *Dimensions of thinking*. Aurora, CO: MCRel.
- Marzano, R. J., Arredondo, D.E., Brandt, R.S., Pickering, D.J., Blackburn, & G.J., Moffet, C.A. (1992). *Dimensions of learning: Teacher's Manuel*. Aurora, CO: McRel.
- McCollister, S. (2002). Developing criteria rubrics in the art classroom. *Art Education*, 55 (electronic), 46-52.
- McCombs, B. L. (2001). Self-Regulated learning and academic achievement: A phenomenological view. In B.J. Zimmerman, & D.H. Schunk (Ed.), *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives* (pp. 67-124). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mehrens, W. A. (1992). Using performance assessment for accountability purposes. *Educational Measurement: Issues and Practice*, 11, 3-9.
- NAEA. (1994). The national visual arts standards. *National Art Education Association News*, June.
- Nelson, J. L. (1999). How we think. *The Social Studies*, 90, 13-14.

- Nicuidula, D. (1997). Picturing performance with digital portfolios. *Educational Leadership*, 55, 26-29.
- No Child Left Behind Act of 2001, Pub. L. (2001).
- O'Neil, J. (1993). Can national standards make a difference? *Educational Leadership*, 50, 4-8.
- O'Neil, J. (1993b). On the new standards project: A conversation with Lauren Resnick and Warren Simmons. *Educational Leadership*, 50, 17-21.
- Paris, S. G., Byrnes, J. P., & Paris, A. H. (2001). Constructing theories, identities, and actions of self-Regulated learners. In B. J. Zimmerman & D.H. Schunk (Ed.), *Self-regulated learning and academic achievement* (Second ed., pp. 253-288). Mahwah, NJ: Lawrence Erlbaum Associates.
- Perfect, T. J., & Schwartz, B.L. (2002). *Applied metacognition*. United Kingdom: Cambridge University Press.
- Perkins, D.N. (1994). *The intelligent Eye: Learning to think by looking at art*. Occasional Paper 4. Santa Monica, CA: The J. Paul Getty Trust.
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, 41(electronic 7/11/03).
- Prawat, R. S. (1989). Promoting access to knowledge, strategy, and disposition in students: A research synthesis. *Review of Educational Research*, 59, 1-41.
- Prawat, R. S. (2002). Dewey and Vygotsky, and the social administration of the individual: Constructivist pedagogy as the systems of ideas in historical spaces. *Educational Researcher*, 31, 16-20.
- Resnick, L. B. (1994). Standards, assessment, and educational quality. *Journal for the Education of the Gifted*, 17, 409-420.
- Sabol, R. (1998). What are we testing? Content analysis of state visual arts achievement. *Visual Arts Research*, 24, 1-12.
- Scholastic. (1999). Understanding the theory of multiple intelligences. *Scholastic Early Childhood Today*, 13, 43-45 (electronic, 08/01/03).
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26, 113-125.
- Schraw, G., & Dennison, R.S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475.

- Schunk, D. H. (1986). Verbalization and children's self-regulated learning. *Contemporary Educational Psychology*, 11, 347-369.
- Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In B. J. Zimmerman, & D.H. Schunk (Ed.), *Self-regulated learning and academic achievement: Theoretical perspectives* (Second ed., pp. 125-152). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schwartz, B. (1996). Communicating students learning in the visual arts. *Yearbook (Association for Supervision and Curriculum Development)*, 1996 (electronic), 78-79.
- Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership*, 46, 4-9.
- Sherry, L., Billig, S.H., & Jesse, D. (2001). Assessing the impact of instructional technology on student achievement. *T.H.E. Journal*, 28 (electronic), 40-43.
- Simmons, W., & Resnick, L. (1993). Assessment as the catalyst of school reform. *Educational Leadership*, 50, 11-15.
- Simpson, D. J. (2001). John Dewey's concept of the student. *Canadian Journal of Education*, 26 (electronic, 07/22/03), 183-200
- Spence, L. D. (2002). Against teaching. *The Educational Forum*, 66 (electronic, August 8, 2003), 330-339.
- Stanford. *Georg Wilhelm Friedrich Hegel*. Retrieved 12/02/02, from <http://plato.stanford.edu/entries/hegel>
- Starko, A.J. (2001). *Creativity in the classroom* (2<sup>nd</sup> ed). Mahwah, NJ: Lawrence Erlbaum Associates.
- Valiant, R. (1998). Growing brain connections: A modest proposal. *Schools in the Middle*, 7, 24-26 (electronic 08/01/03).
- Walker, J. H. (1998, October). *Process portfolios as a means for formative and summative evaluation of student work in the visual arts*. Paper presented at the Annual Meeting of Mid-Western Educational Research Association, Chicago, IL.
- Weidmer, T. L. (1998). Digital portfolios: Capturing and demonstrating skills and levels of performance. *Phi Delta Kappan*, 79, 586-589 (electronic).
- Wellman, H. M. (1985). *The origins of metacognition*. (Vol. 1 theoretical perspectives). Orlando, FL: Academic Press, Inc.
- Wilson, B. (1992). A primer on arts assessment and a plethora of problems. *Design for Arts in Education*, 93, 34-44.

- Winne, P. H. (2001). Self-regulated learning viewed from models of informational processing. In B. J. Zimmerman, & D.H. Schunk (Ed.), *Self-regulated learning and academic achievement: Theoretical perspectives* (Second ed., pp. 153-190). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wolf, D. P. (1987/1988). Opening up assessment. *Educational Leadership*, 45, 24-29.
- Wolf, D. P. (1989). Portfolio assessment: Sampling student work. *Educational Leadership*, 46, 35-39.
- Yussen, S. R. (1985). The role of metacognition in contemporary theories of cognitive development. In M. Forrest-Pressley, & T.G. Waller, (Ed.), *Metacognition, cognition, and human performance* (Vol. 1, pp. 253-283). Orlando, FL: Academic Press, Inc.
- Zemelman, S., Daniels, H., & Hyde, A. (1998). *Best practice: New standards for teaching and learning in America's schools*. Portsmouth, NH: Heinemann.
- Zerull, D. S. (1990). Evaluation in arts education: Building and using an effective assessment strategy. *Designs for Arts in Education*, 92, 19-24.
- Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses. *Contemporary Educational Psychology*, 11, 307-313.
- Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman, D.H. Schunk (Ed.), *Self-regulated learning and academic achievement: Theoretical perspectives* (Second ed., pp. 1-38). Mahwah, NJ: Lawrence Erlbaum Associates.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41, 64-70.



## BIOGRAPHICAL SKETCH

Marcia S. Meale is currently teaching art education courses at the University of South Carolina in Columbia, South Carolina. Her prior teaching experience includes seventeen years in the public schools in Florida, the last fifteen years at a Title I elementary school in Leon County. Marcia holds National Board of Professional Teaching Standards certification in Early and Middle Childhood Art. Throughout her public school teaching career she exhibited personal and student artwork, was awarded a number of grants and awards, and presented at state and national conventions. Topics covered a range of subjects from connecting art and math to using digital resources in the art classroom. Her current research interests focus on the use of digital media in art classrooms and the evaluation of art assessment methods.