Art assessment Policy and Practice at the High School Level: Validity, Reliability, and Resistance

Deborah N. Filbin
db.filbin@gmail.com

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Part of the Art Education Commons

Recommended Citation
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/7032

This Dissertation/Thesis is brought to you for free and open access by the Graduate Research & Artistry at Huskie Commons. It has been accepted for inclusion in Graduate Research Theses & Dissertations by an authorized administrator of Huskie Commons. For more information, please contact jschumacher@niu.edu.
ART ASSESSMENT POLICY AND PRACTICE AT THE HIGH SCHOOL LEVEL: VALIDITY, RELIABILITY, AND RESISTANCE

Deborah N. Filbin, Ph.D.
School of Art and Design
Northern Illinois University, 2020
Douglas Boughton and Kerry Freedman, Co-Directors

National education reform policies that have increased reliance on standardized testing in subjects like reading and math as a way to judge the quality of a teacher’s performance have created challenges for educators in visual art and design when they are required to provide data about student growth. Art teachers who utilize alternate forms of assessment to judge the quality of student artwork as evidence of learning can potentially be in a precarious position because of underlying assumptions that these types of qualitative assessments lack validity. In this context, when compared to colleagues in traditionally tested disciplines, art educators face unique challenges proving their assessments are both valid and reliable. Framed in critical pedagogy, a mixed-methods study was conducted in the state of Illinois to investigate the kinds of assessment strategies high school art teachers found useful in their classrooms to measure student learning and whether their methods differed from the types of assessments their administrators expected. This study brings attention to the important aspects of assessment and how policy can shape teachers’ practice.

This study provided unique insight into Illinois art teachers’ experiences with the current state teacher evaluation policy PERA (Performance Evaluation Reform Act). Participants from throughout the state were surveyed about their understanding of assessment, validity, reliability,
and professional development. To elaborate and contextualize the findings, face-to-face interviews were conducted with eight participants to obtain a deeper understanding of teacher’s actual experiences in the classroom and elaborate upon the role educational policy played in assisting them to meet requirements for their performance evaluation. Assessment at the high school level can present unique challenges when compared to other disciplines because of the complexity of qualitative judgments teachers must make about their students’ work. Applying qualitative assessment methods were particularly troublesome for participants because their methods were not like traditional right-wrong answer choice tests; this created underlying mistrust of art teachers’ data that was derived from professional judgment on qualitative assessments. Deepening the challenge art teachers faced when assessing student art performance, the participants described a lack of professional development specifically for assessment in art. They also struggled with an absence of professional development to establish the validity and reliability of their assessments and sometimes received inappropriate direction about student performance data collection because of a lack of understanding by administrators or supervisors about how student artwork should be judged. In an age of test-based accountability, professional development for teachers, administrators, and pre-service educators aimed specifically for art and design assessment is essential for practitioners in public schools.

Additionally, it was found the socio-economic status of teacher participants’ schools was related to the kinds of data they were asked to collect and the kinds of resources they had available within their departments thus indicating a lack of equitable access to quality art education throughout the state. In response to administrative requests that participants felt were inappropriate to their discipline, multiple forms of resistance were exhibited including covert,
overt, and passive compliance as a way to help them cope with what many participants felt were overwhelming obstacles to teaching a quality visual art curriculum.
NORTHERN ILLINOIS UNIVERSITY
DEKALB, ILLINOIS

MAY 2020

ART ASSESSMENT POLICY AND PRACTICE AT THE HIGH SCHOOL LEVEL:
VALIDITY, RELIABILITY, AND RESISTANCE

BY
DEBORAH N. FILBIN
© 2020 Deborah N. Filbin

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

SCHOOL OF ART AND DESIGN

Doctoral Co-Directors:
Douglas Boughton and Kerry Freedman
ACKNOWLEDGEMENTS

I would like to acknowledge the tremendous wisdom and guidance from my co-advisors, Dr. Doug Boughton and Dr. Kerry Freedman. Their warmth, patience, persistence, and confidence in my abilities taught me to shape ideas and experiences into an argument for the dignity of the field. I would like to thank my committee members Dr. Kryssi Staikidis and Dr. Laura Johnson for their expertise, professionalism, time, and care. Their insight has been truly valued and has helped shape my work. I would like to extend sincere gratitude for my colleagues at Northern Illinois University: Dr. Richard Siegesmund, Gail Jacky at the University Writing Center, and fellow doctoral students. The friendships and professional relationships forged at NIU have changed my life for the better. I would like to thank Dr. Jerry Hausman for his stories and encouragement; he will always have a special place in my memories of my journey as a doctoral student. I would like to thank my long-time administrator and friend Dr. Lenell Navarre for having faith in me and believing the arts are an important part of our students’ academic experience. Finally, I would like to thank my participants for the generosity of their time by contributing to this study. Their honesty, passion, and care for their students is truly inspiring. To all art teachers who work tirelessly for the benefit of their students and strive to promote the arts in schools. Do not lose your spirit and stay true to the discipline—always putting what is best for students first.
DEDICATION

This is dedicated to my husband Michael and my children, Veronica and Liam. Without their patience, understanding, and support I could never have completed my dissertation.

Michael, you have provided continuous unwavering support throughout this journey every moment I needed it. You gave me everything I needed from a partner and soulmate to be successful.

To my children who always understood that sometimes mom could not be there and never once complained: Please know that every time I missed a practice, scout meeting, parent teacher conference, doctor’s appointment, game, field trip, or lost tooth, I was with you every moment in my heart and soul.

I would like to thank my mother for always believing I was smart and capable, even when I did not know it. Somehow, you always knew pursuing a career in the arts was the one thing that would bring me happiness; your guidance shaped me into the strong confident feminist I am today.

I am truly blessed to have a family that has allowed me to grow.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>xi</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>BACKGROUND TO THE STUDY</td>
<td>1</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>8</td>
</tr>
<tr>
<td>Statement of the Problem and Sub-Problems</td>
<td>10</td>
</tr>
<tr>
<td>Context of the Study</td>
<td>18</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>20</td>
</tr>
<tr>
<td>Research Questions and Sub-Questions</td>
<td>22</td>
</tr>
<tr>
<td>Delimitations and Limitations</td>
<td>23</td>
</tr>
<tr>
<td>Definitions of Terms</td>
<td>24</td>
</tr>
<tr>
<td>Theoretical Foundation</td>
<td>28</td>
</tr>
<tr>
<td>Assumptions</td>
<td>33</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>35</td>
</tr>
<tr>
<td>Educational Policy at State and National Levels</td>
<td>36</td>
</tr>
<tr>
<td>Culture of Standardized Testing</td>
<td>42</td>
</tr>
<tr>
<td>Marginalizing Effects of Standardized Testing on Underserved Populations</td>
<td>45</td>
</tr>
</tbody>
</table>
Chapter | Page
---|---
Data Driven Culture | 48
Learning Standards | 50
Teacher Observation and Evaluation Models | 52
Measuring Learning in High School and Secondary Schools with Alternate Assessments in Visual Art and Design | 54
Validity and Reliability of Art Assessment Practices | 61
Characteristics of Adolescent and High School Students | 63
Consideration of Creativity | 65
Forms of Resistance | 68
Overview of Related Studies | 73
Application to the Study | 78
3. METHODOLOGY | 80
Research Design | 82
Quantitative Research Data Collection Process | 83
Survey Instrument | 84
Qualitative Research Data Collection Process | 87
Positionality of the Researcher | 90
How the Quantitative Data Were Analyzed | 91
How the Qualitative Data Were Analyzed | 92
Criteria for Selecting Participants | 93
4. QUANTITATIVE DATA | 95
Survey Participants | 95
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses to Closed-Ended Questions</td>
<td>97</td>
</tr>
<tr>
<td>5. QUALITATIVE DATA</td>
<td>113</td>
</tr>
<tr>
<td>Participants</td>
<td>114</td>
</tr>
<tr>
<td>Summary of Themes from the Qualitative Data</td>
<td>119</td>
</tr>
<tr>
<td>Socio-Economic and Racial Makeup of Schools</td>
<td>121</td>
</tr>
<tr>
<td>Assessment Techniques</td>
<td>123</td>
</tr>
<tr>
<td>Reliability and Validity</td>
<td>131</td>
</tr>
<tr>
<td>Data Collection</td>
<td>134</td>
</tr>
<tr>
<td>Performance Evaluation Process</td>
<td>137</td>
</tr>
<tr>
<td>Analysis of Teacher Evaluation Process Documents</td>
<td>145</td>
</tr>
<tr>
<td>Policy and Standards</td>
<td>153</td>
</tr>
<tr>
<td>Student Performance as Creative or Characteristic of Adolescent Art Making</td>
<td>157</td>
</tr>
<tr>
<td>Professional Development and Need for a Professional Community</td>
<td>161</td>
</tr>
<tr>
<td>Resistance</td>
<td>164</td>
</tr>
<tr>
<td>Job Stress</td>
<td>171</td>
</tr>
<tr>
<td>Conclusion</td>
<td>175</td>
</tr>
<tr>
<td>6. CONCLUSION</td>
<td>176</td>
</tr>
<tr>
<td>Research Question</td>
<td>177</td>
</tr>
<tr>
<td>Sub Question 1</td>
<td>181</td>
</tr>
<tr>
<td>Sub Question 2</td>
<td>184</td>
</tr>
<tr>
<td>Sub Question 3</td>
<td>190</td>
</tr>
<tr>
<td>Sub Question 4</td>
<td>193</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Implications and Recommendations</td>
<td>196</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>211</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>205</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Summative Assessments Implemented</td>
<td>129</td>
</tr>
<tr>
<td>2. District PERA Plan Document Comparisons</td>
<td>148</td>
</tr>
<tr>
<td>3. Manifestations of Resistance</td>
<td>168</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Student artwork sample 1</td>
<td>16</td>
</tr>
<tr>
<td>2: Student artwork sample 2</td>
<td>16</td>
</tr>
<tr>
<td>3: Student artwork sample 3</td>
<td>17</td>
</tr>
<tr>
<td>4: Student artwork sample 4</td>
<td>17</td>
</tr>
<tr>
<td>5: Years of experience</td>
<td>96</td>
</tr>
<tr>
<td>6: Teaching locations</td>
<td>97</td>
</tr>
<tr>
<td>7: Responses to closed-ended question 1 on useful strategies</td>
<td>98</td>
</tr>
<tr>
<td>8: Responses to closed-ended question 2 on useful strategies</td>
<td>99</td>
</tr>
<tr>
<td>9: Responses to closed-ended question 3 on useful strategies</td>
<td>100</td>
</tr>
<tr>
<td>10: Responses to question on administrative directives</td>
<td>101</td>
</tr>
<tr>
<td>11: Responses to question 1 on alignment of lessons to standards for assessment</td>
<td>102</td>
</tr>
<tr>
<td>12: Responses to question 2 on alignment of lessons to standards for assessment</td>
<td>103</td>
</tr>
<tr>
<td>13: Responses to helpfulness of Illinois art standards</td>
<td>104</td>
</tr>
<tr>
<td>14: Responses to administrative directives</td>
<td>105</td>
</tr>
<tr>
<td>15: Responses to administrative emphasis on importance of standards</td>
<td>106</td>
</tr>
<tr>
<td>16: Responses to importance of incorporating reading and math skills into content area</td>
<td>107</td>
</tr>
<tr>
<td>17: Responses to importance of utilizing standardized assessment in practice</td>
<td>108</td>
</tr>
<tr>
<td>18: Responses to implementation of portfolio assessments</td>
<td>109</td>
</tr>
<tr>
<td>19: Responses to professional development time to determine validity of assessment</td>
<td>110</td>
</tr>
</tbody>
</table>
20: Responses to professional development time to determine reliability of assessments. ....... 111
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. QUESTIONS FOR SURVEY QUESTIONNAIRE</td>
<td>219</td>
</tr>
<tr>
<td>B. SURVEY COVER LETTER</td>
<td>223</td>
</tr>
<tr>
<td>C. SURVEY LETTER MAILED TO POTENTIAL PARTICIPANTS</td>
<td>225</td>
</tr>
<tr>
<td>D. IRB CONSENT FORM</td>
<td>227</td>
</tr>
<tr>
<td>E. INTERVIEW QUESTIONS</td>
<td>230</td>
</tr>
<tr>
<td>F. SURVEY QUESTIONNAIRE</td>
<td>232</td>
</tr>
<tr>
<td>G. DISTRICT W</td>
<td>235</td>
</tr>
<tr>
<td>H. DISTRICT X</td>
<td>250</td>
</tr>
<tr>
<td>I. DISTRICT Y</td>
<td>267</td>
</tr>
<tr>
<td>J. DISTRICT Z</td>
<td>293</td>
</tr>
</tbody>
</table>
BACKGROUND TO THE STUDY

I have been teaching art and design at the high school level for over 20 years. Becoming a public-school teacher is the most rewarding choice I could have ever made. However, throughout my career I have perceived a slow and steady erosion of morale, loss of self-efficacy, and a reduction in job satisfaction among my colleagues as educational policies have increased emphasis on testing and databased accountability. Major changes to policy emphasize data collection on student performance, including standardized test scores in the area of reading and math as a required component of a teacher’s performance evaluation. This has presented unique challenges to teachers of visual art and design beyond that of other colleagues in non-performance-based disciplines.

During my undergraduate experience in the late 1980s, I was influenced by the Discipline Based Art Education (DBAE) movement. When I started my career, I felt respected as an educational professional. I was able to obtain the necessary funding for educational activities and materials and was trusted by administrators to make curricular decisions that were in the best interest of my students. Although the school district I taught in had a high minority population from a low socio-economic background and most students performed poorly on standardized tests, there were opportunities to grow the department by adding courses relevant to the contemporary needs of students. The school was aware of its low performance status, but the diverse needs of students were recognized by supporting robust choices of electives to help keep students engaged. As educational policies changed, specifically with the onset of No Child Left
Behind (NCLB), my school was placed on the academic watch list and identified as a failing school. I started to perceive a lack of trust in teachers and a shift away from diverse course offerings for students to a greater emphasis on courses that reflected content in standardized tests.

Test-based accountability was on the rise, and every department was required to create and implement a district-wide written pre-, mid- and post-test in scantron format to collect data on student performance; the art department was no exception. With diverse course offerings that ranged from two-dimensional to three-dimensional and beginning to advanced levels, the only content specific to all art and design classes that could be tested in multiple-choice form was the elements and principles of design and vocabulary questions based on techniques specific to art making. As a department we created district-wide tests to be compliant, although these were not authentic and did not effectively interrogate the learning that occurred throughout the school year. These tests were viewed as a necessary evil that we completed to demonstrate acquiescence; however, no one seemed to pay attention to our test scores, only the assurance they were administered and scored (Filbin & Boughton, 2016). These tests were administered three times a year and were quite different than the normal routine of my art classes that emphasized art making and performance-based learning. My colleagues and I were inexperienced with evaluation and wanted to comply, not wanting to draw negative attention to our department. Needless to say, students complained about why they had to take a multiple-choice test in art because it did not make sense. We tried to pacify their complaints by assuring them it was mandated three times a year and it would not negatively impact their grade. We asked them to just get through it so we could return to our art making activities. As teachers, we felt comfortable with our compliance; after all, the rest of the school year we were left alone. No
one seemed concerned with our assessment practices, and the instructional methods in our discipline were supported. We demonstrated success in our department by participating in multiple art shows and competitions every year, we had students who were awarded scholarships to attend colleges to pursue careers in the arts, and we produced murals or other artwork for the school community whenever it was requested. Generally, the teachers in the art department felt valued. None of my evaluating administrators had any experience in the visual arts and were usually impressed with what was happening in my classes and the artwork my students produced. I received excellent performance evaluations and rarely was I offered suggestions on how to improve my practice.

As NCLB continued, the school never achieved the meets or exceeds benchmarks for Adequate Yearly Progress (AYP) and our subgroups did not make safe harbor, the minimal required level of growth, so more restrictions were put in place. The first noticeable change in the art department in the district was a reduction in both staff and budget. Slowly our supply budget decreased, and if a department member left, the position was not re-filled. In addition, there was a noticeable increase in workload with course offerings staying the same although the number of staff members decreased. In some instances, multiple classes were run concurrently in the same classroom at the same time to maintain full class sizes and as a way to offer upper level courses that generally were smaller in size.

My principal needed a new teacher to lead the School Improvement Plan (SIP) committee and asked me to be the chairperson. I wondered why he chose me over other teachers in tested disciplines. He assured me I was a good teacher, I was well liked and respected by my colleagues, and since I was not an English or math teacher, I was the perfect person to deliver bad news to the faculty in a non-confrontational way. He was right; faculty attendance at SIP
committee was higher than ever. Meetings were filled with caring faculty members who were willing to give up their plan time to try to find strategies to help our students. They were willing to put in extra time to participate in subcommittees with the hope of getting us off the academic watch list. The subcommittees were tasked with creating activities the whole school could implement, and the SIP committee reported the results to the state. One of the plans put in place included school-wide reading. All staff and students were asked to stop what they were doing, put all work aside, and do silent reading for enjoyment for 20 minutes. This was closely monitored, and it was not unusual for administrators to walk through classrooms to ensure everyone was participating. The reading in my art room, however, was usually not very silent or enjoyable. Most of the time was spent fielding complaints and arguments by students who were begging to return to their art activities.

Other plans included school-wide math and school-wide writing. All faculty were given math worksheets or writing prompts to complete with classes, and teachers were responsible for grading and reporting student scores on a bi-weekly basis. The students complained about this too, but I did the best I could to encourage them. Not being a certified English or math teacher, I must admit my grading may have been a bit flawed. Although as the chairperson I understood why this was being done, I also became a bit resentful for having to stop teaching and implement other subject matter in my class time. Wanting to be a good teacher leader and continue to receive excellent evaluations, I did these school-wide activities with my art classes, trying to hide my frustration. Deeper than the frustration with the interruption to my class time, I was becoming resentful of the policy-based disrespect to the value of my discipline. Outside of these activities and SIP meetings, I was also privy to closed door sessions with administrators. I took notes and helped fill out the copious amounts of forms to record test performance data and break
down the data by discipline and sub-group. Our Black and Hispanic sub-groups performed far below benchmarks; however, our group of Black males struggled the most academically. Considerable effort was given to trying to find ways to help this group of students succeed. As I helped administrators with the endless documentation of what our school was doing to help students achieve, it became evident to me there was no way we could raise scores enough to ever get off the watch list. I also began to understand that a lot of decisions that affected students were out of the hands of administrators and teachers who dealt with students every day and in the hands of policymakers who knew nothing about teaching. As the achievement benchmarks continued to rise, our achievement gap widened. I knew it was time for me to step down from this leadership position; I could clearly see there was no place for the arts according to these policies and my time would be better spent on my students in my classroom.

When the movement to College Readiness Standards (CRS) began, there was some hope that we could demonstrate the good things our district had to offer through these standards. When my school sought to increase our AP program, the art department obliged by adding AP Studio Art, yet no additional art teachers were hired to assist in handling this workload unlike other departments in the school that saw an increase in staff members. When we advocated for more teachers, we were told there was nothing in the budget for the arts because they were not part of the core curriculum. Additional funding was to be allocated for reading and math to help raise test scores. Accountability requirements continued to increase, and school-wide professional development on validity and reliability on our teacher generated tests was offered to all departments to help with data analysis on our mandated pre-, mid- and post-tests. After this professional development, the art department had a data meeting with our department administrator who had no background in the arts. We were reprimanded because our students
performed too well on pre-tests and we would not be able to show adequate growth throughout the school year. Students who scored 100% on their pre-test were a detriment to our growth data because they could not show any growth. Under the new student growth policies, we were considered failures as teachers. I became extremely frustrated with this ludicrous comment and blurted out in the meeting, “If we have high schoolers that don’t know what a line or a shape is, I think we have bigger problems.” As a department we explained to our administrator that the language of art and identification of techniques were where our classes began and had little to do with measuring what our students actually learned in our classes or the final outcomes of our courses. Our curriculum was concerned with what the students could do by applying this knowledge to actual art making. The consistently high pre-test scores across all art classes demonstrated reliability; however, we quickly realized the scores were not valid because they had nothing to do with what our students were actually learning in our classes (Filbin & Boughton, 2016). By using the model of portfolio assessment that was introduced to us when we added AP Studio Art, we were able to present evidence to our administration and the district-wide curriculum committee that the multiple-choice tests we were administering in the department were not valid. We were able to provide an alternate assessment format the art department could implement that utilized only performance-based assessments scored with rubrics and a student portfolio in lieu of a written right/wrong final exam for all art classes, thus eliminating our multiple-choice tests.

Because our district offered AP Studio Art courses and the AP student exams were portfolio based, we were able to advocate for portfolio assessments to be implemented department-wide in lieu of any right/wrong types of tests. In reflection, what seemed like winning the war was only a small victory. I realized the problem my art department faced was a
symptom of a much deeper systemic problem in American public schools. It was quite clear to me the importance of assessment practice in visual art and design education, and the inherent flaws in educational policies that were shaping the teaching profession. I sought answers to questions that no one in my school could provide. I made the decision to pursue my doctorate to find what assessment practices were best for the field so I could help other art educators who might be struggling with similar questions.
CHAPTER 1
INTRODUCTION

When I started teaching in public high schools over 20 years ago, I felt respected and trusted as an education professional; however, throughout my career I witnessed the swing of the pendulum of educational policy. As an art teacher in an Illinois high school with a high population of underserved students, I have experienced firsthand all these changes. Unfortunately, my story is not unique; instead it is a small part of a widespread nationwide trend.

The effort to reform American public schools by increasing accountability, promoting standardization that can be easily measured by rote right-wrong learning, narrowing the curriculum to reading and math, and labeling public schools and teachers as failures is part of a nationwide attack on public education (Ravitch, 2014). Without much evidence of these claims, Americans have been led to believe our schools are failing and non-tested disciplines, including the visual arts, have been moved to the periphery of education (Kuhn, 2014; Ravitch, 2014). Although this movement has recently seemed to gain widespread attention in the media, reform efforts have slowly built momentum over decades. Chapman (2000) noted that in 1965 when the National Endowment for the Arts (NEA) and the National Endowment for the Humanities (NEH) were established it led to the separation of arts and humanities from the schools. Subsequently, art education in the schools is viewed as an enrichment, not an obligation and the arts are perceived as a luxury, making it difficult to promote art as a worthy academic study in the public schools.
Since then, in the book *Reign of Error*, Ravitch (2014) meticulously explained the succession of federal educational policies, including the publication of *A Nation at Risk*, the No Child Left Behind act (NCLB), and Race to the Top (RTT), which have systematically deceived the American people into believing public schools are lowering their quality of education, thus creating a false crisis mentality. These changes have ushered in a movement toward the privatization of education and cut a deep divide into the equity of American public schools.

Ravitch explains how policy makers and educational reformers have created a never-ending thirst for data. This has led to judging the quality of teachers by test scores, meaning non-tested or hard-to-test subjects like the arts are most likely ignored, limiting opportunities for risk-taking and creativity in school (Boughton, 2013; Ravitch, 2014). Performance-based disciplines, including the arts, have been a victim of budget cuts, and students in less economically advantaged districts with high demographics of racial minorities have suffered the most. Students enrolled in schools in higher affluent areas are afforded the luxury of arts in their schools. However, in a majority of public schools, mainly in high poverty areas, the emphasis remains narrowly focused on achievement in reading and math, leaving little time for learning that promotes critical thinking or any enjoyment in school (Ravitch, 2014).

As a doctoral student and practicing teacher, I found it unconscionable to sit back and watch the discipline of art and design education become undervalued because it is not included on standardized exams. Throughout the course of my career I witnessed changes in policy that altered the judgment of the quality of teaching from the act of teaching to ranking schools and teachers by scores completely unrelated to the quality of a teacher or a school (Kuhn, 2014; Ravitch, 2014). The inception of Common Core State Standards (CCSS) has accelerated the perceived definition of data to the notion of numerical, pre-determined scores derived from a
standardized test, becoming the accepted indicator of teacher performance (Ravitch, 2014). The current climate of educational reform has changed teacher performance evaluations and systems of tenure, and excluded visual art and design from the mainstream of the core curriculum (Freedman, 2007) thus motivated my research.

Statement of the Problem and Sub-Problems

Current educational policy for K-12 public schools in the United States emphasizes achievement in core subjects, including reading and math, performance on standardized tests, increases in accountability measures, and expectations for teachers and schools to continually produce data that demonstrate student growth while taking time and attention away from classroom instruction. Many states, including Illinois, have adopted the CCSS as their state standards to be applied across all areas of the curriculum. Because the visual arts have been left out of what is considered the core, there is a climate of uncertainty for art teachers when addressing assessment, data, and meeting expectations to align with CCSS. While no one can argue with the importance of schools to be accountable for providing students with a quality education, public school teachers, regardless of discipline, are being asked to adapt to a one size fits all model for the sake of data collection and raising scores rather than tending to the unique learning needs of each individual child. In 2010, Illinois implemented a law known as PERA, the Performance Evaluation Reform Act (PERA; ISBE, 2016). In addition to teacher observations, this act requires student performance data on school wide standardized tests, and student growth data to be collected by classroom teachers to be included as indicators of teaching effectiveness within a teacher’s performance evaluation. This act also determines the performance categories in which teachers are ranked. Based on each district’s interpretation of PERA requirements,
school-wide data for reading and math become part of the measurement of a teacher’s effectiveness regardless of the discipline they teach. Subsequently, many teachers in performance-based classes, including visual art and design, are being held accountable for incorporating reading and math activities into their classes, taking valuable class time away from teaching the content of their discipline. Additionally, teachers need to collect their own student performance growth data using assessment tools in their classrooms through valid and reliable means. Because of the current culture that holds an overreliance on standardized testing, a narrowed definition of data has been implied as something numerical that should be collected from right/wrong answer choice tests. This creates additional challenges for teachers in the visual arts that judge the growth of students using open-ended tasks and a more holistic means of assessment. To help identify the specific problems for this research, I conducted multiple pilot studies.

Knowing there was a potential problem related to the usefulness of assessing students with right/wrong answer choices and perceptions of authentic assessment, I conducted a pilot study in the spring of 2015 as a case study of four of my former advanced level art students who were adults at the time of the study. I asked them to reflect on their high school experience. Each of the participants was interviewed about what they considered the most memorable learning experiences from high school and what skills they developed in school that they found most useful in their adult lives. They were also asked about assessments they remembered most from high school that they felt were applicable to their adult lives. The data suggested the participants learned to apply many skills from their studio art course to other adult life issues, including how to hold a critical conversation with peers, the ability to take criticism, and the ability to be a self-directed learner. The assessment strategies they felt were the most useful included critiques and
portfolio presentations in their studio art course. None of the participants described any experiences related to any standardized test or right/wrong answer format as useful or applicable in their adult life. Instead the participants found those assessments to be the least useful for adult life skills.

It is important to note that the participants of this pilot study described above were former students from my classroom and the results of the study were specific to students who took an upper-level art class. The results of this pilot study inspired me to further critically analyze the current model of schooling and the policies that increased the movement toward standardization, thus limiting opportunities for creative exploration.

Understanding there was a strong connection between assessment data and teacher evaluation, I conducted a pilot study in the fall of 2015 about the requirements to collect student growth data, the perception and definition of data, and how PERA was perceived by teachers and administrators. I interviewed seven teachers from multiple disciplines and one administrator, analyzed documents shared with the staff related to changes in the teacher evaluation process, and made observations during faculty meetings and professional development related to the school’s adoption of PERA. In the interviews, staff members overwhelmingly defined data as something quantitative because that was what they were instructed and something derived from test scores, although the data they found most useful from their daily professional practice were qualitative experiences with students. When discussing how the teachers understood the new evaluation process, many described experiences of receiving inconsistent information from evaluating administrators, frustration with multiple changes to the process mid-year, and lack of clarity about how their observations would affect the outcome of their final evaluation. There were multiple indicators of elevated stress among the staff, and observations during faculty
meetings and professional development sessions revealed frustration and confusion with the increase in paperwork and the time it took away from actual teaching and learning.

I followed up this pilot study with a mini study in the fall of 2016 to examine this issue with other art and design teachers from five school districts throughout the state of Illinois. I also wanted to gather data to see if PERA was being implemented consistently in other schools throughout the state. I conducted informal interviews to get a better understanding of the fidelity with implementation of this policy. These interviews revealed drastic inconsistencies with implementation of PERA across different schools and even within districts as well as inconsistencies with ways information about assessment and student data collection had been disseminated to teachers. Further inconsistencies were revealed about the teacher evaluation process and the rubrics and evaluation models schools used for teacher observations and data collection. The art and design educators I spoke to also expressed concerns about the validity of their evaluations, since none of their evaluating administrators were trained in the field of art and design education. Many stated a concern that their evaluator had little understanding of the assessment methods employed in art classrooms and often asked for inappropriate forms of data or to implement inappropriate assessments in the art classroom for the purpose of collecting student growth data for evaluation. As I reflected on the intended purpose for teacher evaluation, which should ultimately improve student learning, I also considered what students learned in visual art classes.

Thinking deeply about student learning in the art room, I conducted a pilot study in the spring of 2016. I examined issues of creativity and risk-taking with students who were currently enrolled in my upper-level art studio class. Using ethnographic methods, students who were willing to be interviewed were asked questions about their views on creativity and risk-taking in
school. Data collected in this pilot study from classroom observations, student artwork, and the interviews suggested the students had concerns about the lack of time in their school experience to take risks and explore creative solutions to problems in other classes throughout their school day. The participants appreciated and enjoyed their art class because it was the one time during the school day when they felt they were allowed to express themselves. They also appreciated the opportunity to take risks in their art class and recognized that time as one when mistakes were allowed and failure would not be penalized. The participants expressed an understanding that learning through art making and academic risk-taking was a valuable learning experience that would help prepare them in adulthood. This series of pilot studies helped me gain clarity about the problems I wanted to research.

Through the analysis of these pilot studies, it became evident the many layers of the problems that were directly related to the Illinois educational policy of PERA, and how it influenced instruction and assessment practices in schools. PERA was interpreted differently by each school district, so art and design teachers in the state of Illinois were experiencing a different level of difficulty than their peers from other academic areas with compliance. Often the evaluating administrator was not certified in the visual arts, creating an additional burden on the art teachers to explain their practices to their evaluator. Additionally, instructional and assessment practices in the art room looked different from strategies utilized in other disciplines. Frequently, the assessment practices employed by art teachers were not be recognized by their evaluating administrator as a valid form of assessment, thus negatively impacting their performance evaluation.

To help illustrate the complexity of qualitative assessment, Figures 1, 2, 3, and 4 are four different examples of high school student artwork. These were produced by students in their
senior year taking the same class, yet each student used different media and technique but were judged against common generic criteria. Although they all look very different, they all received high marks when they were assessed. To arrive at such a determination, evaluators must possess widely differentiated experience of multiple artistic modalities sufficient to detect fine grained distinctions of quality irrespective of media or technique. Making accurate qualitative judgements about art work requires one to be a professional in the field. Eisner (2002) refers to this as connoisseurship, or the ability to make a fine-grained judgment about what a student has accomplished. Several dilemmas are presented in this scenario: How can anyone outside the field of art education be qualified to make judgements about these works? The challenge art teachers have is explaining to an outsider how the visual qualities in each of the works were first identified and then assessed. The problem administrators have with art teachers making these judgments is that the ways in which validity and reliability of assessment are established in artistic judgement are considerably different than the methods used in quantitative assessments, like multiple choice testing, with which most educators are familiar.
Figure 1: Student artwork sample 1.

Figure 2: Student artwork sample 2.
Figure 3: Student artwork sample 3.

Figure 4: Student artwork sample 4.
Adding another layer to the problem, Illinois adopted the CCSS. These standards do not address the visual arts directly or desired outcomes in the arts including creativity, so many teachers in the visual arts have been asked to align to standards inappropriate for their discipline. Furthermore, Illinois faces a statewide problem of unequal access to arts education. Districts with limited financial resources had reduced budgets or eliminated the arts completely to allocate school funds for core classes like reading and math. A study conducted in Illinois in 2005 about the inclusion of the arts in public schools found that “on average, only 25% of Illinois public high school students are enrolled in any art courses in a given year” (Illinois Creates, 2005, p. 11), illustrating the disparity in Illinois with curricular offerings in the arts. With the movement toward standardization, CCSS, reduced budgets, and school-wide test and student growth data are being included in a teacher’s performance evaluation, so it is imperative to examine the assessment practices in high school art and design classrooms. Attempting to understand the origins of these challenges has inspired my inquiry into the roots of current educational policies and opened my eyes to the importance of promoting authentic assessment strategies in the field of visual art and design education.

Context of the Study

This mixed-methods study was limited to Illinois public high school visual art and design teachers to gather data on the assessment methods implemented in high school art classrooms. Special attention was paid to the current Illinois educational policy known as PERA and how information is disseminated to art and design teachers about expectations to successfully fulfill requirements under this policy – in particular an examination of the consistency in the way information about PERA was disseminated throughout different school districts, if art and design
teachers received consistent information, and if they felt directives related to standards and assessments implemented for their performance evaluation were appropriate to their area of academic expertise. This study further examined the assessment methods employed by art and design teachers and whether they were consistent with best practices in the field with regard to validity and reliability. I examined how art teachers collected student data and what assessment data were shared with students, parents, and administrators. Furthermore, I examined possible forms of resistance art teachers may have exhibited if the directives from their evaluating administrators differed from best practices for art and design classrooms.

Concerns for the future of the field of art education were raised by Freedman (2014) when she described the educational crisis created by policies that have stripped the curriculum to reading and math because of the emphasis on testing, thus eliminating creative activities from schools. She urged the field of art education to produce “research to demonstrate its importance” (p. 2) and explained how current educational policy statements do not include terms that suggest creativity because it is difficult for test makers to assess creativity. This places art education at a disadvantage in this political time. Freedman urges researchers to “respond to the relationship between art and [the] politics of education, [and] new forms of accountability” (p.7).

Specific to the practice of classroom high school art educators, Boughton (2018) urges vigilant awareness of the significant pressure to increase accountability in education in the United States. A significant factor related to this study is the requirement for teachers to produce valid and reliable data-based evidence on student performance. Unlike other countries, assessment of the arts in the United States has relied on the judgment of the teacher alone. This has caused difficulty in producing valid and reliable data, suggesting a likelihood their judgments could be self-serving and raising the possibility of an equity issue if teachers can be dismissed
from their position because they have not produced convincing data-based evidence about student growth. Producing research within this context will help provide evidence of what practices are employed by practicing teachers in art classrooms.

Considering recommendations for future research, Eisner (2002) recognized contemporary changes in research methods encouraging the addition of qualitative methods to enhance the interpretation of educational problems and as an appropriate method for uncovering what strategies art teachers actually use in their classrooms. Eisner stated: “If we know little about the process teachers employ in classrooms, we will be in a poor position to improve teaching” (p. 216). To examine these issues on a large scale, a mixed-methods approach was appropriate to address such a purpose. Quantitative methods were used to collect data from a large population of teachers from throughout the state, and qualitative methods collected data from a smaller population to provide an opportunity to reveal detailed accounts of teacher’s actual professional experiences.

Significance of the Study

PERA requires teachers in all disciplines to assess their students and produce valid and reliable data that demonstrate student growth as part of their performance evaluation. This can potentially have a direct effect on teacher tenure. This is a significant change in the teacher evaluation process; however, this type of evaluation reform is not isolated to Illinois. Other states throughout the United States have implemented similar policies. As of 2015, 43 of the 50 states require student growth as a component of teacher evaluations (Boughton, 2018). Teachers are required to collect and document student growth data within their own classes known as Student Learning Objectives (SLO) plans. Data are only one component of a teacher’s performance
evaluation and are in addition to administrative observations. Student performance on schoolwide or state-mandated standardized assessments are also part of the evaluation process based on each school district’s PERA plan. While the state of Illinois provided options for non-tested disciplines in PERA, if a local school district PERA committee chose not to implement variations of assessments for non-tested disciplines, teachers in these areas of academics are at a disadvantage to their colleagues in tested disciplines (Goe & Holdheide, 2011; Marion & Buckley, 2011). According to Goe and Holdheide (2011), the percent of non-tested subjects and grades nationally is 69%. This is significant since this accounts for more than one half of the educators being evaluated, yet with the movement toward CCSS, most educators are bound by requirements in reading and math and not the discipline they are actually certified in. I am not asserting teachers should not be accountable for the quality of education of their students. It is important to ensure there are effective teachers in classrooms and they are helping their students learn. Instead I believe the means for collecting data to demonstrate student growth and the data that are reflected in a teacher’s evaluation should be appropriate to the discipline in which the educator is licensed to instruct.

The most appropriate and valid means for assessing learning in the visual arts is a portfolio or performance-based assessment (Beattie, 1997; Boughton, 2013; MacGregor, 1992). Research and literature have been generated about how to achieve reliability in portfolio-based assessments; however, because of the great variations in curricula and access to visual art education throughout the United States, there are limited opportunities for art teachers in public schools to achieve reliability as compared to those in other countries (Boughton, 2018; MacGregor, 1992). This study explored what methods high school visual art and design teachers in Illinois are implementing in their classrooms to assess their students, whether their
assessments are valid and reliable, and what differences are experienced among teachers from different school districts.

Additionally, the observational component in the current teacher evaluation model that assesses teacher classroom performance in most districts is largely based on the Charlotte Danielson (2007) method that promotes a teacher portfolio to demonstrate classroom practice. This model and the domains for teacher evaluation have been adopted by most districts for administrators to use; however, individual school districts are allowed to adopt their own model, rubric, or other variation, leading to inconsistency throughout the state. Most professional development provided school-wide or district-wide caters to meeting CCSS in reading and math. Without adequate resources, or appropriate professional development specifically for art educators, teachers in non-tested disciplines are left vulnerable to policy that seemingly excludes the significance of performance-based learning that is essential to the visual art and design field.

Research Questions and Sub-Questions

To address the problems I have outlined, one main research question guided this study, and four sub questions helped examine the nuances of problems experienced with assessment more closely. This set of research questions guided the quantitative and qualitative data collection for this mixed-methods study:

1. How are high school visual art and design teachers expected to utilize assessment practices by their school districts in the context of a nationalized shift toward standardized assessment?
   a. What types of assessment strategies do Illinois high school visual art and design teachers report using in their classrooms?
b. What directives are Illinois high school visual art and design teachers given by their school district administrators about reporting assessment data?

c. What professional development about assessment is provided for Illinois high school visual art and design teachers by their school districts?

d. What professional development about determining validity and reliability in assessment is provided for Illinois high school visual art and design teachers?

Delimitations and Limitations

This study was influenced by national trends in educational policy; however, one of the delimitations of this study was that the participants were selected from active public secondary or high school art and design teachers to limit the scope of experience with grade level and assessment strategies. Participant selection only included teachers working in the state of Illinois to keep the study specific to educational policies and standards enacted in Illinois. Another delimitation was when generating the contact list for survey research, participants were selected from counties that had available online information about their high schools and had staff listings for their visual art departments. Further delimitations for this study included the selection of eight participants for 45-minute face-to-face interviews because of time limitations to complete the research.

An effort was made to select interview participants who would represent multiple geographic regions of the state of Illinois; however, a limitation was the disproportionate number of high school art teachers who work in the northern, central, and southern parts of the state. Because of this limitation, the only participants who volunteered for interviews were from the northern third of the state. The researcher intended to interview participants who were ethnically
diverse; however, another limitation was the only participants who volunteered to be interviewed were white, which does not reflect the diversity of the population.

The quantitative data collected during this study may be generalizable to art and design teachers in Illinois. However, as a limitation, the data may not be generalizable to other populations in other states. A benefit to using a web-based survey is the ease of distributing and collecting data to a large population (Mertens, 2010). A total of 160 participants responded to this survey; however, not all respondents answered every question creating a limitation of the instrument. Frankel and Wallen (2006) suggest possible reasons for item nonresponse may be the respondent did not know how to answer a question, may have found the question irrelevant, or may have failed to record an answer. According to Mertens (2010) and Frankel and Wallen (2006) this creates a potential threat to validity. The number of respondents who answered each question is provided with the description of the quantitative data. Some questions in the survey were contingency questions and did not require all respondents to answer the question.

A limitation of the qualitative data is it is only generalizable to the interview participants, and the number of participants for this data set was smaller than the number of respondents in the quantitative data. A benefit of this method of data collection, however, is the ability for the researcher to provide thick description with multiple participants and allow for more attention to the voice of individual participants (Mertens, 2010).

Definitions of Terms

For the purposes of this study, the following terms are defined:

**Authentic Assessment**: Student performance is judged in terms of the capacity to complete a real-life performance task.
**Common Core**: An educational initiative in the United States that determines the level of learning each student should attain at grade level in math and English language arts.

**Common Core State Standards (CCSS)**: Illinois state learning standards based on Common Core standards that measure student performance in mathematical practices and literacy.

**Covert Resistance**: Inwardly objecting to a directive; however, outwardly complying to pacify the one making the request.

**Critique**: A judgment of student work that can be used as a formative assessment, by providing information to assist students in seeing work in a different way to improve performance, or it can be used as a summative assessment as part of the judgment or merit of work.

**Discipline Based Art Education (DBAE)**: An approach to art education that drew upon four components to art instruction: art production, art history, art criticism, and aesthetics. This movement emphasized museum/fine art.

**Formative Assessment**: A way of providing students feedback about their progress of their own work.

**Goals 2000**: A 1994 federal education act that mandated schools provide equitable learning opportunities for all students of all levels.

**High Stakes Tests**: A test that carries important, sometimes life-changing, consequences for the test taker and/or the educational system that administers the test.

**Moderation**: The process of adjusting grades to achieve equivalence across multiple contexts in which the same content is taught.

**No Child Left Behind (NCLB)**: A 2002 federal education law that increased accountability of schools by rating them according to the results of school-wide standardized testing. This policy set a graduated timetable for all schools to demonstrate that all students would meet or exceed
performance standards. Ratings for schools were proficient or failing, and compliance with the policy coupled with the rating determined the amount of federal funding a school would be awarded.

**Overt Resistance**: Outwardly objecting to or defying a directive.

**Performance-Based Assessments**: An assessment that determines a students’ level of ability on a prescribed task. For the purposes of art and design assessment, this assessment can be open ended.

**Performance Evaluation Reform Act (PERA)**: An Illinois law passed in in 2010 that created statewide reform of the teacher evaluation process. This law provides guidelines for school districts to follow on how to conduct teacher evaluations using four rating categories: excellent, proficient, needs improvement, and unsatisfactory. The information used to determine a teacher’s rating is calculated by administrative observation, schoolwide performance on standardized tests, and teacher collected student growth data. These categories can affect teacher retention.

**Portfolio Assessment**: The process of judging students’ art learning based on the data present in a portfolio of work. Assessment is based on the whole body of work and is not constituted as an aggregation of individually graded projects.

**Qualitative Assessment**: Making a judgment about the degree to which qualities are present in a performance or object.

**Quantitative Assessment**: Assessing student achievement based on things that can be counted.

**Race to The Top (RTT)**: A federal policy enacted in 2009 that offered competitive grant funding to schools willing to show growth and progress toward college preparation.

**Reliability**: The extent to which an assessment yields the same or consistent results when repeated.
Rubric: A set of statements describing performance levels that may be achieved in relation to a criterion. Rubrics will typically have a numerical scale aligned with performance levels.

School District: A zone determined by local and state officials that sets geographic boundaries for students attending a specific school based on their residence. Public school districts in the state of Illinois are overseen by a local school board that has voting rights over personnel and specific school district policies, which can include grading and assessment procedures.

Standardized Tests: A standardized test is administered and scored in a consistent or standard manner. Standardized tests are designed in such a way that the questions, conditions for administering, scoring procedures, and interpretations are consistent and are administered and scored in a predetermined and standard manner (Popham, 1999).

Student Growth: A measurement of student learning data that teachers and school districts are responsible for collecting and tracking throughout the school year to measure how much students learned in a specific class or over a pre-determined length of time.

Student Learning Objectives (SLO): A teacher-generated plan that must be approved by the evaluating administrator that is used to plan and track the growth of each student and projects the percentage of growth for a class in accordance with PERA.

Student Led Portfolio Examination: An authentic assessment method that allows students to make choices of which artworks are submitted for a portfolio exam and are encouraged to verbally defend or explain their choices and how it demonstrates learning or growth in visual art or design.

Summative Assessment: An assessment used to measure student learning or mastery at the conclusion of the unit, semester, or specific learning period that can be used to determine a grade for a marking period. Summative assessments can also be used as gatekeeping assessments to
judge or determine access for students to educational opportunities such as continuance in any program or for university/college entry.

**Type I Assessments:** Assessments that are state mandated standardized assessments or assessments issued on a large scale such as the ACT or SAT. According to PERA, schoolwide performance of these assessments is included in a teacher’s performance evaluation.

**Type II Assessments:** Department-wide or district-wide assessments implemented within a school or school district that can be issued as pre-, mid-, or post-test. These are most often written in a multiple-choice or right/wrong answer format. According to PERA, Type II assessments are used for teachers to gather student growth data regarding a teacher’s SLO plan and are part of a teacher’s performance evaluation.

**Type III Assessments:** Teacher-generated assessments gather student growth data regarding a teacher’s SLO plan. These can be formative, performance-based, rubric-scored, or portfolio-based and must be agreed on between a teacher and the evaluating administrator. According to PERA, Type III assessments are used for teachers to gather data as part of the teacher’s performance evaluation.

**Validity:** The extent to which a test or an assessment measures what it is supposed to measure.

**Theoretical Foundation**

The theoretical foundation of this dissertation is rooted in critical pedagogy (Freire, 1970/2000). Critical research that analyzes the effects educational policies have had on teachers’ working conditions, assessment practices, and students’ learning environments is necessary to understand the context of the problematic circumstances art and design teachers are experiencing today. Critically examining the many complex layers of the act of teaching and how one knows
that students are learning is necessary to understand how educational policy and actual classroom practice either work in harmony or are at odds with each other (Freedman, 2014). To understand the nexus of critical research and examination of educational policy, a contemporary context of critical pedagogy is explained by Nikolakaki (2012):

Critical pedagogy refers to educational theory and teaching and learning practices that are aimed to raise learner’s critical consciousness regarding oppressive social conditions. Critical pedagogy not only focuses on personal liberation through the development of critical consciousness, but has a strong political component in that critical consciousness is positioned as the necessary first step of a larger collective political struggle to challenge and transform oppressive social conditions and to create a more egalitarian society. (p. xi)

Early in my career, I embraced the philosophy of the progressive movement. Wanting to be a caring teacher who served the needs of my students through the visual arts was the primary motivation for my career; however, throughout my career I witnessed stifling conditions imposed on teachers that seemingly limited the capacity for educators to promote criticality. Responding to these shifts in policy and practice, I adopted a more contemporary paradigm of critical pedagogy.

The progressive movement was a time in American schooling that reacted to the problems in education and political policy that was caused by industrialization and was committed to promoting mobility and equality. Dewey (1916/1997) believed education to be a process that nurtures students’ ideas rather than directs and stifles their interests. He contended an education that fosters enlightenment and guides students through an authentic learning process was necessary for a democratic society to thrive. His philosophical perspective is just as poignant today, as it was over 100 years ago. The underlying belief that education is the process of nurturing, coupled with my personal belief that individuals flourish when they are allowed to learn at their own rate, has led me to question the motives of current educational policies that
demand all students perform at the same rate through a standardized curriculum. Dewey (1934/1980) advocated for meaningful learning through experience and suggested creative and artistic approaches as appropriate ways to do this. He respected the thinking and knowing of the artist as an important part of human experience. This should not be separated from education but fostered as a part of the learning process.

Encouraging teachers and students to empower themselves by questioning the status quo and learning to critically navigate complex power structures has foundations in the writing of early critical pedagogue Paulo Freire (1970/2000). He advocated for students to become co-constructors of knowledge with the teacher, allowing the teacher and students to learn together, breaking the cycle of oppressive education that controls the flow of information to the learner. In Pedagogy of the Oppressed, Freire is well known for his description of popular public education as a banking model in which teachers provide knowledge by filling their students as if they were empty vessels. The signification of learning requires students to simply regurgitate back the facts, thus successfully demonstrating what they have learned. There is no challenge and no thoughtful dialogue between student and teacher. It is an efficient superficial form of learning that does not promote deep meaningful questions from the students. This becomes a way to perpetuate a capitalist agenda, dehumanizing a population for a lifetime of mindless work. Students are conditioned to become passive receptors with no exchange of deep meaningful learning. Critical pedagogy, however, promotes knowledge that “emerges only through invention and reinvention, through the restless, impatient, continuing, hopeful inquiry human beings pursue in the world, with the world, and with each other” (p. 72). Freire (1974/2008) explains that the organization of human labor is one of the most powerful byproducts of a technologically mass production minded world. Man becomes domesticated by learning to behave mechanically,
requiring no critical thoughts. Education becomes the key to reclaiming humanity. He describes the promotion of teaching critical thinking as an act of love and courage.

Looking deeper into the control exemplified in this type of education, Foucault (1975/1995) critically examined the structure of schooling, likening it to the prison system, and critiquing the limiting nature of how schools were structured. He was critical of how school personnel are often expected to keep the status quo by transmitting mediocrity and low-level knowledge rather promoting criticality in students. In a contemporary context, this can be compared to the promotion of a standardized curriculum that does not promote higher levels of learning: the fear that knowledge becomes power that fuels resistance and resistance threatens power, thus creating the cycle of power and resistance. Critical thinking and risk-taking are necessary components in an art class that promotes visual culture (Freedman, 2003), and these skills cannot be fostered if a teacher clings to prescribed and predictable outcomes that do not challenge thinking and authentic learning. Similar to Foucault (1975/1995), Nikolakaki (2012) also likens schools to prisons in a culture of fear that will not resist dominant forces. “In other words, they want to kill the soul of the Youth before it becomes expressive or offensive” (p. 19). As policymakers and education reformers seemingly focus on creating better test takers, they lose sight of the purpose of a well-rounded education by disregarding any individuality on the part of the student.

Encouraging teachers to challenge a predetermined curriculum, Greene (1995) sees the arts as a way to provide insight into the world. She encourages teachers to promote a pedagogy that engages all member of the class in meaningful dialogues; making the classroom an active place is likely to deter apathy, which opens possibilities to expand education beyond the school and resist meaninglessness. Rejecting standardization and fostering an atmosphere for students
that promotes purpose in learning and having an imagination of what their hopes can become is where the arts have a place of importance in the school curriculum. Greene advocates for teachers to model a willingness to transform thinking and shape vision, becoming more wide-awake to the world. Similarly, hooks (2010) encourages engaged pedagogy that promotes active participation of all members of the class, including the teacher, in a working relationship. This type of classroom eliminates hierarchy, stimulates independent thinking on the part of the students to help them find their voice and, as a result, engage in critical thinking. This type of classroom space and learning environment is seemingly opposite to the kind currently promoted through a standardized curriculum, and philosophy of a one size fits all education. However, hooks notes teaching critical thinking is difficult because students are more comfortable with passive learning, “Students do not become critical thinkers overnight. First, they must learn to embrace the joy and power of thinking itself” (p. 8). With mutual respect and the teacher demonstrating a commitment to courage and imagination, students can develop confidence in their thoughts. hooks reminds teachers that it is okay to not always know the right answers and to acknowledge what they do not know, thus everyone in the classroom can recognize “they are responsible for creating a learning community together, learning at its most meaningful and useful” (p. 11). This kind of exchange is crucial to an effective art studio class that promotes critical thinking, creativity, and risk-taking.

These theoretical lenses have laid the foundation for resisting pre-determined curriculum and embracing critical pedagogy. In Giroux’s (2001) book Theory and Resistance in Education, he unequivocally argues for the position that criticality and resistance are imperative for educators today. He warns that teachers and schools are under siege unlike ever before and corporations now define schools. He is aware that teaching criticality is no easy task and
recognizes the reality of teaching in overcrowded classes, working with limited resources, and constantly fending off media driven remarks from hostile legislators. Giroux (2012) describes educational reform as the new push against critical education. He cites the purpose of education is to produce citizens who can think critically and use knowledge to make judgments necessary for a democratic society, criticizing the current paradigm of educational reform as a way to reduce students to cheerful robots. In response to the commodification and privatization of public schooling, Giroux explains critical pedagogy seeks to open possibilities for resistance and social change. He supports Freire’s (1970/2000) criticism of current educational reform as a passive absorption of knowledge. Political policy has worked to reduce education to a narrow focus of rote learning, memorization, through high-stakes testing to serve a market-driven culture in which young people are treated as consumers. Giroux (2012) encourages teachers and students to “actively participate in narrating their identities through a culture of questioning [and to] act on the knowledge, values, and social relations they acquire by being responsive to the deepest and most important problems of our times” (p. 14). Educators, like me, who have rooted their educational philosophy in critical pedagogy view their students as individuals who have unique learning capabilities and know they cannot be standardized.

Assumptions

This study rooted its research questions on the following assumptions: The current movement of standardization in public education is reducing opportunities for students to be creative and decreasing the chances for students to practice independent activities in school that promote critical thinking skills. Overreliance on predetermined answers limits the possibilities of what students can learn in the art classroom. Students at the high school level should be allowed
to take educational risks to test knowledge and find ways to apply their personal experiences to a variety of media. This type of learning welcomes a lack of predictability in the students’ outcomes and opportunities for teachers to make informed judgments when assessing students’ learning, thus widening the scope of assessment beyond right/wrong answer choices. Performance-based learning in visual art and design is best assessed by portfolio and authentic assessments.

There is an assumption that most current art and design teachers are moving beyond a DBAE curriculum and embracing a visual culture curriculum (Freedman, 2003) that fosters opportunities for students to participate in creative activities, critical thinking, and educational risk-taking. This moves the art curriculum beyond a white-male Western-centric view of art to a more contemporary understanding that art teachers help students navigate the complex visual world students inhabit. Additionally, with the current requirements for data collection, an increase in standardized testing, and pressures for CCSS in reading and math to be implemented school wide, I am making an assumption that art teachers are choosing to implement best practices like performance-based assessment for their students’ art and design work. I am further making an assumption that art teachers who are asked to implement multiple-choice or right/wrong format tests to collect student performance data are responding with some level of resistance, either covert or overt, as part of satisfying requirements mandated by PERA, with the motive of preserving their positions as teachers.
CHAPTER 2
LITERATURE REVIEW

This study is concerned with the educational policy known as PERA and is limited to high school visual art and design teachers throughout the state of Illinois. Within the context of the study, it is worth noting PERA imposes guidelines on school districts for teacher performance evaluations that include test-based data and teacher-generated assessments as a part of the determination on the quality of a teacher. To establish the context for understanding how the American educational policy has shaped teaching and learning conditions in schools, it is necessary to examine historical patterns of educational policies to understand the context of current policy and the state level. Part of the marginalizing effect of standardization is a movement to Common Core that has narrowed the curriculum to reading and math, creating unequal access to the arts with underserved populations.

Because this study is also concerned with the teacher evaluation process, observation models and practices are examined in the context of the current policy, and attention is given to how art teachers are evaluated with respect to their curriculum. Because teachers are required to collect student growth data though self-generated assessments, there is a review of best practice for art and design education, including performance-based, portfolio, and alternate assessments. Since this study is concerned with high school, it is necessary to understand the characteristics of adolescent aged students and what is appropriate in art learning for this age group. A desirable
outcome of art learning with adolescents is creativity, so literature that examines creativity and
the conditions that can foster it is also included.

Finally, this study collected data from practicing teachers about their performance
evaluation, which ultimately has an influence on a teacher’s working conditions. This presents
possibilities for compliance or resistance. An examination of resistance and the various
manifestations of resistance is necessary to understand working teachers’ experiences and why
they may choose to comply or resist. The literature review concludes with recent research
relevant to this study.

Educational Policy at State and National Levels

To understand current American educational policies and the effect they have had on
classroom teachers in visual art and design, it is important to begin by examining the roots of
these policies and how they have evolved. Maranto (2015) described some of the earliest known
educational legislation. In 1642, the first law pertaining to education in America required
children to learn to read so they could understand religion and laws. Parents were required to
uphold this duty. However, as soon as a public-school system was established in the U.S., a push
for educational reform quickly followed. The first standardized test was administered by Horace
Mann in 1845. The results alarmed school officials by showing only 35.5% of students were able
to answer questions correctly. Maranto further explained that standardized testing gained
popularity during World War I to help identify the abilities of military recruits. The interpreted
poor results put education as a priority when it came to issues of national security. The report
bolstered the implementation of achievement testing, and in the 1930s, due to technological
advances, multiple-choice became the preferred test format because of cost efficiency and ease of scoring.

Under the Reagan Administration, the publication of *A Nation at Risk: The Imperative for Educational Reform* in 1983 called for additional testing, citing underachievement and mediocrity in the American educational system, creating a push for data collection in reading, science and math. President Clinton introduced legislation that encouraged a test-based retention policy known as *Goals 2000*, but it was the No Child Left Behind Act (NCLB) under President George W. Bush in 2002 that fueled the reliance on the use of standardized testing, with its focus on reading and math to measure the success of schools (Maranto, 2015; Ravitch, 2014). Chapman (2005) carefully outlined the components of NCLB, translating the acronyms associated with it to help art teachers navigate the complex set of mandates and explain how these reforms could have a potential negative effect on art and design programs across the country. She explained why NCLB was never attainable, foreshadowing the movement to charter schools and business opportunities created from testing and education describing an “enterprise of schooling into a free marketplace where education service providers compete for customers” (p. 12). Ravitch (2014) concurred that the 100% proficiency timetable outlined in NCLB was never intended to be attainable. Instead it was designed to punish schools to open the door for entrepreneurial opportunities. While the original intention for charter schools was for the benefit of students, the way NCLB was designed allowed misappropriation of charter schools, turning them into corporate run businesses. Chapman (2007) questioned why school districts became so beholden to federal policies like NCLB when funds allocated to states were “only about 7 to 13 percent of state education budgets” (p. 25).
With a national emphasis on reading and math skills, it has been documented in research (Maranto, 2015; Ravitch, 2014) how policies have been narrowing the curriculum, thus leaving other disciplines out of the educational focus. To examine this issue in further detail, Ravitch’s 2014 book, *Reign of Error: The Hoax of the Privatization Movement and the Danger to America’s Public Schools*, critically analyzes over 30 years of U.S. educational policy and cites evidence that the American public is purposefully being misled into believing public schools are failing, and private and corporate-run charter schools are the solution by promoting a crisis mentality driven by fear. She described how the corporate reform movement is destroying public education; it is making way for privatization and exploitation; “It is not meant to reform public education but is a deliberate effort to replace public education with a privately managed, free-market system of schooling” (p. 4).

Reform policies like NCLB created a pattern of punishing schools that were not performing well according to tested standards and thrived on humiliation tactics as a motivator or schools to step up. McMurrey (2014) explained that it made the achievement gap even greater. Following NCLB, President Obama enacted Race to the Top (RTT), which continued the cycle of punitive policies that focus heavily on standardized testing in reading and math as a way to measure school success (Maranto, 2015; McMurrey, 2014; Ravitch, 2014). A study by Berryhill, Linney, and Fromewick (2009) found that the pattern of political policies focused on test-based educational accountability have had more negative effects than positive ones.

RTT furthered the push for charters and punished schools by promoting competition among school districts. Ravitch (2014) contended RTT was actually worse than NCLB since it created bi-partisan support of more testing, accountability, and choice by allowing consultants and vendors to profit from schools. Furthermore, she explains that since the onset of NCLB and
continuance with RTT, there have been significant cutbacks in non-tested subjects, especially in lower economic areas negatively affecting opportunities for all students to have access to the arts. Affluent communities continue to have a robust curriculum including the arts; however, schools in less affluent communities have schools that cater to the tested skills demonstrating undeniable inequality. Quinn (2013) concurred that funding for arts education and the number of art programs in public schools are down. A decade of punitive federal policies and economic recession, combined with the increased emphasis on reading and math, “have contributed to the weakened state of art education in public schools” (p. 37). She further explained how private and charter schools can hire uncertified educators to teach the arts, perpetuating a pattern that continues to undervalue the arts by not paying equitably for the education and not providing an optimal learning opportunity to students.

The United States enacted Every Student Succeeds Act (ESSA; U.S. Government Publishing Office, 2015) to modify RTT with the intention of scaling back the amount of required standardized testing in schools; however, there is still a strong component of measuring and comparing schools. The Illinois adoption of ESSA (Illinois State Board of Education, 2017) is based on the U.S. Department of Education’s plan and outlines the percentages of performance categories by which each school will be measured and ranked. Like the U.S. model, Illinois’ plan has a lessened emphasis on standardized tests and considers other indicators for determining school performance. The fine arts were supposed to be added to the calculation of school performance starting in 2019 at a value of 0-5%. The exact value and ways to calculate this performance indicator were undetermined, and inclusion of the arts as a component of school success has been delayed because of the lack of equitable access to the arts throughout Illinois schools. While it is encouraging that the arts are potentially being considered as a performance
indicator of schools, it continues to take a back seat because of its undetermined value and has a potential of being repealed due to unequal access to art and design. Furthermore, Section 6.1 of the document states the “ISBE is dedicated to providing resources that enable schools to support the development of the whole child” (p. 118); however, nowhere in this section is visual art and design or funding for the arts specifically addressed. Many other performance-based learning indicators that address the whole child are identified; however, references to art learning or creativity are not included, demonstrating a lack of regard for the arts as an integral part of addressing all children’s learning needs.

More policy-based changes from many states now correlate student test data with teacher performance evaluations. Empirical research from Ballou and Springer (2015) found good intentions behind the changes in the evaluation system, however uncovered “a number of problems in their design and implementation that need to be addressed” (p. 6). Because of the high-stakes nature of these policy changes and no sign of an easy fix, Ballou and Springer foresee legal action that could ensue as a result of these changes. These policies are relatively new, and the research is just beginning to unfold on the impact they have on education. Walker (2018) explained how the series of policies that have promoted test-based accountability have tainted the environment of teaching and learning. Despite grassroots movements that have spoken against the misuse of testing led by educators, parents, and students, many policymakers continue to rely on high stakes testing as a way to evaluate and improve schools. A statement released by the American Educational Research Association (AERA) in 2015 cautioned against the use of high stakes testing as a way to evaluate educators and suggests this as a platform for future research.
Regardless of negative public opinion and research-based concerns, Illinois has created a policy to include test data as part of a teacher’s performance evaluation known as PERA. According to the Illinois Association of School Boards (2016), the plan to roll out PERA requirements for teacher evaluations was on a staggered schedule. Chicago Public Schools were the first to implement it in 2013; schools covered by a School Improvement Grant (SIG) would specify a date in their grant agreement for implementation. Districts that received RTT grants would have alternate phase-in dates according to grant agreements, and districts ranked in the lowest 20% of performance in the state based on 2014 test scores would have begun in 2015. All remaining school districts were to implement PERA in 2016. Each district was to have developed a PERA committee of teachers and administrators to interpret the requirements of the law and develop a plan for student growth measures for their school district. Support materials for non-tested disciplines were provided by the state of Illinois to help those disciplines with appropriate alternate assessments. However, if a district’s PERA committee chose not to recognize the suggested alternate methods, teachers in non-tested disciplines could possibly be measured by inappropriate instruments. When PERA began its rollout in 2013, complete information and details were not available to all school districts. However, some schools in the state were held accountable for implementation. Because different districts began implementation at different times and each district could interpret and implement PERA differently, there was potential for inequality in implementation throughout the state.

Pressure on schools and teachers to quantify student growth as part of teacher retention is not limited to Illinois, but instead it is part of a nationwide reform movement. Boughton (2018) described how pressure on accountability has increased significantly in the U.S., with many states requiring teachers to produce valid and reliable data-based evidence on student
performance. As of 2015 43 of the 50 states required student growth as a component of teacher evaluations. Since assessment in the arts is predominantly qualitative, one of the greatest challenges for arts teachers in the current culture of standardization is to quantify student achievement in a way that is easily understood by parents, administrators and legislators.

One of the most threatening consequences of polices like PERA is the potential to eliminate teacher tenure. Ravitch (2014) asserted that eliminating tenure and structuring teacher evaluation to focus on test scores does not create a better educational environment:

Without tenure, teachers would be wise to stick to the blandest, least controversial books and topics, or to the textbooks, which have already been carefully screened by review panels to eliminate anything remotely controversial. So, if reformers succeed in eliminating tenure, they will eliminate teacher’s academic freedom as well. (p. 129)

This would lead to not taking educational risks essential for a robust arts curriculum. Ravitch further described the value of learning in the arts, and how overreliance on multiple-choice tests does not help foster a culture of learning:

It teaches students that questions have one right answer, and in life that is seldom correct… many questions that people encounter… in real life have answers shaded in gray… In real life, people do not always agree on the right answer. The tests we now value don’t teach what matters most, which is the ability to think for oneself. (p. 266)

As educational policies have increased structure on public school curriculum, one of the primary means of measuring school success has been through standardized testing.

Culture of Standardized Testing

There is no doubt that teachers and schools should be held accountable for the quality of the students’ education; the disagreement lies in how to accurately measure student learning (Boughton, 2013; Popham, 1999). Popham (1999) stated, “Standardized achievement tests should not be used to evaluate the quality of education. That’s not what they are supposed to do”
(p. 10) and do not accurately reflect what is being taught in the classroom. They are simply a measure of how a student is performing compared to their peers on a large scale. This has very little to do with teacher or school performance, yet the cycle of overreliance on standardized testing in American education is being perpetuated.

Many other nations have adopted performance-based assessments to measure their students’ academic achievement; however, American school districts continue to rely heavily on high-stakes and multiple-choice tests as a measurement of students’, teachers’, and schools’ performance. Schaffer (2012) noted, “The United States is the only economically advantaged nation that relies heavily on multiple-choice tests” (p. 40). Many researchers have concurred that standardized tests are still popular in the United States for political and economic reasons (Chapman, 2005; 2007; Ravitch, 2014; Schaeffer, 2012). Boughton (2004b) described how many classroom teachers feel pressured to use multiple-choice assessments because of educational policies that demand data about student learning, which can result in students becoming accustomed to providing short response right/wrong answers rather than extended responses that require more independent thought on the part of the student. Subsequently, teachers and students succumb to the pressure of high performance on standardized tests, leaving few opportunities for individualized learning or working on real-world problem-solving skills.

Because of the importance placed on standardized test performance, valuable learning activities are becoming overlooked in many courses to make way for test preparation activities. Chapman (2005) stated that only seven months out of the school year are spent on teaching, with the remainder of the school year spent on test preparation and testing activities, thus significantly reducing instructional time. Subsequently, teachers in core courses are taking fewer academic
risks with instructional activities, which reduces the opportunity for students to take educational risks for fear of jeopardizing test scores (Berryhill et al., 2009).

Going deeper into the issues resulting from standardized testing, Kearns (2011) examined how standardized testing negatively impacts students’ educational experience and self-image when they do not perform well on tests. Berryhill et al. (2009) conducted a mixed methods study on the effects educational policies have had on teachers. Their findings included empirical evidence of teacher burnout, reduced self-efficacy, and increased stress. There was a higher incidence with teachers who worked in low achieving schools and those who were in non-tested disciplines. Further highlighting the misappropriation of testing, at the National Art Education Association in 2018 Chapman presented findings that demonstrated current examples of how private corporations, policy makers, and corporate-run charter schools are systematically working together to standardize education and mine data from student performance to digitize teaching and learning for profit.

Providing further evidence that contradicts media-driven reports that American schools are failing, Ravitch (2014) cited evidence that school achievement is actually improving in the U.S. and the material students are learning is actually more difficult than in the past. She contends there is no crisis as reformers would lead the public to believe; however, a crisis mentality is profitable for testing companies. Schaeffer (2012) further emphasized the disconnect between the views of public officials and educators by asserting many politicians are beholden to heavy financial supporters in favor of high stakes testing because of the financial connection, although general public opinion supports cutting back on testing. This brings our attention to the needs of the learners, as they are the real victims in this cycle.
Marginalizing Effects of Standardized Testing on Underserved Populations

There is much evidence that the cycle of standardized testing has had an adverse effect on marginalized and underserved populations. Kraehe and Acuff (2013) define the term, underserved as synonymous “with terms such as urban, inner city, diverse, at risk, and low SES” (p. 296). They concur that most rhetoric about closing the achievement gap refers to core subjects such as reading or math, which ignores the arts and other non-tested disciplines and creates a larger problem for the neediest populations having the fewest opportunities in arts education by widening the quality-of-service gap. Gonzalez Stokas (2016) supports this by stating:

Most troubling, yet known anecdotally by many educators, was the clearly reported statistical equity-divide in access to the arts. The report showed that the higher the percentage of students receiving free and reduced lunch, the lower the percentage of students receiving arts instruction (p. 140).

There is an indisputable correlation between socioeconomic status and access to education in the arts. This disparity of access to the arts creates an equity gap that reaches across social class and ethnicity. Katz-Buonincontro (2018) expresses concerns about equitable access to art education by providing data that demonstrates spending on arts education has decreased and access to arts education has been reduced, especially with African American students:

The data point toward a lack of opportunities for students to experience the development of creativity in arts-based learning. And, these data point toward a disproportionate effect of decreased arts-based creativity for African American students and students living in high poverty (p. 35).

Illustrating how the equity gap has driven a wedge between classes, Quinn (2013) explains wealthier communities offer a rich variety in curriculum, and students living in poorer communities attend schools with a narrow curriculum, thus limiting access to the arts. Many
schools deemed low performing risk losing government funding if they do not comply with testing mandates, which continues to magnify the inequities in American schools (Chapman, 2005, 2018; Ravitch, 2014). Through his experience as an administrator, Kuhn (2014) saw how policy-based reliance on test scores alone as a way of advancing students to the next grade level or as a graduation requirement leaves behind students who are gifted in hands-on skills or artistic ability and who may not necessarily test well but can demonstrate their knowledge and skills in performance-based tasks that are not tested.

Chapman (2007) explained how punitive educational polices have laid the groundwork for school choice, vouchers, charter and for-profit schools and systematically undermined the public schools, which actually exacerbates segregation. These policies have acted as a distraction to the real issues in education (i.e., inequity with funding and poverty), thus driving the deeper divide between the rich and poor. To debunk the necessity for charter schools, Chapman states “seventy percent of parents are satisfied with the schools their children attend. The idea that competition among 'education service providers' will guarantee better and more equitable funding is a fraud” (p. 28). Chapman asserted that charter schools have not proven themselves to be any better at retaining teachers: “About half the teachers who leave charter schools opt out of teaching altogether” (p. 28). Chapman further questions if it is wise to organize American public education “in a manner that bears some resemblance to both Soviet-style state planning and free marketeering” (p. 31) with management of curriculum and teacher training handed over to a market-driven system. Quinn (2013) concurs that free market or school privatization may seem like a compelling idea that supports the American way; however, she contends this has driven a wedge between classes. Wealthier communities offer a rich variety in curriculum, and poorer communities’ schools have a narrow curriculum with limited access to the arts. Many
private/charter schools hire uncertified educators to teach the arts, continuing to undervalue the arts by not paying equitably for the education and not providing an optimal learning opportunity to students.

Cautioning against reliance on standardized assessment by reformers, Ravitch (2014) explains research-based evidence is not considered before applying mandates. Reform efforts have narrowed the curriculum, especially in schools with high poverty and Black and Hispanic students, thereby failing to narrow the achievement gap. All schools should promote equity offering a diverse curriculum to all students, including the arts. Ravitch reminds present day critics of education who long for the good old days that the old days were times when schools were segregated, did not serve non-English speaking students, did not include students with disabilities, and most students did not complete high school or go to college. Poverty and high concentration of racial minority are the leading factors in lower test scores; “the schools for the poor are the likeliest to sacrifice time for the arts and other studies to make more time for standardized testing and test prep… for fear the staff will be fired and the school will close” (p. 295). The achievement gap goes far beyond the reach of schools; social, economic, health, and external reasons are part of it, yet the schools are held completely accountable for most things that are out of their control.

Kuhn (2014), who was a superintendent of a school district in a high poverty area in Texas, described in detail from his perspective as an administrator how policies that relied on test-based accountability actually did more harm to his school and student body than good. Educational leaders initiating expensive mandatory testing as a means of educational accountability intentionally designed this premise to mislead the public and denigrate teachers. He clearly saw how reformers ignored poverty as a factor in low test scores: “Poverty wasn’t an
excuse; it was an ironclad guarantee” (p. 11) for failure. Based on his experience, test score gaps had more to do with poverty and racial segregation than with bad teaching: “It’s easier on all of us if we blame bad teachers for poverty instead of race-clumsy lawmakers. After all, we didn’t elect the teachers” (p. 124).

Finally, Giroux (2001) contends schools provide a dramatically different curriculum and hidden curriculum based on socioeconomic groups, with each socioeconomic group requiring a different form of education. Public schools have become a factory for the working class reproducing existing class structure. Students of vocational or general education track are seen as working-class adolescents who must be taught subordination and have very little control over the choices in their curriculum or the activities in their school, thus preparing them for service work jobs. These choices are contrary to the students from the middle to upper class suburban schools and colleges who are being prepared for white-collar work. The vast differences among public school districts based on socioeconomic status, race, and access to the arts puts teachers who work in less advantaged schools in an especially difficult situation when trying to defend the work they do when met with demands to produce data that demonstrate student growth and learning.

**Data Driven Culture**

Literature demonstrates how teachers and school administrators are being bombarded with demands to provide data and make data driven decisions about curriculum while utilizing research-based strategies. Chapman (2007) warns that this is a rigged game, causing school officials to search “frantically for those ready-to-use, scientifically based interventions for the coming year with little to no time for staff development” (p. 26). It is rigged because schools are
never at the same level and never start at the same level, so it is unexpected and unlikely they will ever finish at the same level. Ravitch (2014) supports this by stating policymakers have created a climate with a never-ending thirst for data, causing school officials to focus time and effort on data collection rather than efforts that will help students.

Kuykendall (2004) describes a historic pattern of incorrect and low expectations for schools, especially those with high minority populations who have low test scores. Quick fix commercial programs are offered to help these schools improve data; however, rarely do these programs produce high student achievement. Kuykendall cautions that educational reform efforts that present packaged curricula can never be a substitute for good teachers.

Teachers have been inundated with policies, terms, acronyms, and mandates and are being asked to use data to drive instruction. To help illustrate the complex and sometimes confusing array of information passed on to teachers with the increase of mandates, Goodwin and Webb (2014) conducted a study on teachers’ perceptions of the term research-based strategies. The researchers found this term was being so frequently overused in education that teachers no longer truly grasp the meaning. Their study found a major communication breakdown and a “disconnect between teachers’ understanding of research-based strategies and the expectations on teacher evaluation tools” (p. 6). Some legitimate concerns in the field of education from a teacher’s perspective regarding how policy related to teacher data collection and test scores could potentially promote dishonesty in data reporting as well as a reduction in educational risk-taking on the part of both teacher and student and create stressful learning environments. Stevens (2014) infers the definition of data in the current school climate is primarily numerical. Lack of professional development on how to collect performance-based data leaves teachers in the arts at a disadvantage. Issues arise in the field of art and design
education with data collection on student performance when educational policy and mandates are coupled with educational standards that do not support qualitative assessment.

**Learning Standards**

Much of the data schools and teachers must produce is how students have met learning standards. This becomes a complex terrain to navigate in the arts. Through the language in the Illinois State Board of Education platform, it is clearly seen why so many art teachers are being asked to meet inappropriate learning standards, namely the state standards that have been modeled after Common Core. Although creativity is addressed in the national and state standards, there are no models provided for its assessment. Assessing creative learning cannot be assessed in a standardized way, so art teachers who promote creativity are struggling with this problem (Boughton, 2013). Freedman (2015) described how separating creative learning from the Common Core standards promotes core knowledge that was chosen arbitrarily and narrowed the “curriculum to reading and math, and largely ignoring the creative branches… of school subjects” (pp. 47-48). Thomas (2014) and Wexler (2014) have also expressed criticism of Common Core, describing the frustration it has caused with teachers, parents and students as well as the detrimental impact it has had on the arts. Ravitch (2014) also criticizes the intentions of Common Core as a means for promoting profitable opportunities for businesses in schools because of the technology needed to implement national assessments online. “This is the first time in history that the U.S. Department of Education designed programs with the intent of stimulating private sector investors to create for-profit ventures in American education” (p. 17).

If an art teacher looks to the National Core Arts Standards for visual arts to find assistance in meeting the standards (NCAS, 2014), as extensive as they are, the teachers will find
they have been aligned to Common Core Standards (The College Board, 2014). Language in the new standards moves away from aligning with the elements and principles, thus widening the scope of curriculum to more critical and concept-based art education experiences. Adding concepts such as big ideas and essential questions seems to be a step in the right direction for the standards (Walker, 2014). At first glance, it may seem to be a relief to arts teachers, suggesting the arts have a perceived value in the educational environment for students and have a perceived worth if they have been aligned to Common Core. Upon closer inspection, when considerations must be made to demonstrate student learning and achievement, as extensive as the standards are, there are no performance indicators or benchmarks to guide high school teachers’ judgment of the levels of student performance. This sets the stage for a rather difficult problem for guiding judgement of the quality of student performance, ultimately leaving it inherently to each individual classroom teacher, school or school district. The standards have been sharply criticized by London (2016) for not being research based, too far removed from practicing art teachers’ needs, too complex to align at developmentally appropriate age levels, and for utilizing professional jargon that is too difficult to effectively communicate to non-art teachers. He cautions that the standards read as a national curriculum guide although it does not claim to be one, which is inappropriate because of the vast diversity throughout our country and the lack of equitable access to arts education.

As demonstrated in the Illinois State Board of Education (2016) platform, this can become a dangerous issue if there is inconsistent or low support from the state for the arts. The Illinois Art Education Association has released new art standards for the state of Illinois; however, they do not help teachers solve the problem of assessing performance based on standards. Rather than creating art learning standards and performance indicators or benchmarks,
they are modeled after the National Core Arts Standards. The concepts of validity or reliability are not addressed in the document, which potentially sets art teachers up for failure in their ability to determine valid assessments and establish reliability necessary for a teacher’s performance evaluation.

Teacher Observation and Evaluation Models

Among current models for teacher evaluation, the Charlotte Danielson (2007) framework is currently the most widely applied method. Applying a constructivist view of the act of teaching, the model divides the act of teaching into four domains: 1) planning and preparation, 2) classroom environment, 3) instruction, and 4) professional responsibilities. Within the four domains, the act of teaching is further divided into components and additional descriptors and indicators with rubrics that describe the levels of teacher performance. The system is designed to help improve teacher performance and student learning. While many school districts have adopted this format, only two of the domains: classroom environment and instruction are observable. The remaining two domains are not easily observed and often rely on the teacher to provide evidence; however, depending on how a school district interprets this model and how teachers are expected to provide evidence, teachers are burdened with a considerable amount of documentation and paperwork. Additionally, while there is much ancillary information available (ISBE, 2016), such as lesson plan formats and professional development, to assist both teachers and administrators with how to apply and observe the numerous indicators within each domain, the amount of professional development to implement is up to each individual school district.

Aside from student growth data, a significant component of teacher evaluation is still the classroom observation process; however, few art teachers are actually evaluated by an
Eisner (1996) discusses teacher evaluation as a means of providing useful feedback on the quality of their work in detail. Because teaching is mainly an isolated profession, it is important to get useful feedback to improve classroom performance. Eisner relates this to his idea of connoisseurship; however, it is difficult to get useful feedback to improve the quality of teachers’ work from someone who does not understand the nuances of teaching visual art. As Eisner stated, “It is difficult, even impossible to make adequate judgments about the intellectual merits of content that one does not understand” (p. 92). Eisner (1998) described the evaluation of educational practice with a qualitative lens, and while much of this is generalizable to the field of education as a whole, there are specifics to art education practice that are often left out of other teacher evaluation methods. In his argument for applying qualitative methods, Eisner further supported his argument for connoisseurship and explains how this is aimed at broadening the lens of understanding what is going on as well as the context of teaching and learning.

Regardless of teacher evaluation methods that promote qualitative judgments of the effectiveness of a teacher, policy requires the inclusion of data related to student test performance, although this practice has been refuted as an accurate measure of teaching and school success (Popham, 1999). The book, Linking Teacher Evaluation and Student Learning by Tucker and Stronge (2005), argues for high stakes testing as part of the teacher evaluation process. Rather than considering any research that disavows this practice, Tucker and Stronge stated that “today, standardized testing has become a political reality in mandated programs that exist in almost every state… Despite the concerns and criticisms that often are leveled at today’s high-stakes testing, it is a reality” (p. 17). They further ignore any evidence of visual art-related learning as part of student evidence of learning. The section labeled “Other Types of Student
Assessments” offers a generic description of examples of authentic assessments, including “writing samples, student portfolio entries, and other performance-based assessments” (p. 19). The text does not provide any examples of these types of assessments or possible rubrics for performance-based learning. For further details about student portfolios, the reader is directed to the book’s Appendix, which provides the following information:

Portfolio: Collection of artifacts and running records of performance that can be used for reading, math skills, and writing skills. Exhibitions: Work products that are judged by a teacher or panel of experts. Performances: Demonstrations of knowledge and/or skill in a ‘natural’ manner. (p. 112)

According to this text, portfolios are only for reading, math and writing skills and there is no specific connection to or further description of the judgment of performance-based learning or information on who the panel of experts would be or how a teacher could seek out these resources. This non-specific language does not provide usable information to help any teacher of a non-tested discipline prepare for their evaluation. This brings into question the reliability and validity of teacher performance evaluations conducted by administrators not adequately trained in the academic domain they are evaluating. Without clear direction on performance-based alternate assessment tools, art teachers are left defenseless if they do not have a strong background in assessment.

Measuring Learning in High School and Secondary Schools with Alternate Assessments in Visual Art and Design

Secondary and high school teachers are under great pressure to demonstrate student success through assessment because of the current national trend for college and career readiness. The field of art and design education lacks a history of testing, and most teachers have utilized teacher-constructed forms of assessment to determine student learning and success
Historically the field of art and design education started addressing the problem of how the arts could be assessed in 1967, and the challenges associated with assessing art on a national scale in a consistent and cost-effective manner have been well documented (Wilson, 1970; 1975). Some of the initial challenges Wilson described when the field started assessing the arts on a national level were the lack of cost-effective ways to score open-ended responses from students, assessments were time consuming, and while the information from large scale assessments may have been useful at the national, state, and local levels, it was difficult for classroom teachers to see any utility with the process. Wilson (1992) continued to track the problems and internal debates in the field surrounding the idea of assessment stating:

Comprehensive assessment and evaluation programs will not be developed until they are required, until they are adequately funded, until arts educators are provided with the enormous amount of time needed to develop assessment programs, and until individuals acquire the expertise they need to establish comprehensive assessment programs. (p. 34)

Because there are no formal agreed upon standards for conducting school art assessments in the United States, art and design teachers must now become more cognizant of assessment issues as Dorn, Madeja and Sabol (2004) explain:

Art teachers must now become experts in another field of knowledge, … They must have in-depth knowledge of assessment’s terminology, methods, and processes. They must be able to interpret assessment data and communicate their meaning to students, colleagues, administrators, school boards, and other stakeholders. Moreover, as the field of assessment continues to expand, and questions and issues about its use in art education arise, teachers must keep their knowledge current. (p. 16)

Educational evaluation research methods have been greatly influenced by scientific methods of prediction and control for the purposes of simple quantification (Boughton, 1996). However, qualitative methods are necessary to tell the rest of the story. When art teachers become astute in the language of assessment, they can appropriately advocate for assessments that measure what their students are learning in the art classroom.
The field of art and design education and other performance-based disciplines have sought alternative assessments to accurately measure student learning in response to the educational reform movements. When developing alternate assessment methods, Ryan (2006) recognizes the valuable input of teachers: “Teachers are in the best position to assess, interpret, document, and report, due to their personal contact and knowledge of each student throughout an entire school year” (p. 102). However, Ryan concedes the professional experience of educators is often overlooked by most test developers. This argument remains true in the art and design classroom. What makes authentic assessment in visual art and design different than assessment and other educational fields is the allowance for unexpected outcomes, multiple interpretations, and the possibility of creativity. Boughton (2004a) noted, “When we think about assessing the arts, the words ‘standardized’ and ‘art’ do not sit comfortably together in the same sentence” (p. 267).

Another challenge for many high school art teachers is the requirement of regularly posting grades. If districts require teachers to grade students on a regular basis as an indicator of progress, it can be challenging to quantify student progress in a meaningful way. Evidence of learning and progress may look different depending on the level of course in which a student is enrolled as well as their previous experience in visual art and design. Not all students in the United States are required to take art classes and not all will submit a final portfolio at the end of their high school career, so clearly there is a distinction between grades for projects that indicate completion and meeting learning objectives and a portfolio submitted for summative assessment. Boughton (2013) raises a concern for students who are in the formative process of art learning and cautions that a sure way to destroy possibilities of creativity is to always link the completion
of an assignment with an immediate grade, depleting the possibility for students to collect work over time and select work that is representative of their learning process.

Innovation and creativity cannot come from a test and drill mentality. Testing should only be used diagnostically rather than to drive the structure of a school. Excessive testing actually limits students’ opportunities for creativity and innovation. Ravitch (2014) raises an important question: “Do we want to honor and reward only those teachers who excel at teaching to the test? Or do we want to honor those teachers who are best at getting their students to think and ask good questions?” (p. 111). When teaching is limited to testing, it narrows the curriculum, thus limiting any creative potential for students. Taylor, Carpenter, Ballengee Morris, and Sessions (2006) explain one of the complexities art teachers face is that teaching art is essentially media driven; however, in addition to the many duties related to teaching art, they are also expected to assist in delivering content related to high-stakes tests. Because of administrative requests, many art teachers feel their program is not viewed as an equal to other academic subjects. Regardless, the important issue remains; in the climate of high-stakes testing, it is important for high school art teachers to find ways to report student growth to administrators through assessment. Taylor et al. describe the importance of assessment as “the process of placing a value on and judging what students have done, made, said, or crated. Simply, assessment explains what has been done and evaluation describes the degree to which it was accomplished” (p. 133). Drawing from the work of John Dewey, Taylor et al. describe the high school art classroom as an experience that should create connections of experiences. Students must see how to apply what they are learning.

Learning in the visual arts requires students to create visual artifacts that demonstrate complex thought processes (Eisner, 2002). To complement the performance-based learning that takes place in the arts, research strongly advocates the measurement of student performance with
a variety of performance-based and qualitative assessment methods (Boughton, 2004a; 2004b; Dorn & Sabol, 2006, Eisner, 2002). This can include, but is not limited to, formative assessments like critiques and student-selected artifacts of learning that can include sketchbooks, visual research journals, and portfolio presentations. Beattie (1997) describes different strategies art teachers use for critique, including informal non-threatening diagnostic strategies to help identify students’ strengths and weaknesses and more formal methods as part of a summative evaluation.

There are programs available for American secondary and high schools that specifically address adolescent learning in the arts that foster imagination, promote creativity, and risk-taking. Such programs include the Advanced Placement (AP) Studio Art program and the International Baccalaureate (IB) Studio Art program (Filbin, 2019). Both of these programs are well established in literature, and while each program is unique, they similarly utilize an examination method of scoring student portfolios by professionals who are not the students’ classroom teachers. This examination process demonstrates reliability and is respected by higher education institutions. Although these programs are highly respected, they are not available in all schools in the United States, demonstrating a lack of equity for all American students.

The assessment process in IB high schools in which students in visual arts culminate their school experience with a portfolio examination is advocated by Boughton (2004b). Through this process, students select the works to be considered, provide a written statement about their learning process, and provide documentation of their learning in their research workbook. The examiner’s use of visual benchmarks that are established prior to grading is based on standards set by previous students’ work. Boughton (2013) recommends this type of holistic summative assessment because of its validity to the learning that takes place in the visual arts and how the moderation process sets a standard of reliability.
Similarly, the College Board’s AP (2015) program, which is open to American public and private high schools who choose to participate, also conducts its exam through submission of a portfolio. Similar to the IB exam, students select and submit their portfolio to external graders and provide a written statement about their work. The readers who score the AP portfolios are trained in visual benchmarks determined by previous years’ student’s work. One of the differences between the IB and AP exam is students taking the AP exam are not able to provide additional documentation like a sketchbook or workbook. Similarly, the high-stakes examinations for the visual arts in AP or IB do not use any multiple-choice or question/answer format to grade student work; instead they use a visual benchmarking process (Filbin, 2019). The adolescents must visually demonstrate what they have learned as an authentic means of assessment.

Programs like AP and IB have a summative exit exam at the end of secondary or high school, which means only students who choose to participate in these programs will submit a portfolio exam. The issue of classroom teachers addressing day-to-day classroom assessment throughout the school year needs to be considered. MacGregor (1992) explains the pressure to add testing to the arts is not new and has been problematic for decades. It is not problematic because of the expectation to test but because of the means or methods used to test. The issue becomes a question of the quality of the test. Paper and pencil tests are generally promoted; however, they are not the best method to test the arts. Many other countries have practiced assessment of art in a way that demonstrates reliability; however, in the United States teachers are often asked to employ methods that are not always the one best for the visual arts. Teachers should look beyond the scope of what is offered in the United States. To support this, Boughton (2018) has provided many models of authentic assessment in visual art and design that do not
rely on standardized exams, but are regarded as valid and reliable, from many other countries. Beattie (1992) also provided an example of a Dutch model; however, for her example to be implemented in the United States, students would need to have foundational art experiences, which is inconsistent throughout the nation. Also because the United States is not a homogeneous population and higher education teacher preparation programs differ, consistency of implementation may be an issue. Beattie recommends accommodations to make this a viable option.

Another alternative for art assessment is provided by Davis-Soylu, Pepplar, and Hickey (2011) in response to the mandates of NCLB. They propose an Assessment Staging Theory that creates art assessments at three levels: an assessment used by classroom teachers and students, an assessment implemented at school and district levels, and a larger scale assessment conducted for government agencies and researchers interested in trends in the arts. They propose this theory by making the argument that portfolio assessments will not work; however, they do not address any empirical evidence or existing literature about AP and IB portfolio examinations that have long established credibility for successfully conducting large-scale portfolio assessments. Furthermore, they do not address details about how this theory would work in schools on a large scale as an additional responsibility on the part of the classroom art teacher who would need to implement this or the funding that would be required to back a three-tiered program. Regardless of the model of art assessment, an integral component to assessment is that it needs to be both valid and reliable.
Validity and Reliability of Art Assessment Practices

Unlike other academic fields, the field of art education does not have a long history of testing, and because of this, teachers have often had to create their own assessment techniques to determine student learning. Beattie (1997) has asserted assessment should be integrated into instruction. There are many techniques art teachers use to determine student understanding that are informal formative assessments; however, “summative assessments are more formal than formative assessments; they require more preparation, cover more content, [and] demand that issues of validity and reliability be addressed” (p. 104).

This leads to the problem faced in quantifying the qualities present in student work and methods to ensure the validity and reliability of assessment strategies. Boughton (1996) explains the components of an effective assessment model: “Validity (do the methods of assessment reflect the aims and objectives of the course?); reliability (does the methodology produce similar results for similar students under similar conditions?); and utility (is the assessment process convenient and cost-effective?)” (p. 299). Formatively assessing the progress of student artwork and assessing student learning in art through a portfolio assessment are two examples of valid forms of assessment, since they reflect the aims of the course. However, if the classroom teacher is alone in the responsibility for grading, this is a threat to reliability. The process of utilizing visual benchmarks and the cooperation of multiple teachers or examiners making judgments on student portfolio work based on the benchmarks can create the conditions for reliability. However, this is often not the case (Boughton, 2018). The possibilities for reliability in art assessment in the public schools are present, but the reality is public schools that do not have an
IB program or do not have students participating in the AP Studio Art exam have little to no opportunity for reliability in general art and design curriculum courses.

One of the greatest threats to reliability in complying with PERA is the lack of time or attention allotted to art educators for professional development or professional time allotted specifically to confer with one another on their local assessments. Allocating time for teachers or multiple judges to make holistic assessments of their colleagues’ student portfolios or to create visual benchmarks or visual learning standards even within districts is limited at best (Boughton, 2018). If art teachers are indeed conducting portfolio assessments with their students, these may be valid assessments, but reliability is not established if art teachers are left to fend for themselves. The possibility exists if art teachers in the same content area, within the same school district are allowed to work together to determine visual benchmarks for art learning to assess their students’ work. However, if this professional development time is not available or in cases when an art teacher is working alone in a district or region, there is no realistic capability to establish any reliability. Concurring with Boughton regarding threats to reliability, Beattie (1997) explains that “reliability can be defined as the consistency of assessment scores…” Interestingly, an assessment can be reliable, that is, repeated again and again with the same score results, without being valid - but an assessment can never be valid without being reliable” (p. 127). This highlights the intrinsic problem visual art and design teachers face: “unreliable results are never valid” (p. 128). In addition to validity and reliability, visual art assessments for high school students should be appropriate to the developmental stages of adolescence, so it is critical to understand the period of development that leads to adulthood.
High school students are at an important time in their social and emotional development, known as adolescence. They are at a time in cognitive development when they are ready to construct and deconstruct visual and verbal metaphors (Davis, 2012). Engaging in the arts is an excellent way for many students to demonstrate such a level of development. While standardized tests measure right/wrong answers, the arts can help make sense of the answers that fall somewhere in the gray areas. Davis describes how much of adolescence is spent in the gray areas of trying to make sense of the parameters put in place by adults, so exploring in the arts helps high school students find clarity. There is some debate among art education professionals about using right/wrong testing in the arts (Brewer, 2008; Davis, 2012; Siegesmund, Diket, & McCulloch, 2001); however, Davis (2012) contends right or wrong is not the point of what should be achieved in the arts. Instead it is more about making the connection between art learning and characteristics such as independent critical thinking skills and developing ownership of one’s own learning. Freedman’s (2003) extensive research in visual culture supports the idea that critical thinking is involved when processing the myriad of visual images adolescents encounter in daily life.

Researchers have identified this developmental period as a turbulent time, marked by feelings of conflict, crisis, and the desire to test their surroundings (Csikszentmihalyi & Larson, 1984; Davis, 2012; Graham, 2003; Michael, 1983; Watson & Gable, 2013). Csikszentmihalyi and Larson (1984) explain the mood extremes most associated with adolescence include frustration, anxiety, sadness or feelings of hopelessness, sharply contrasted with positive feelings, happiness and positive self-image. These extremes are a necessary part of adolescence.
These contradicting emotions are the result of conflicting messages young people navigate when they are asked to act like mature adults; however, they are still viewed as children without the same privileges as adults. They are often treated like children at school and home. These feelings can be overwhelming as they realize their impending transition away from home (Csikszentmihalyi & Larson, 1984; Michael, 1983). Watson and Gable (2013) describe how adolescents with learning or emotional disabilities can experience exacerbated feelings of pressure in school from academic and social environments. Frustration levels are magnified in academic classes that require constant repetition of question/answer practices and long sequences of directions. They provide suggestions for educators that include visual cues, visual modeling, creation of visual images to help students remember information and allow choices on assignments – all strategies supported in art learning environments.

Artistic expression can be nurtured during this developmental time to help adolescents learn to convey meaning. Artistic activities can be ways of expressing rebellion that can potentially help students deal with emotional issues or have a cathartic effect because these activities require concentration and can be motivating. These forms of artistic expression can have a positive effect on development while also being intrinsically rewarding and a socially acceptable way for adolescents to let off steam while aiding with emotional stability (Csikszentmihalyi and Larson, 1984; Davis, 2012; Michael, 1983; Watson & Gable, 2013). Davis (2012) recognizes the art room as a place where adolescents can also have meaningful contact with their peers and adults that can also positively aid in their social development. She describes how many art educators may regard their students as colleagues or collaborators. This also models positive adult interactions by showing students possible applications for what they are learning. Unfortunately, the arts are not always a part of course offerings for all high school
students. Regardless of this lack of equity, Freedman (2015) describes ways many students today resolve the lack of formal art experiences; they find artistic outlets and meaningful contact with others through online creative outlets or other online forums to create a community to share and/or create a live production of music, dance, graffiti, or self-generated visual art forms. This can fill the void of not having a formal class setting to connect knowledge with emotion and can help fulfill the human need for creative production.

Graham (2003) noted that “adolescent artistic developmental potential is often underestimated, ignored, or dismissed as irrelevant” (p. 162), so it is not always incorporated as required course offerings for high school students. As a result, students’ school schedules consist only of core classes that are routinely assessed through standardized testing. The lack of an art requirement and increase in requirements for students to prepare for classes that are tested exacerbates the limitations some adolescents experience with access to arts adding to the perception that art is not a worthy academic subject. Davis (2012) believes art has been devalued because the outcomes are not easily quantifiable. She describes the stress associated with high-stakes testing and how the arts can help students through emotional times. She values the potential of a creative outlet and opportunities for creative problem solving for high school students. To understand the potential of high school student work, a closer examination of creativity is necessary.

Consideration of Creativity

Freedman (2010) explains that including creativity in the art curriculum can be a delicate balance of encouraging students to produce thoughtful work and helping them learn skills to improve the quality of their work. She argues that one purpose of high school education is to
prepare students for their future as adults, so creativity should be part of the curriculum in a democratic society. Including creativity helps prepare students for a future that will need professionals who can think creatively for the development of the economy. However, this type of learning must be cultivated because true creative thinking on the part of the student can extend far beyond the classroom walls and often involves auto-didactic thinking or independent learning and problem solving. While this can present a problem for quantifying this type of learning, flexibility on the part of the art teacher that recognizes and honors student learning that takes forms in unpredictable or unexpected ways helps prepare students for a creative workforce. In response to concerns raised from the economically driven workforce, Freedman (2010, 2015) discusses the demands for creativity from multiple economic sectors, advocating for the many opportunities art and design education can offer. She contrasts the real situation in schools, describing limited opportunities for creativity and evidence that current schooling seemingly does not value the levels of learning involved in creativity.

In a report from Lichtenberg, Woock, and Wright (2008), creativity was cited as a necessary 21st century skill. They recommended arts are a way that creativity can be incorporated into the curriculum as a required course. Katz-Buonincontro (2018) concurs that educational policy is becoming interested in creativity. She describes how fostering creativity in the art class provides students with an opportunity to imagine and question through art making, requiring teachers to expand the art curriculum beyond technical application of art materials. However, for art teachers to incorporate this and assess it well, teacher training and professional development are needed. Similarly, Noonoo (2018) suggests that the conditions for creative exploration space and time need to be provided along with allowing the time and space for possible failure, but this rarely happens in schools because of the structure of time and schedule in school. Time is the
biggest enemy of creativity. When understanding what creativity should look like in the schools, students need time to ask questions, struggle with getting to the answers, and take risks. This requires a different kind of assessment that allows students to be a part of the process that takes time and nurturing.

Creativity does not need to be limited to the art room. The book, *The Rise of the Creative Class* (Florida, 2012), describes the positive contributions creative people make to the United States workforce and economy and how this creativity needs to be nurtured. Florida does not limit his definition of creativity exclusively to the arts, but he expands his definition “of the Creative Class to include people in science and engineering, architecture and design, education, arts, music, and entertainment” (p. 8). He essentially includes anyone who engages in creative problem solving that requires higher levels of thought and can contribute to human capital. His definition of the creative class is not unique; it is similar to the way Csikszentmihalyi (1996) defined creativity as an act, product, or idea that creates change and is viewed as useful. Florida (2012) applies this definition to contemporary society. Csikszentmihalyi (1996) found multiple instances in which creative people struggled in their traditional schooling. While he found no one way to teach creativity, he did cite many ways that creativity had been fostered in the individuals he interviewed. He explains how creativity is necessary to economic development. While there is no way to teach someone to be creative, there are clearly conditions that can allow possibilities for creativity, and similar to Noonoo (2018), Csikszentmihalyi (1996) describes how creativity can be nurtured, which takes time. He describes this as flow, or a sense of not caring about much else other than the creative pursuit in which one is interested. This is not something that is conducive to a traditional school setting. A traditional predetermined curriculum limits opportunity for critical thinking and creativity, so it is essential to resist standardized types of
learning to foster conditions for more independent thought. Understanding forms of resistance and ways it can manifest is essential to understanding teachers’ reactions to school policy.

Forms of Resistance

Ravitch (2014) contended that American public education is not as broken as policy makers would lead us to believe nor can the existing problems be completely blamed on teachers and administrators. Instead she argued public education reform efforts are a “deliberate effort to replace public education with a privately managed, free-market system of schooling” (p. 4). Giroux (2001) supported this argument when he asserted that teachers and schools are under siege unlike ever before; corporations now define schools and compares pre-made curriculum (like Common Core) and use of computerized tools for teaching as a way of silencing resistance in teachers. He contends that the “objectives, knowledge, skills, pedagogical practices, and modes of evaluation are built into and predefined by the curriculum program itself. The teacher’s role is reduced to merely following the rules” (p. 71). He places responsibility on teachers as a starting point for resistance to this model of schooling since students have little to no control over their educational experiences in schools. Students need to learn to question and need to be taught how to think critically. Giroux warned, however, that the promotion of criticality and resistance is not an easy task. There is no easy solution or easy formula for resistance, but teachers, including those in higher education, are responsible for teaching criticality through a radical pedagogy and teaching resistance. Progressive educators should actively reject marginalization and find ways to make school inclusive versus exclusionary.

Since teachers are the ones who must manage overcrowded classes, work with limited resources, and are the target of hostile legislation, Giroux (2001) places responsibility on
teachers as a starting point for resistance to this model of schooling since students have little to no control over their educational experiences in schools. Teachers need to be the models for resistance so students can learn to be active participants in their own learning. They need to learn to question and need to be taught how to think critically. However, the promotion of criticality and resistance is not an easy task; some view it as much easier to accept their position than to mobilize and change position. Giroux (1998) explained, “Ironically, emancipatory forms of knowledge may be refused by those who could most benefit from such knowledge…the oppressed may take the form of actively resisting forms of knowledge that pose a challenge to their worldview” (p. 49). For practicing teachers at various points in their career, participating in overt resistance can be risky.

Taking a more realistic view of resistance, sociologist Erving Goffman (1959) uses a dramaturgical metaphor to explain self and daily behaviors in comparisons to theatrical performances. We are performers, and in this analogy, he illustrates how we change our roles depending on the situation we are in or who is in one’s presence. When we are around someone who is intentionally observing our behavior, they are considered our audience. These roles we play become adaptations of self-defense mechanisms that help us through our day to day workings. This can be best exemplified in the relationship of teachers and administrators conducting observations and the way teachers may change their performance when they know they are being observed rather than the way they normally perform when the teacher is in a typical class routine with students. More specific applications of times teachers would change the role they play with is in the current climate of standardization and pressures to conform to aligning art and design curriculum to Common Core and collect data that demonstrate improvement in school wide test scores. Simply complying with these kinds of requests goes
against best practice in the field of art education and is counterintuitive to the goals of a classroom curriculum that promotes creativity and educational risk-taking. However, flat out refusal or overt resistance could result in disciplinary action or even termination of employment. Thus, Goffman describes a harmless form of deception or pragmatic assimilation to keep harmony in social workings, concealing one’s own real opinions or thoughts as a form of courtesy to those around them. Goffman’s form of resistance manifests in a more socially acceptable way. Keeping decorum as polite behavior is a means of self-preservation; the survival mechanism that helps people adapt to the situations they are placed in, so they can survive in the roles that they must play. Goffman suggests a more covert form of resistance that is non-confrontational for pragmatic reasons.

Michael Schwalbe (1993) explains the levels within the roles Goffman (1959) describes as forms of deception and how they become a means of self-preservation. This is sometimes necessary in the workplace, especially if asked by superiors to do things that go against what is in an individual’s best interest or in the best interest of the field. This applies most directly to the expectations placed on art and design educators in the current climate of standardization and pressures to align the curriculum to Common Core and provide evidence of collecting data that demonstrate improvement in school-wide test scores. Complying with requests to assess art learning with multiple-choice or promoting reading and math activities over the art curriculum goes against best practice in the field, but the reality for some teachers is choosing to comply with the request. In this case, Schwalbe describes how a covert form of resistance takes place:

We may need to conceal our true feelings to get a job done, to get along, to avoid being punished, or to save our skins… Under these conditions, the self has become so chameleon-like that, even if it is still there, it is hard to see, since it has been well-programmed to blend in with its surroundings. Instead of being visible in its resistance, it is now invisible in its complicity. (p. 341)
Within the deception is also an attempt to communicate authenticity to attain a practical goal of actually achieving what you want and generally protecting the feelings of others. The arts seem to be a great liberator and promoter of resistance. It makes sense then that mindful art educators actively resist educational policies that stand in the way of promoting creativity and critical thinking. The field of art education needs to remain vigilant in resisting standardized forms of assessment and promoting the kinds of assessments that value careful and attentive judgment of authenticity in thought, risk-taking, and creativity. There is a delicate balance between overt resistance and covert resistance within the lens of critical pedagogy.

Analyzing the social mechanisms that fuel forms of resistance more deeply, Hall looks more critically at the ranking of cultures and the positioning of dominant and subordinate cultures and its manifestations of domination and subordination (as cited in Clarke, Hall, Jefferson & Roberts, 2006).

This does not mean that there is only one set of ideas or cultural forms in a society. There will be more than one tendency at work within the dominant ideas of a society. Groups or classes which do not stand at the apex of power, nevertheless, find ways of expressing and realizing in their culture their subordinate position and experiences. (p. 5)

This means there is a negotiation of space among subordinate cultures and dominant culture; a coexistence. Comparing these negotiations to hegemonic control over the subordinate classes, power shapes experiences and creates consent on the part of the subordinate class as a way of granting legitimacy to the ruling class. Hall explains how this appearance of normalcy prevents conflicts:

It works *primarily* by inserting the subordinate class into key institutions and structures that support the power and social authority of the dominant order… Often, this subordination is secured only because the dominant order succeeds in weakening, destroying, displacing or incorporating alternative institutions of defense and resistance thrown up by the subordinate class. (p. 29)
Within the context of the push for standardization in the schools, the subordinate class is represented by teachers. Policy is designed to control the teachers, like accountability for school wide test scores in reading and math that are out of the control of an art teacher but make the teachers believe it is in their control because it becomes part of their professional responsibility. This creates a relationship of need/control/submission.

Building on this portrayal of dominant and subordinate cultures, Hall (1998) deconstructs popular culture identifying popular as the masses being controlled by the dominant culture that controls media. He associates this with manipulating and debasement of the people based on an assumption that people will blindly consume because they are passive cultural dopes. He sees this as a necessary manipulation to propagate the dominant culture. Hall, however, does not see ordinary people as “cultural dopes” because most people can recognize what is happening, so industries and those who control the dominant culture are constantly re-shaping and re-working popular culture to re-package it to make it popular. As the people resist, there is a response of suppression, creating a constant struggle with no victories only strategic positions. Relating this cycle to the promotion of educational policies and trends in promoting the next new thing in education is not new. It seems there is always a re-packaging of ideas, as seen in the pendulum swings in trends in education. The questions become: “Who is promoting?” “Who is believing and accepting?” “Who is resisting?” “And how do teachers push back or resist the commercialization of education?” In this context of resistance, the cultural power of educational institutions, “along with the many positive things they do, also help[s] to discipline and police this boundary” (p. 450). This reaches far beyond the discipline of the students; in this context, it moves toward policing the teachers and creating boundaries around the kind of education that is
promoted. This is best illustrated in the current practice of disciplining teachers, schools, and sometimes entire districts for failure to meet set standards, or achieve success in prescribed programs – for example, the state of Illinois’ policy for teacher evaluation includes standardized test data as part of a teacher’s performance evaluation.

Hall (1998) notes the cycle of resistance and domination and the necessary dependence one has on the other, so there is exertion of discipline and punish within the schools to maintain dominance. This context of the school’s position of dominance in this cycle is well illustrated in the writing of Michel Foucault (1977):

> It must also master all the forces that are formed from the very constitution of an organized multiplicity: it must neutralize the effects of counter-power that spring from them and which form a resistance to the power that wishes to dominate it: agitations, revolts, spontaneous organizations, coalitions. (p. 219)

Pickett (1996) describes this resistance that fights against larger political struggles – an empowerment for those who are constrained. Because of Foucault’s (1977) lack of limitations on resistance, it is not the students resisting but the school personnel for refusing to be the transmitters of mediocrity and low-level knowledge. Instead they are promoting thinking critically and questioning the ways education are promoted. Knowledge becomes power that fuels resistance and resistance threatens power, thus creating the cycle of power and resistance. This kind of resistance is essential in an art class that promotes critical thinking and risk-taking. These skills cannot be fostered if a teacher clings to prescribed and predicable outcomes that do not challenge thinking and authentic learning.

Overview of Related Studies
There are some recent studies in the field of art and design education that directly relate to issues of assessment and educational policy, including those that specifically researched assessment methods and the effects of policy on K-12 education. One study included school districts in the state of Florida and Illinois from 51 classrooms in 15 school districts (Dorn, 2003). The study examined portfolio assessment, the reliability of that form of assessment, and whether student performance could be quantified. The findings showed support of the claim that portfolio assessment can quantify expressive behaviors in students. There was a high level of reliability between raters’ scores for student portfolios through a moderation process. Authentic assessment supported in student portfolios “are not designed to compare teachers and schools with one another but rather to assess student progress within a given classroom as a guide to improving instruction” (p. 367). Dorn states the findings suggest school environments and teacher abilities are unequally distributed throughout the United States, so using performance evaluations to compare students from different school districts is inappropriate. Dorn’s finding is important to the context of my study, since Dorn suggests that if teachers are appropriately trained in portfolio assessment, there is a viable alternative to paper tests in the field of art and design.

In a study conducted by Sabol (2010), a 55-item questionnaire was given to over 3,000 art and design educators nationwide regarding the effects of the No Child Left Behind Act on their staffing, teaching load, enrollment in art programs, and funding among other related issues. He also asked about the perceived positive and negative effects of NCLB on programming. The findings showed staffing in art education programs decreased 25%, teaching load in number of courses taught had increased 26%, 58% of the art teachers felt their workload increased, and 21% of teachers reported enrollment in their programs decreased. Sabol’s study addressed
concerns, including art education advocacy, the implementation of creativity in schools, creation of Common Core standards, new national standards in the arts, and arts achievement with at risk youth. Specific to this study, Sabol addressed concerns about the need for student assessment and assessment practices appropriate to visual art education, and teacher evaluation practices. He cites concern for the field and recommends appropriate professional development for teachers in the area of visual art and design, especially for assessment.

Jackson Goodwin (2015) conducted a qualitative study about how art teachers from multiple levels gathered information from assessments related to local accountability requirements in Colorado. She identified “themes, commonalties, and areas of divergence that present a way for art teachers to describe and justify how, what, and why they assess student learning and provide insight into the world of visual art assessment in the age of accountability” (p. ii). In describing the need for her study, she explained that art teachers have always assessed student learning; however, they are not always taught the language of assessment in a way that helps them describe their assessment practices to others outside the field to help them defend their practices. Similar to Illinois’ requirements for teachers to demonstrate student growth, Jackson Goodwin describes how Colorado teachers are required to demonstrate student growth as a part of their teacher performance evaluation; however, many of the art teachers do not know how to do this and most of the administrators are not able to help their art teachers either. She identifies underlying problems: there is no standard curriculum for art education in the United States and not every school in America has an art department. If a school does have an art department, not all students take art classes. She cited the lack of equitable access to visual art education and how this leads to a problem of equitable assessment. For the study, Jackson Goodwin collected her data by interviewing six teachers from five different school districts.
ranging from elementary to high school levels. She also interviewed their principals and parents. She asserted all teachers were using assessment methods recommended for art and design education practice; however, some teachers found it very stressful and time consuming to adapt requirements to provide student growth data. Some struggled with aligning their assessment practices to what was asked of them by administrators but considered unrelated to art and design education. Some of the teachers in this study were in districts that required all teachers, including art teachers, to include testing data in math and language arts as part of their student growth component for their evaluation, which they felt took time away from art making. Some teachers with more years of experience in the study did not express as much stress about these mandates, since they had seen different reforms come and go. Possible implications for this study included developing educational literature specifically for the field of art education that could assist teachers, a need to add professional development for art teachers in assessment, and a need for teachers to promote advocacy to demonstrate what is done in school art programs.

Recently Hanawalt (2018) conducted a study of why new teachers entering the profession continue to promote formalism, teacher directed projects and more traditional art education curriculum rather than the methods they were taught in their university education program that included social justice, contemporary issues, big ideas, visual culture, material culture, digital culture, queer theory, community and place, and learner-centered curriculum. She found the teachers in her study reported an overwhelming repressive nature in public schools due to the test-based accountability environment and felt art was treated as a frill rather than a serious subject. Thus, the teachers’ art curricula were less about what they learned in their university courses and “more a manifestation of an educational climate focused on data, standardization, and accountability” (p. 93). Factors that influenced curricular decisions for the art teachers were
teacher accountability measures put in place with current teacher evaluation models that 
emphasize student growth and the pressure to collect data to demonstrate student growth coupled 
with evaluations plans that included school wide test score data from subjects unrelated to art. As 
new non-tenured teachers learning to navigate the evaluation system and learning to please 
administrators, compounded with the lack of resources or models available in the schools to 
assist teachers on how to demonstrate Student Learning Objective (SLO) measures, they often 
implemented some kind of multiple choice or recall type of assessment because it was the easiest 
way to collect data, document, and convey results to evaluators. Compliance and following the 
path of least resistance was a way for many new teachers to survive their first few years in a new 
school and a new profession even if it was not fulfilling what they knew was best for the field. In 
her conclusion, Hanwalt suggests:

There is significant work to be done to provoke systemic change with regard to the audit 
culture of public schools. Resistance within existing forms of institutional governance 
may remain merely reactive… While structural change may seem out of reach, without 
pursuits for change, existing cultures of accountability and compliance may continue to 
function as blockages to new, provocative visions for art education. (p. 101)

In these studies, there is an indication that teachers in the field of art and design education 
need further professional development and training in assessment methods to help prepare them 
for the reality of having to relay assessment data about student learning to administrators in a 
way that is clearly understood. The studies by Sabol (2013), Jackson Goodwin (2015), and 
Hanawalt (2018) suggest national educational policies have had a negative effect on art and 
design programs and the pedagogical expectations of art and design teachers. The studies by 
Jackson Goodwin (2015) and Hanawalt (2018) further suggest that Illinois art teachers are not 
isolated in the problem of having to address standards inappropriate to art and design learning 
but are part of a nationwide trend responding to demands to narrow the curriculum to reading
and math. Hanawalt’s study suggests the climate of test-based accountability has trivialized the importance of art and design education as an academic subject that can promote critical thinking skills, and the practicing art teachers in her study felt the arts were not viewed as equal to other subjects. Her study further suggests the risk newer teachers perceive concerning resistance. Because of art teachers’ inconsistent knowledge of assessment practices as discussed in the studies by Dorn (2003) and Jackson Goodwin (2015) and the difficulty many teachers in the United States have in attaining reliability, the field is in need of further research into educational policy and assessment practices in visual art and design education.

Application to the Study

Because this study examined educational policy, it was necessary to review literature related to policy at state and national levels and to understand the overall research problem. Literature that explained the effects of a culture of standardized testing in the United States revealed possible reasons some teachers implemented certain assessment techniques or why they were asked to utilize assessment techniques inappropriate to art and design education by administrators. Manifestations of job stress were revealed in the qualitative data, which were also a response to standardized testing. Depending on the socio-economic status of a participant’s school or the ethnic makeup of the student body, some participants described how they experienced feelings of marginalization at their school or their students of underserved populations experienced marginalizing effects of educational policy. To help interpret findings about other requirements that were part of the participants’ evaluation process, literature on data driven culture, learning standards and requirements of the teacher observation, and the evaluation process was necessary.
Because the participants for this study were limited to secondary or high school visual art and design teachers, it was important to understand the literature that explained best practice for assessing art learning with students at this developmental level. Understanding the characteristics of adolescents was also necessary when the participants discussed the learning artifacts of their students, how they were assessed, and discussions of creativity and creative outcomes of student work. The literature on validity and reliability of art assessment was also necessary to understand and interpret the data from participants when they were responding to questions related to these necessary components of assessment.

Finally, because this study is framed in critical research, the literature that explained multiple forms of resistance, the different ways resistance manifests, and the reasons one would choose to resist or conform was necessary to interpret the qualitative data from participants. Each of the interview participants was at a different level of experience and was in a unique school setting, the manifestations of resistance were individual to each participant.
CHAPTER 3

METHODOLOGY

The purpose of this study was to examine how Illinois high school visual art and design teachers fulfilled requirements for assessment. In the context of a nationalized shift toward standardization in education, Illinois teachers are required to produce data on student growth from valid and reliable assessments as a required component of their teacher’s performance evaluation. Because standardized assessments are inappropriate for judging the quality of student learning in art and design and high school teachers may employ alternate forms of assessment, a mixed-methods design was appropriate for this study. Quantitative data were useful to analyze the kinds of assessments high school visual art and design teachers were using in their classrooms and whether they differed or complied with directives they received from their evaluating administrators. Qualitative analysis was helpful for interpreting why there may have been discrepancies with answers in the quantitative data, the types of assessment tools used, and whether these were helpful to teachers in art education. This chapter includes an explanation of the research questions and the following sections: A description of the research design, how the quantitative data were collected, how the qualitative data were collected, the positionality of the researcher, how the data were analyzed, and the criteria for selecting participants.

The research questions for this study were developed to examine multiple layers of a problem related to assessment in the visual art and design classroom under the current context of a shift toward standardized assessment. To design research questions to gather data on this
problem it was important to contextualize the conditions in which teachers make decisions about their classroom practice: They are not alone in determining what types of assessments or instructional strategies they use; instead they are part of a public educational community that includes students, parents, school district administrators, school board members, and policy makers. To allow for the possibility some of these community members may have potential influence on participants’ experiences, the research questions considered the context of art teachers working in a public-school environment. One main overarching research question and four sub questions were written to guide this research. Although it is ideal to develop two sets of research questions for mixed methods research (Creswell, 2014), a single set of research questions was used to examine both the quantitative and qualitative data, which helped me merge the data from multiple sources.

The following research questions guided this study:

1. How are high school visual art and design teachers expected to utilize assessment practices by their school districts in the context of a nationalized shift toward standardized assessment?
   a. What types of assessment strategies do Illinois high school visual art and design teachers report using in their classrooms?
   b. What directives are Illinois high school visual art and design teachers given by their school district administrators about reporting assessment data?
   c. What professional development about assessment is provided for Illinois high school visual art and design teachers by their school districts?
   d. What professional development about determining validity and reliability in assessment is provided for Illinois high school visual art and design teachers?
Although not part of the original research questions, resistance was an unexpected and a relevant thematic finding in this study. When analyzing the data sets to answer the original research questions, patterns of resistance emerged in both the quantitative and qualitative data. Resistance, as contextualized in this study, refers to how participants responded to any job-related directive they did not agree with, wanted to avoid or ignore, or otherwise felt was not in the best interest of their students. Although resistance was not initially evident in the survey questionnaire that was part of the quantitative data collection process, optional open-ended responses provided by some respondents demonstrated some emergence of resistance. None of the interview questions was written to ask participants directly about resistance; however, when analyzing the qualitative data, resistance also emerged in responses from all interview participants, and because it manifested in both data sources, it was a relevant finding.

Research Design

The model used to collect data for this research was an explanatory sequential mixed methods design. Creswell (2014) describes this method as the researcher conducting the quantitative portion of the data collection first, analyzing the results, then using the qualitative phase of the research to explain the quantitative findings in more detail. The first phase of this research used a survey questionnaire to collect data from a large number of high school visual art and design teachers in the state of Illinois. All survey respondents had an option to provide anonymous comments at the end of the survey. Because names and email addresses were not linked to the survey, anonymity of survey participants was protected. Any survey respondents who were interested in participating in a face-to-face interview could volunteer in a separate email to ensure their survey responses remained anonymous. This was used to create a purposive
sample of participants for the qualitative portion of the study. Eight face-to-face semi-structured interviews were conducted to help explain unexpected results from the quantitative analysis (Creswell & Plano Clark, 2007). To triangulate the data, any artifacts shared during the interview, any documents related to teacher evaluation plans of the participants’ schools, and researcher’s notes taken during the data collection process were coded and analyzed with the quantitative and qualitative data to gain a better understanding of the problem. The mixed methods approach helped develop a more complete picture of a complex situation (Mertens, 2010).

Quantitative Research Data Collection Process

The target population for this survey was public high school art and design teachers working in the state of Illinois. To find an accessible population for this study, purposive sampling was used to specifically select participants from public high school districts throughout the state of Illinois that had visual art and design teachers listed on district websites. The quantitative data were collected using a cross sectional survey to collect information from the specified population at a specific point in time (Frankel & Wallen, 2006). A survey link was emailed to potential participants, and a survey link was mailed in a hand addressed envelope to any potential participants who did not have an available email address. Participants that chose to complete the survey used a link that directed them to Survey Monkey ensuring anonymity of respondents. The frequency and percentage of responses from the survey are displayed in bar graphs and pie charts to help look at the data and illustrate the results and interpret the data through descriptive statistics, which was analyzed for expected or unexpected results (Frankel &
Wallen, 2006). These results were useful to structure the questions asked during the semi-structured face-to-face interviews (Creswell, 2014).

**Survey Instrument**

The survey questionnaire was designed with a limited number of questions to have a positive effect on the response rate (Merriam, 2010). The survey contained a total of 20 questions: one to ensure the respondent was a high school art or design teacher, two demographic questions, 16 closed-ended questions, and one open-ended optional comment or response (Frankel & Wallen, 2006; see Appendix A). After the survey, a message asked respondents who were willing to participate in a face-to-face interview to send a separate email to the researcher to ensure all survey responses remained confidential and no personal or identifying information appeared in the survey data or open-ended responses. These contacts were kept confidential and used to create a list of potential interview participants for the qualitative data collection.

I received feedback from my university supervisors on the questions before piloting the instrument to ensure the questions were not leading and used the feedback to edit the question order. I chose a web-based questionnaire format and used email with a link to the survey to promote a higher response rate (Merriam, 2010). Survey Monkey was selected as the service to send the survey questionnaire because it provided a link that ensured anonymous responses, and the data collected from the survey could be statistically analyzed by frequency of responses, number of respondents that answered each question, percentages of responses, and display of responses in charts and graphs to visually demonstrate the data (Frankel & Wallen, 2006).

Survey questions were examined to avoid negative wording, did not ask multiple concepts in one item, and avoided jargon to keep the questions specific (Mertens, 2010). The
survey was pre-tested by sending it to a pilot group similar to the target population that consisted of eight colleagues and the researcher. The colleagues who piloted the survey were selected because they were familiar with performance-based or art teaching practices. Those participants included two humanities teachers, two reading teachers who specialized in visual literacy, one English teacher, one music teacher, one AP language arts teacher, and one special education teacher who applied art making activities in her classes. These colleagues consented to analyze the pilot survey and provide detailed feedback about the validity of questions, readability of the questions, proper wording and clarity of questions, ease of completing the survey, user-friendly format, and the length of time it took to complete the survey. I also took the survey to understand how the survey would look to a respondent and to ensure the link was working properly.

Colleagues who taught art were purposefully not selected to test the instrument so it would not potentially tamper with the data that would be collected when the revised actual survey was sent to high school art and design teachers throughout the state. When feedback was received about the pilot survey, any data collected from the pilot group were deleted to protect the validity of the quantitative data.

Based on the feedback from colleagues who piloted the questionnaire, the survey instrument was edited to improve readability, multiple responses were allowed for a question in which the respondents could potentially choose more than one answer, and a required response needed to be removed because of a contingency question that not all respondents would potentially need to answer. The pilot group provided feedback that the questions were clear, the format was user-friendly, the number of questions were appropriate, and it took between three and five minutes to complete the survey questionnaire, which they felt was appropriate. If a
respondent chose to provide a response to the optional open-ended question, a few extra minutes were needed.

To create the email list, I attempted to locate a database of high school art teachers in the state of Illinois through the Illinois State Board of Education. A database existed for all school districts in the state of Illinois; however, it was not separated by high schools, high school districts, or which high schools offered visual art. I contacted the Illinois Art Education Association for a list of all high school art and design teachers in the state of Illinois who were part of the active membership. A list of active members separated by grade level taught did not exist. Since no pre-existing database was available, I had to create a database to search for listings of high schools in counties throughout Illinois. Each high school that had a website available was searched for art faculty. If art faculty members were listed and email addresses were listed, the email addresses were copied into a separate address book for the sole purpose of contacts for research. The following counties were included in the first search: Lake, Kane, Will, Kankakee, Champaign, Adams, Ogle, Winnebago, Sangamon, DuPage, Cook, Charleston, Rock Island, Jackson, and DeKalb. When 700 contact emails were collected, the survey was sent to these contacts. When the survey was sent via email, a separate email account was used for the sole purpose of research, and the email included a brief cover letter explaining the purpose of the study and the link to the survey through Survey Monkey. (Appendix B)

My goal for survey responses was 100. After three weeks 151 responses to the survey were received. There were very few respondents from central or southern Illinois as compared to the northern third of the state. In an effort to obtain a more representative sample of respondents from throughout the state, the researcher contacted the Illinois Art Education Association and requested an announcement about the survey and the link in the quarterly email newsletter. The
announcement appeared; however; no additional survey responses were received. To continue the effort to expand the number of respondents, I conducted an additional search of the following counties: Pike, Brown, Montgomery, Shelby, White, Edwards, Pope, Clinton, St. Claire, and Madison, and repeated the same process of searching high school websites to search for additional art teachers. As a result of this search, 14 additional email addresses were obtained, and six mailing addresses for art teachers in high schools that had no email addresses listed. The survey was re-sent to the new contacts, and six letters were mailed with the link to the survey (Appendix C). After the second attempt to receive more responses, a total of 160 respondents to the survey was achieved.

Qualitative Research Data Collection Process

Participants for the qualitative portion of the research were purposefully selected from the survey questionnaire participants who volunteered to be interviewed. This included eight semi-structured interviews, analysis of any documents shared by the interview participants, notes about conversations with participants before, during, or after each of the interviews, and analysis of documents located online (Merriam, 2009). Questions for the interviews were based on the results from the survey questionnaire. In an explanatory sequential mixed methods design, the quantitative data were collected and analyzed first. To explain the quantitative results and any inconsistencies in the data, the qualitative phase of the research was conducted (Creswell, 2014). To help me understand the meaning the interview participants assigned to their experiences and the context of each of the participants’ teaching situation, it was necessary to collect multiple forms of qualitative data (Merriam, 2009). Through triangulation of data, I was able to find
patterns in the data from multiple participants and several evidence sources to build internal validity (Mertens, 2010).

The intended time for the face-to-face interview was 45 minutes to elicit enough information about the research problem from participants without letting the interview exceed one hour. Questions for the face-to-face interview were written and checked by my university supervisors to ensure the questions were not leading. Unnecessary questions that were repetitive or did not directly relate to the research were removed. A pilot interview using the proposed questions was conducted with a colleague who taught high school visual art and design to check for clarity, answerability, and timing to ensure the interview was within the projected 45-minute time. After the pilot interview, feedback was given about wording of questions, to ensure questions were not leading and were clearly stated, and enough information was obtained from the interview to address the research problem (Merriam, 2009). The data collected in the pilot interview and any observations made during the process were not included in the data to avoid bias.

I conducted eight semi-structured interviews. I talked to the participants before the interview process to explain the purpose of the study and to informally build rapport; however, I limited the amount of personal information was revealed to the participants so it would not bias the participants’ responses (Mertens, 2010). The total length of time spent with each participant varied depending on the time it took to introduce the purpose of the study, review and sign IRB consent and agreement forms, (Appendix D) individual participant’s length of responses, and the number of follow up questions. The planned interview consisted of seven questions and three potential follow-up questions if they were needed (Appendix E). In an effort to avoid biasing the results of the face-to-face interviews, the same interview questions were asked of all participants;
However, depending on responses from participants, clarification questions were asked if needed and follow-up questions were asked when it was necessary to explore unexpected responses. The actual length of each of the interviews varied. Not including the time for introduction and IRB consent, the longest interview lasted 48 minutes, the shortest lasted 20 minutes, and the average length of an interview was 37 minutes.

Notes were made in a research notebook before interviews about the participant and the location of the interview. Before the start of an interview, participants were given IRB informed consent forms signed by the participant and myself. A copy was given to the participant. The participants were ensured confidentiality and informed they had the right to withdraw from the study at any time. The purpose of the interview was explained and preliminary questions were answered. Consent to record the interview was obtained from each participant. Each of the interviews was audio recorded with participant consent. When the interview concluded, the participants were asked if they had any additional comments or questions and the recorder was shut off.

Immediately following the interviews, I wrote notes about observations of any documentation or artifacts shared by participants for analysis during the interview or any additional observations of the participant. Some participants asked additional follow-up questions after the recorder was shut off. Notes were made in a research notebook if the questions or comments were pertinent to the study. As suggested by Saldana (2016), everything that transpired related to the interview process was documented for possible coding. I personally transcribed all interviews in a timely manner after each interview. If participants shared documents or artifacts during the interview notes were added if needed when transcribing, or notes were made if participants sent any follow-up information to the researcher. Follow-up
questions, if needed to clarify any information in the transcripts, were obtained via telephone call with participants. Notes were made about any follow-up email correspondence from the participants.

Positionality of the Researcher

I have been a practicing high school art and design teacher since 1996 and have continued to work as a teacher throughout the duration of the research. The researcher has worked in three schools throughout her career; however, most of her career has been spent in a school that serves most students from a SES and has a high population of students who are Black and Hispanic. The researcher has been asked by administration at different times in her career to meet standards not related to visual art and design education, experienced reduction in funding, had an increase in workload, was asked to use standardized assessment tools in her classroom, experienced changes in her performance evaluation, and has seen fewer opportunities for students in her school to enroll in art classes. These experiences have motivated her interest in research to understand if these experiences are shared by other high school art teachers throughout the state of Illinois and, if so, understand the implications for the field of art and design education. This has positioned the researcher as an insider to the field of art and design education at the high school level, allowing the researcher to develop rapport with participants and help put the interview participants more at ease. However, in the role of researcher conducting the interviews and not working at the same schools as the interview participants also positions me as an outsider. The transitions between insider and outsider are something I was aware of when developing rapport with the interview participants (Mertens, 2010). I was aware of my biases based on personal experiences, and to reduce bias, removed any of my personal
experiences related to my teaching experience from the data. I also attempted to examine and code all data in an unbiased lens.

How the Quantitative Data Were Analyzed

The results from the survey questionnaires were analyzed looking for frequency of responses. Data were segregated by question number. The indices from the sample of the population who responded to the survey questionnaire were provided (Frankel & Wallen, 2006). For each item, frequency and percentage of responses were used to analyze the responses to questions, which included the number of respondents who selected an answer choice, the percentage of respondents who selected an answer choice, and a bar graph or a pie chart that illustrated the differences in proportions of the answer choices (Frankel & Wallen, 2006). The anonymous open-ended responses were coded and included in the qualitative data.

I analyzed the demographic information about the population who completed the survey to check if the respondents represented an equitable sample of years of experience and represented multiple geographic locations throughout the state. Frequency and percentages of responses were reviewed from the survey and used bar graphs and pie charts to “illustrate the difference in proportions” with descriptive statistics (Frankel & Wallen, 2006, p. 201). Descriptive analysis was preferred because the number of participants were too small for inferential statistics (Creswell, 2014). Each item was analyzed by looking for expected or unexpected responses, and was examined for patterns which helped understand how the results aligned with the research questions. Nonresponse to questions was also considered in the analysis, which helped draw inferences about the responses. Analysis of the data helped guide
the interview questions for the qualitative portion of the study to help explain findings in the quantitative data (Creswell, 2014).

How the Qualitative Data Were Analyzed

The qualitative data were analyzed through a phenomenological lens to interpret the individual participant’s experiences and to understand their points of view (Mertens, 2010). Interview transcripts, researcher notes, and document analysis were coded in three cycles of manual open coding using Microsoft Word to note all patterns and themes that emerged (Saldana, 2016). Based on the coding methods Saldana suggests, the first cycle of coding used holistic coding to get an overview of the data for the first set of codes, listing salient themes or subjects the participants discussed. All codes were noted by highlighting the text and inserting a comment in the margins of Microsoft Word. A list of common codes, including words and phrases that emerged after the first round of open coding was created. Each common code was written on a sticky note so I could move and re-arrange the visual diagram to assist in the second cycle coding. The second cycle utilized eclectic coding to refine the first cycle codes and organized the codes into categories. Refined codes were noted in Microsoft Word. The third cycle of concept coding to further refine the data by merging categories to determine themes and assertions (Saldana, 2016).

To merge the results for the mixed methods design (Mertens, 2010), a fourth round of coding included a review of researcher notes from analysis of the quantitative data, i.e., coding the open-ended responses to the survey identifying themes, for a final holistic interpretation of the data. This were arranged into categories and subcategories using an outline format to help organize all the findings (Saldana, 2016). This helped me review both open and closed ended
responses from the participants, use multiple forms of data to analyze multiple perspectives of the participants, and blend the interpretations of participants (Creswell, 2014). The literature discussed in Chapter 2 helped me code and analyze the data collected from the survey and understand the context of the responses shared by the interview participants. The literature also clarified analytical notes related to observations and document analysis.

Criteria for Selecting Participants

The following criteria were set for the selection of participants for this study: All participants needed to work at a public secondary or high school in the state of Illinois and teach visual art and/or design classes. The participants for the quantitative portion of the study were selected by purposive sampling through a researcher-created database to meet the specific criteria of population the researcher wanted to contact (Frankel & Wallen, 2006). I attempted to sample respondents equitably from throughout the entire state; however, public high schools with visual art programs were not equitably distributed throughout the state of Illinois. Counties in the northern part of the state had significantly more visual art departments than the remaining two thirds of the state.

The participants for the qualitative portion of the study were selected by purposeful sampling based on voluntary willingness to participate after responding to the survey (Merriam, 2009). Potential participants contacted the researcher via email. Follow up correspondence was sent to ensure the potential participant met the criteria and to set up meeting dates and time. Phone calls or text messages were sent to confirm interview time, date and location when a potential participant shared that information. Some participants were referred through snowball
sampling by others who knew someone who might be interested in participating in the study (Mertens, 2010). These potential participants were emailed for initial contact.

The goal was eight interviews. Although there was an attempt to select interview participants throughout the state, only participants from the northern part of the state were interviewed because no volunteers from the central and southern part of the state contacted me. At the start of the qualitative data collection process, there was a total of 11 potential interview participants. Two of the 11 stopped responding to my emails. One potential participant withdrew from the study because she was not re-hired at her school and no longer wanted to discuss her experiences. Eight participants completed face-to-face interviews: three white males and five white females. There was an attempt to represent other ethnicities; however, only white participants volunteered. Four of the participants worked in upper-middle to upper SES schools with low numbers of minority students, and four worked in lower-middle to low SES schools with high numbers of minority students. Pseudonyms were assigned to seven participants to ensure confidentiality; one participant selected her own. General descriptions of the participants’ school locations were also made to protect anonymity. More description of each of the interview participants will be provided in Chapter 5.

This chapter described the explanatory mixed methods design used for this research and how the survey questionnaire was used to collect the quantitative data, and how the interview process, field notes, document analysis and notes made by the researcher were used to collect the qualitative data. The criteria for how the participants were selected and my positionality were explained. This chapter also described how the quantitative data and qualitative data were analyzed. The next chapter will provide a detailed description of the responses to the survey questionnaire and the quantitative data.
CHAPTER 4
QUANTITATIVE DATA

The purpose of this study was to examine how Illinois high school visual art and design teachers fulfilled requirements for assessment. In the context of a nationalized shift toward standardization in education, Illinois teachers are required to produce data on student growth from valid and reliable assessments as a required component of their teacher’s performance evaluation. Because standardized assessments are inappropriate for judging the quality of student learning in art and design, this study used a mixed-methods design to collect data on the kinds of assessments high school visual art and design teachers were using in their classrooms and whether these differed from or complied with directives they received from their evaluating administrators. This chapter includes a description of the quantitative data collected from the survey, including a description of the respondents and a summary of responses to the closed-ended questions and the open-ended optional responses.

Survey Participants

A total of 160 individuals responded to the survey questionnaire. Sixteen percent of the respondents indicated they had 0 to 5 years of experience, 20% indicated they had 6 to 10 years of experience, 21% had 11 to 15 years of experience, 20% had 16 to 20 years of experience, and 23% indicated they had 20 or more years of experience (Figure 5). A total of 67% of the survey respondents selected they taught in the Chicago area, 23% selected northern Illinois but not in
the Chicago area, 7% taught in central Illinois, and 3% taught in southern Illinois (Figure 6).

Two respondents did not provide demographic information about where they taught. Of those who responded, 90% of the respondents represented the northern Illinois or Chicago area, and 10% represented the remainder of the state, which is a geographically larger portion of the state.

Figure 5: Years of experience.
Responses to Closed-Ended Questions

Survey respondents were asked a series of three questions from the same question stem.

1. How useful are the following strategies in improving your own classroom art and design instruction? How useful is analyzing student data from standardized or school wide tests? A total of 159 of the 160 respondents answered this question. Of those responding, 9% identified extremely useful, very useful, or useful, and 91% responded somewhat useful or not useful or not used (Figure 7).
Figure 7: Responses to closed-ended question 1 on useful strategies.

2. How useful are formative assessments such as one-on-one critique, class discussion, or group critique? All 160 respondents answered this question, with 88% identifying extremely useful or very useful; 12% considered these techniques useful or somewhat useful. No respondents selected not useful or not used (Figure 8).
3. How useful are summative assessments such as authentic assessment, portfolio review, or rubric based assessments? All 160 respondents answered this question, with 84% of the respondents considering these techniques extremely useful or very useful, 16% considering them useful or somewhat useful, and no participants selecting not useful or not used (Figure 9).
Considering administrative directives about standards, respondents were asked: If you have been instructed by your administrator to align your planning and assessments to standards, which standards do you align your planning and assessment to? Please check all that apply. A total of 159 of the 160 respondents answered this question. Fifty-one respondents selected Common Core State Standards, 97 selected National Art and/or Media Art Standards, 93 chose the new Illinois Art Standards, and 13 selected “I am not instructed to align to standards,” or “I do not align to standards” (Figure 10).
Survey respondents were asked a series of three questions from the same question stem:
If you align your lessons to standards, how helpful are the following standards in assisting you with assessing what your students learn in art or design classes?

1. How helpful are Common Core State Standards? A total of 156 of the 160 respondents answered this question. Five percent of the respondents considered Common Core State Standards very helpful, 42% considered Common Core State Standards somewhat helpful, 38% considered Common Core State Standards not helpful, and 15% chose “unclear how standards help with art assessment” (Figure 11).

Figure 10: Responses to question on administrative directives.

<table>
<thead>
<tr>
<th>ANSWER CHOICES</th>
<th>RESPONSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Core State Standards</td>
<td>32.08%</td>
</tr>
<tr>
<td>National Art and/or Media Arts Art Standards</td>
<td>61.01%</td>
</tr>
<tr>
<td>The new Illinois Art Standards</td>
<td>58.49%</td>
</tr>
<tr>
<td>I am not instructed to align to standards or I do not align to standards.</td>
<td>8.18%</td>
</tr>
<tr>
<td>Total Respondents: 159</td>
<td></td>
</tr>
</tbody>
</table>
2. How helpful are National Art and/or Media Art Standards? A total of 158 of the 160 respondents answered this question with 34% of the respondents considering National Art and/or Media Art standards very helpful. 51% of the respondents considered National Art and/or Media Art standards somewhat helpful. 10% of the respondents considered National Art and/or Media Art standards not helpful. 5% of the respondents selected “unclear how standards help with art assessment” (Figure 12).
Figure 12: Responses to question 2 on alignment of lessons to standards for assessment.

<table>
<thead>
<tr>
<th>ANSWER CHOICES</th>
<th>RESPONSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very helpful</td>
<td>34.18%</td>
</tr>
<tr>
<td>Somewhat helpful</td>
<td>51.27%</td>
</tr>
<tr>
<td>Not helpful</td>
<td>10.13%</td>
</tr>
<tr>
<td>Unclear how standards help with art assessment</td>
<td>4.43%</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
</tr>
</tbody>
</table>

3. How helpful are the new Illinois Art Standards? A total of 157 of the 160 respondents answered this question. Twenty-five percent of the respondents considered the Illinois art standards very helpful, 53% considered them somewhat helpful, 13% considered them not helpful, and 9% selected “unclear how standards help with art assessment” (Figure 13).
Figure 13: Responses to helpfulness of Illinois art standards.

Survey respondents were asked a series of four questions about direction they received from their administrator: Over the course of a school year, how often has your administrator conveyed information about the importance of school wide test data analysis? A total of 159 of the 160 respondents answered this question. Fifty percent of the respondents selected “very often” or “often,” 25% selected “occasionally,” and 25% “once or twice” or “never” (Figure 14).
Over the course of a school year, how often has your administrator conveyed information about the importance of teaching in your content area? A total of 159 of the 160 respondents answered this question. Thirty-four percent of the respondents selected “very often” or “often,” 22% selected “occasionally,” and 44% selected “once or twice” or “never” (Figure 15).

Figure 14: Responses to administrative directives.
Over the course of a school year, how often has your administrator conveyed information about the importance of incorporating reading and math skills in your content area? All 160 respondents answered this question. Thirty-three percent of the respondents selected “very often” or “often,” 24% selected “occasionally,” and 43% selected “once or twice” or “never” (Figure 16).
Over the course of a school year, how often has your administrator conveyed information about the importance of utilizing standardized assessment in your practice? All 160 respondents answered this question. Twenty-seven percent of the respondents selected “very often” or “often”, 26% selected “occasionally” and 47% selected “once or twice” or “never” (Figure 17).
Respondents were asked specifically about implementation of portfolio assessments as part of their classroom assessment practice: Not including AP or IB portfolio examinations that are graded outside your school, do you implement student led portfolio examinations in any of your classes? All 160 respondents answered this question. Fifty-four percent of the respondents answered “yes”, and 46% answered “no”. If yes, in which classes are your portfolio examinations conducted? This was a contingency question, and 148 of the 160 respondents selected a response. Thirty percent of the respondents selected “all classes”, 31% selected “only
upper level classes”, none selected “only beginning level classes”, and 39% selected “this does not apply”.

Survey respondents were asked a series of three questions about the amount of professional development time allocated for assessment per school year.

1. How often are you given professional development time specifically for the development of visual art and design assessments? A total of 159 of the 160 respondents answered this question. Twenty-one percent of the respondents selected “6 or more times a school year”, 34% selected “5 to 2 times a year”, 24% selected “once a year”, and 21% selected “never” (Figure 18).

![Figure 18: Responses to implementation of portfolio assessments.](image-url)
2. How often are you given professional development time specifically to determine the validity of assessment? A total of 158 of the 160 respondents answered this question. Fifteen percent of the respondents selected “6 or more times a school year”, 35% selected “5 to 2 times a school year”, 23% selected “once a year”, and 27% selected “never” (Figure 19).

Figure 19: Responses to professional development time to determine validity of assessment.

<table>
<thead>
<tr>
<th>ANSWER CHOICES</th>
<th>RESPONSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 or more times a school year</td>
<td>14.56%</td>
</tr>
<tr>
<td>5 - 2 times a school year</td>
<td>34.81%</td>
</tr>
<tr>
<td>Once a school year</td>
<td>23.42%</td>
</tr>
<tr>
<td>Never</td>
<td>27.22%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>158</td>
</tr>
</tbody>
</table>

3. How often are you given professional development time specifically to determine reliability of assessments? A total of 158 of the 160 respondents answered this
question. Thirteen percent of the respondents selected “6 or more times a school year”, 32% selected “5 to 2 times a year”, 18% selected “once a year”, and 37% of respondents selected “never” (Figure 20).

<table>
<thead>
<tr>
<th>ANSWER CHOICES</th>
<th>RESPONSES</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 or more times a school year</td>
<td>12.66%</td>
</tr>
<tr>
<td>5 - 2 times a school year</td>
<td>32.28%</td>
</tr>
<tr>
<td>Once a school year</td>
<td>18.35%</td>
</tr>
<tr>
<td>Never</td>
<td>36.71%</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
</tr>
</tbody>
</table>

Figure 20: Responses to professional development time to determine reliability of assessments.

After the closed-ended questions, respondents were welcome to provide an optional open-ended response or comment. Out of the 160 respondents, a total of 24 open-ended responses were received (Appendix F). These responses were coded and analyzed with the qualitative data looking for patterns that emerged in the responses. Respondents who were
interested in getting more information about the study or participating in a face-to-face interview were asked to contact the researcher via a separate email to keep all survey responses anonymous. A total of 11 respondents expressed an interest in participating in an interview however, one potential participant did not have their teaching contract renewed so declined participation, one potential participant forgot to provide contact information, and one potential participant did not respond to follow-up emails.

This chapter described the results from the survey questionnaire. All results were explained using descriptive statistics with bar graphs or pie charts to illustrate the responses. The next chapter will describe participants who were interviewed and the themes and patterns that emerged in the qualitative portion of the study.
CHAPTER 5
QUALITATIVE DATA

The purpose of this study was to examine how Illinois high school visual art and design teachers fulfilled requirements for assessment. In the context of a nationalized shift toward standardization in education, Illinois teachers are required to produce data on student growth from valid and reliable assessments as a component of their performance evaluation. This study used a mixed-methods design to collect data on the kinds of assessments high school visual art and design teachers were using in their classrooms and whether they differed from or complied with directives they received from their evaluating administrators. To interpret the findings from Chapter 4 that presented the quantitative data, this chapter presents the findings from the qualitative data collected from the eight interview participants and information about the schools in which they worked, PERA documents from four Illinois school districts, and written anonymous comments from survey questionnaire respondents.

This chapter describes the themes that emerged from the data: the socio-economic and racial makeup of the participants’ schools and how it shaped their experiences, the kinds of assessment techniques the participants used in their classrooms, the reliability and validity issues participants experienced, how participants met the requirement for data collection, the participants’ experience with their performance evaluation, how participants understood educational policy and meeting standards, descriptions of student performance in the art room, the teachers’ experiences with professional development and professional communities, their
manifestations of resistance, manifestations of job stress, and an analysis of school district teacher evaluation plans from four high school districts from throughout the state of Illinois.

Participants

A total of eight participants consented to face-to-face interviews. Five were white females and three were white males. Four of the participants taught in schools they identified as upper middle class or affluent with a majority white student population. Some participants from affluent schools identified Asian or Indian students as part of their school population in addition to white students. None of these participants suggested a significant population of minority or underserved students. The remaining four participants taught in schools they identified as middle to low income and described a significant number of their students as Hispanic, Black, or underserved. All of the participants worked in high schools outside of the city of Chicago in the northern third of the state of Illinois. Pseudonyms were assigned to all interview participants to protect their anonymity.

The first participant was Charlotte, a white female with over 25 years of art teaching experience. She started her career in a catholic school; however, she has worked in a public school the majority of her career. She was interested in participating in this study because she was interested in educational legislation. She was involved with her school district’s PERA committee when it first started because she thought it was important that art was part of the decision making. Charlotte expressed concern for the future of the arts in schools because throughout her career she had witnessed budget cuts to the arts. This inspired her to become involved in the teacher’s union at the local and state level. She currently works in a high school in a south suburb of Chicago. She described her school population as high minority and middle
to low income, indicating a significant number of students are in poverty. Charlotte was very easy to talk to and willing to freely share information about her students and work history and articulate about her opinions.

Charlotte teaches the studio level courses and AP studio art courses in her high school. She recently started integrating digital technology into her classes by creating online forums. She described how the adoption of digital archiving of artwork has made her job managing the volume of student artwork easier. In addition to teaching high school, she also teaches at a community college, which has helped her integrate art history into her advanced level classes by providing contemporary and historical art references to enhance her student’s work. Charlotte spent considerable time with me after the interview discussing some of her concerns and sharing multiple images and artifacts.

The second participant was Lisa, a white female with 19 years of art teaching experience. Lisa has worked in the same school in the far south suburbs of the Chicago area throughout her career. She identified most of her students as middle to low income with some in poverty and a high number of minority students. Lisa was very knowledgeable about the teacher performance evaluation process in her school because she was a member of her school district’s PERA committee. She was motivated to become a member of the committee to ensure the arts had a role in district-wide decision-making. She is the only art teacher at her school and instructs all levels of classes from foundation level through AP. When she first started at her school, she had two colleagues; however, when they left, their positions were not re-established due to budget cuts and lower enrollment. Lisa followed up by telephone and email with the researcher to provide additional documentation and artifacts.
The third participant was Nathan, a white male with 20 years of teaching experience. Nathan started his career in a high school with a low socio-economic and racially diverse population. He moved to a more affluent middle school where he taught for five years. Realizing this age group was not the right fit for him, he returned to teaching high school, and currently teaches in a small partially rural suburb in northwest Illinois. He described his district, which includes three high schools, as large, and he was one of four full-time art teachers at his school. He described his student population as middle class to affluent and mostly white. The curriculum at his high school was robust, and he taught foundation level art and introductory digital arts courses. Nathan was polite, easy to talk to, and expressed a willingness to provide follow-up information or artifacts if needed.

The fourth participant was Cyndi, a white female teacher with five years of experience. She started at a middle school in a small city in northwestern Illinois and recently had moved to the high school in the same city. She described her school as the one in her district with the highest number of students in poverty, one of the lowest performing schools in her area, and most of her students as Black. When she moved to the high school, she knew many of the students from the middle school, which was helpful for her to have an established rapport with students. She was one of three art teachers in her school and instructed the foundational level courses; however, her colleagues had considerably more years of experience than she did. She stated these teachers have shown no interest in working in a professional community with her. To help build her curriculum and pedagogy, she found art teachers from neighboring schools who were willing to collaborate with her because she felt having a professional community of teachers was important to her practice. Regardless of her limited years of experience, Cyndi was very confident, energetic, and willing to voice concerns about her students’ futures.
The fifth participant was Noah, a white male with four years of teaching experience who was proud to have recently received tenure. He taught in a suburban high school in an upper-middle to affluent area in northwest Illinois. He described his student population as not having a lot of racial diversity, being mostly white. He was one of three full time art teachers at his school, and one teacher traveled between schools. He described the art rooms at his school as well equipped with plentiful supplies and access to newer technology. It was evident the arts were valued as part of the school community because he described multiple display areas throughout the school, including a small gallery. The school hosted an annual fine arts fair showcasing work from the year. He taught ceramics and computer graphics and was part of a team to write curriculum for a new course, 3D printing and design. Noah was very upbeat and easy to talk to and willing to provide examples of student work.

The sixth participant was Vivian, a white female with 18 years of teaching experience. She had worked in two different high schools in her career. Both schools were located in the southwest suburbs of Chicago, both schools had high numbers of minority students, and the school populations were mainly low-middle class to low-income. She was one of two teachers at her school and described her district as large 12 art teachers. She had been at her current school for 12 years. She identified a majority of her students as Hispanic and said the remaining student population was an equal mixture of black and white. She expressed great concern for a significant number of her students who were undocumented and noted that many of the students were struggling with the current immigration problem although her school was considered a safe harbor. She taught all the two-dimensional studio art courses and had one colleague at her school who taught the digital and AP classes. Vivian was very easy to talk to, friendly, and used animated body language. She talked at great length about concerns she had for her school and the
field after the interview was completed. Despite these challenges, she had a good sense of humor about many of the problems her school faces. She shared many artifacts related to rubrics, student examples, and her school district’s evaluation plan.

The seventh participant was Charles, a white male with 24 years of teaching experience. He taught elementary school art to grades one through six at the beginning of his career and moved to a high school for the remainder of his career. He described the high school he taught in as very affluent in a far western suburb of Chicago. Most of his students were white, and the minority populations who attended his school were Asian and Indian. He described this population of students as very high achieving because they came from professional families who held them to high standards. Charles also mentioned the school he worked at is often rated as one of the top schools in the nation. He was one of five full time art teachers at the school, taught ceramics and drawing, and was the Professional Learning Community (PLC) leader in the art department. In addition to teaching, he described his personal art practice as a very important part of his classroom instruction and said he shared much of his personal work with his students. His artwork was related to popular culture and cult films, which provided contemporary art references his students enjoyed. Charles was very easy to talk to and willing to answer questions; however, he was unaware of most current educational policy or any current legislation about PERA. He admitted he was in an unusual situation because of the high performance of his school and that the art department is left out of any requirements for standardized testing, or most school wide test preparation.

The eighth participant was Lily, who had taught high school art for seven years. Her first year of teaching high school was in a northern suburb where she was split between two schools. One was an alternative K-12 school with violent and high-risk students. She described this as an
excellent learning experience. Class sizes were small and there was lots of security at the school; however, it was a physically and emotionally demanding position. Lily changed to the high school where she is currently teaching. The school is located in a far northwest suburb in Illinois that also has a rural area. She identified most of the students as white and upper-middle class. She teaches introductory level classes and AP and described photography as her passion. Lily is one of three full time art teachers at her school, who she felt worked together as a cohesive group. She was very confident and articulate. She was very interested in participating because she has become very interested in the politics of education. After the interview, she discussed at length how she wished in her undergraduate experience she had been taught about political policy, the workings of school boards, or the truth about how political legislation shapes education so she could have been more prepared for reality.

Summary of Themes from the Qualitative Data

The following themes emerged from the data: The kinds of overall experiences the participants described with assessment had a relationship to the SES and racial demographics of their student body. Most teachers in this study valued one-on-one critique with their students as the most effective and informative assessment technique. Other assessment techniques implemented in their classrooms included performance-based strategies and some right/wrong assessments. All the participants described the requirements to collect student growth data; however, the methods and requirements among the participants’ schools varied. Very few of the participants understood the reliability of their assessments, while most described valid assessment techniques, they were not all able to articulate why they were valid. The participants all described different requirements for their performance evaluation and all cited inconsistencies
in how they were evaluated based on differences between administrators. Very few of the participants mentioned art standards when they explained their assessments; however, a significant number of the participants, especially those who taught in schools with higher populations of students from low SES, described requirements to align to Common Core State Standards in reading in their visual art and design classes. Themes of resistance manifested with all the participants in very different ways in response to policy, standards, or requirements for assessment-based data collection.

All participants had a very good understanding of appropriate student performance for high school aged students and offered many examples of student art making appropriate for adolescents. More of the participants who taught in affluent areas were able to describe classroom conditions that fostered creativity with their students, and all the participants expressed an understanding of importance of creativity for high school students. Finally, all participants expressed an appreciation for the value of professional communities and their importance for professional development. Overwhelmingly patterns in the data suggested a great need for further professional development, specifically for art and design teachers, to help them with data collection and responding to policy-based expectations on assessment. Underlying themes of job stress emerged throughout the data. Finally, related to Illinois’ PERA, four teacher evaluation plans were selected from high school districts throughout the state of Illinois and were analyzed to collect data on consistencies or inconsistencies of implementation among school districts.
The participants in this study who taught in schools with a majority of students from low SES and a higher numbers of minority students tended to struggle more with budget, smaller departments, and being asked to implement inappropriate assessments. They expressed more frustration and job stress than the participants who taught in more affluent schools. Cyndi worked in an underserved school with a high population of minority and at-risk students. She described how students in her school are discouraged from taking art classes and said she has to advocate for her students to have equal access to the arts: “They should have [art] and not be penalized for it just because of their neighborhood.”

Vivian worked in a struggling school with a high number of minority students. In addition to perceiving a general disrespect for her discipline, she described most of her students as Hispanic and had great concern for them because a lot of them are undocumented. She said her school was a safe harbor, but she still worries about them. Charlotte also expressed concern for the future of many of her students who are from low socioeconomic backgrounds. She described how she saw funding inequities throughout the state and believed many students could not realistically afford to continue their education beyond community college, which put them at a disadvantage by making the equity gap wider. She also described struggling with funding issues for her program. Knowing many of her students cannot always afford art materials, she described how she spends money out of her own pocket or she applies for mini grants to supplement the supplies her students need when her budget is expended. In addition to these budget and equity issues, the participants who worked in low SES schools described having to
take time out of their art curriculum to implement reading and writing strategies that align to CCSS to help students in their school improve test scores on standardized assessments.

Contrary to the participants who worked in low SES schools, participants who worked in more affluent schools did not describe any requirements to implement CCSS and did not feel any pressure to use standardized tests. They had autonomy to teach their art curriculum without any external pressure to meet other standards. They did not describe budget issues or cuts in their departments. Noah worked in an affluent school and described working in a very supportive department with a robust curriculum and the opportunity to expand the curriculum by adding new classes. The arts were respected and supported in his school in multiple ways. He explained how during a reflection with his administrator after an observation, an administrator asked him if there was anything more that could be done to help support him. The administrator was aware that they only saw a snapshot of his lessons. Noah described the kind of support his department received:

Because administrators only see a day of the lesson and usually our projects take a couple of weeks, they never see the end result. But every once in a while, they will see the end result when we have our art shows. Whether it’s in our gallery at the school or our massive Fine Arts Fest that we have at the end of the year, they can see the finished work. They will say, ‘Oh, I remember seeing them working on that’, then they can see the end result of part of that lesson they saw earlier.

This type of regular support from the administrators and the whole school community helped Noah and his colleagues feel like respected members of the faculty. Charles also expressed his gratitude for working in a fortunate situation in the arts, but he also expressed concerns for some of his students who were from affluent families. He could see they were under considerable pressure from their families to excel in school, which he could see as very stressful to some. He was also relatively unaware of any major legislation that affected Illinois schools. During the
interview, when I asked about PERA, and its implications on assessment, he had no response. He asked me to clarify what PERA was. When I did, he still did not have a response:

Interviewer: Is this something familiar to you?
Charles: No.
Interviewer: So, this has not been conveyed to you at all?
Charles: No, and thank goodness. But if I had to, and I could see the importance and that, we could do that. But it would just infringe on what I think is a pretty good system right now in our art department. So, no I’m not familiar with it.
Interviewer: So, you feel you’re trusted to do your job?
Charles: Yeah.

Charles and Noah expressed gratitude for the fortunate situation they were in and knew other schools around the state were not in these kinds of situations. To explore further similarities and differences in participant’s experiences, assessment techniques used in art classrooms will be discussed.

Assessment Techniques

The interview participants described ways they varied assessments in their courses based on the level of class and the purpose of the assessment as well as how they balanced this based on expectations placed on them by their school. Qualitative assessment methods were the ones all participants discussed using most often in their classrooms, especially formative one-on-one critiques. Each of the participants described ways they used critiques with their students. Vivian described how learning in art is an ongoing process, and critique is useful because “assessment should be in the process; [It’s] not always [about] the end product.” She explained how she finds critique more useful than grading with a rubric: “Actually, I prefer to not use a rubric necessarily, but I prefer to use actual artwork and critiques. That’s my favorite assessment, I don’t know if you could call it formative or summative, but that’s what I usually do.”
Charles described the satisfaction he gets from conducting personal critiques with students:

I’m happiest when I’m speaking to the student one-on-one and we evaluate the work together… this is in a perfect world when I get them to critique themselves and I submit my opinion… In reality, this model takes way too long to meet with every single student every day. We just don’t have that kind of time… So frequently I will address the whole group with very typical beginning problems and then evaluate one-on-one… It’s a balance of what I can realistically do.

He further described how he prefers to devote his time to one-on-one assessment with students rather than pencil-paper formats:

A lot of teachers do a lot more formal assessing than I do. It’s one of those things where I really admire them for that, because I think for some students that might work better. But for me, I like dealing with the individual student. I like interacting with them one-on-one. I do large group demonstrations, but the evaluation is number one, establishing a teacher-student relationship; a good one hopefully and I think that’s a good thing. Now if I had to do some sort of, here’s a piece of paper, fill this out, and you learn this lesson. [huffs at that] Honestly that takes up so much of my time and its time spent in isolation, not with the student, but evaluating with this piece of paper.

Cyndi worked with high risk students in a high poverty school in a small city. She described why formative critiques are so valuable to her students’ learning and an important part of building a positive rapport in the classroom:

I think it’s important for them, because I feel like they feel they fail so often in that community, that we need to build trust to let them know that they still have time to fix it [their artwork] … I don’t penalize them for it. They just have to put in the time. So, I really try to make it as low risk as possible for them.

For Cyndi’s student growth assessments, in addition to a portfolio exam at the end of the year, she described how she documents her students’ growth with visual artifacts and explains to her students what she expects them to do using visual benchmarks:

What I do is photograph benchmarks so they can see and get a visual… [as the students go through the process of learning to draw] I photograph all the steps they go through. Then I evaluate it and I give them scores on the rubric. It’s the same rubric for each step,
and then, I give them a little spreadsheet afterwards with the [visual] benchmarks for each step for each data point. So, this is what a 1 looks like, this is what a 4 looks like… I make it work for me because I’m not doing unnecessary grunt work for something that isn’t going to make me better and isn’t going to make my kids better.

Lisa described the importance of formative assessments. She said she likes walking around the room while students are working to see how they are progressing and provides constructive criticisms that can help her students improve before the completion of an assignment. When it is time for the summative assessment, she utilized a student portfolio and felt the most important part is for students to record their learning over time to show growth and self-assess:

For me that would mean I’m teaching them process, I’m teaching them product… a way to think critically, a way to assess themselves, and to create…So, when we discuss work, or look at their portfolio, we talk about their work … It’s different than what is being done in other classes… So, for me it means teaching my students how to internalize the idea of creation as a personal process… that’s how I show growth.

In her opinion, these assessment methods were the most valuable; however, they were also the most difficult to assign numerical values to for purposes of grading and data collection on student growth. Lisa said her growth data needed to be given to her evaluating administrator before the end of the school year, so it was hard to work that into the natural flow of the assessment cycle when the summative portfolio assessment for the class came after her student growth data was due for her evaluation meeting.

When the participants discussed the need to assign numerical values to their assessment of students’ artwork to meet the practical needs of assigning grades, all of the participants described department-wide rubrics implemented throughout their school district. However, many participants described challenges in creating universal district-wide rubrics. There were variations from district to district, and most participants discussed how each teacher had some
autonomy in adjusting the rubric for individual assignments. Vivian for example, described her struggle with utilizing a pre-determined department-wide rubric for all assignments; “I can’t wrap my mind around assigning numbers to artwork aligned to this whole assessment thing; I just can’t when it comes to art. So, what I do is fit the project into the rubric.” An anonymous survey respondent described how they felt assessing art was too complex to be done in only one way:

Art is so very broad. Standardized assessment - even a rubric - does not really cover the many components that go into making a work of art. In many ways art can be “un-assessable” in the traditional manner. Art is more of an experience than a close ended lesson. Assessing art is a very complex thing.

Lily, who taught in a more affluent district, described how she adjusted her grading criteria according to the level of class in which the students were enrolled. For example, she had multiple levels of a digital photography course that included introductory through AP. Beginning level students’ grades were based on points for completing assignments on time, correctly, or demonstrating they learned specific techniques, “So, we’re not assessing them on if they are a good artist.” However, as the students progressed to higher level classes, they were expected to develop their artistic skills, so the assessments become more critique based. The students could apply feedback to revise or edit their work before it was graded following a rubric. She found this especially helpful for the students who progress up to the AP level.

Charles, who worked in an affluent school, felt trusted as a teacher and appreciated that the administrators allowed him academic freedom to use his best judgment in the classroom; “Their approach is to let us do what we think is important. They let us make decisions… They trust us as professionals to decide on… what is the best practice.” He did note this was an exceptional situation and he was very appreciative to be working in an environment like this.
Unlike the trust and academic freedom Charles experienced, most of the other participants described pressures to develop common student learning criteria they had to assess for school-wide student growth data.

One common component all the participants discussed was some form of vocabulary assessment. This seemed to be the one part of the visual arts that was easily assessed and the easiest to produce numerical data that could be understood by an administrator or evaluator who did not have a background in visual art and design education. Many described using scantron, multiple-choice or survey methods to do this. They felt art vocabulary was practical classroom knowledge, although they also described it as one of the most superficial and lowest indicators of learning in the visual arts. Noah talked about assessing student vocabulary with his ceramics class; however, he only used it as a check for understanding before proceeding with a project. He felt the more meaningful assessment he used was critique and made sure every student received feedback from him at least once a week. His district implemented summative portfolio exams for all art and design classes by creating a common districtwide rubric to help teachers with collecting appropriate data for the visual arts. Unlike many of the other participants’ districts, Noah’s district provided the teachers in the art department with professional development to help them develop rubrics.

Lisa explained how at one time her department was required to implement multiple-choice tests for the art department final, but requirements for data collection to demonstrate student growth sparked a change to department-wide portfolio examinations:

We were asked to do multiple-choice tests as department exams before we did portfolio examinations. We asked to change… because the multiple-choice tests really didn’t fit what we were doing anyhow, and they became laughable with what we were asking students at the end. It didn’t gauge what our students were learning in the classroom… In the end the students were laughing at the final.
Similar to the transition Lisa’s district moved away from multiple-choice tests to assess vocabulary to a portfolio, Charlotte also described how her district evolved from assessing vocabulary in a paper-pencil test to having students answer reflection questions as part of their portfolio as a way of assessing student command of art vocabulary. She felt this assessed student learning beyond images by requiring students to provide commentary about their work as a way of documenting their own learning and growth. Charlotte described her district’s summative form of assessment as a student portfolio:

Our main form of assessment is digital portfolio at the different levels. Intro type courses, third year, and fourth year is AP art. They’re different requirements and different levels of achievement; compositional, technique, and so on, that we take into account. We’ve been doing digital portfolios now probably about 10 to 12 years… [with the] photos of their work [students respond to] reflection questions, use vocabulary, and terms. So, we feel like we’re ahead of the ball doing this which includes written essays along with their work.

Vivian also used portfolio examinations; however, she described a conversation with an administrator who was trying to encourage her to use more assessments that were easier to score like written vocabulary tests. She strongly disagreed with the administrator and tried to explain that if a student could identify specific vocabulary related to art, it was not necessarily an indication of success in an art class because it did not demonstrate any ability to create, which is what is at the essence of learning in the visual arts. The administrator pacified her with a response she understood but tried to insist she needed to place a significant focus on vocabulary in her classes like other disciplines do:

We actually had administrators tell us, well, what science does is, blah, blah, blah; and what English does is, blah, blah, blah; until I finally stopped listening because I just got tired of telling them we’re not English and were not science. We’re not. It’s different content. It’s performance-based.
Nathan was the only participant who did not rely on a portfolio-based summative assessment. He supported use of multiple-choice or short right/wrong answer format assessments for his students. Part of his rationale was that he taught several semester long prerequisite courses as gatekeeping classes before students moved up to more advanced level art or design classes. Since he was the only one in his school teaching these classes and they were semester long, he preferred the ease and time efficiency of scantron forms of assessments. He knew he was the only one in his department who preferred this format; however, he did describe a performance-based component to his test that required students to demonstrate the ability to apply art techniques such as value blending.

To summarize the different kinds of summative assessments each of the participants utilized, Table 1 shows whether the participant utilized a district-wide assessment or created a teacher-led assessment. The table also identifies the socio-economic status of the participants’ schools.

Table 1

Summative Assessments Implemented

<table>
<thead>
<tr>
<th>Participant</th>
<th>SES of School</th>
<th>District-Wide or Teacher-Led Assessment</th>
<th>Details of Summative Assessment Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charlotte</td>
<td>Low SES</td>
<td>District-Wide</td>
<td>District-wide rubric for end of unit assessments that considers an element or principle of design and composition. Digital portfolio assessed by a rubric that considers composition, technique, student extended responses to reflection questions about their work, and use of vocabulary.</td>
</tr>
</tbody>
</table>

Table continued on next page
<table>
<thead>
<tr>
<th>Name</th>
<th>SES Level</th>
<th>Grade Level</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lisa</td>
<td>Low SES</td>
<td>District-Wide</td>
<td>Student portfolio and student written formal analysis of work. Rubric assesses demonstration of growth with composition and technique, and student written reflection of work.</td>
</tr>
<tr>
<td>Nathan</td>
<td>High SES</td>
<td>Teacher-Led / District-Wide</td>
<td>Upper-level courses have performance-based teacher-led exams at teacher discretion. The district-wide exam is only utilized for the foundation level courses as a gatekeeping exam before a student progresses to higher level courses. 50% of the exam is right-wrong answer choice. 50% of the exam is a performance-based task for students to demonstrate proficiency.</td>
</tr>
<tr>
<td>Cyndi</td>
<td>Low SES</td>
<td>Teacher-Led</td>
<td>Teacher generated rubric for performance-based tasks utilizing written and visual benchmarks for performance levels designated 1 through 4. Does not have a district-wide summative assessment.</td>
</tr>
<tr>
<td>Noah</td>
<td>High SES</td>
<td>Teacher-Led / District-Wide</td>
<td>Teacher-led assessment based on levels of student performance. District-wide assessment has two components and is still under development: Component one: Vocabulary assessment based on the language of art. Component two: Portfolio assessment that will use a performance-based rubric to determine levels of student performance and include visual and written benchmarks.</td>
</tr>
<tr>
<td>Vivian</td>
<td>Low SES</td>
<td>District-Wide</td>
<td>District-wide rubric for common unit assessments and summative portfolio exam which scores achievement on elements and principles of art, craftsmanship and composition.</td>
</tr>
<tr>
<td>Charles</td>
<td>High SES</td>
<td>Teacher-Led</td>
<td>Performance-based summative assessment that determines levels of performance. Currently developing visual rubrics that visually demonstrate expected levels of performance.</td>
</tr>
<tr>
<td>Lily</td>
<td>High SES</td>
<td>Teacher-Led</td>
<td>Introductory and lower level courses utilize teacher-student created online digital ‘mini-portfolio’ as a record of growth with a demonstration of successful completion of required skills. Teacher submits students’ AP scores as artifacts of summative assessment for her performance evaluation for higher level courses.</td>
</tr>
</tbody>
</table>
Reliability and Validity

Most of the participants had a clear understanding of validity. They provided many examples of valid assessments for art and visual design, and some even produced examples of student work and rubrics for me to look at. However, most of the participants had a very vague understanding of reliability. Often when the participants were specifically asked about the reliability of their assessments, they had a hard time answering the question or asked me to repeat the question so they could think about it. Although some participants could not define reliability, they alluded to it when they described some of their assessment practices if they were allotted professional development time to work on assessment with colleagues. Participants who were AP teachers had a better understanding of reliability than the participants who did not teach AP. They understood the student portfolios were scored under reliable conditions, so they felt those scores held more weight than work they graded in their classrooms.

One of the problems consistently described by the interview participants was the lack of formal professional development provided for them by their local school district. One survey respondent provided this comment:

Looking at the reliability and validity of assessments are not specifically mandated activities, but we are given plenty of time to review and tweak our assessments, which should include those evaluations if we’re doing our jobs well. Since it is informal, some educators are probably not driven to be critical of their assessments.

Nathan, who relied on multiple-choice tests, had a difficult time articulating whether his exams were valid or reliable. As he reflected on the question, he discussed how he wished his department had a leader who could help him with this and could guide his department through the process of implementing district-wide assessments. Since he was the only teacher in his school who taught the foundation-level art courses and did not have colleagues to confer with
about the same subject, he realized understanding of reliability and validity was a weakness for him:

You know, I think if I had a leader right now, or over these past five years, when we made that art fundamentals summative test, I think we had a leader who was to tell us what to do with this test, and tell us how we could use it across the whole district, so all three high schools, I think it would make so much more sense. Um, but I’m not the only teacher who struggles with those art fundamentals students, I think every teacher in the district who teaches that class has the same problems I have.

Similarly, Cyndi had been given no professional development or support on determining validity and reliability. She explained:

That definition really hasn’t been given to me… What I turned in to my principal was benchmarks, photographs, and the spreadsheet of student scores, and how they grew over time. I don’t know if that’s valid but I do know I was honest about how I evaluated my students, how I introduced it, how I gathered it, and that’s when I turned in. Until they give me a hard and fast definition about what is valid and reliable that’s what I mean to keep doing because that’s what helps the kids grow… Until they really decide to crack down and say that’s not reliable, I’m going to do what works… Because I think what I’m doing is more valid than some stupid bubble test or some vocabulary test that isn’t very worthwhile.

Noah suggested the potential of valid and reliable assessments through his description of a district-wide rubric to assess summative portfolio exams and professional development that his school provided to help all the teachers. He was able to describe the model they used to create portfolio examinations for a large district that met everyone’s individual needs. He described the rubric as specific and tangible, yet it yielded flexibility. He explained how the teachers in the district attended a professional development session and were able to work together to set visual benchmarks that matched the benchmarks in the rubric. He gave me an example by showing me an artwork and giving a hypothetical example; “This is C work. How do you know it’s C work? We tried to put all the little intuitive things we knew as art teachers into a format that was easily understood by someone reading the rubric.”
The practices in Charles’ district also demonstrated possibilities for reliability.

Professional development time was allocated weekly for teachers to meet in their department:

Every Wednesday we have a late start and we have about an hour or so. We get together and we work sometimes on common assessments. Most of us do not teach overlapping classes… We work on rubrics in isolation but, when we do meet, we are able to get together and share that information with each other… We also started developing visual rubrics and verbal rubrics. For example, in ceramics class they get a picture of what a bad pinch, looks like, medium, and then advanced. And it’s also explained.

He described regular collaborative professional development time to help each other fine-tune these kinds of visual and verbal benchmarks in rubrics.

Vivian shared the technique she used when grading student work. She said she disregarded the rubric, did not look at student names, and visually benchmarked the work by sorting the work. She said, “I know this sounds crazy, but I take all the projects, I make them into separate piles. Here’s my A pile, my B pile, my C pile. I start with the concept for the unit and see if they understand it… And I will do little checkoff sheets.” Vivian disregarded the rubric because the direction she was given by her administrators for using the district-wide rubric assessed only the elements and principles of art. What she was looking for in her students’ work was technique and craftsmanship as well as how well and at what level a student performed. Vivian demonstrated potential for reliability if she could do visual benchmarking in her professional development time with her colleagues. Instead her district emphasized that time should be spent on data collection for vocabulary, so she worked on assessing student work alone. When I asked her if she thought her assessments were valid and reliable under PERA legislation, she did not think so and expressed her frustration about how legislation is shaping what is happening in classrooms:

I think, I feel like what this is really. It’s not, to me it’s not valid at all. It is not a true assessment of what I’m doing as a teacher at all. It’s not at all valid. What it’s doing is
just trying to like I said, give them something concrete to use to show that students have learned. Because, I feel like teaching and learning is so abstract. But they’re trying to make it concrete. And, I don’t think giving us these new evaluation tools, that’s making it more concrete, when it should be more, teaching is an abstract thing. There is no right thing to do every day. So, I don’t feel like this new PERA legislation, all I feel like that’s doing is satisfying the politicians. Because really that’s what it comes down to. The politicians are the ones that have to show people that were doing our job so the only way to schools can do that is appeasing the politicians. So, the politicians can go to the public and say okay, our schools are doing good because test numbers. It’s like were at the bottom of the shit hill. So, I feel like all PERA is doing is making it easier for them to show that yes, our teachers know what they’re doing. It’s not really a reflection of if we are really good teachers or if there’s even learning going on. I don’t even think it’s a reflection of learning even happening. Because to be honest I could fudge all those numbers. I could be like, I could give Susie a three this time, so then give Susie a four next time. Teaching is so, this is taking the human component of teaching out of the equation on my evaluation which I don’t think is cool. And it’s because of the subject we teach, it’s not like math where 2+2 = 4, were teaching a subject that’s an abstract subject. So, it’s just like where this square peg in a round hole. But I feel like they’re starting to take those corners off of us, we’re just about in there now.

Lily implemented valid teacher-led assessments in her room; however, she also taught AP Studio Art and AP Art History. Unlike the other participants, she was the only participant who told me she could submit her students’ AP scores as part of her student growth data for her performance evaluation. She felt this demonstrated valid and reliable evidence of student performance.

Data Collection

Many of the participants struggled with data collection for their performance evaluation. This manifested when it came to assigning a numerical value to what their students could do in art because most of the participants were instructed by their evaluator that data were something numerical collected from a test. Charlotte struggled with collecting numerical data for her student growth. She considered data and what her students produced in their portfolio as two profoundly different things: “I don’t see a strong relationship between the data… and I can’t give
you a number or somehow quantify that. I can’t do that… I can’t really see a direct major
relationship between the data and the actual visual art portfolio students produce that
demonstrates student growth.” Lily did not think her data were valid because the only data she
could submit for her student growth component were based on superficial, easily numerated
variables unrelated to student artwork. She struggled with assigning numerical values to
creativity and critical thinking that she felt better represented what her students learned in art:

    I do not feel like they’re valid because I create a survey. And it is true, they do know
very little and at the end they know so much. But the data that I’m collecting is just so,
it’s just so, it’s not what I’m teaching. They learn so much more about creativity and
critical thinking and all these different things that can’t be data. I just don’t know how.

Vivian described frustration with collecting data using a mandated district-wide rubric.
This became a point of contention because her district was more concerned with CCSS data than
considering student performance in art:

    I’m not gonna lie, we make it up. We make up the data. We have rubrics that we use for
every unit which is principle and element of art… We were using that one rubric that gets
fed into mastery manager. All the numbers on that rubric we have to transfer those
numbers onto a bubble sheet onto a Scantron sheet… And then it tells you how many
kids got better or worse. It’s all BS.”

Lisa, who works in a school with high numbers of underserved students, described the
lack of regard for learning the arts and the priority of learning aligned to CCSS: “We are asked to
provide data on reading and math performance in our classrooms. We are not asked for data
regarding student proficiency in art.” Lisa further described how her district tried to impose a
standardized test in the art rooms to collect data because it was promoted through an outside
company. When Lisa and her colleagues investigated it, they realized it had little to do with
actual student art learning. They discovered it was a program that assessed students’ digital
design knowledge and product tools and did not have a design or art making component.
Although Noah worked in a district that did not have high demands on him to collect data and he felt he had a good relationship with his administrators overall, he struggled with understanding the relationship between art learning and collecting data:

Art unfortunately is very subjective which makes it very difficult to get what they call concrete data. Where, if you were in math, they could say okay here is a problem you got the right answer you understand; done. Art is very different; you do a task; you did it; you kinda learned; you did it at this level; there’s multiple ways of doing a task. So, the whole idea of having that concrete data for art is going to be very difficult.

Lily utilized multiple performance-based strategies, including a digital portfolio her students develop throughout the school year to measure her students’ growth. Students had their own web page they could use to turn in work and have online peer critiques. She shared these strategies with her principal and showed the websites her students developed to document their growth; however, her principal had a difficult time understanding visual evidence of performance-based growth and wanted numbers instead. Lily confessed:

We have to fake it. And it’s so fake I don’t know how else to say it. So, what I do at the start of a school year is give a survey to my students through Google about what they learn in class. Like, one through five, how confident to you feel using Photoshop? One through five, how confident do you feel using a specific tool? I’ll do that for about 20 questions, and then I’ll get a pie chart from Google that tells me where everyone was at, and then I’ll do it again the last week of class, and they can’t not grow… I’ll just give surveys randomly to get numbers.

Nathan worked in a district that was more affluent. He described a summative test he gave to all students for data collection purposes. He described it as 50% a multiple-choice scantron type of test and 50% a performance component. Each of the teachers in the department varied in how much this was weighted toward the students’ final grade. Compared to other departments, the art department issues the test to comply with data collection; however, they are not held accountable for this data:

Other departments [non-art] have this summative assessment and they are keeping very specific information. I’ve seen their spreadsheets about what students know and they
have that categorized. And we [the art department] should have that categorized but nobody has taken on that task… We have our scores on paper, but we don’t keep track of our scores. We haven’t had a leader to tell us what to do or how to keep track of these numbers… so it has been a struggle to have to issue a scantron test.

Nathan found the data collection process contrived to sway the numbers in the teacher’s favor, so he did not see much use for it;

I mean the process is a little silly. You test ‘em [the students] in the beginning but they’re brand new to the class. You say, okay you don’t know much… Then in the post-assessment you’re asking ‘em all the same questions but in a different format, and then they are able to answer them. You can see that growth and you can validate what you’ve been doing for the past 16 weeks in the classroom which makes sense on paper. It’s just kind of obvious.

Charles, who also worked in a more affluent school that performed very well, suggested that as teachers they were trusted to manage their own data and were not required to present data formally. He described how as the PLC leader he could work collaboratively to assist teachers in his department if they had a struggling student. “We would give each other feedback on different situations... they were healthy conversations, and if we didn’t get to what we were supposed to, that was fine… I think that’s the freedom we’re given from the administration.”

Performance Evaluation Process

The performance evaluation documents described by the participants demonstrated differences in the kinds of forms teachers needed to complete in their school districts. It was evident some of the participants’ school districts had changed forms from year to year. Observation rubrics and professional development the participants received to help them with their teacher performance evaluation process also varied among the schools. This is further described in the Analysis of Teacher Evaluation Process Documents section of this chapter.
One of the survey respondents shared insight on test-based accountability measures and teacher evaluation practices:

Teaching to the test and the personal evaluation associated with makes it somewhat pressure filled. We work in a district that is pretty supportive. I know in most of the country there are draconian evaluation practices in place for teachers that do not serve education in general or the students teachers work with. What I work with is pretty good, but it is little by little being eroded.

Although the participants expressed a responsibility for demonstrating student growth as part of their evaluations, there were notable differences between the kinds of experiences teachers had with their administrators based on the socio-economic demographics of the schools. Charles, who worked in an affluent school, was not expected to turn in data as part of his performance evaluation. He felt his administrators trusted him and were not aware of PERA. What he did find notable about recent changes to the teacher evaluation process, however, was an increase in paperwork. “It’s the most demanding it’s ever been, the amount of paperwork we have to do… You know, filling out stuff online, the number of times we are observed… but honestly, I understand. I’m sure it’s needed.” One unique thing Charles mentioned that the other participants did not was that in his evaluation meeting with his administrator, he was asked about how he was doing in the artistic community. Although his evaluator did not have any experience in the arts, that administrator took an interest in his personal artwork. He told me his evaluator believed that if he was a practicing artist his students would be motivated to make art.

All participants were evaluated by administrators with no visual art background, creating a notable hardship on the teachers. One exception was Noah, who worked in a more affluent school. He was observed by multiple administrators who did not have any visual art background; however, one of his evaluators had a music degree. He described how this administrator was understanding of an art classroom environment, but he also had positive interactions with his
other administrators as well. “Overall, I think the experience is great in the way our admin works. They have a very realistic mentality.” He described how all the administrators emphasized reflection and understanding how you can always continue to improve your lessons or practice. He said they kept a realistic attitude that everyone has bad days; they were concerned with growth but did not overwhelm their teachers with paperwork. In a follow up question, I asked him if he felt the current evaluation system accurately reflected his performance as an art teacher. He was one of the few participants who did feel it did and described how he felt lucky the administration at his school is understanding and supportive and knows the evaluation process is to help teachers get better. However, he did note, he was aware not all other teachers had as positive experience as he did:

I’ve heard from other teachers at other schools, it’s not how they look at it. What they see that day shows everything. There is no up there is no down; this is what it always is; which then hinders the evaluation system because some days it might be the worst day and that now says you are the worst teacher ever because that’s how your class was that day.

Lily worked in an affluent school. She said the district used the Danielson framework for teacher observations, but she described her initial disappointment with her performance evaluation rating:

I’m at the point now where I’m tenured… But forever I was terrified that I was at the bottom of the totem pole. The first two years no matter how good you are, you are not going to get an excellent rating. You’re going to get marked proficient to protect the tenured teachers. No one physically said that to me but that’s what other teachers implied. And I was like, I don’t understand, I’m doing all these things and I am two points short of excellent every time, how could that be?... Now that I’m in my sixth year, I’m ranked excellent.

She also discussed how her evaluating administrator had no experience in the arts, which is something she feels is unfair. “You have to argue during your evaluation [to get a fair rating] and I do that a lot.” She also knew when administrators were coming to class for an observation, they
wanted to see a full lesson start to finish. “That’s hard with high school art to do a one-day project, it almost never happens. So, I have to make these fake projects that are one-day projects where they see a beginning and then they see a result.”

Analysis of the PERA plans for different school districts revealed inconsistencies in the teacher evaluation process in the participants’ school districts. A few of the interview participants described how rules and directives for teacher evaluations kept changing as the school year progressed, demonstrating a lack of consistency and mid-school-year changes made to evaluation processes in response to changes with interpretations to PERA requirements. Some of the participants cited specific inconsistencies in the implementation of PERA in other school districts. When Lisa described her participation in her school district’s PERA committee, she discussed the inconsistencies she observed with implementation of the Danielson model for teacher observation:

We have found that things are being asked from different teachers in different places and different schools, that aren’t being asked of us. The criteria are different. People misconstrue evaluation rubrics. It’s like interpreting the Bible; everybody has their own interpretation of the [Danielson] evaluation model. And when schools are compared to other schools, or teachers are compared to teachers, how can you tell if one is better than another if they’re being evaluated on a system… where everyone has their own take?

Lisa described frustration with how her district kept changing the rules about the teacher evaluation process mid-school year. As a member of the PERA committee, she found inconsistencies within her district, stating their evaluating administrator asked teachers of different disciplines were to complete different forms and the kinds of forms varied depending on who the evaluator was. Specifically, because the evaluating administrator was not an art teacher, her department had to complete additional forms to justify their teacher-led assessments for the student growth component because they were performance-based and not right/wrong
answer format. Lisa also described humiliating experiences when her evaluating administrator came to her classroom unannounced during non-observation times and made negative criticisms in front of her students because the administrator did not understand the context of art lessons.

Vivian said her school did not use the Danielson model or any other existing model for teacher observations. Instead they created their own district-generated rubrics that emphasized specific observable indicators related to CCSS. What her observing administrator looked for when observing her classes was completely unrelated to her art teaching. She said her evaluating administrator never looked at student artwork during classroom observations and never came to any of her students’ art shows. “It was challenging because my idea of performance evaluation and their idea of performance evaluation was different. When they would come in to observe me for my evaluation, I had to do really elementary things, like they had me do a word wall.” Vivian also described how the rating categories shifted and administrators were instructed not to give teachers an excellent rating but instead a proficient rating: “Proficient is the new excellent. That’s what they keep telling us.” A teacher really had to perform unusually well to achieve an excellent rating. The same kind of statement resonated with most of the other participants regardless of the socio-economic status of their school. Many participants described a proficient rating was acceptable and the norm of what was expected.

When Vivian was asked if she thought the current evaluation process accurately reflected how teachers actually performed, she responded:

Oh hell no. No, it has nothing to do with it. Because I feel like, and I keep telling them this, because you know when you’re up for evaluation, all these pop ins. And I have gone down to administration when I am starting a project and I have said I need you to come in. [taps on table for emphasis] Don’t come in just once. Can you come in when I’m actually starting a project and you can actually see and understand what I’m doing? I’ll even see them in the hallway, and I’ll say hey, why don’t you just pop into today? No. So they’re just getting that little snippet, and because I know they’re coming all I do is go to
my evaluation tool so I know what they’re looking for that day, my little evaluation tool in the computer. All right, I know what they’re looking for and I make sure that happens. It really doesn’t, and I don’t think there’s anywhere on that evaluation tool that they can document how I’m connecting with the kids as a teacher and an instructional coach. There isn’t anything in there. They’re looking for very specific things. Can a student verbalize back to you what they’re learning? Um, is she maintaining classroom management? And I’m like this has nothing to do with what’s actually going on in here. It’s crazy.

Charlotte described how her district’s PERA committee made its own version of the Danielson model. As a member of the PERA committee, she was able to incorporate portfolio data into the performance component of the evaluation rubric. She described push back from other teachers in other disciplines that the evaluations were not valid data; however, she was able to successfully advocate for her department. She also expressed frustration about the lack of consistency in the rules and procedures for the performance evaluation process because her district kept changing them every year. Charlotte, who had over 25 years of experience, described what it was like to work on her evaluations with administrators who had no background in the visual arts:

I’ve never been evaluated by [anyone] with any background, degree, or professional certification in my field… And so, the evaluator needs a great deal of information and instruction on actual visual art teaching and education; the vocabulary, everything related to the field of art education. It’s a difficult thing. And I’ve had some bad experiences… over the years I think because… of their lack of background in the visual arts.

She continued to describe other experiences with observations for her evaluation when observers were late and missed integral introductions to the class, missed scheduled dates, or were not paying attention during the classes, which resulted in negative comments on her evaluation because of their lack of knowledge with art or only seeing partial classes and not understanding the context of her lessons.

Nathan never had an administrator with art experience evaluate him. He works in a more affluent school and described the time his assistant principal evaluated him:
She was very curious about the SLO and PERA scores that we were doing in the classroom… while it came up in conversation during the evaluation process, it wasn’t a focus. It was just a matter of, ‘How did your students do based-on what you thought they should do?’ and ‘Okay congratulations. Looks like they learned what they were supposed to learn.’

Although none of his administrators had any knowledge of art, he generally felt trusted and generally had good evaluation experiences although he had been considered proficient on a few of his evaluations. He has accepted it as part of the current educational culture. He described how his district used the Danielson framework for observations and noted that he found it helpful:

It’s been great about telling you what to expect. Our professional development at the beginning of the school year is always on evaluation. They [make sure all teachers have a copy of] the rubric… they tell you about the four domains… and each administrator says on different visits what they are going to look for before they come in… so the person who’s going to read through the rubric [knows what to show the administrator] what they are looking for to get the marks you want. It is all very standardized.

Cyndi felt she has received good ratings on her evaluations; however, because she is a newer teacher, she feels she is her toughest critic, is very reflective about her own practice, and is always thinking of ways she can improve. Her evaluators have had no background in art, so she was confused when she was criticized by an observing administrator who was concerned her art room was too quiet, perceiving it as a lack of engagement in the room. According to the administrator’s rubric, she wanted to see the students interacting verbally as an indicator of engagement. Cyndi described how she defended her classroom environment:

I think for the most part that the engagement piece is vague and engagement in an art room is very different than in a classroom that utilizes verbal questioning every day. If you’re in a humanities class there’s a lot more conversation. Engagement in an art class that is studio based could really mean that they are super into what they are drawing and don’t need to have a conversation. When you’re being creative sometimes, you’re just silent and you’re into what you’re doing and you’re not talking. So that part I’ve had to explain to my evaluator, this is what engagement looks like when you’re using your brain. The administrator told me, ‘Well I think engagement looks like this, this, and this.’ I again told her no they’re not talking because their fuckin busy. They’re into what
they’re doing. There’s two different ways to be engaged in an art room, critique, and work time, and they look different. So that part I’ve _had_ to defend myself on.

Many of the participants described frustration with the labeling of teacher performance levels with the new evaluation process. A point of contention raised many times with multiple participants was the difficulty of attaining an excellent rating and what the teachers perceived as an unspoken expectation for evaluators to label teachers as proficient. Cyndi described how issuing a proficient rating is an unspoken rule:

So, they’re seeing it, so I think it did, and they just don’t give anyone an excellent because you know it’s like one of those things, one of those golden things that you just never hand out. Which is fine. It doesn’t hurt my feelings. I think proficient is enough. And I think it’s probably accurate. I’m not twenty-year experienced teacher being a superstar. I am just doing a good job. And that’s fine. I do think that it depends on who it is and who the administrator is that sometimes it’s a tool used to evaluate people out.

Vivian described her frustration with the addition of data collection to earn an excellent for her performance evaluation:

Proficient is the new excellent, that’s what they keep telling us. So now the way to get the excellent is with the student growth data. And I missed it twice because I didn’t have check-ins between the beginning of the project and the end. I said, yes, I did. And they said but you didn’t document it on paper and turn it in as data. You can’t use that. I’m like, every day in that classroom is a check in; every time I look at a project, every time I look at the students work, every time I talk to the students. But all you have to do is just make a little paper, make little sheets and make it worth 10 points. So, because I didn’t have those little check-ins, I missed an excellent by a fraction. [rolled her eyes] And then the little teaching coach came in and showed me my word wall, and boy, I got my excellent. Now I can be an excellent teacher, thank you for that. [laughter] Man, it felt like I was in first grade, I was like is this really happening? It was incredible, even the kids were like what’s going on? Why do we have to do this? To prove to them that you learn something. Even one of the kids said, couldn’t they just come and look at my work Miss? I said that to one of my administrators, I said, can’t you just look at the work? I’ll make a portfolio of kid’s work for you. Come into drawing and painting, and I’ll show you each drawing and each painting from project to project, I’ll give it to you and if you just look through that, you’ll see how the students improve. And she said, well yeah but that should be easy for you because that’s what you are teachers do. But we just need you to quantify it and put it in numbers for us. That’s when I wanted to cry.
Nathan also described how his district changed the ranking categories for the performance evaluation:

So, the first administrator that I had that gave me a hard time actually I had someone before them for the first two or three years of my working there, and she loved what I was doing too. So, it was weird to go from getting what we called distinguished, so the highest rating as a teacher to all of a sudden go to proficient, one step down, without the best evidence. Maintaining that proficiency through the third evaluator. And now the last one gave me proficiency but I’m tenured at that point so I wasn’t really going argue it.

To help illustrate further the inconsistencies many of the participants described about implementation of the evaluation process, the burden of paperwork, and inconsistencies in the observation process, the next section provides an analysis of teacher evaluation plans from multiple school districts.

**Analysis of Teacher Evaluation Process Documents**

Several of the participants described changes with their evaluation process, a significant increase in paperwork, and a lack of consistency among school districts with the onset of PERA legislation. Charlotte described her perspective as a member of her school’s PERA committee: “There have been changes to the evaluation plan; evaluation system changes. With experienced teachers being evaluated every two years, her district revises and changes ‘the rules’ and procedures every year. It constantly changes and there is no consistency.”

Based on her perspective as a member of her district’s PERA committee, Lisa also recommended examining other districts’ evaluation plans to look for inconsistencies within the same state:

It may be worthwhile as a member of an evaluation team for you to find rubrics and PERA documents from other districts, and how it’s collected. Because we have found that things are being asked from different teachers in different places and different schools, that aren’t being asked of us. The criteria are different; people misconstrue
evaluation rubrics. Um, it’s like interpreting the Bible. Everybody has their own interpretation of the evaluation model. Uh, and when schools are compared to schools, or teachers are compared to teachers, how can you tell if one is better than another if they’re being evaluated on a system where everyone has their own take on it?

To support claims from participants about inconsistencies among school districts, evaluation procedures changing from year to year, and the burden of paperwork, PERA and teacher evaluation plans were obtained from four different Illinois school districts with high schools. These documents were analyzed, and the findings are included. Excerpts of the documents are included in Appendices. It is worth noting when searching for PERA and teacher evaluation plan documents many school districts that were selected from the state of Illinois did not have their plans accessible to the public via their website. Many school districts’ websites required a staff member log in or staff credentials to obtain access to the evaluation plans and forms, thus removing accessibility of these documents from public domain. The evaluation documents were from school districts that will be identified as District W (Appendix G), District X (Appendix H), District Y (Appendix I), and District Z (Appendix J).

Each of the district’s teacher evaluation plans and documents provided a mission or summary statement, an outline of the teacher evaluation process, a timeline of the evaluation process, and a list of required forms, student growth information, and documentation that needed to be completed by teachers and their evaluating administrators. With the exception of District X, the plans also included scoring rubrics for how a teacher’s performance ranking would be determined as outlined by a detailed breakdown of how student assessment information would be weighted in a teacher’s performance evaluation. It is worth noting that none of the documents from the four sampled districts provided any information that specifically addressed assessments for teachers in the visual arts or any other performance-based discipline. Each of the district’s teacher evaluation plans differed with the number of required documents teachers needed to
complete for their evaluation, number of pages of documents, and the types of assessments teachers were expected to use to calculate student growth. Table 2 summarizes the differences in the districts.

**District W**

The professional evaluation plan for District W (Appendix G) is a 27-page document that outlines the district’s overall professional evaluation plan commitments and themes, the district objectives, description of evaluation cycles detailing observation, performance ranking, and charts that describe how a teacher’s summative rating will be determined. The evaluation plan was revised in 2015 and outlines the plan through 2017. A more current document was not available online. The document identifies 12 additional forms in the appendix, including scoring rubrics, observation notes, pre-observation form, formative feedback form, post observation form, final evaluation of professional practice, teacher self-reflection, mid-year meeting record, in formal observation form, professional development plan, and remediation plan; however, none of these were included in the PDF. These additional forms must be obtained separately through human resources on the district website.
### Table 2

District PERA Plan Document Comparisons

<table>
<thead>
<tr>
<th>School District</th>
<th>Type I Assessment</th>
<th>Type II Assessment</th>
<th>Type III Assessment</th>
<th>Number of Forms to be Completed by a Teacher for Evaluation</th>
</tr>
</thead>
</table>
| District W      | X                | X                 | X                   | A total of 5 required forms:  
1 pre-observation form*  
1 post-observation form*  
1 teacher self-reflection*  
1 mid-year meeting record*  
1 professional development plan*  
*it is unknown the number of pages per form |
| District X      | X                | X                 | X                   | The number of required forms vary from 4 to 6:  
1 lesson plan template (optional)*  
1 teacher performance evaluation rating*  
1-2 teacher evaluation instruments**  
1 evidence of practice rubric*  
1 professional development plan*  
*it is unknown the number of pages per form  
** tenured teachers have one instrument, non-tenured have two |
| District Y      | X                | X                 | X                   | A total of 3 required forms:  
1 3-page rubric for professional growth  
1 10-page rubric for teacher performance  
1 additional document for statistically calculating student growth*  
*it is unknown the number of pages for this document |
| District Z      | X                | X                 | X                   | A total of 5 required forms:  
1 23-page rubric for self-evaluation  
1 goal worksheet*  
1 pre-observation form*  
1 post-observation form*  
1 additional document for calculating student growth*  
*it is unknown the number of pages for this document |
| Districts W, X, Y and Z | X | X | X | All districts require additional forms for any teacher not receiving a proficient or excellent rating. |
District W’s measures for student growth identify student growth accounting for 25% of a teacher’s summative rating during the 2015-2016 school year and increases to 30% for the 2016-2017. Type II and Type III assessments are factored into the percentage, also varying the weight of the percentage each year: The 2015-2016 school year weights Type II pre and post exams at 80% of student growth and Type III student learning objectives at 20% of student growth. The 2016-2017 school year weights Type II pre and post exams at 66% of student growth and Type III student learning objectives at 33% of student growth. In addition to the documentation a teacher must provide for Type III student learning objectives, District W identifies additional external factors that may impact student achievement, including student attendance. This district does not include Type I assessments that are state-wide standardized tests in the calculation of student growth. The district identifies rubric for teacher observation as the Framework for Teaching (FfT) domains, which is an interpretation of the Charlotte Danielson Framework (Danielson, 2007).

District X

The Teacher Evaluation Process for District X (see Appendix H) is a 20-page document that offers pertinent definitions, provides a philosophy of education, instruction, and personnel evaluations and the required components of teachers; and explains the evaluation cycle, the student growth component, and due process. The evaluation process was most recently updated in 2017. Additional forms include a lesson plan template, teacher performance evaluation ratings and teacher evaluation instruments for tenured and non-tenured teachers, evidence of practice rubric, professional development plan, and remediation plan that were not included in this
document. These documents must be obtained separately from the school district; however, there is a timeline provided for teachers and administrators to follow to submit required paperwork.

The plan for District X describes professional practice components for teachers and the evaluation cycle; however, it does not specify how the evaluation is quantified. A teacher’s evaluation in District X can also include information from parents or the community; however, this document does not specify the form. The percentage of the student growth component and weight it has on a teacher’s evaluation is listed as 25% of the performance evaluation in school years 2016-2017 and 2017-2018 and increases 30% in years 2018-2019 and beyond. The assessment menu for teachers to use for the student growth component does not include a Type I assessment as an option, which is a state wide or nationwide assessment.

District Y

The faculty evaluation and development system for District Y (Appendix I) is a 144-page document that outlines five domains that will be included in a performance evaluation, including professional practice through classroom observation under four domains described as design and planning, learning environment, instructional delivery, and assessment. Additionally, there is a professional growth component outlined through goal setting, student survey, self-reflection, and professionalism. District Y’s plan also details the summative evaluation cycle, the components of each cycle, and how the summative rating will be calculated. A published or revised date for this document is not listed; however, it states the development of this plan began in 2010 and expected implementation was the 2012-2013 school year. The student growth component was to begin in 2014. The section of the plan that details the student growth component was developed
by an educational consulting firm. What is provided to assist teachers on calculating the student
growth component, which represents 30% of a teacher’s evaluation, utilizes technical statistical
language requiring classroom teachers to have the ability to read, understand, and calculate
statistics. The stipulation about which student growth data teachers are allowed to calculate for
their student growth component uses standards-based criteria and it must be based on core
academic reading and math with no provision made for any art or performance-based learning.
To calculate the first 15% of student growth, a Type I or II assessment that measures reading or
math must be used. The second 15% of student growth is measured by a Type III assessment that
measures reading or math. This plan does not provide visual art or other performance-based
teachers with the opportunity to utilize a Type III assessment in their content area; rather it
assesses reading or math skills in the classroom. This component of the plan for District Y
necessitates teachers being well versed in assessment language and procedures.

Overviews of the domains used to observe teacher performance in the classroom were
adopted from the Charlotte Danielson framework (2007); however, the language was modified
by the district. In a component of the evaluation process District Y utilizes student surveys that
become a part of the record of a teacher’s summative evaluation. Stipulations about the
summative evaluation states multiple school personnel may conduct observations and a faculty
member must have at least two different observers. This guarantees a teacher will be observed by
someone outside of their field who will make judgments about the quality of their performance.
Additionally, detailed documents, evaluation and observation rubrics, forms, framework for
instruction, procedures for tenured and non-tenured staff and a detailed scoring system overview
are included in this document. Additional information and scoring rubrics are included
specifically for counselors, library media specialists, school nurse, and other support staff.
There are three versions of the teacher performance evaluation system for District Z (Appendix J); a 29-page version dated 2015-2016, a 34-page version updated in 2018, and a 45-page version from 2019. Separate documents that pertain to the pre-observation, post-observation, goal setting, calculation of student growth, teacher self-assessment, rubric for classroom observation, and summative rating must be downloaded separately through the human resources department on the district website. The most recent 2019 evaluation system document was analyzed. The document provides a statement of purpose, timelines, and processes for the evaluation cycle; descriptors of level of performance, calculations for domain and summative ratings, student growth ratings, assessment tools and procedures for calculating student growth, and a listing of the required forms and documentation that must be submitted for the performance evaluation.

District Z identified working with an educational consultant to understand and interpret the teacher observation process and based the rubric and domains on the Charlotte Danielson Framework (2007), which accounts for 70% of a teacher’s professional practice. The remaining 30% of the performance evaluation is based on assessment. This district utilizes a Type I assessment, which is a state or nationwide exam on which 12.5% of the teacher’s performance evaluation is based. Additionally, throughout the school year teachers are expected to monitor students’ growth goal, which can be measured by a Type II or Type III assessment worth 17.5% of a teacher’s performance evaluation. The Type III assessment can be teacher created, which provides an art teacher the opportunity to count 17.5% of their evaluation on actual student art
practice and growth. Additionally, this document provides tables and rubrics for calculating a teacher’s summative rating.

**Policy and Standards**

Based on the documents related to the participants’ district’s PERA plans, there was inconsistent direction on how individual teachers were to collect data and provide data-based evidence for student growth. There were inconsistencies among districts on the percentage weights teachers were held accountable for overall school performance data, and the guidelines for completing the teacher evaluation process for teachers in non-tested disciplines differed. The participants knew PERA was being implemented differently within their school and across other school districts.

The interview participants from schools in low-socioeconomic areas were very aware of PERA and were more informed of policy-based directives, even if they changed mid-year. Some of the participants from schools in higher socioeconomic areas were aware of PERA; however, they were not concerned about it. These participants viewed it as more of an inconvenience because of added paperwork as part of their performance evaluation rather than something that shaped their practice. Overall, many of the interview participants were aware of political policy and the ways they have changed education. Charlotte expressed concern for the way things have changed in education and especially the way the arts are viewed. She noted the series of budget cuts to the arts in the United States was a sign that although the public wants the arts, policymakers are unwilling to fund it. These actions have motivated her to become active with her teachers’ union at the local and national levels. She considers it important as an art teacher to feel empowered about decisions being made for education.
Vivian expressed concern about the way political policy has negatively shaped education. “It was terrible when it first started with No Child Left Behind… but now with technology it’s just getting more sterile.” She described the way PERA is being interpreted by her district and contended it is not helping teachers improve their practice. Also because her district emphasizes data collection based on CCSS in literacy as part of implementing policy, she felt her teacher evaluation process had nothing to do with any of the art learning that took place in her classroom. She did not see how a new teacher evaluation tool driven by political policy could improve teaching. “There is no one right way to do things every day when you are working with kids… All I feel like [PERA] is doing is satisfying the politicians… The politicians are the ones that have to show people that we’re doing our jobs.” Vivian described the priority in her school was to collect data on Common Core standards because of policy based changes that have extended beyond her performance evaluation:

And now what’s happening with, Common Core, administrators are starting to come into my room when I’m not being evaluated because they want to see if everybody is on task with Common Core. That’s what at least they’re saying. They say want to just get a snapshot of what’s going on in classrooms, and it’s not going to be used on your evaluation. We need this data so we can be better administrators, blah, blah, blah. There was one point where they would have three administrators at a time come into the room with their little iPads, and they would say, “Oh don’t worry, were just looking for specific things, were just looking for Common Core.” Oh, and we had to have our objectives written on the board every day. I told them, you know, that’s going to say the same thing for about three weeks? That’s about how long it takes us to get through a unit. Just so you know. But they kept looking for those specific things, and they kept saying were going to go to each department, working to go to an elective, were to go to math, science, so it’s nothing specific. But in one week I had at least 9 to 12 administrators in my classroom and I just lost it! Because they weren’t going into any other art rooms. So, I asked is there a problem? The kids are starting to freak out. The kids are like, is something wrong Miss? Are you in trouble? There’s a lot of people coming in here. I said, why are there so many administrators coming in my room? And they said, you are just the one we happened to draw when we were elective. I said, and I the only elective in the whole building? Because, you know there’s other art teachers down the hall. Do you know what I found out? It was because my classroom is the closest one to the office. It was my proximity to the office. [laughter] They were able to come right to my room, check and elective off
their list, and they didn’t have to walk that far. They all fuckin chose my elective class because I’m the closest to the office. And I’m like, this is not a reflection of what’s going on in the whole school, and I kept telling them that. I kept telling them you need to talk to the kids. That’s who you should talk to. Talk to the kids, and look at their work, and that will tell you what’s going on in here.

One of the survey respondents provided a comment about state and national standards but noted they are not held accountable for implementing them in their school nor are they required to utilize only one department-wide rubric:

Assessments are great to have in place in the classroom, it validates what is being measured for art projects. State and National standards are great to have the recognition from government about the work we do as educators, though I have never needed to actually demonstrate the relationship between classwork and those goals. My school has a designated rubric to use for projects, though I prefer my personal rubrics as it caters to exactly what the students are expected to demonstrate, and I am not held accountable for using the district rubric.

Noah, who worked in a large affluent K-12 unit-district, described how the entire art department has over 40 teachers. When his district knew the PERA legislation was coming, they were very proactive about trying to create standard rubrics that could be implemented at different levels in the arts. They piloted these internally for two years to ensure they were collecting accurate performance-based data from the teachers, so when they had to incorporate these data in compliance with PERA, it would not create additional stress on the staff. “We felt ready when it became implemented and we were prepared for the evaluation system. We weren’t all freaking out because we had been doing this for two years, so there really was no issue.”

Charles, who worked in a high performing school, had never heard of PERA. When he asked me to explain what it was, he told me none of it sounded familiar. He did not describe any requirements from any political policy that shaped any of his duties as an art teacher. His department was not asked to implement any standardized assessments and was not asked to align
to any standards. He did know other teachers in other school districts were having a harder time; however, none of that affected him in his professional experience.

Lily also worked in an affluent school and was unfamiliar with the term PERA. When I explained it to her, she realized it related to her school’s plan for data collection. She could not see how PERA could help her improve her skills as an art teacher. Her colleague who teaches music agreed that data collection is misused and misinterpreted for the performance-based disciplines and they have no choice but to make up the data to survive in their jobs. “It’s a double edge sword because I’ve learned how to beat the system… It’s hard when you’re a small department and a small part of the school to find someone who’s an art person to evaluate you accurately.” Although Lily did not feel she had been negatively affected by educational policy, she was aware of the politics of education. She described how as a new teacher it was very disarming and that she had a lot to learn about politics for her job. She did not feel her undergraduate experience prepared her for understanding the politics of school boards, teachers’ unions, or how much political legislation actually affects the classroom.

Most of the interview participants who discussed aligning to standards cited CCSS in language arts as the ones their evaluating administrators looked for. The emphasis placed on standards varied based on the schools and the perceptions of each administrator. Vivian stated, “It depends on which administrator. Sometimes an administrator says, okay, we’ve got Common Core… So, then we make it all fit Common Core. And then the next administrator comes in and says, oh, now let’s work to unpack the standards. Now we’re on whole child.” She expressed frustration with the lack of consistency, and because of that she said she does not pay attention to the standards anymore. She believes in teaching what she feels is most engaging for her students, but to pacify evaluators, she makes sure her lesson plans fit into whatever standards her
administrators are looking for. She is also required to write objectives on the board and only lists objectives related to Common Core because she was criticized by an administrator who was collecting walkthrough data because she had performance-based art objectives written on the board and never wanted to make that mistake again.

There was some brief mention of national or state standards for the visual arts with a couple of participants when they discussed creating rubrics; however, since it was not a priority in their final evaluations, they were not discussed in detail. Most discussion about standards with the participants referred to CCSS because that is what evaluators were looking for. Nathan’s district worked at aligning their department-wide rubrics to the national standards in art over 10 years ago; however, they did not re-align them to the new NCAS when they were released.

**Student Performance as Creative or Characteristic of Adolescent Art Making**

Overall, all the interview participants were aware of the appropriate performance levels to expect from their adolescent-aged students socially and through art making activities. The participants described student interactions that suggest they care about the development of their students and try to instill diverse opportunities for art making. A notable difference, however, was participants working in schools with a high population of underserved students had more difficulty incorporating lessons that fostered creativity than those who worked in more affluent schools. These participants attributed this to restrictions placed on them by administrators who had no knowledge in art education. The participants who worked in more affluent schools had more freedom to foster creative opportunities for their students and generally felt fewer restrictions were placed on their classroom practice. Charlotte had a realistic perspective of the overall purpose of monitoring and evaluating student performance most appropriate to art
development of school-aged adolescents. “It’s not about us. It’s about students of all levels being able to perform and … being able to show visual art development over a period of time. [It’s not something that] should be reduced to numbers or data.”

Lily shared how she developed good relationships with her students, and those connections became vital to getting positive performance evaluations. Because her lessons normally lasted a week or two, she knew during a scheduled observation her evaluator would never get to see one lesson from start to finish. She changed her normal routine to accommodate an observation; however, for this to be successful, she let her students know in advance about an observation:

Mr. So-and-So is going to be in the class to observe and it’s going to be the fakest day. And they’re like, okay. And they will sabotage teachers if they don’t like you. So, I never had that thankfully… But the evaluator will write down every single word that the kids say. Kids will come in and they will be like, ‘This is my favorite class!’ Oh my God, one time I heard the kids say, ‘I learned so much today!’ And the observer is just laughing.

Cyndi shared the importance of developing a relationship with adolescent students built on mutual respect and trust: “Once they’re in, man, they’re in for life. They love you forever. I like that about them.” Sensitive to the stressful conditions she knew some of her students were experiencing outside of school, Cyndi described the care she put into building rapport with her students to try to get the best work from them as well as understanding the balance of pushing them to take risks and respecting their stress levels and limitations:

It’s all about low risk management so they’re not intimidated. That group of kids will stop before they start if they feel like [what I’m asking them to do] is too risky. You know? Because they have enough stress. They’re going through crazy amounts of adrenaline as it is. They could come in from having a shooting happen the night before. They don’t need that. I try to de-escalate it…because I don’t want them to be afraid to learn.
Cyndi had confidence in her art ability and skills as a teacher; however, she described the importance of classroom management with teenagers as an important component she needed to strengthen so she could be more successful in her classroom:

Everybody struggles with [classroom management] when they’re new. I’m 5’3” and I look like a kid, you know? But I’ve got grown teenage boys apologizing to me for swearing in front of me… but what they’re really saying to me is you’re really good at keeping me on task… when they come in the room no matter how late they are, you’re making sure the expectations are clear to get their stuff and sit down. They know the system. They know consistency. So that was really positive for me to know that I had laid the foundations of what was expected in the room.

Noah respected his students’ creative process and described how he balanced critique and feedback with knowing it was okay to let his students struggle a little so they could figure some things out for themselves. By promoting a creative process, “I don’t want to give them the answer. I want to see them come up with solutions that might be different from something I would suggest.” Being responsive to the contemporary interests of his students, he also described collaborating with colleagues to develop new curriculum for the art department that incorporated more digital technology in both 2D and 3D classes.

Vivian conveyed frustration about not being able to exercise as much creative art making opportunities with her students because of the emphasis placed on assessing vocabulary and writing. She noticed a decline in the creative quality of her students’ work when her whole school increased emphasis on CCSS. She wanted to be able to show her best student work and let others see the good things that can come from the art department; however, “I can’t get good show quality work out of them because of all the element and principal stuff, I feel like it’s too elementary for high school kids.” She described the delicate balance of doing what she is told and what she knows in her heart is the right thing for her students:
You know I have all these projects because keep a lot of them; the stuff kids don’t take home. I can remember going through a pile about four years ago and I started to think, wow, I used to do such fun stuff. The kids used to make really good artwork. That I look at the current stuff: Boring, boring, boring, it’s terrible! I think it’s [CCSS] taking the creativity out. I really feel sad for the kids; I’m trying to ride the balance line between satisfying my administrator with data without losing the integrity of teaching art and what art is meant to be. Art is the best subject in the world, and I feel like the kids are really losing what it could mean to them.

She described a disagreement with her department administrator about having to create a multiple-choice test that would fill 90 minutes on the last day of school because it was the designated exam day rather than what Vivian planned for that day, which was a student-led portfolio presentation:

And she said we had to have a test to last 90 minutes because the finals were an hour and a half long. So, I said okay, I understand you want us to assess to collect data, but earlier in this meeting you were talking about having kids be creative thinkers, and critical thinkers. You are telling us to teach all the way up to the top of Bloom’s taxonomy. I get that. Do you see what’s at the top of that triangle? It says ‘creative works’. We’re the ones that get the kids to the top [of Bloom’s taxonomy] but you’re asking us to assess at the bottom knowledge level. It just doesn’t make sense. Do you see how the assessments change as you move up the triangle? But you’re asking us to assess at the bottom. I can’t reconcile this. She shut me down and told me it was a conversation for another meeting.

Despite the many restrictions placed on her teaching practice, Vivian shared great insight into knowing how to tap into her students’ creativity and interest in art. She shared a conversation she had with her superintendent when he was doing a building visit:

I had the superintendent in our building one day and he was standing right across from my room when all this new assessment business started. So, he was out there greeting the kids… And when the kids come in my classroom what do they do? They grab their stuff immediately. Do they wait for the bell? No! And in the morning our kids sometimes come in 15 minutes before the bell. So here they come, and I said to the superintendent, ‘Come here, I want you to see something’. And he goes, ‘Oh, what’s going on?’ And he looks at the kids and goes, ‘Oh wow! Look at them go! Do you have a bell-ringer on the board?’ [Vivian laughed] Look at them go right to work!’ And I said no, I don’t have a bell-ringer on the board. I don’t have to tell them to do this. This is what it looks like when you let students take charge of their own learning, this kind of self-sufficiency, this desire for creativity; this is what it looks like, and it’s learned in an art room. This is
what happens without data, so you keep that in mind next time you want us to quantify it and put it on a bubble sheet.

Charles was able to share many ways he was able to instill creativity in the classroom and was able to connect with his students through popular culture to help get his students more engaged in art making and foster a willingness to participate in critique:

In drawing the whole group doesn’t always get to see what everyone else is doing. I think it’s important for them to see what their classmates are doing because I value what they’re doing… I put on a show called *Ink Master*, it’s a typical kind of elimination reality show but with tattoo artists… I have the students watch the critiques because this show does critiquing awesomely! And then I have them critique the tattoo on the show. Because that’s drawing too, but it’s just on people’s skin. And through the shows we are able to get a great conversation going among the students because it’s something they can relate to. And they can learn how a critique is done, and they can learn to evaluate other people’s work… So, having stuff like that really helps a good classroom environment. It makes it fun. And thank goodness I have the freedom to do stuff like that.

Charles shared stories about former students who are currently working in the arts. While he knew not all his students would become artists, he wanted to give them the chance to find what made them happy in a particular subject. “Because at the end of the day that’s what I want to inspire kids to do. I want them to thoroughly enjoy what they’re doing.” At the conclusion of our interview he shared a letter from a former student who told him he had inspired him to become an art teacher. The letter meant a lot to Charles because it helped him realize the incredible influence he has had on young adults.

**Professional Development and Need for a Professional Community**

Many of the interview participants expressed needing the professional community of fellow art teachers as a means for support, member-checking on rubrics and lesson planning, and moral support. Participants explained various kinds of interactions and professional development experiences with their PLC. Some of the participants described district-wide allocated
professional development time specifically for working with colleagues in their discipline; however, this was not consistent with all participants’ experiences. Charles was the PLC leader for his school’s art department. He described how they were allocated weekly time built into the school schedule with a late start every Wednesday. His administration trusted him to allocate the time wisely to meet the needs of his department, so they had the time to work together as colleagues. He appreciated that his department was not micromanaged, and as the PLC leader, he could make decisions of what his group needed to work on week to week.

Lisa’s district was using multiple choice tests as district-wide assessments. She provided an example of how her PLC in art worked together as a support system to develop performance-based rubrics and advocate for portfolio assessments more appropriate for their discipline:

As a team of art teachers, we’ve had to fight for our voice to be heard. We asked to have a meeting with district administration so they understood how we were going to assess because we were asked to do something that was not the best for measuring art learning. I’m still not sure if they understand, but we’re sort of doing what we feel is best. We’re asked constantly… to justify how our data was collected and to provide proof that we’ve actually done it when other disciplines didn’t get the same request. It’s strange to have to verify learning constantly when all that needs to be done is to take a walk into my classroom to see that it exists.

Also the participants described the need for further professional development, especially with rubrics, portfolio assessment, and calculating numerical values to student work, as a demonstration of data to help them since there was little to no professional development specifically for the arts. When she started as a new teacher, Lily appreciated her teacher mentor who helped her navigate the teacher evaluation process and calculate her data. She did not think she could have successfully made it through her first few evaluations without it. They have continued a supportive professional relationship.
Sometimes, however, not all the professional development experiences went smoothly for the participants. Vivian described the difficulties of the entire art department working together in a district-wide professional development to come to consensus on a department-wide rubric in a large district:

Four buildings using the same rubric. We had to get together and write that? We have like 12 of us. One crazier than the other, I call it the crazy train. All of us together to write one rubric. Can you imagine all the teachers sitting around one table to write that? It was insane!... And we were all freaking out because we have to put our projects on rubrics now for the first time ever.

She explained one of the problems with their professional development time was lack of direction when the whole art department met because none of the administrators over the art department had art experience. The art teachers were often asked to implement rubber stamp directives for the whole district and did not make sense for art teaching:

In the 12 years that I’ve been there, I’ve had four different department chairs and curriculum directors. None of them had any knowledge of art. It’s kind of like a stepping stone for someone who wants a principal job. We really are the redhead stepchild; nobody knows what to do with art.

Nathan discussed the teacher evaluation process under PERA when I asked him if he thought the changes to the evaluation policy helped make him a better teacher. He felt learning from his peers and developing a professional community with other teachers actually helped him more

Honestly much of the process hasn’t helped me as a teacher. I get the best feedback from my co-workers and seeing what they do. Not just in the art department, but in other classes that I might sit in or substitute for. I can see how their classes work and I get good ideas… My best way of growing is by watching other teachers and have them give me feedback rather than somebody who’s not familiar with the class.
Cyndi had a difficult time when she started at her school. She was a newer and younger faculty member and tried to reach out to her colleagues at her school for support; however, she was not welcomed or helped by them:

I have colleagues who are not very big team players… They’ve been there a really long time, really jaded, don’t want to work together… I never really went in [with a let’s change this attitude], it was more like, can you share with me what you’re doing? How are you doing this? And there was nothing, crickets. They just completely ignored me.

To fulfil her need for a professional community of supportive colleagues, she found other art teachers from other schools who were willing to work with her, give her guidance, and share information about curriculum.

Resistance

The interview participants displayed different forms of resistance; some were more overt than others. Generally, the newer participants with fewer years of experience were less likely to overtly resist any directives from an administrator they did not agree with to maintain their jobs or attain tenure. Also, the participants who worked in affluent schools generally felt respected and trusted as professionals and had fewer reasons to resist. The tenured participants with more than five years of experience seemed to be more comfortable with forms of overt resistance. The participants with more years of experience, like Charlotte and Lisa, demonstrated a more activist form of resistance by becoming members of their school district’s PERA committee to ensure the arts were not ignored when decisions were made about teacher evaluations. Charlotte actively participated in school governing through the teachers’ union to assert the importance for the arts, especially on issues like budget.
As a member of the PERA committee, Lisa described how she saw resistance to meeting the data collection requirements of the new policy manifest in other faculty members who taught in performance-based classes not related to art. “Teachers skew the data, whatever they want to. They know what they’re doing; test and retest… [they] are choosing to just give tests that document reading and math, rather than their subject area so it doesn’t accurately depict what they’re doing in their classroom.” Lisa also explained how she had to defend herself during her evaluation meeting when she felt she was being penalized for things that were not on the rubric.

In reading the anonymous comments from the survey, it became evident that some of the respondents were venting frustrations that were also manifestations of resistance. One of the survey respondents lashed out about the pressure to collect assessment data and how it was not for the benefit of the students:

I actually just teach my subject area, keep my kids interested, and constantly push them to improve their skill sets and knowledge. I abhor testing. Look at the student work, and watch the teacher teach, and you will know if the Visual Arts teacher is doing their job. All this other business is government taking over our classrooms with paperwork and drawing attention away from our actual jobs so someone can point at a chart for a meeting.

Vivian shared many stories that described her overt resistance to directives she did not feel were in the best interest of teaching art. Vivian recounted a conversation with an administrator who suggested she change to passing the grade of a student who did not complete a drawing assignment. The administrator gave her an option to allow the student to describe in words how to apply the art skill rather than demonstrate the skill and said she should count that as an equivalent passing grade.

So, an administrator said, well let’s say there is a kid and he just can’t shade. He can’t do the gradation shading. But what if he can write a paper for you and he’s got all the steps down, and he can explain gradation, doesn’t it mean that he actually learned the concept? And shouldn’t that be what you’re evaluating and assessing? So, I said, you know what?
I’m going to go along with you on this. Let’s just say Bob and I were in carpentry school. Bob can use those tools and cut every piece of wood perfectly straight. Every piece of board is perfectly cut, but I really can’t do it. But I could tell you how to build a house on paper perfectly. And we both get A’s in carpentry school. According to your theory who do you want to come build your house? Do you want me to show up? I got it on paper perfectly… No, we’re not gonna grade like that.

Nathan’s school is required to give department-wide post-tests for all departments. The art department has a multiple-choice test they give; however, no one checks the department’s data and no one manages the data. He described a covert resistance this faculty exhibited:

“Teachers don’t even want that to begin with, and because we don’t know what to do with the information it has caused a point of contention. It’s like we’re doing it to just do it and that’s it.” Nathan chose to comply with what was asked of him to not draw any negative attention to himself or his department.

Cyndi, as a newer younger teacher, tried to implement things to benefit her students and help her department grow. She faced resistance from her more experienced colleagues who preferred to maintain the status quo. She received push back from them because “they’ve never been asked to do more, and they’ve flown under the radar for a long time. They’ve gotten away without doing any extras, and you know those extras matter.” She also described how she overtly resisted directives from her principal about how to collect data for her growth goal. She was told to collect data about vocabulary; however, she advocated for performance-based targets to document students’ growth:

I told my principal, I’m not doin’ this the way everyone else is doin’ this because it doesn’t make sense for me, it’s not gonna help me grow as a teacher, and it doesn’t really show you what I’m doing… Vocabulary is important, but my real job is to teach how to draw.

Lily worked in an affluent district; however, because she is still a newer teacher, she was concerned about attaining tenure to keep her job. She described, “playing the game” to do what
was asked of her to satisfy her administrator so she would be left alone to teach and assess the way she felt was best for her discipline. Lily described learning how to “beat the system” by giving administrators what they wanted during observations and how to use online programs to create data to satisfy that requirement of her performance evaluation. Lily could see how her colleagues in other tested disciplines had an easier time producing numerical data, so her covert form of resistance became a survival mechanism so she could earn tenure.

Multiple forms of resistance manifested throughout the qualitative data with all the participants. Table 3 summarizes three different kinds of resistance that became evident patterns: covert resistance – resistance that was less obvious, overt resistance – outright defiance of an administrator or directives, and passive compliance – succumbing to what has been asked out of frustration or wanting to pacify administrative requests even if the participants did not agree but needed to maintain a good evaluation or good working relationship.
Table 3
Manifestations of Resistance

<table>
<thead>
<tr>
<th>Covert Resistance</th>
<th>Overt Resistance</th>
<th>Passive Compliance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I’ll get a pie chart from Google that tells me where my students were at during</td>
<td>Art is in a weird spot with data… We have to fake it. And it’s so fake I don’t</td>
<td>I’ll do surveys throughout the school year that show data. I’ll just give surveys</td>
</tr>
<tr>
<td>the beginning of the class, and then I’ll do it again the last week and they</td>
<td>know how else to say it.</td>
<td>randomly just to get numbers.</td>
</tr>
<tr>
<td>can’t not grow.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In [teacher] evaluation if you look at the model we use you have to play the game.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>You have to argue during your evaluation, you can argue your number during the</td>
<td>We were using one rubric that gets fed into Mastery Manager. And then it tells you</td>
</tr>
<tr>
<td></td>
<td>evaluation and I do that a lot.</td>
<td>how many kids got better or worse, all that BS.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>You need to know how to put on a show to get a good evaluation.</td>
<td>Administration asked us to give a 50% floor on grades which I don’t do. We fought</td>
<td>I’m trying to ride the balance line between satisfying my administrator with my</td>
</tr>
<tr>
<td></td>
<td>that and said no. We’re not going to do that, I’m sorry.</td>
<td>data, without losing the integrity of teaching art; and really what art is meant to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>be.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>They [the students] know its fake because I tell them. So, I tell them the day</td>
<td>All they want [is for] us to process the concept because administration says that’s</td>
<td>Now it’s the most demanding it’s ever been. The amount of work you have to do, you</td>
</tr>
<tr>
<td>before that I’m having an observation and it’s going to be the fakest day. We</td>
<td>the most important thing. There is no ‘I didn’t do it concept’. We keep trying to</td>
<td>know, filling out stuff online, the amount of times we are observed, and on, and</td>
</tr>
<tr>
<td>need to do all the things for the observation… And they’re like, okay.</td>
<td>tell administration because we’re product based it’s completely different how we</td>
<td>the amount of work we’ve got to do. The reflecting, filling out all this stuff. So,</td>
</tr>
<tr>
<td></td>
<td>have to assess.</td>
<td>the evaluation is more of a pain than anything else to me but, it’s a formality.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>You really have to play the game.</td>
<td>They wanted us to do a multiple-choice and fill in the bubble and we were able</td>
<td>This just doesn’t make sense. But we have to do that for our evaluation. That’s what</td>
</tr>
<tr>
<td></td>
<td>to fight that and not have to do it.</td>
<td>we have to do to prove that were doing our jobs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What’s in the final in reality is worth 10% [of the student’s grade] but in</td>
<td>I’m not going to test them on some vocabulary that doesn’t really apply to the</td>
<td>We get directed; we have a model that we’re supposed to be following… And if we didn’t</td>
</tr>
<tr>
<td>practice I just don’t do it. Nobody checks up on that. Everyone’s way too busy.</td>
<td>skill they’re learning. Vocabulary is important, but my real job is to teach them</td>
<td>get to what we’re supposed to, that was fine.</td>
</tr>
<tr>
<td></td>
<td>how to draw.</td>
<td></td>
</tr>
</tbody>
</table>

Table continued on next page
To be honest I could fudge all those numbers. This is taking the human component of teaching out of the equation on my evaluation which I don’t think is cool.

Did you ever see what subjects at the top of Bloom’s taxonomy? It says ‘creative works’. We’re the ones that think the hardest in this building. We’re the ones at the top. But you’re asking us to assess down at the bottom. So, you’re asking us to assess down at the base knowledge level, but you’re asking us to teach up here? I can’t reconcile those two, I can’t. I can’t make sense out of this; I don’t know who they want me to be.

When we get evaluated, we have to show student growth using the rubrics we made. The way we do it is because we’re assessing the same thing from project one to the end of the year which is craftsmanship and composition. As long as the student has grown one number, we have met growth… now you’ve shown growth. He could suck as an artist, but Johnny showed growth.”

Until they really decide to crack down, and say that’s not reliable, I’m just to do what works. I think what I’m doing is more valid than some taking some stupid bubble test, you know.

I’m going to tell you something, this kind of taking charge of your own learning, this kind of self-sufficiency and learning, this is what it looks like, and it’s learned in an art room. This learning happens without data, without me even telling them to do anything. They just do it, so you keep that in mind next time you want us to quantify it and put it on a bubble sheet.

Those kinds of reflection things are included after the lesson has happened, although a lot of that is the feedback both to myself and them, that there is; because they only see a day of the lesson and usually our projects take a couple of weeks.

And it really wasn’t like super official because there really wasn’t much guidance on how to do it. So, I kind of go rogue.

I told my principal; I’m not doing this the way everyone else is doing this because it doesn’t make sense for me. It’s not gonna help me grow as a teacher, and it doesn’t really show you what I’m doing. I’m going do it a little different, that’s what I told him.

The days that they come in and observe; I try to include pieces like a quiz just so they can see and understand what is happening especially if it’s a day I normally wouldn’t do it.

On paper we don’t keep track of scores. We haven’t had a leader to tell us what to do or how to keep track of these numbers so we can see this through. It has caused a point of contention.

I’m not gonna lie, we make it up. We make up the data.

The pre-and post-test; we had those implemented two years before they were required. We were able to practice without it influencing any part of our evaluation system so we could fine tune the exams and tests and components; whatever we were using at the time, find tune them.
The students are the proof that you’re teaching and you’re doing it well. Rather than somebody deciding that they don’t like where you stand in the classroom, or you move around, determining that you’re a good teacher or not a good teacher. It doesn’t really make sense.

I’ve had to explain to my evaluator this is what engagement looks like when you’re using right brain. No, they’re not talking, because their fuckin busy. Then I’m like come on in on critique day if that’s what your definition of engagement is. Because I’m gonna tell you there’s two different ways to be engaged in art room, critique, and worktime, and they look different. So that part I’ve had to defend myself on.

I think if I had a leader when we made that summative test who was to tell us what to do with this test, and tell us how we could use it across the whole district, I think it would make so much more sense.

When are you ever going to do a one-day project? Never. So, I have to make these fake projects that are one-day projects where they see a beginning and then they see a result.

I mean the process is a little silly. You test them in the beginning but they’re brand new to the class. So, you’re testing them to say, you don’t know much. And then in the post-assessment you’re asking them all the same questions, but in a different format, and then then they are able to answer them. So, you see that growth and you can validate what you’ve been doing in the classroom which makes complete sense on paper.

The data and the digital portfolio that the students actually produce are two profoundly different animals.

The teachers skew the data, whatever way they want to. If they know what they’re doing. Test and retest.

And now the last one gave me proficiency but I’m tenured at that point so I wasn’t really going argue it.

Because it’s not about us after all. It’s about students of all levels being able to perform, being able to produce, and being able to show visual art development over a period of time. Which really should not be reduced to a number or data.

As a team of art teachers, we had to fight for our voice to be heard. And we asked to have a meeting with district administration so that they understood how we were going to assess because we were asked to do something other than what fit for us.

We were asked to do multiple-choice tests as department exams before we did portfolio examinations. We asked to change. The multiple-choice tests really didn’t fit what we were doing anyhow, and they became laughable with what we were asking students at the end.
We hear people say how important the arts are in education. I just think that people talk about it, but the lack of funding and level of importance has caused me to get more involved with my union because there’s more of a voice sometimes outside the district.

I did have to be very vocal about the inclusion of the performance aspect of portfolio part of it. It took a lot! It took a few of us on the committee to really push and to make sure that was included in our revised Danielson evaluation tool.

I don’t see a strong relationship between the data; I can’t assess that and give you a number, or somehow quantify that. I can’t do that. I can’t see a direct relationship between the data and the actual visual art portfolio that students produce that demonstrates student growth.

And this is where being too honest sometimes can cause more problems. I can claim anything is growing.

<table>
<thead>
<tr>
<th>Job Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some of the interview participants demonstrated and/or described high amounts of job stress. This was more frequent with participants who taught in schools with high numbers of minority students from low socio-economic demographics. These participants described micromanaging by administration, feelings of mistrust, more pressure regarding data collection, more expectations to emphasize reading skills, and greater expectations to assess low-level learning. Vivian described how she was assigned a reading coach to create an elementary-school-style word-wall for her high school art classroom. Her coach suggested she do this when she was being evaluated to win the favor of her administration.</td>
</tr>
</tbody>
</table>

So, I gave all the kids little envelopes full of all the words from the vocabulary and had them put the vocabulary words on the artwork wherever it applied. [pauses with a look of disgust and frustration] And oh my God, the administrators were so impressed I was the star of my building… Even the kids were like, why are we doing this? And they
[administration] were so impressed... And then when the kids were talking to each other about [their artwork] ... they were so impressed that the kids were talking to each other using the vocabulary. Of course, [the students] learned that without that word wall. They learned it just from being an art class. But because we were doing it with those little words like we were in first grade, the administrators were blown away… They didn’t even look at the kid’s artwork… [the administrators] kept asking us how do you know the students have learned what you want them to? And I keep saying… I can see it in their artwork. That’s how I know if they learned it. And all they want is to quantify that in numbers.

Lisa described considerable amounts of job stress related to her evaluating administrator not having any background in the arts. She never had an administrator with any arts background throughout her career; however, since the changes in the teacher evaluation process took place, she has had to complete more paperwork than ever and felt she was treated unfairly when compared with her colleagues who teach in other disciplines. She described a humiliating experience when she was meeting with her PLC to discuss data; this was supposed to be a private time for art teachers to meet alone and work as a team independently. An administrator walked into the room unannounced and sat in the back of the room taking notes on their meeting. This led to feelings of hostility, and she viewed it as being babysat because they could not be trusted to do her job.

Cyndi expressed frustration with the way her district kept changing initiatives every year, making it difficult for her to align her teaching practices and assessments to what the school wanted. Because of the lack of consistency and minimal direction she was given, she often felt lost or left out. When she tried to participate in school-wide meetings to get more information about initiatives, she said she got the “brush off” because she was “just the art teacher” and not welcomed into the conversation of improving things for the benefit of the students. This made her feel like she was not respected as an equal member of the faculty.
Similarly, one of the anonymous survey respondents expressed her concern about how hard she and her colleagues work in the art department, however, does not feel respected as an equal member of the faculty:

The visual arts do not carry a high value in our school. We are an elective that is often used as a “dumping ground” of sorts to put students in who need to fill a schedule. We are not valued or respected by other department areas. We work hard to create a strong and viable program, work with the community, enter competitions, put up major exhibits, and hold ourselves to very high standards. We are innovators in the school in our assessment approaches. Our administration values the arts when we make the school “look pretty” or “sound great” but ignores us otherwise. This attitude permeates the rest of the school and makes it hard to push forward. Despite our inherent worth and significance, we are fighting to stay alive. A lack of worthy teaching candidates is one part of the problem, as is the overemphasis on STEM, without the A. I am concerned that as we veterans retire, we will not have valid candidates to take up the fight and keep the visual arts alive in the schools. Not to be too melodramatic, but it is a dark world for the visual arts at this time.

Lily worked in an affluent school but discussed unique situations many art teachers experience that teachers in other disciplines do not.

The hard thing about art is sometimes you teach so many different classes. I had one evaluator ask me how many preps I had, and I told him I had six. He said to me if I were a chemistry or math teacher, I would only have up to two. I really didn’t know what I was getting into when I added a new class, AP Art History. I had to lecture and teach about 250 different works of art, and that was on top of all my other preps. I don’t know how other art teachers who have families juggle this the whole school year. When I talked to my union rep and told her I felt like I was being taken advantage of she said there was nothing in the contract that could help me. She said, ‘What do you want to do, ruffle feathers?’ So, I got scared and backed down.

She did not feel any of the extra things art teachers do – extra assignments, putting up art shows or displays – are recognized in the performance evaluation process, which added to how she felt the arts were underappreciated.

Contrary to the stress and frustration many of the other participants expressed, Charles described working in a good situation, having confidence in his fellow art teachers, and appreciating that the administration treated them like professionals. He said his administrators
told the staff that they are great at what they do and the teachers are respected. He described an example of how trust was demonstrated by administration:

> It’s not that they ignore it, [student growth data] but you know what’s most important to this administration? What we’re told on a daily basis? It’s more important to give the students a comfortable place where they are made to feel welcomed, loved, and cared for, and create an environment where they feel safe, and everything else comes second. And I’m so happy to have that.

Charles was aware this was an unusual situation and was very appreciative of it. He said most of his experience at this school was usually like this; however, he had one principal who caused tremendous amounts of stress and damage to the curricular offerings. He said she had no background in the arts and used her authority to change everything in the art department from the course offerings to the way the art office was laid out. Without consulting the art teachers, she collapsed all the advanced level classes to just one general art class. “To me this was a horribly backwards move… To this day all of my fellow teachers, we agree that she was a horrible administrator. It was a dark time for the school.” He was very happy that administrator was gone, but it took his department over five years to recover from the damage she caused and to build back their upper level classes.

This chapter described the qualitative results from the study. The themes and patterns that emerged from the data indicated each of the interview participants had very different experiences with their teacher evaluation process. There were many consistencies in the kinds of assessments they used in their classrooms; however, the directives they were given by administrators differed greatly. Most of the participants struggled with the reliability and validity of their assessments, and although most of the participants described having professional communities they could work with, the quality of professional development they received varied greatly. Each of the participants had a different understanding of the policies that shaped the experience with their
performance evaluation. Finally, many of the participants’ overall experiences were shaped by the socio-economic and racial makeup of their school. These experiences also produced manifestations of resistance and job stress. The next chapter presents the conclusion to this study, including application for the findings and recommendations for future research.

Conclusion

Through analysis of the qualitative data it became evident that the SES of the participants’ schools shaped the kinds of experiences these teachers faced related to fulfilling requirements for their performance evaluation. Under the same policy within the same state, there were variations in the way participants were judged for their performance evaluation, differences in the kinds of assessment data teachers were allowed use to demonstrate student growth, and variations in the rubrics school districts utilized for teacher performance evaluations. All participants, regardless of their schools’ SES, struggled with demonstrating the validity and reliability of assessments as well as limited access to quality professional development specifically for art and design educators. As educators in the field of art and design, many of the participants expressed doubt about their own professional expertise because of an expectation in many of the participants’ schools to emphasize CCSS as a measure of student learning. Although patterns of resistance frequently emerged throughout the data, it was most evident in ways interview participants expressed work-related stress; forms of resistance manifested overtly, covertly, or in instances when participants felt completely worn down by directions they felt were not in the best interest of their discipline, they sometimes responded with passive compliance. The next chapter will synthetize the data collected from this research and provide recommendations for the field.
CHAPTER 6

CONCLUSION

The purpose of this study was to examine how Illinois high school visual art and design teachers fulfilled requirements for assessment. In the context of a nationalized shift toward standardization in education, Illinois teachers are required to produce data on student growth from valid and reliable assessments as a component of their performance evaluation. Best practice in visual art and design education suggests authentic and alternate forms of assessments as the most appropriate means for judging the quality of student learning in art and design. Therefore, the researcher used both quantitative and qualitative methods to analyze the participants’ views and practices of assessing high school student art work within the context of each participant’s unique educational setting. This chapter includes a brief overview of the study, analysis of the findings, and how these assertions helped answer the research questions. This chapter also includes possible implications and suggestions for future research.

The participants in this study were limited to high school visual art and design teachers in Illinois to help understand how the nationalized shift toward standardization had an influence on Illinois policy and teachers’ assessment practices. This study analyzed issues related to quantitative forms of assessment to judge student artwork used by participants. Furthermore, this study considered teachers’ responses to educational policy and administrative guided by teacher evaluation policy that includes student performance data. To fully examine the many layers related to this problem, it was necessary to use a mixed methods approach to collect data from
multiple sources to help answer the questions (Mertens, 2010). The quantitative research was able to collect information from a large number of participants throughout the state of Illinois; however, the respondents chose from closed-ended responses. Although survey respondents had an opportunity to provide an open-ended comment at the end of the questionnaire, the quantitative data was limited to response choices. To help explain the context of responses and provide rich description of teacher experiences, a smaller number of participants was included in the qualitative research, allowing for rich descriptive data. Unlike many current studies in the field of art and design education that have emphasized qualitative research, applying a mixed methods design honored the rich descriptive attributes of qualitative analysis yet was able to apply a quantitative method to collect data from a larger population of participants. Merging these two methods and triangulating data sources helped neutralize the weaknesses, and the qualitative data were able to explain the quantitative data (Creswell, 2014). This study examined one research question and four sub-questions. Answers to the research questions and additional problems or suggestions that emerged as a result of this research will be posed to the field.

Research Question

How are high school visual art and design teachers expected to utilize assessment practices by their school districts in the context of a nationalized shift toward standardized assessment?

Because of shifts in educational policy, standardized assessment and teacher performance are becoming inextricably linked. In this context it is important to understand the precarious position art teachers may be in related to the high stakes placed on assessment. The movement toward Common Core Standards has reduced the curriculum to easily assessed content like reading and math, which leaves little time for higher level or creative thinking (Ravitch, 2014). To understand how these nationalized shifts have influenced state policies, this study limited its
focus to the state of Illinois and local policy. In Illinois, Common Core Standards were adopted as the state standards. Illinois also adopted a teacher evaluation policy known as PERA that requires teachers to provide assessment data on student growth, and teachers of all disciplines are held accountable for school-wide student performance on standardized tests (PERA, 2015). With regard to the visual arts, Illinois also adopted a version of the National Core Arts Standards (NCAS) as the new Illinois Art Standards (Illinois Art Standards, 2016). However, although the fine arts are listed in ESSA (ISBE 2017) as an indicator of a school’s performance, they do not currently have a determined value; leaving the arts with an undermined status of importance in the schools. To corroborate this, something that was discovered when trying to find equitable representation of participants for this study is that not all high schools in Illinois have visual art education. To help uncover answers to this research question about how the nationalized shift toward standardized assessment has guided expectations of visual art and design teachers to assess students is to understand that not all students have access to quality art education. Thus, an underlying assumption exists that even though there are standards for the visual arts; the CCSS seem to be the overarching standards that are generally accepted when discussing standardized assessment. This presents the context of the problem art teachers may be experiencing when assessing student artwork during a nationalized shift toward standardized assessment.

When analyzing the data collected for this study, results from the survey questionnaire suggested some participants were expected to assess using standardized methods, while others cited performance-based methods that presented a contradiction that needed to be explained. When the interview participants discussed having to meet requirements for assessment, it was sometimes explained they were asked to assess things that were easily quantified through quizzes or checks for understanding or other knowledge level indicators like vocabulary, which signified
this was the emphasis of learning in their school. Ironically, despite the use of these kinds of assessments, the participants did not consider these a good indicator of meaningful learning in art. Instead they were used to satisfy requests for data collection of student growth in a way that was easily understood by an administrator who had no background knowledge in art education. This was consistent with Ravitch’s (2014) assertion that movements in public schools emphasize CCSS, meaning the non-tested disciplines that allow for creativity and innovation are ignored. None of the eight interview participants mentioned the NCAS when discussing how to generate art assessments. This had much to do with the participants reporting administrators valued CCSS more than standards in the art, explaining some of the contradictions in the quantitative data about the usefulness art teachers found with analyzing school wide standardized assessment data versus the emphasis administrators placed on this.

The demographics of the schools in which the participants worked was related to the degree of emphasis placed on standardization. Interview participants who worked in low SES schools were asked to produce standardized types of evidence of student growth that aligned to CCSS more often than participants in affluent areas. To explain this, the schools that had low SES also struggled with achievement on school-wide standardized assessments and the faculty were under pressure to close the achievement gap. These participants reported lost instructional time and resources to help make up for school-wide accountability deficits. Because participants who worked in schools with high SES populations exhibited better performance on school-wide standardized assessments, they experienced less pressure to lose instructional time to meet accountability measures and were encouraged to focus their time on teaching and assessing within their own art and design curricula. These findings support assertions made by Chapman
180

(2018) and Ravitch (2014) and further illustrate the problem with inequitable access to quality art education in the state of Illinois.

Regardless of the demographics of the participants’ schools, the interviewed teachers did not feel the changes to the teacher performance evaluation process that included assessment data as a measure of their job performance helped any of them become better teachers, nor did they feel it was an accurate reflection of what actually happened in an art classroom. This was clearly indicative of policy-based decisions made about public education assume everything students learn can be measured with a test and have ignored the complex process of teaching as a whole, especially in a performance-based discipline like the arts that embraces novel student responses (Eisner, 2002). Throughout the qualitative data, the participants expressed frustration with assessment techniques in visual art and noted struggling to fit into standardized ways of assessing or feeling less respected when compared to teachers in other disciplines. This was largely attributed to an underlying assumption that assessing student achievement through a written test or right-wrong answer choice was a more respected measure of student learning than a form of assessment that requires teachers to make judgments about their students’ work. It became evident that the expectation to standardize assessment in the arts was highly problematic for the participants, especially because assessment in the arts should not be standardized (Boughton, 2013).

Going deeper into the data and considering the analysis of school district’s PERA plans became pivotal to understanding the communication breakdown of assessment and what could be considered acceptable methods. None of the teacher evaluation plans analyzed for this study addressed the assessment needs of art teachers or of any other performance-based discipline. Subsequently, the underlying assumption that assessment must be in a right/wrong question-
answer persists from the NCLB era. Freedman (2014) expressed a concern for the repercussions from the No Child Left Behind Act (NCLB) and how it shifted the curriculum to reading and math; this study has provided evidence that the discipline of visual art and design education is still experiencing negative effects of test-based policies. From the experiences of the participants in this study, a narrow view of the curriculum is still evident today in schools that continue to struggle with test scores, placing a priority on learning what is considered the basics and ignoring the kinds of learning the arts offer students.

**Sub Question 1**

What types of assessment strategies do Illinois high school visual art and design teachers report using and their classrooms?

Literature suggests performance-based assessments that rely on the judgment of the quality of student work are the most appropriate forms of assessment in the field of visual art and design education (Beattie, 1997; Boughton, 2013; MacGregor, 1992). However, as part of the educational reform movement, the United States has increased its dependence on standardized assessments to measure student learning (Kuhn, 2014; Popham, 1999; Ravitch, 2014). The movement to rely on standardized assessments has diminished the value placed on authentic learning experiences and decreased opportunities for creativity in the classroom, further limiting teachers’ opportunities to promote academic risk-taking (Chapman, 2018; Ravitch, 2014; Freedman, 2015). With regard to these disparate positions on how to measure student learning in art, this question examined the kinds of assessment strategies Illinois high school art teachers described using in their classrooms and how useful they found these assessments.

A majority of survey questionnaire respondents and interview participants described performance-based assessments as the form they preferred to make judgments about the quality
of student work. Among the many kinds of assessments visual art teachers in this study utilized, interview participants described critique as the form of assessment they used most often and felt was most valuable for measuring student learning. Although each interview participant varied slightly on the ways they conducted critiques, they found this method helpful when monitoring their students’ growth. They all agreed that providing students with usable feedback while they were in the creation process was something their students could apply to their work to improve it before the work was complete. One participant stated, “Assessment is about the process, not always about the end product.” In the context of producing assessment data for their performance evaluation, most of the participants struggled with quantifying formative assessments as part of student growth data because they did not feel critique was something they could or should quantify. Allowing students room to improve their work as a result of critique before it was submitted for a grade was important to many of the participants because it was a way to build rapport with students and help them grow without students feeling the pressure of being judged and potentially penalized for that judgment which could thwart student’s creative process. Boughton (2013) supports the position the interview participants held in that always linking a grade with the artistic learning process is a way to destroy the possibility of creativity. Much of the problem quantifying this kind of formative assessment seems rooted in administrators not understanding the difference between formative assessments from other disciplines that can use quizzes or bell ringers and a check for understanding, like a critical conversation between an art teacher and student, that is part of the creative learning process and should not be quantified in a gradebook.

Regardless of the position art teachers held about assessment, there were external demands they were required to meet as faculty members that guided some of the choices they
made about assessment strategies. Furthermore, pressures to regularly monitor students’ grades and school performance has stifled the participants’ opportunities to promote creativity or risk-taking especially in schools with struggling with test scores for fear of being penalized on their performance evaluation. Although the participants seemed to understand that not all learning had to result in a grade; heightened pressure on school accountability and requirements to quantify student progress, the participants described many instances in which they used formative or superficial assessment data to satisfy requests from their administrator.

Some examples the interview participants gave about assessments related to superficial knowledge rather than meaningful learning in the arts usually related to vocabulary assessment in some way. These assessments were useful as a check for understanding at the start of a unit but not appropriate for a summative assessment. Many interview participants shared these types of assessment with their administrators – not because they felt it assessed any quality of art performance, but because it was easily quantified and seemed to be something administrators understood. This was indicative of an underlying problem that prioritizes easily calculated numerical data collection over more thoughtful forms of assessment that consider multiple indicators as evidence of student learning. Eisner (2002) cautioned against perpetuating this type of practice: “A school system designed with an overriding commitment to efficiency may produce outcomes that have little enduring quality” (p. xiii).

Other kinds of assessment techniques the participants described included rubrics like self-generated, department-wide, or rubrics with visual benchmarks. Although rubrics were the assessment method interview participants described as most useful to quantify finished student work, the participants felt this method was either not valued by administrators or judged by administrators as being too subjective. Some participants had department-wide rubrics they felt
were more trusted by administrators; however, they experienced difficulty implementing them because they did not always provide the flexibility they needed when assessing creative or unexpected outcomes with student work. One of the interview participants explained that since the implementation of a department-wide rubric, she noticed the creative outcomes of students’ assignments had noticeably diminished over the years.

Overall the participants in this study described the importance of performance-based assessment strategies as most useful to judge the quality of student work, but these sometimes conflicted with what they felt could be used as data for the student growth component of their evaluation. Many participants seemed overwhelmed and frustrated with the number of tasks to complete for their performance evaluation, and assessment seemed to be at the pinnacle of an endless workload that has taken joy out of their careers. Tired of feeling as if they were fighting an endless battle, some participants found it difficult to assert they were in the best position to choose the assessment method that best measured student art learning. Within the context of answering this research question, further questions arose about assessment and grading that may need future research. Although the variety of assessments the participants used was generally qualitative, these were in conflict with the expectations of administrators regarding the form of assessment required to satisfy their performance evaluation. The next sub question examined in more detail the types of assessments administrators looked for from art teachers.

Sub Question 2
What directives are Illinois high school visual art and design teachers given by their school district administrators about reporting assessment data?

The reason for examining the kinds of assessment directives administrators issue to teachers is rooted in the current Illinois teacher evaluation policy that holds all school personnel
accountable for student assessment data (PERA, 2015). Although it is known that test scores are not an accurate reflection of the quality of a teacher’s performance (Popham, 1999; Ravitch, 2014), the policy-driven mandates that guide school officials’ decisions seem to ignore the nuances involved in quality teaching, causing them to direct their efforts to data collection and reporting results rather than to staff development efforts that will ultimately help student performance (Chapman, 2007; Ravitch, 2014). Creating a problem for visual art and design teachers at the high school level is the assumption that student performance data must be numerical and derived from some type of right/wrong answer is symptomatic of the data-driven culture promoted under a series of American educational polices that rely on test to measure school performance (Ravitch, 2014). Illinois policy has inferred data should be derived from numerical results of tests; however, school districts have local interpretations of PERA to guide their teacher performance evaluation guidelines.

This option has presented school districts with the opportunity to determine what forms of assessment are best to demonstrate student growth. None of the evaluation plans analyzed for this study addressed any performance-based or qualitative assessments that could be used by teachers in performance-based disciplines, including art. From this, it is clear that the directive about required assessment data is that right/wrong answer choices are valued and assessments that easily collect numerical data are preferred.

When seeking an answer to this research question, a pattern in the quantitative data suggested administrators had conveyed information about the importance of utilizing standardized assessments, which contradicted the participants’ preference for performance-based forms of assessments. More information needs to be collected to explain this discrepancy. Interviews helped explain why the kinds of assessments art teachers find useful in their content
area differed greatly from administrators. One potential reason for the difference in responses is interview participants explained their evaluating administrators had no experience in the visual arts. This seemed to cause difficulties from an administrator’s perspective understanding what appropriate assessment for art learning should look like. This presented problems for some participants as they tried to assess creative student outcomes or received conflicting messages about data contrary to assessing the quality of students’ artwork. Keeping in mind the primary reason for performance evaluations is to help teachers improve their practice, and it is most useful to receive feedback from another professional who understands the discipline (Eisner, 1998). The lack of administrative guidance from a professional in visual art and design education is clearly an issue that emerged.

When analyzing the teacher performance evaluation documents from four different school districts from Illinois, there were no provisions for teachers in performance-based disciplines, including the visual arts, or specific explanation about qualitative judgment of student work. This helped explain some of the uncertainty the participants had about the forms of assessments they should use in their practice, especially when trying to meet administrators’ expectations for data collection. Additionally, if no provisions or explanations were provided about performance-based assessments in the teacher evaluation plans, the administrators were likely to not understand what the teachers intended by using an alternate assessment or may have been unable to support these teachers through lack of knowledge about the assessments.

When the participants received direction from administrators that suggested using some form of right/wrong answer choice assessment to report data, some participants tried to explain and/or demonstrate what assessment strategies worked best in the art room. However, depending on the situation of each individual school and the administrator’s perspective, this sometimes led
to contentious situations between the teacher and administrator that led to finding the patterns of resistance described in Table 3 and in more detail later. Although some of the participants were able to successfully communicate performance-based assessment information with their evaluator and it was well received, other participants had to outright fight with their administrator, which led to participants describing how they overtly resisted administrative directives. Alternatively, they utilized deceptive tactics like the ones described in their descriptions of covert resistance. In short, inappropriate assessment requirements were used for the sake of their performance evaluation. This became a point of contention for many participants when they expressed deep frustration about not being valued or trusted as the professional in their discipline. Ultimately these participants had to fulfill requirements for their jobs, which included reporting assessment data, which out of practicality led to patterns of passive compliance. Within the context of how resistance manifested throughout the study, some participants began to doubt themselves and their own professional judgment about appropriate forms of assessments for visual art when those forms of assessment differed from what their administrator directed, further suggesting a loss of self-efficacy.

A few participants suggested their administrators did support their performance-based assessments and felt their administration supported creativity in the art room; however, these participants worked at high SES schools and there was less contention about administrative direction about assessment. This brings up the issue of creativity and appropriate ways to assess it. Although student learning in the visual arts can help foster the possibilities of creativity (Freedman, 2010; Katz-Buonincontro, 2018), it cannot develop instantly; it must be nurtured over time (Csikszentmihalyi, 1996). Boughton (2013) cautions that creativity is not something that can be assessed in a standardized way. The disconnect that emerged with some
administrative directives about assessment and the way some participants felt they should assess
created a point of contention for the teachers who wanted to encourage creative solutions in their
student’s work but felt stifled when right/wrong answer choice or assessments related to CCSS
were promoted by the administrator.

Historically the field of art and design has struggled with assessment issues (Wilson, 1970), but now in the context of a movement toward standardization, what art teachers are
instructed to do by administrators may be different from what they know as best practice for their
discipline. When trying to reconcile the differences between what administrators were looking
for in assessment data and what art teachers knew they should be assessing in student work, it
became evident that school wide test data related to CCSS were placing the visual arts in a lesser
role because it is a non-tested discipline. This pattern suggests a larger problem in the American
educational system that has narrowed the conception of assessment to testing and student growth
data to something derived from students choosing from a predictable outcome rather than
allowing open-ended responses from students to glean a more comprehensive picture of learning.

However, the pattern that emerged connecting the SES of the school in which they worked to the
kind of assessment directives they received from administrators is notable. The interview
participants who worked in economically struggling schools with higher numbers of minority
students reported concerns with overall low school wide test scores. Because of this, the faculty
in the school were under pressure to report assessment data related to reading and math by their
administrators in an effort to demonstrate support of raising student scores regardless of the
discipline the teachers taught. This also helped explain unexpected patterns that emerged in the
quantitative data when the survey participants described how often evaluating administrators
emphasized test data related to CCSS rather than data in their own discipline. The participants
who worked in more affluent schools did not have the same problem as the other participants related to school wide test performance and were actually able to devote more of their school day to teaching and assessing art. The participants who worked in high SES schools expressed less stress and pressure related to administrative direction about assessment, and generally felt their professional judgment was respected by their administrator. This points to the achievement gap in student populations and how it may shape the professional experience of teachers in Illinois. These drastic differences in the teachers’ professional experiences based on the SES of their school is well documented on a national scale (Kuhn, 2014; Ravitch, 2014).

A further challenge many participants explained regardless of the SES of their school was their evaluating administrator did not know how to interpret assessment data related to creative art production, so the data from performance-based assessments were sometimes inadmissible for teachers’ performance evaluations. Although some teachers were trusted and supported by their evaluating administrators, others explained that when these forms of data were disregarded by their administrator for their performance evaluation they felt mistrusted or less of a professional than other teachers in traditionally tested disciplines. This presents a dilemma for administrators who are required to uphold policy and are held equally accountable to implement their individual school district’s teacher performance evaluation plan. It is difficult for administrators without experience in alternative assessments to understand assessment data that are not presented in a standardized format. Illinois teachers and administrators are not alone in this dilemma. Similar studies of art teachers in other states that also have assessment-based data accountability measures in place (Hanawalt, 2018; Jackson Goodwin, 2015) described how their participants felt pressure from evaluating administrators to collect assessment data unrelated to art. This indicates a potential need for professional development to help teachers and...
administrators work together on assessment strategies appropriate for the discipline. The next question examined professional development opportunities provided specifically for visual art and design teachers.

Sub Question 3
What professional development about assessment is provided for Illinois high school visual art and design teachers by their school districts?

Professional development is important for enhancing the quality of a teacher’s practice, but as outlined in this study, visual art and design teachers have specific assessment needs that are different from colleagues in tested disciplines. Professional development directed specifically for the kinds of performance-based methods used in visual art and design education must be addressed. The survey questionnaire indicated that about half of the respondents received some kind of professional development for assessment specifically focused on their discipline. However, when analyzing the qualitative data, most of the art teachers who participated in the interviews clearly lacked professional development specific to assessment in visual art education to help them successfully address student growth data collection requirements related to their performance evaluation. This discrepancy indicated a gap related to performance-based assessment and appropriate professional development to support teachers specifically in art and design. The resources provided by the Illinois State Board of Education for school districts to support teacher evaluation data collection (ISBE, 2016) offer guidance about the need to provide teacher training:

Teachers may need training or practice in analyzing student data and setting student growth goals… Also, teachers may need additional training and support in analyzing student data and setting student growth goals for specific groups of students, taking into account student baseline scores. If teachers are writing their own assessments to measure student growth, they may need training on writing quality assessment. (p. 6)
This document also addresses a requirement to ensure the instruments teachers are using to assess student growth is collecting correct information and to check that “evaluators may be showing unconscious biases toward certain groups of teachers. Evaluators may benefit from training and support in countering their implicit biases” (p. 6). The document also lists resources from the ISBE for teacher and evaluator training; however, the hyperlinks from this PDF were not working when they were accessed at the time of writing this analysis. Based on the participants’ experiences, their needs were not met with adequate support for professional development in assessment nor did their evaluating administrators receive necessary support.

With consideration to an entire faculty, the number of art teachers in a school accounts for a small percentage when compared to those from other disciplines. Financially it makes sense for a district to provide professional development on assessment for the majority of staff members; however, within the context of current policy, when assessment data could be a determining factor for teachers’ employment or tenure, it becomes an equity issue to ensure all faculty of all disciplines are equally prepared to assess within their discipline. Unfortunately, the problem of schools not providing professional development for assessment in visual art is not a new issue. Over 25 years ago Wilson (1992) urged the field of art education to become more astute about assessment, citing the challenges that would be faced with providing adequate professional development for arts educators because of the time and expense involved. It appears the urgency of this issue is becoming more apparent for all art teachers because of the high-stakes nature of assessment.

As explained by the interview participants, any professional development they were provided related to assessment in art was minimal, if any at all. This became a point of
contention for many of the participants and an equity issue for teachers judged on assessment data as part of their performance evaluation: If the professional development in assessment only addresses one form of assessment that works best for easily tested subjects, the teachers in performance-based disciplines are at a disadvantage. Only one interview participant who worked in an affluent district described receiving actual professional development specifically for the art department from his district. Most of the interview participants expressed a sincere desire for more direction or help with assessments, guidance on how to collect assessment data, or how to take what they knew as evidence of learning in the art room and express it in a way that could be understood by others to improve their practice. However, this was not always provided and there seem to be limited resources to meet this need in some of the participants’ districts. Some of the interview participants described seeking professional development outside of their schools because what was offered to them in their schools did not apply to the assessment of art. One participant who taught AP Studio Art suggested the professional development they received to teach AP was helpful with that level class, but difficult to apply with introductory level courses. Not all teachers in this study taught AP, so this kind of external professional development was not equitably offered to all teachers and not all schools offer AP Studio Art courses. The IB program also provides professional development in art assessment for their teachers; however, none of the interview participants taught in an IB school and, similar to AP, not all schools had an IB program, further limiting opportunities in professional development specifically for teachers in visual art and design education.

Some participants mentioned attending IAEA (Illinois Art Education Association) or NAEA (National Art Education Association) conferences for additional professional development specifically for art; however, this was usually at the expense of the participant and
required them to miss school days. A systemic concern about professional development related to assessment was also evident in the analysis of teacher evaluation plans for the school districts. Although the plans addressed assessments, none of the plans specifically addressed art discipline-specific forms of assessment; instead they suggested assessments that produced numerical data from right/wrong answer choices. This signifies a serious need for quality professional development for not only art teachers, but also evaluating administrators so they can appropriately support their teachers and ultimately benefit the students. This is not as daunting as it may seem. By looking to the field of art education in an international context, there are many assessment models available from other countries that employ alternate forms of assessment American schools could learn from (Beattie, 1992; Boughton, 1996; 2004b; 2018; MacGregor, 1992). There is much that can be learned from international counterparts in art education with relatively easy accessibility to this kind of global information. Related to concerns about a lack of professional development in assessment, another concern some of the interview participants expressed was uncertainty about the validity or reliability of their assessments, and the absence of professional development to establish validity and reliability, which is described in the next sub question.

**Sub Question 4**

What professional development about determining validity and reliability in assessment is provided for Illinois high school visual art and design teachers?

Similar to the findings in Sub Question 3 that demonstrated a lack of professional development for assessment in the field of art and design education, there was also a lack of professional development provided to art teachers to help them determine the validity and reliability of assessments. The components of an effective assessment involve validity, which
requires using methods that actually assess what the students should learn, and reliability, which means the method would produce similar results if repeated in similar conditions with a similar group of students and examiners (Boughton, 1996). This study revealed establishing validity and reliability for assessments is problematic for high school visual art and design teachers in Illinois because, according to policy, this is a necessary requirement for their performance evaluation (ISBE, 2016; PERA, 2015).

When analyzing the quantitative data, about half the respondents stated they had received some kind of professional development related to validity or reliability. However, the qualitative data revealed a lack of confidence in the participants’ understanding of validity and reliability. Because the participants described the ways administrators promoted forms of assessment that are inappropriate for art education, it also suggests evaluating administrators lack understanding of conditions that foster validity and reliability with assessments in visual art and design education. A majority of the interview participants were able to articulate some understanding of valid assessments by describing the kinds of art assessment strategies they used regularly in their classrooms; however, some of the participants who described implementing valid assessments doubted their validity because evaluating administrators had difficulty understanding why these assessments were valid or would not accept performance-based assessments as evidence of student growth data. Looking deeper than the participants’ understanding of validity, this is also a clear indication evaluating administrators are also unclear about valid assessments for visual art and design education.

Reliability was much more difficult for the interview participants to explain because they often worked by themselves when assessing student work and did not receive appropriate professional development specifically for art. Thus, they were the sole arbitrator of scores, which
diminishes any possibility of establishing reliability of assessments. In a few instances, a couple of interview participants who worked in more affluent schools were provided with professional development by their school districts to work with colleagues specifically on these strategies. They described how groups of teachers helped establish visual benchmarks for grade rubrics demonstrating the potential for reliability. Only one of the interview participants who taught AP courses understood her students’ AP scores were scored outside of her school and had established reliability. Although she had the opportunity to submit data from AP scores as part of evidence of her professional practice, the data for student growth for her evaluation were due before the end of the school year. She did not receive AP scores until the summer, subsequently limiting the opportunity to submit valid and reliable data for her performance evaluation.

The lack of professional development specifically to help art educators establish reliability is the most problematic for the future of art and design educators in the context of this study. Beattie (1997) and Boughton (2018) strongly asserted that without reliability, an assessment can never be valid. Thus, without any professional development or support for art teachers or their evaluating administrators from school districts specifically for determining validity of assessments and promoting the conditions for reliability, the field of art and design education will continue to experience difficulties with assessment. Similar to the results to Sub Question 3, the American educational system is missing opportunities to learn from international models that have well established systems of reliability of visual art assessment (Boughton, 2018). AP and IB programs within the United States are already utilizing a moderation process to score student work that demonstrates reliability (Filbin, 2019), and in an international context, programs have implemented these kinds of models for many years, so there is an existing foundation that the American educational system can build on. Incorporating professional
development that could help local school districts establish reliability would not require creating a new method but could build on and/or adapt existing models for local use.

The lack of teacher and administrative understanding of validity in art assessment and an understanding of how to foster conditions to create reliability places the field of art education in serious jeopardy under current policy-based conditions; this further places teachers in an uncertain position for maintaining their employment if requirements to evaluate teachers based on assessment data continue. Furthermore, the lack of equitable access to art education throughout the state of Illinois demonstrates the low priority of providing visual art and design teachers with the tools necessary to be successful, indicative of systemic problems valuing quality art education for all students. Art and design teachers will continue to be at a disadvantage compared to their colleagues from tested disciplines (Wilson, 1992), which can ultimately mean an art teacher could lose their job because of the inability to establish validity and reliability and the lack of district support to develop an understanding of these conditions. The themes that emerged in this study are worth noting as implications for the field and recommendations for future research.

Implications and Recommendations

This study examined issues regarding policy, assessment, practice, validity, reliability, and how they are interpreted by practitioners in the field of art education. Although these issues are relevant and need attention, unexpected findings surfaced that warrant further discussion. Unlike other recent studies in the field, this research was unique because of the utilization of mixed-methods to uncover the complex layers of a problem. The quantitative data collected information from a large number of participants throughout the state, while qualitative methods
were able to closely examine and honor the personal experiences of practicing art teachers and give voice to a population overlooked under policies that promote standardization. The depth of personal experience the interview participants openly shared humbled me as a researcher. Their honesty, passion, the obstacles they sometimes faced, detailed descriptions of how they tried to do what was best for their students, and their emotional outpours brought me to tears many times. When reviewing my qualitative data, I had to stop to compose myself and revisit the data later with fresh eyes. This experience kept me mindful that teaching is a very human act and any attempt to standardize the process of teaching art is simply counterintuitive to the discipline.

The following implications and recommendations for future research are being made for the field of art and design education. Some of these implications include the need to address the lack of equitable access to quality art education throughout the state; the need for strong professional communities in the field that can help foster continual professional development for current professionals and support pre-service teachers in the area of assessment; providing more opportunities to help stimulate conditions that promote creativity and academic risk-taking in art classes; building strong advocacy and leadership in the field; and addressing the pattern of resistance and its implications for professionals.

1. The lack of equitable access to quality art education for all students needs to be addressed.

The differences in the quality of teaching experiences among the participants based on SES were undeniably. Resonant of national trends that indicate a widening achievement gap in the United States (Chapman, 2007; Kuhn, 2014; Ravitch, 2014) is also a gap in equitable access to quality arts education for minority students and those from low SES (Krahe & Acuff, 2013; Gonzalez Stokas, 2016; Katz-Buonincontro, 2018). This pattern was closely linked to the
experiences the Illinois participants of this study shared when they discussed the varying opportunities students had for arts in their schools and the different conditions they described that were part of their performance evaluation guided by current educational policy.

All the interview participants were at different stages in their careers and had varying approaches to art education; however, half of the participants worked in high SES schools and the other half worked in low SES schools. An overarching pattern that emerged throughout the data from the face-to-face interviews was a difference between the SES demographics of the schools the participants worked in and the conditions of their work environment. Specific to the interview participants in this study, teachers who worked in schools that performed better on school-wide standardized tests also had populations of students from high SES demographics. These participants described an overall good working relationship with their administration, described feeling trusted as a professional, had more art educational resources, felt valued as a member of the school community, and were allowed more autonomy in their art teaching and assessment practice. Interview participants who worked in schools with low SES demographics more often described feeling mistrusted to carry out the duties of their jobs, had fewer resources and teaching staff in their art department, expressed frustration about their lack of autonomy in implementing their art curriculum, and described more stress when communicating with administrators during their performance evaluation.

Without confronting the differences in class and the equity gap that includes access to the arts, no educational policy will help alleviate disparities (Ravitch, 2014). Previous educational policies including NCLB and RTT ignored visual art and design education as a core subject and, as a result, has not been considered a vital part of American education. Furthermore, because it cannot be easily assessed in a cost-effective manner it has been devalued, subsequently widening
the equity gap (Chapman, 2005; 2007). These policy-based issues have significant implications for the field of art and design education: Under current educational policy, ESSA contains a component that includes fine arts as an indicator of a school’s performance (ISBE, 2017).

However, because of the inequitable access to the arts as described by both SES and schools that offer art in high schools throughout the state of Illinois, there is uncertainty about how to value fine arts as part of a school’s performance in Illinois. This became apparent when there was difficulty in locating potential participants for this study from rural and less affluent areas around the state. Because not all schools in Illinois have fine arts programs, the state is currently valuing the weight of the arts as part of school performance at 0% and will re-visit this in the future (Illinois State Board of Education, 2017).

It is important to understand what the equity divide means to students’ educational experiences when it comes to equal access to quality art education. Most of the attention directed to closing the achievement gap refers specifically to reading and math and ignores the arts and other non-tested disciplines. This creates a larger problem for underserved populations who have fewer opportunities in arts education (Kraehe & Acuff, 2013). There is a correlation between SES and access to arts; the lower the SES of a school, the lower the percentage of students who receive arts instruction, with this gap reaching across social class and ethnicity (Gonzalez Stokas, 2016). Chapman (2005, 2018) and Ravitch (2014) explained some of the reasons for the inequity of American schools are schools are deemed low performing on standardized tests risk losing government monies if they do not comply with testing mandates, so schools in this risk will direct funds to improve test performance and reduce funds directed to the arts. The result is wealthier communities can offer a rich variety of curricula, and schools in poorer communities often must cut the arts (Quin, 2013). Furthermore, there is a need to communicate that the arts
are an academic discipline that should be offered to all students: The arts require complex thinking and are not less important than other disciplines because they are measured with qualitative methods (Eisner, 2002). Being mindful of existing research about the equity gap, the results of this study and the experiences of the participants suggest more research needs to be done about the lack of equitable access to art education and what that means to students, teachers, and schools.

2. There is a need for continuing professional development through strengthened professional communities in the field of visual art and design education with an emphasis on alternate assessments and how to establish validity and reliability of those assessments.

In response to policy-based changes in education that affect all educators, there is a clear need for continual quality professional development for all visual art and design educators including veteran teachers, pre-service, and administrators who evaluate art educators. Although the data collected for this study were related to current policy for teacher performance evaluations in Illinois, similar policies are in effect in other states that include test-based accountability measures (Boughton, 2018). Katz-Buonincontro (2018) concurs with the needs for teacher training and professional development for art educators. It is important to remember, however, the United States is not a homogeneous population and there are variations in teacher preparation programs (Beattie, 1992). This indicates a need for strengthening the professional community as a whole, especially to assist teachers in rural areas or those who are working in situations in which they are isolated or have limited contact with other art teachers to help create conditions for reliability for those professionals. This task is not insurmountable: For example, there is an existing American model, like the AP Studio Art exam, that utilizes digital portfolios
and has established reliability. The IB program, which is in some American schools, uses a moderation process for reliability (Filbin, 2019), and there is much to be learned from other countries that have provided art teachers with ways to establish reliability (Boughton, 2018).

Teaching has often been described as an isolated profession; many of the interview participants expressed an appreciation for having colleagues to turn to for support or assistance, and those who did not shared a sincere desire to have opportunities to build professional communities with other art educators. When trying to contact current high school art teachers in the state of Illinois for this study, it became evident there was no database for art and design teachers specifically identified by state or county. This indicated a limited ability to network directly with art teachers as a professional community outside of those who choose to join professional organizations like the NAEA. Often the cost of membership to professional organizations or expenses to attend conferences for professional development are not always funded by school districts, and school funded professional development days are not always provided by all districts for conferences, thus potentially isolating some members of the professional community. Some schools or districts have a smaller number of art teachers compared to teachers in other disciplines or are located in isolated rural areas, a database or online community for this specific population would be helpful for increasing communication about assessment.

As demonstrated in this study, other concerns the participants expressed related to their performance evaluation and data collection requirements were inconsistent direction, copious amounts of paperwork, and few colleagues to turn to for assistance with assessing student artwork, understanding how to quantify it, and determining the validity and reliability of those assessments. When analyzing teacher evaluation plans, many of the documents used complex
technical assessment language, and expectations for teachers to perform statistical analysis of data, which can be challenging if a teacher is not adequately prepared to understand the assessment techniques of their field. Although the participants in the study were experienced professionals in their field, it was evident many of the evaluators did not understand their assessment methods and some lacked the language to articulate how their assessment methods best served their students, causing some of the teachers to doubt their own professional expertise.

Since art educators in the United States have a limited history with assessment as compared to those in other countries (Wilson, 1992), professional development in assessment for art education is becoming essential during a time of test-based accountability.

An essential part of this ongoing professional development, as suggested above, is a deep understanding of the language of assessment. One of the biggest issues the participants in this study faced when communicating with evaluating administrators was a lack of confidence with their understanding of assessment methods they were implementing, and the professional technical language to communicate with their evaluator why their assessments were valid for their discipline. Furthermore, participants in the study needed a way to advocate for their discipline effectively ways to establish reliability with their assessments. This was especially challenging for teachers who had few colleagues or were isolated in their professional community as the sole art teacher. Strengthening a community that provides art teachers with the confidence in their professional expertise, and promote confident self-efficacy in the field is necessary for its survival in a data-driven policy school environment. Including evaluation in the professional development process helps build a trusting and supportive relationship among teachers and administrators and promotes a healthy professional community and a more productive work place rather than an adversarial situation. The kind of professional development
in schools that supports methods for qualitative judgments in student artwork will help establish a strong foundation to support the reliability of assessment. Furthermore, it will help support possibilities for creative student work rather than narrowing student outcomes to predictable results.

3. More opportunities for creativity and academic risk-taking for students in art classes should be fostered.

Overall, fostering creativity is a desirable outcome in art education (Freedman, 2010). In the future, jobs in the creative sector are expected to rise (Florida, 2012) and the arts become a place that can lay a foundation for students to experiment with a multitude of possible outcomes. Csikszentmihalyi (1996) supports the idea that creativity it is not something that can be taught but rather fostered by providing conditions that support it. Thus, to prepare students for their future that may require thinking creatively to solve problems (Florida, 2012), the arts become a way to teach students possibilities of what can be done (Eisner, 2002). Providing these conditions is unrealistic in an educational environment that promotes standardization and predetermined outcomes of learning. Instead Boughton (2013) advocated for allowing students to take chances and risks in their work and not penalize students who take risks and fail; educators should instead find ways to encourage students to keep trying and foster conditions for creative growth.

In response to standardization, the participants emphasized the importance of access to the arts for all students so they had an opportunity to be creative during the school day or take risks. Even if a student was not likely to pursue the arts in their future, an art class was viewed by the participants as an important way to help prepare students for opportunities in their future by learning in ways that did not always have a predictable outcome. The interview participants
described the value of allowing students to have a chance to be creative and think critically. Some explained that a heightened emphasis on testing, reading, and math, necessitated a performance-based class like art as a valuable learning opportunity in the school day.

The interview participants who worked in schools with a high population of underserved students had more difficulty incorporating lessons that fostered creativity. Some of the reasons were few art course offerings limiting the chances students had to enroll, limited budgets for supplies, few resources, and fewer staff members. Some participants expressed concerns students had fewer foundational art education experiences in earlier grade levels, creating challenges for adolescents to have confidence in their art abilities. Because these participants worked in low performing schools, they felt they were given less academic freedom to emphasize student art making and faced more pressure to demonstrate learning in their classes that was aligned to CCSS in literacy because of data needed for their performance evaluation. Many of these participants also felt external demands to help students achieve on school-wide standardized tests stifled the creative process for students and further marginalized them. In schools in which the arts are more likely to be cut, the opportunity for creativity is perhaps most valuable (Katz-Buonincontro, 2018).

Interview participants who worked in schools located in more affluent communities had more opportunity to devote their teaching time to encouraging creativity with their students. These participants also experienced less pressure to collect student performance data, thus allowing more academic freedom and opportunities for diverse art making activities appropriate to adolescents. They also described more support from the school community, and some students had opportunities to seek additional arts experiences outside of the school. Even if their evaluating administrators did not have any background in the arts, they felt their administrators
were still able to share an appreciation for the value art had for the students. The different experiences regarding opportunities to foster creativity in the classroom based on the SES of the school demonstrated further evidence of inequitable access to arts education for underserved students (Gonzalez Stokas, 2016; Katz-Buonincontro, 2018; Krahe & Acuff, 2013). Even though creativity cannot be assessed in a standardized way and academic risk-taking is not encouraged in the context of high-stakes test-driven accountability (Boughton, 2004b), it does not make it any less important to the success of students’ educational experiences. To successfully implement a model that promotes creativity and risk-taking requires courage. Similar to the manifestations of resistance by some of the participants, their actions were not for the sake of disobedience, but instead to promote what was best for student learning in the arts. This further highlights the need for strong leadership and advocacy for the field.

4. Build strong leadership and advocacy in the field.

An important pattern that emerged in the data was none of the interview participants were observed or evaluated by an administrator with any background in visual art education. They felt this created a notable hardship on their performance evaluation when compared to other colleagues in tested disciplines. Many of the participants described having to take the time to teach evaluators about the arts, spend additional time with evaluators to explain what should be assessed in the art room, and in some instances, defend best practices, which not only was an additional burden but sometimes led to confrontational situations with administrators. This added to the stress levels of some of the participants by making them feel underappreciated or not valued as a professional. Freedman (2007) firmly advocated for art education professionals to take on leadership roles. Based on this study, this is still an urgent need in the field today.
Not all leadership roles in the field require becoming an administrator. Some of the participants described taking on leadership roles in their schools, which included positions like PLC leader, being a member of the school’s evaluation committee, or pursuing active participation in the teacher’s union. All the participants who held these roles explained how they found it was important they took on these additional duties to ensure the arts and other performance-based disciplines had a voice in school decisions. Other leadership roles that do not require becoming and administrator can include membership or participation in state-wide, national, or international professional art education associations; providing professional development for colleagues as a way to share the value of the arts with others outside the discipline; publishing in professional journals and periodicals; participating in the visual art community by exhibiting artwork; and/or seeking opportunities for advanced degrees in art and design education through universities.

The information in this study is limited to the experience of a few of the participants. The absence of administrators with experience in art education surfaced as an unexpected finding; however, is clearly an important issue that needs to be addressed. A suggestion for future research is to examine how many art educators have administrative certification, if they are actually practicing as an administrator in schools, or if they have remained in their teaching roles. Furthermore, surveying a large pool of administrators nationally would build a more comprehensive picture of the need for leadership in the field and how advocacy for leadership roles in the field of art education could be promoted. Finally, it is important to examine how the lack of administrators with knowledge of art education contributed to patterns of resistance.

5. Resistance and its implications for professionals.
Resistance was an unexpected finding and manifested in multiple forms as demonstrated in Table 3. There seemed to be many factors that contributed to this pattern. Goffman (1959) explained this pattern of behavior using a dramaturgical metaphor to explain the many different parts of our self and how it shapes our daily behaviors. He compared us to performers and illustrated how the roles we play may change depending on our situations. When we play our parts at our jobs, subtle forms of resistance can become self-defense mechanisms to help us adapt when we must do required things that may go against what we may feel is actually the right thing to do. Acting out our true feelings becomes risky and could potentially result in the loss of a job. The patterns of resistance the participants exhibited, especially the covert or passive compliance manifestations were indicative of how Goffman’s *Presentation of Self in Everyday Life* still has significance today.

More contemporary scholars like Giroux (2001), on the other hand, advocate for more overt forms of resistance. However, this is riskier and tended to emerge out of extreme frustration or feeling the participants had no other choice. Acting out of frustration and/or feeling mistrusted about the assessment techniques the teachers were using for visual art education were the most evident reasons for participants to exhibit forms of resistance. This is consistent with a study by Berryhill et al. (2009) that found higher incidences of teacher burnout, reduced self-efficacy, and increased stress with teachers who worked in low achieving schools in non-tested disciplines. None of these teachers intended to be insubordinate, or were they bad teachers. Instead it was sometimes a reasonable response to a less than desirable situation.

Sometimes resistance was productive because the participants were able to make positive change in assessment practices to improve conditions for the students. Quinn (2013) explains how public-school teachers must sometimes resist to take the lead in countering the status quo.
For example, a couple of the participants described how the members of their departments were able to advocate for performance-based assessments and eliminate the department-wide multiple-choice types they were originally expected to do. The teachers had to initiate some pushback on the way they were told to conduct assessments; however, this also required these participants to assume a leadership role to initiate the change. Although there was no original research question that directly addressed resistance for this study, it emerged as a relevant finding. The phenomenon of resistance and how it manifests in the professional context of education today is worthy of future research.

Conclusion

This mixed-methods study examined issues related to assessment of student learning by high school visual art and design teachers in the context of a nationwide shift toward standardization and educational accountability. In 2010, Illinois enacted a policy that requires teachers to provide valid and reliable assessment data as part of their performance evaluation (ISBE, 2016; PERA, 2010). This new context provides significant challenges for teachers of art and design given the prevailing educational zeitgeist that values quantitative multiple-choice testing in schools throughout the nation. This study was premised on the notion that visual art and design teachers may be at a disadvantage when compared to colleagues in traditionally tested disciplines given that student performance in the visual arts is best assessed with performance-based qualitative assessments (Beattie, 1997; Boughton, 2013; MacGregor, 1992). Although this study was limited to teachers in the state of Illinois, educators in other states certainly have experienced the same kind of difficulties because similar policies have been enacted throughout the United States (Boughton, 2018).
The findings in this study have confirmed that high school visual art and design teachers experience difficulty with assessment practices in a time of a national shift toward standardization. Art and design teachers preferred performance-based methods to assess students’ art learning. They reported difficulties with their administrators and supervisors because the assessment strategies they implemented in art and design classes differed from those used by teachers in traditionally tested disciplines. In cases in which the administrators were unfamiliar with performance-based assessments and more knowledgeable about right-wrong answer choice tests, it was difficult for them to provide support for those art and design teachers who implemented their preferred qualitative assessment methods. Similarly, participants described their supervisors as valuing numerical data as more acceptable. Significantly, the findings of the study revealed that school districts provided little to no professional development to help teachers to determine the validity and reliability of the more applicable assessments.

To help improve the field of art and design education, there is a pressing need for teacher and administrator education about performance-based assessments in the field of visual art and design, supported by resources to help teachers determine the validity and reliability of those assessments. This kind of professional development and support extends to pre-service teachers, existing professionals, and evaluating administrators.

Additionally, because creativity is a desired outcome of art (Freedman, 2010), it is important to foster conditions that help promote it during a time when education emphasizes standardization. Creativity can be inhibited when there is an emphasis on teachers to produce numerical data that demonstrates student growth ( Boughton, 2018). Although creativity and risk-taking cannot be tested, it certainly can be judged to be present using qualitative assessment methods, which is one of the major arguments for rejecting quantitative assessment techniques in
the arts. Furthermore, it is imperative for visual art and design teachers to take on leadership roles and advocate for the field. It was noted in this study that none of the interview participants were evaluated by an administrator with experience in art education. This created a hardship on some of these teachers when communicating about best practice in the art room. The participants who did hold leadership roles in their schools found it important to ensure their discipline had a voice in their school’s decision-making process.

Finally, the kinds of resistance the participants exhibited demonstrated a need to further examine such behavior in the professional educational community. This was an unexpected revelation and, consequently, one of the most powerful findings of the study. Evidence of deep-seated rejection by art and design teachers of quantitative measures is testament to an important disconnect at play in arts education today. If the assessment measures expected by uninformed administrators are so directly at odds with the most valued tenants of art teachers that both covert and overt resistance behaviors are clearly evident, something is fundamentally wrong with the assumptions underpinning the assessment of teacher performance in the arts. Since many of the directives the visual art and design teachers in this study resisted were policy-based, it is imperative that policy directives be examined to help promote a more appropriate climate for education in the visual arts in particular as well as the arts as a whole in public school education.
REFERENCES


APPENDIX A

QUESTIONS FOR SURVEY QUESTIONNAIRE
Demographic Information

a. Are you an Illinois visual art and/or design teacher at the secondary or high school level?
   Yes     No

If yes, please continue. If no, you do not need to complete this survey. Thank you for your time.

b. How many years of experience do you have teaching visual art and/or design?
   0-5 years   6-10 years   11-15 years   16-20 years   20+years

c. Where in the state of Illinois do you teach?
   Chicago Area
   Northern Illinois/not in the Chicago area
   Central Illinois
   Southern Illinois

Survey Questionnaire

How useful are the following strategies in improving your own classroom art and design instruction?

1. Analyzing student data from standardized or school wide tests:
   Extremely useful   Very useful   Useful   Somewhat useful   Not useful or not used

2. Formative assessments such as one on one critique, class discussion, or group critique:
   Extremely useful   Very useful   Useful   Somewhat useful   Not useful or not used

3. Summative assessments such as authentic assessment, portfolio review, or rubric based assessments:
   Extremely useful   Very useful   Useful   Somewhat useful   Not useful or not used

4. If you have been instructed by your administrator to align your planning and assessments to standards, which standards do you align your planning and assessment to? Please check all that apply.

Common Core State Standards
National Art and/or Media Arts Art Standards
The Illinois State Goals
The new Illinois Art Standards

I am not instructed to align to standards or I do not align to standards

If you align your lessons to standards, how helpful are the following standards in assisting you with assessing what students learn in art or design classes?

5. Common Core State Standards
   Very Somewhat Not helpful Unclear how standards help with art assessment

6. National Core Art Standards
   Very Somewhat Not helpful Unclear how standards help with art assessment

7. Illinois Art Standards
   Very Somewhat Not helpful Unclear how standards help with art assessment

Over the course of a school year, how often has your administrator conveyed information about:

8. The importance of school wide test data analysis.
   Very often Often Occasionally Once or twice Never

9. The importance of teaching in your content area.
   Very often Often Occasionally Once or twice Never

10. The importance of incorporating reading and math skills in your content.
    Very often Often Occasionally Once or twice Never

11. The importance of utilizing standardized assessment in your practice.
    Very often Often Occasionally Once or twice Never

12. Not including AP or IB Portfolio Examinations which are graded outside your school, do you implement student led portfolio examinations in any of your classes?
    Yes No

12a. If yes, which classes are your portfolio examinations conducted in?
    All classes Only upper level courses Only beginning level courses

How often are you given professional development time specifically for:

13. The development of visual art and design assessments?
6 or more times a school year  5-3 times a school year  1-2 times a school year  Never

14. To determine validity of assessments?
6 or more times a school year  5-3 times a school year  1-2 times a school year  Never

15. To determine reliability of assessments?
6 or more times a school year  5-3 times a school year  1-2 times a school year  Never

Optional: Any additional comments:

If you are interested in being contacted for a face to face interview on this subject, please contact Deborah Filbin in a separate email at [db.filbin@gmail.com]. You will be contacted with follow up information. This will be a face to face interview that will last approximately 45 minutes and will be scheduled at your convenience at a mutually agreed upon time and location.
APPENDIX B

SURVEY COVER LETTER
Dear Participant:

My name is Deborah Filbin and I am a Doctoral candidate in the Art and Design Education department at Northern Illinois University. I am conducting research on high school visual art and design teacher’s implementation of assessments in the art room and how data is collected for evaluation. You were selected for participation in this study because you are a visual art and design teacher at the secondary level. I invite you to participate in this research study by completing the attached survey. You may choose to not participate by disregarding this survey questionnaire. If you are not an art and design teacher at the secondary level, you should disregard this survey.

The following questionnaire will require approximately 10 minutes to complete. There is no known major risk to your participation in this research study. There is no compensation for completing this survey. All information will remain confidential. If you choose to participate in this project, please answer all questions as honestly as possible and submit the survey as soon as it is completed. Participation is voluntary and you may refuse to participate.

After the survey data is collected, one on one interviews will be conducted to gain further information from participants. If you are interested in being interviewed please email me directly at db.filbin@gmail.com to indicate your interest. You will be contacted with further information.

All data collected will be used to determine current assessment methods used by secondary visual art and design teachers, and recommendations for future professional development.

If you have any additional questions you may contact Deborah Filbin, Doctoral Candidate, Northern Illinois University, or faculty advisors Dr. Douglas Boughton, or Dr. Kerry Freedman. If you have questions about your rights as a research participant, you may contact the NIU Office of Research Compliance at 815-753-8588 or email at researchcompliance@niu.edu.

Thank you for your consideration to participate,

Deborah Filbin
APPENDIX C

SURVEY LETTER MAILED TO POTENTIAL PARTICIPANTS
My name is Deborah Filbin and I am a Doctoral candidate in the Art and Design Education department at Northern Illinois University. I am conducting research on high school visual art and design teacher’s implementation of assessments in the art room and how data is collected for evaluation. You were selected for participation in this study because you are a visual art and design teacher at the secondary level. I attempted to email you the link for the research, however I was unable to contact you via email. I am interested in your input. I invite you to participate in this research study by completing the survey by going to the link below. The following questionnaire will take approximately 5 minutes to complete.

[link]

If you would like more information, please contact me at [email].

Thank you for your consideration to participate,

Deborah Filbin

You may choose to not participate by disregarding this survey questionnaire. If you are not an art and design teacher at the secondary level, you should disregard this survey.

There is no known major risk to your participation in this research study. There is no compensation for completing this survey. All information will remain confidential. If you choose to participate in this project, please answer all questions as honestly as possible and submit the survey as soon as it is completed. Participation is voluntary and you may refuse to participate.

After the survey data is collected, one on one interviews will be conducted to gain further information from participants. If you are interested in being interviewed please email me directly at [email] to indicate your interest. You will be contacted with further information.

All data collected will be used to determine current assessment methods used by secondary visual art and design teachers, and recommendations for future professional development.

If you have any additional questions you may contact Deborah Filbin, Doctoral Candidate, Northern Illinois University, [email], or faculty advisors Dr. Douglas Boughton, [email] or Dr. Kerry Freedman, [email]. If you have questions about your rights as a research participant, you may contact the NIU Office of Research Compliance at 815-753-8588 or email at researchcompliance@niu.edu.
APPENDIX D

IRB CONSENT FORM
Informed Consent to Participate in Interview Research

Study: Assessment in Visual Art and Design: Policy and Practice

I am willing to participate in a research project conducted by Deborah Filbin, Doctoral candidate in the Art and Design Education department at Northern Illinois University. I understand this research is designed to gather information about high school visual art and design teacher’s implementation of assessments in the art room. I will be one of approximately 8 people interviewed for this research.

My participation is voluntary and I will not be paid for my participation. I may withdraw and discontinue participation at any time without penalty. I understand I have the right to decline to answer any question during the interview.

The face to face interview will last approximately 45 minutes. Notes will be taken during the interview by the interviewer. The interview will be audio recorded and later transcribed for the purpose of data analysis. I may be asked follow up questions in the future for purposes of clarification.

I understand the researcher will not identify me by my name in any reports, or in any information obtained from this interview, and that my confidentiality as a participant will remain secure. All audio recording and transcripts will be kept secure by the researcher and will not be shared with anyone to ensure anonymity. No identifying names will be on the transcripts to ensure anonymity. Names of participants will not be published in any papers, professional journals or presentations that may result from this study.

I understand that this research study has been reviewed and approved by the Institutional Review Board (IRB) for Studies Involving Human Subjects at Northern Illinois University. There are no known risks or discomforts that are anticipated from participating in this study. The anticipated benefit of participation in this study is to share your expertise about your professional practice, contribute to a better understanding about assessment in the field of art and design education, and reflect on my own practice.

I have read and understand the explanation provided to me and have had all questions answered to my satisfaction. I voluntarily agree to participate in this study. If you have further questions after completion of this interview, you may contact Deborah Filbin, Doctoral Candidate, Northern Illinois University, db.filbin@gmail.com or faculty advisors Dr. Douglas Boughton, dboughton@niu.edu or Dr. Kerry Freedman, kfreedman@niu.edu. If you have questions about your rights as a research participant, you may contact the NIU Office of Research Compliance at 815-753-8588 or email at researchcompliance@niu.edu.
I agree to be audio recorded during the interview. I understand it is for the purposes of transcription and data analysis by the researcher and will not be shared with anyone. The audio recording and transcripts will kept secure and my identity will be kept confidential.
APPENDIX E

INTERVIEW QUESTIONS
45 minute semi-structured interview

Before beginning, introduce the purpose of the study, confidentiality, and consent to interview and audio recording. Answer any participant questions.

1. How long have you been an art teacher?
2. What are some of the different ways you assess your students?
   a. Which of these methods do you find most useful for giving feedback to your students?
3. Can you describe the kinds of information or data you give to administrators about your student’s performance?
4. Have you ever been asked to implement a standardized test in your class, and if so can you describe that experience?
5. Describe your experience with your performance evaluation.
   a. Explain how you incorporate your measures of student performance into the documentation for your evaluation?
6. Can you tell me if you feel the current evaluation process reflects teacher performance?
   a. As an art and design education professional, do you feel this process accurately reflects the learning that takes place in the art room?
7. The PERA legislation requires Illinois teachers to provide valid and reliable data indicating student learning. What is your understanding about what this actually means and how do you go about achieving this?

Thank you for your time. Do you have any questions for me?
APPENDIX F

SURVEY QUESTIONNAIRE
We are given professional development time and can choose how to use that time as we see fit in terms of assessment and aligning to state standards.

Assessment in art is ridiculous. Grades should be abolished in art and all subjects. I support going gradeless!

While we are not "given" time to make assessments, we do have a PLC model in which we can formulate formative assessments and analyze data from that as an art department.

Honestly...Our program exceeds the state art mandate etc....We try to initiate art into all of the other curriculums...sometimes with much success. It is difficult at time for other teachers to make time in their day for any differentiation what so ever. Teaching to the test and the personal evaluation associated with makes it somewhat pressure filled. We work in a district that is pretty supportive. I know in most of the country..their are draconian evaluation practices in place for teachers that do not serve education in general or the students teachers work with. What I work with is pretty good, but it is little by little being eroded.

Happy to be interviewed. Our school places a premium on academic achievement, with the arts as a secondary priority. We are valued but not the same way. Our school is still very interested in maintaining it's mission as a comprehensive high school.

Our program is all about formative assessment strategies that rely on local standards (derived from National/State standards)

Even IB Exams have an internal assessment that I do. And all classes, even beginners are creating process journals or sketchbooks that reflect their ideas and their thought process and more and more options for materials. These are assessed through portfolio and critique. The Standardized tests help me understand the reading and writing levels of my students, and how to modify the reading/writing assignments. Sometimes a student might have to tell me their thoughts and I could type them in for her.

I actually just teach my subject area, keep my kids interested, and constantly push them to improve their skill sets and knowledge. I abhor testing. Look at the student work, and watch the teacher teach, and you will know if the Visual Arts teacher is doing their job. All this other business is government taking over our classrooms with paperwork, and drawing attention away from our actual jobs so someone can point at a chart for a meeting.

Formative assessment is a hot topic at the moment. The idea that a teacher can use PearDeck or other assessment; see data and make changes in the moment is appealing to administration. I think that it really distorts how learning occurs and is abusive to learners.

Assessments are great to have in place in the classroom, it validates what is being measured for art projects. State and National standards are great to have the recognition from government about the work we do as educators, though I have never needed to actually demonstrate the relationship between classwork and those goals. My school has a designated rubric to use for projects, though I prefer my personal rubrics as it caters to exactly what the students are expected to demonstrate, and I am not held accountable for using the district rubric.
Art is so very broad. Standardized assessment - even a rubric - does not really cover the many components that go into making a work of art. In many ways art can be "un-assessable" in the traditional manner. Art is more of an experience than a close ended lesson. Assessing art is a very complex thing.

The visual arts do not carry a high value in our school. We are an elective that is often used as a "dumping ground" of sorts to put students in who need to fill a schedule. We are not valued or respected by other department areas. We work hard to create a strong and viable program, work with the community, enter competitions, put up major exhibits, and hold ourselves to very high standards. We are innovators in the school in our assessment approaches. Our administration values the arts when we make the school "look pretty" or "sound great" but ignores us otherwise. This attitude permeates the rest of the school and makes it hard to push forward. Despite our inherent worth and significance, we are fighting to stay alive. A lack of worthy teaching candidates is one part of the problem, as is the overemphasis on STEM, without the A. I am concerned that as we veterans retire, we will not have valid candidates to take up the fight and keep the visual arts alive in the schools. Not to be too melodramatic, but it is a dark world for the visual arts at this time...

I do have the worry that our content is not as valued as STEM, etc., nor is it as easy to evaluate by multiple choice, which leads it to being undervalued.

My current administration does not micromanage instruction if there is not a problem. They understand the importance of arts education and believe the subject matter experts should make the decisions about the curriculum.

Looking at the reliability and validity of assessments are not specifically mandated activities, but we are given plenty of time to review and tweak our assessments, which should include those evaluations if we’re doing our jobs well. Since it is informal, some educators are probably not driven to be critical of their assessments.
APPENDIX G

DISTRICT W
PROFESSIONAL EVALUATION PLAN

AUGUST 2015
## COMMITMENTS and COMMON THEMES

### Commitment: Equity

The Professional Evaluation Plan must support equity with clear and consistent guidelines through a common understanding.

In order to embed **equity** into practice, the district commits to evaluating educators in a reliable, on-going, evidence-based manner utilizing multiple data sources which are differentiated based upon roles and responsibilities.

### Commitment: Professional Growth

The Professional Evaluation Plan must support continuous professional growth through engagement and reflective practice.

In order to embed **professional growth** into practice, the district commits to fostering a culture that promotes training opportunities, goal setting, and reflective practice through the allocation of resources. Supports Certified Staff Members in taking ownership of their professional practice as it impacts student growth and influences the broader learning community.

### Commitment: Student Outcomes

The Professional Evaluation Plan must support student outcomes through highly effective instruction, professional development, and data supported resource allocations.

In order to embed **student outcomes** into practice, the district commits to aligning this process to student growth.

### Commitment: Collaboration

The Professional Evaluation Plan must support collaboration among all stakeholders and will dedicate time for ongoing refinement to ensure a continuous cycle of improvement.

In order to embed **collaboration** into practice, the district commits to fostering a culture that promotes trust, reflection, and open professional dialogue.

### Commitment: Expectations

The Professional Evaluation Plan must support clearly defined expectations.

In order to embed clearly defined **expectations** into practice, the district commits to integrating consistent expectations into professional practice through performance standards which are evidence-based.
DISTRIBUTION PROFESSIONAL DEVELOPMENT PLAN OBJECTIVES

The Professional Evaluation Process commits to...

**Professional Development that will:**
- Educate and guide stakeholders in the paradigm shift from the previous evaluation system to the Framework.
- Create different levels of support for Certified Staff Members.
- Establish Inter-rater reliability among Evaluators.
- Equip Evaluators with the strategies that will foster a supportive environment through collaborative, open communication during professional conversations.

**A Culture that will:**
- Embed Commitments into the District culture.
- Foster professional dialogue and collaboration within the professional learning community.
- Dedicate time and resources for professional growth.
- Empower Certified Staff Members to use research based best practices and innovative strategies that will challenge students and enhance their learning.
- Embrace and utilize data for making sound educational decisions.

**Resources that will:**
- Explain the Framework, evaluation process, and all tools as they relate to each position.
- Provide a meaningful and manageable timeframe for the implementation of the evaluation process.

**A Continuous Learning Process that will:**
- Use a variety of data sources to guide reflection and decision-making.
- Develop procedures to review the evaluation process and adapt as appropriate.
- Identify areas of need for district-wide professional development for Certified Staff Members.
FORMAL AND INFORMAL OBSERVATIONS CONDUCTED DURING THE EVALUATION CYCLES
OBSERVATION OF TEACHERS
HIGH SCHOOL DISTRICT REQUIREMENTS

<table>
<thead>
<tr>
<th>Years 1-2 Non-Tenured</th>
<th>Year 3-4 Non-Tenured</th>
<th>TENURED Needs Improvement</th>
<th>TENURED Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>A minimum of four (4) observations shall be required each school year: Two (2) must be formal observations (formal observations include both a pre- and post-observation conference)</td>
<td>A minimum of three (3) observations shall be required each school year: Two (2) must be formal observations (formal observations include both a pre- and post-observation conference)</td>
<td>A minimum of three (3) observations shall be required each evaluation: One (1) must be a formal observation (formal observations include both a pre- and post-observation conference)</td>
<td>A minimum of three (3) observations shall be required each evaluation cycle: Two (2) must be a formal observation (formal observations include both a pre- and post-observation conference)</td>
</tr>
</tbody>
</table>

OBSERVATION OF PROFESSIONAL PRACTICE
Evidence of professional practice is collected through the use of multiple observations that include formal and informal observations and focus upon acquiring evidence of the planning and preparation, classroom environment, instruction, and professional responsibilities.

FORMAL OBSERVATIONS
As part of the formative process, formal observations are an effective way of gathering information about the teacher’s practice:

- Full-time, non-tenured teachers will be formally observed at least once each year
  - Formal observations are a minimum of 45 minutes at a time, or a complete lesson, or during an entire class period
- Tenured teachers will be formally observed at least once in the course of every two (2) school years
- Tenured teachers who receive a ‘needs improvement rating’ will be formally observed at least once in the school year following the Professional Development Plan (PDP)
- Formal observation will include a pre-observation conference. In advance of the conference, teachers will submit the Pre-observation Form (Form 3) and may be asked to submit any appropriate artifacts that pertain to the observation. Evaluator will discuss and make recommendations for areas of focus during the observation. A copy of the Pre-observation Information Request Form can be found on the district website.
- Together, teachers and evaluators will determine a focus for the observation and any specific areas of feedback.
- Observation notes will be provided to the teacher within two (2) working days. For example, if an observation is completed on a Monday at 11:30am, the observation notes are due by Wednesday at 11:30am.
- A post-observation conference will be held within five (5) working days after the formal observation unless both parties mutually agree on another time to meet.
POST-OBSERVATION CONFERENCE
The Post-Observation Conference is a designated interactive meeting time to follow the classroom observation period. The conference time, a minimum of 30 minutes in length, is intended to be a useful tool to advance mutual understanding of the mastery of the FFT Domains. The purpose of the meeting is interactive discussion between the teacher and the Designated Administrator.

- The post-observation conference will focus on teacher reflection on the lesson and evidence of student learning. Teachers will be asked to submit a self-reflection form, post-observation form and/or any appropriate artifacts that pertained to the observation.
- The evaluator shall meet with the educator to discuss the evidence collected.
- The evaluator shall provide feedback to the educator about professional practice, including evidence specific to the areas of focus that may have been identified in the pre-conference.
- Copies of observation notes will be given to the teacher prior to the conference. The conference should be completed no later than five school days after the classroom observation. Each Post-Observation Conference will be recorded through a written summary by the Designated Administrator and a copy of that summary will be filed with the teacher within five days after the Post-Observation conference.
- The educator shall reflect upon instruction and may provide additional information or explanations to the evaluator.
- If the evaluator determines that evidence collected to date (at the time of the observation/post conference) may result in the educator receiving a "needs improvement" or "unsatisfactory" rating, the evaluator must notify the educator of that determination. Additionally, the teacher will be referred to the PAR Panel.
- The educator shall work with the evaluator and others as determined to identify areas for improvement.

INFORMAL OBSERVATIONS
As part of the formative process, informal observations are an effective way of gathering ongoing and authentic information about teacher practice with the intent of focusing on teaching, learning, and the interactions between teachers and students.

- Evaluators shall make as many informal observation visits as deemed necessary to properly assess teacher effectiveness.
- These observations need not be pre-announced nor do they require pre or post-observation conferences.
- Informal observations are not intended to supplant formal observations; rather, they are intended to broaden opportunities for professional growth.
- Following an informal observation the evaluator shall provide feedback either orally or in writing.
  - If it is in writing, then the evaluator shall provide the educator with an opportunity to have an in-person discussion.
  - Evidence gathered during an informal observation may be considered in determining the performance evaluation provided it is documented in writing and will be provided within (five) 5 working days.
- Teachers are always encouraged to seek feedback and/or initiate conversations from their informal observations. Informal observations can provide opportunities for evaluators and teachers to strengthen relationships and foster an atmosphere of trust and respect, to motivate teachers to engage in self-reflection in order to move their practice forward, to allow evaluators to be accessible and provide support.

CLASSROOM OBSERVATION
- Formal observations shall be one of the following: 45 minutes, a complete lesson, or an entire class period.
• Evaluator shall acquire evidence of planning, instructional delivery, and classroom management.
• Evidence of professional practice shall be collected through the use of multiple observations that include formal and informal observations.
• Any of the teacher's regularly scheduled classes may be observed.

DESIGNATED ADMINISTRATOR
A designated evaluator may designate another qualified evaluator to conduct observations in situations where it cannot be completed in a timely manner and the teacher will be notified.

EVALUATION CONFERENCES

Mid Year Evaluation Conference
The Mid Year Evaluation Conference is a documented meeting time scheduled by the Designated Administrator with the probationary teacher for the purpose of reviewing mid-year progress on the teacher's mastery of the FFT Domains. The teacher shall be given a copy of the Preliminary Evaluation at least twenty-four (24) hours before the conference. The evaluation is to be completed by November 15. No final ratings as designated by the State of Illinois will be given at this conference. Preliminary evaluation documents, with original signatures, will be filed with the teacher, the designated administrator and the principal. In the event it is determined a probationary teacher is in danger of having a final rating of 'needs improvement' or 'unsatisfactory', the teacher will be referred to the PAR Panel.

Summative Evaluation Conference
The Summative Evaluation Conference is the documented meeting time scheduled between the teacher and the Designated Administrator(s). The teacher shall be given a copy of the Final Evaluation with ratings at least twenty-four (24) hours before the conference. The Final Evaluation, complete with Summary Ratings as required by the State of Illinois, is to be completed by March 1 for all teachers. Final evaluation documents will be filed with the teacher, the designated administrator and the principal. The original copy filed with the Director of Human Resources will be placed in the teacher's official personnel file. Original signatures are required for the teacher records and the Human Resources Office.

SELF REFLECTION
The self reflection is a narrative document prepared by the teacher being evaluated, which provides the teacher with the opportunity to reflect on the Teaching Essentials or to inform the Designated Administrator of accomplishments and contributions to the school which may not be evident in a classroom observation. Information from the Self Reflection shall be attached to the Final Evaluation document. Self-Reflections will be provided to the Designated Administrator on or before the Summative Evaluation Conference. A template for the self reflection is located at the end of the Professional Evaluation Plan.
STATE OF ILLINOIS PERFORMANCE RATINGS
The Illinois School Code recognizes four performance ratings: Excellent, Proficient, Needs Improvement and Unsatisfactory. For tenured teachers rated as "needs improvement", within 30 school days of the completion of an evaluation rating, the evaluator, in consultation with the teacher, must create a professional development plan that is directed to the areas of needed improvement. Tenured teachers rated as "unsatisfactory" places a teacher in remediation. The procedures for remediation are delineated in the Remediation Procedures section of this document. For details of the performance ratings see page 10 of the Professional Evaluation Plan.

FRAMEWORK FOR TEACHING (FFT) DOMAINS
The Domain are defined as the summation of the examples listed in each of the certified professional’s four Roles: Domain 1: Planning and Preparation, Domain 2: The Classroom Environment, Domain 3: Instruction, Domain 4: Professional Responsibilities

TEACHING DOMAIN PERFORMANCE RANKING
These rankings accompany each Domain. The goal of the individualized rankings is to assess the level of performance in each Domain.

Excellent
A State rating of Excellent means that the teacher has received Excellent rankings in the FFT Domains and no rankings less than Proficient on his/her final evaluation. The teacher demonstrates a consistent mastery of each FFT Domain. The teacher displays knowledge of and exhibits behavior that shows a commitment to the Domain above and beyond the minimum expectations.

Proficient
A State rating of Proficient means that the teacher has received rankings of Proficient on his/her final evaluation. The teacher displays knowledge of and exhibits behavior that is consistent with the Domains.

Needs Improvement
A State rating of Needs Improvement means the teacher has received rankings of Needs Improvement. A state rating of Needs Improvement for tenured teachers will initiate the Professional Development Plan.

The teacher demonstrates a basic understanding of the requirements for meeting this Domain but is inconsistent in their application. The teacher may be in need of additional reinforcement, guidance or information concerning a particular Domain in order to progress to the Excellent or Proficient rankings.

 Unsatisfactory
A State rating of Unsatisfactory means that the teacher has received rankings of Needs Improvement and/or any ranking of Unsatisfactory on his/her final evaluation. A State rating of Unsatisfactory for tenured teachers will initiate the Remediation Process.

The teacher does not display a clear understanding of the requirements for meeting this Domain. The teacher is in need of additional monitoring, guidance and support in order to progress to the Proficient or Excellent rankings.

Not Applicable
The Domain may not apply or be relevant on the date of a formal observation.
OPERATING PRINCIPLES

**DOMAIN** Ratings in [High School District] Professional Evaluation Plan
- **Excellent** – No more than two components rated Proficient, with the remaining components rated Excellent. The anchor component must be rated Excellent.
- **Proficient** – No more than one component rated Needs Improvement, with the remaining components rated at Proficient or Excellent.
- **Needs Improvement** – Two or more components rated Needs Improvement, with the remaining components rated as Proficient or Excellent.
- **Unsatisfactory** – Any component rated as Unsatisfactory.

**OVERALL** Ratings in [High School District] Professional Evaluation Plan
- **Excellent** – No more than one domain rated Proficient, with the remaining domains rated at Excellent.
- **Proficient** – No more than one domain rated Needs Improvement, with the remaining domains rated at Proficient or Excellent.
- **Needs Improvement** – Two or more domains rated Needs Improvement, with the remaining domains rated as Proficient or Excellent.
- **Unsatisfactory** – Any domain rated Unsatisfactory.

**Non-Tenured Staff Contract Renewal** - Each non-tenured Staff member will receive a final summative rating and a recommendation for renewal or non-renewal of his/her contract. It is understood that non-tenured Staff in years 1 and 2 may receive a final summative rating of Needs Improvement as they are emerging towards proficiency. Non-tenured Staff in years 3 and 4 are expected to maintain a final summative rating of Proficient or Excellent.

**Tenured Staff** are expected to maintain an overall Summative Rating of Proficient or Excellent.

- If a Tenured Staff member receives an overall Summative Rating of Needs Improvement, a Professional Development Plan (PDP) will be developed. A tenured Staff member whose performance is not Proficient or Excellent after the completion of a PDP will be rated Unsatisfactory.
- If at any point in the appraisal cycle a Tenured Staff member exhibits evidence of Unsatisfactory practice, an overall Summative Appraisal may be conducted at any time during the contractual school year. An overall Summative Rating of Unsatisfactory will result in the development of a Remediation Plan in accordance with the law.

*Please see the following page for an example of how domain and summative ratings are determined.*
EXAMPLES of DOMAIN and FINAL SUMMATIVE RATINGS

DOMAIN Ratings in [High School District] Professional Evaluation Plan

- **Excellent** – No more than two components rated Proficient, with the remaining components rated Excellent. The anchor component must be rated Excellent.
- **Proficient** – No more than one component rated Needs Improvement, with the remaining components rated Proficient or Excellent.
- **Needs Improvement** – Two or more components rated Needs Improvement, with the remaining components rated Proficient or Excellent.
- **Unsatisfactory** - Any component rated Unsatisfactory.

<table>
<thead>
<tr>
<th>Domain 2 – The Classroom Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
</tr>
<tr>
<td>2a</td>
</tr>
<tr>
<td>2b (anchor)</td>
</tr>
<tr>
<td>2c</td>
</tr>
<tr>
<td>2d</td>
</tr>
<tr>
<td>2e</td>
</tr>
<tr>
<td>Domain Rating</td>
</tr>
</tbody>
</table>

OVERALL Ratings in [High School District] Professional Evaluation Plan

- **Excellent** - No more than one domain rated Proficient, with the remaining domains rated Excellent.
- **Proficient** – No more than one domain rated Needs Improvement, with the remaining domains rated Proficient or Excellent.
- **Needs Improvement** – Two or more domains rated Needs Improvement, with the remaining domains rated Proficient or Excellent.
- **Unsatisfactory** - Any domain rated Unsatisfactory.

<table>
<thead>
<tr>
<th>Final Summative Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
</tr>
<tr>
<td>Domain 1</td>
</tr>
<tr>
<td>Domain 2</td>
</tr>
<tr>
<td>Domain 3</td>
</tr>
<tr>
<td>Domain 4</td>
</tr>
<tr>
<td>Overall Rating</td>
</tr>
</tbody>
</table>
MEASURES OF STUDENT GROWTH

It is the expectation of High School District that all students will grow in their knowledge, skills, and aptitude during their time in our school district. The Performance Evaluation Reform Act (PERA) states that student growth will be measured and incorporated into the summative evaluation rating of all certified staff. Student growth will account for 25% of the summative rating during the 2015-2016 school year and 30% of the summative rating for the 2016-2017 school year and beyond.

High School District schools will measure student growth using two different types of assessment. Students will be measured through the final exam assessments (Type II) and Student Learning Objectives (SLO) (Type III). The Type II assessment will account for 80% of the student growth score and the Type III assessment will account for the other 20%. The district will be utilizing the Illinois State Board of Education (ISBE) Student Learning Objective template and resources located at http://www.isbe.net/assessment/htmls/balanced-asnt.htm.

![Diagram of assessment structure]

Final exams assessments (Type II) will be administered at the beginning and end of the semester. Growth scores will be calculated on average per pupil growth from pre-exam to post-exam, for the curricular team within that building. For example, all teachers of English 9 at a given campus, will earn the same growth score for that course. This method of measurement was designed by the Joint PERA Committee to reinforce the standard practice that students within the building are everyone’s responsibility. The Committee wanted to support the collaborative nature that is a hallmark of High School educational practices. For a given teacher, the Type II growth score will be weighted on their teaching schedule. For example, Teacher A has a class schedule with three (3) English 9 and two (2) English 10 Honors classes. Therefore, 60% of the Type II rating will come from the English 9 growth score and 40% of the Type II rating will come from the English 10 Honors growth score.

Student Learning Objectives (SLO) are designed for teachers and their evaluator to collaborate about expected growth in the individual teacher’s classroom. Working together, the teacher and the administrator will work to develop a predicted growth goal for a specific sub-set of students on that teacher’s roster. This goal will identify a percentage of students that the teacher feels will achieve a particular growth benchmark. After establishing baseline data and creating a rate of achievement for each student, the teacher and administrator will measure the number of students who met or exceeded that rate of achievement. The SLO rating will be determined by calculating the ratio of students that met or exceeded the achievement rate compared to the predicted growth goal. The SLO should focus on a particular unit or chapter of study and is meant to be of
shorter duration than the Type II assessment. For example, Teacher A predicts 60% of their sub-set will meet the growth benchmark. At the end of that unit or chapter it is measured that 70% meet or exceeded the benchmark. Based on this information, Teacher A had a 117% achievement rate.

The district recognizes there is a potential need for conversations revolving around student attendance and its impact on student achievement. These conversations should take place between the teacher and their evaluator before the summative rating meeting. During these student achievement meetings, it is critical that the conversation be ruled by data and an accurate representation of what is actually happening in the classroom. For example, if Teacher A had student in their class who happens to be on homebound instruction for a large portion of a unit or chapter, it would be reasonable to discuss if this student should be included in the student growth calculation. In the event a student is unable to achieve the goal because it is mathematically impossible, the evaluator and teacher must have an agreed upon remedy. For example, in an SLO with a growth goal of 25%, if a student scored 87% for the baseline assessment, that student would have to score over 108% on the end assessment. It would be reasonable, in this case, to create goal that maintains the baseline level of achievement. The evaluator has final approval in each case.

Specific measures of growth, from pre-test or post-test or baseline to end of unit assessments, will translate into evaluation ratings. Evaluative ratings are described in the Type II and Type III rubrics shown on the next page.
### INTEGRATION OF STUDENT GROWTH AND OBSERVATION RATINGS

<table>
<thead>
<tr>
<th>SLO (Type III)</th>
<th>Unsatisfactory</th>
<th>Needs Improvement</th>
<th>Proficient</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 40% of students meeting established goal</td>
<td>40% - 59.999999% of students meeting established goal</td>
<td>60% - 84.999999% of students meeting established goal</td>
<td>85%+ of students meeting established goal</td>
<td>Percentage of students meeting goal (Calculated on individual teacher basis)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final Exams (Type II)</th>
<th>Unsatisfactory</th>
<th>Needs Improvement</th>
<th>Proficient</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than or Equal 20%</td>
<td>Greater than 20%, but less than or equal to 30%</td>
<td>Greater than 30%, but less than or equal to 40%</td>
<td>Greater than 40%</td>
<td>Average percentage point increase per student on final exam pre/post test (Calculated on course team basis by building)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Student Learning Objectives (Type III) 20%</th>
<th>Unsatisfactory</th>
<th>Needs Improvement</th>
<th>Proficient</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>Unsatisfactory</td>
<td>Unsatisfactory</td>
<td>Unsatisfactory</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>Needs Improvement</td>
<td>Needs Improvement</td>
<td>Needs Improvement</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Proficient</td>
<td>Proficient</td>
<td>Proficient</td>
<td>Proficient</td>
<td>Proficient</td>
</tr>
<tr>
<td>Excellent</td>
<td>Proficient</td>
<td>Excellent</td>
<td>Excellent</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

First, determine performance levels for final exams and Student Learning Objectives (SLO) based on the performance indicators in the top two charts. Next, use the bottom chart to determine the overall student growth rating at the intersection of the two ratings.
INTEGRATION OF STUDENT GROWTH AND OBSERVATION RATINGS

Take the student growth score based on the intersection of the rating for the final exams and SLO and use the chart below to find the intersection of the student growth rating and the observation rating to determine your overall rating for your evaluation.

<table>
<thead>
<tr>
<th>Student Growth (25% of Total Score)</th>
<th>2015-2016</th>
<th>Observation Rating (75% of Total Score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>Unsatisfactory</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>Unsatisfactory</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Proficient</td>
<td>Needs Improvement</td>
<td>Proficient</td>
</tr>
<tr>
<td>Excellent</td>
<td>Needs Improvement</td>
<td>Proficient</td>
</tr>
</tbody>
</table>
INTEGRATION OF STUDENT GROWTH AND OBSERVATION RATINGS

Take the student growth score based on the intersection of the rating for the final exams and SLO and use the chart below to find the intersection of the student growth rating and the observation rating to determine your overall rating for your evaluation.

<table>
<thead>
<tr>
<th>Student Growth (30% of Total Score)</th>
<th>2016-2017</th>
<th>Observation Rating (70% of Total Score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>Unsatisfactory</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>Unsatisfactory</td>
<td>Needs Improvement</td>
</tr>
<tr>
<td>Proficient</td>
<td>Needs Improvement</td>
<td>Proficient</td>
</tr>
<tr>
<td>Excellent</td>
<td>Needs Improvement</td>
<td>Proficient</td>
</tr>
</tbody>
</table>
APPENDIX H

DISTRICT X
DISTRICT X

Teacher Evaluation Process

(Updated 8.1.17)
Definitions

In an effort to ensure all participants have a thorough and common understanding of the terms being used, this section defines key terms that are associated with Part 50 Evaluation of Certified Employees – Under Articles 24A and 34 of the School Code.

Performance Evaluation Plan - a plan to evaluate a teacher that includes data and indicators that measure the teacher’s professional practice.

Formal Observation – a specific window of time that is scheduled with the teacher to directly observe professional practices in the classroom or in the school.

Informal Observation – observations of a teacher that are not announced in advance of the observation and not subject to a minimum time amount.

Performance Evaluation Rating – the final rating of a teacher’s performance using the rating levels of “excellent”, “proficient”, “needs improvement”, or “unsatisfactory” regarding their professional practice.
Philosophy of Personnel Evaluations

The purpose of evaluation of the professional teaching staff is to improve the quality of instruction. By improving the quality of instruction we are allowing all students the ability to learn at high levels. This goal can best be achieved through teacher evaluation systems which accomplish the following objectives:

1. Improve specific skills.
2. Recognize proficient and excellent performance.
3. Communicate expectations to the teacher.
4. Facilitate two-way communication between evaluator and teacher.
5. Develop trust and confidence between evaluator and teacher.
6. Provide a needs assessment for staff development programs.
7. Motivate the professional development of individuals.
8. Provide information needed for personnel decisions such as tenure, promotion, transfer, and termination.

Evaluation judgments and decisions are limited by the quality of data collected. Data sources include work, observation by the evaluator and others, and interviews with the teacher. As the significance of the evaluation decision increases, the number of sources of data expands. The teacher may request reasonable, additional data collection. The teacher evaluation system shall include:

1. Research-based criteria reflecting the current findings in effective instructional and administrative practice.
2. Clearly stated standards of performance.
3. Evaluation procedures incorporating the due process rights of the teacher and evaluator.
4. Training for evaluators and teachers which provides for the effective accomplishment of the objectives of personnel evaluation.

This evaluation process should be continuous, constructive and cooperative in nature.
Plan Components Required for the Evaluation of Teachers

The district has stated in its philosophy of Personnel Evaluation that "the purpose of evaluation of the professional staff is to improve the quality of instruction." The evaluation procedures outlined below provide the employer the opportunity to assess and evaluate the professional performance evaluation criteria. This evaluation process is intended to be continuous, constructive and cooperative in nature.

A. On or before the first day of student attendance, the school district shall provide a written notice (either electronic or paper) that a performance evaluation will be conducted during that school year to each teacher affected or, if the affected teacher is hired after the start of the school term, then no later than 30 days after the contract is executed. The written notice shall include:

1. A copy of the rubric to be used to rate the teacher against identified standards and goals and other tools to be used to determine a performance evaluation rating;

2. A summary of the manner in which measures of student growth and professional practice (see Teacher Evaluation Performance Rating Form) to be used in the evaluation related to the performance evaluation ratings of "excellent", "proficient", "needs improvement", and "unsatisfactory"; and

3. A summary of the district's procedures related to the provision of professional development in the event a teacher receives a "needs improvement" or remediation in the event a teacher receives an "unsatisfactory" rating to include evaluation tools to be used during the remediation period.

B. Teachers in contractual continued service (i.e., tenured) shall normally be evaluated, in accordance with the following policies, formally at least once every two years. If a tenured teacher has obtained a "needs improvement" or "unsatisfactory" Performance Evaluation Rating on the previous year's evaluation, the teacher shall be evaluated in the next school year after receiving that rating.
C. Teachers not in contractual continued service (i.e., non-tenured) shall have a minimum of two formal evaluations each year.

D. Any teacher may be evaluated on a more frequent basis for the following reasons:

- A significant change in the teacher's program or assignment.
- An unsatisfactory or needs improvement rating on the Performance Evaluation Rating.
- The teacher's current performance has significantly declined since the last evaluation.
- A principal new to a building may choose to evaluate any teachers within a school during their first year as principal.

There shall be a minimum of one formal observation during this evaluation cycle. Additional observations may be scheduled at the discretion of the evaluator or the request of the teacher.

Professional Practice Components for Teachers

In order to assess the quality of the teacher's professional practice, the evaluation plan shall include an instructional framework developed or adopted by the school district that is based upon research regarding effective instruction: addresses at least planning, instructional delivery, and classroom management; and aligns to the Illinois Professional Teaching Standards. The evaluation plan shall consider the teacher's attendance and his or her competency in the subject matter taught.

1) The instructional framework shall align to the roles and responsibilities of each teacher who is being evaluated.

2) The evaluation plan shall contain a rubric to be used in rating professional practice that aligns to the instructional framework developed or adopted by the district.

...
3) The district shall quantify the relative importance of each portion of the framework to the performance rating.

4) The teacher's strengths and weaknesses and the reasons for identifying the areas as such will also be specified.

A. Informal observations will be conducted and the qualified evaluator shall provide feedback to the teacher either orally or in writing (electronic or paper) and if the feedback is in a written format, also provide the teacher with an opportunity to have an in-person discussion with the evaluator. Evidence gathered during the informal observations may be considered in determining the performance evaluation rating, provided it is documented in writing.

Input from supervisors and other district administration familiar with the teacher's work may be used to assist in getting a total picture of the teacher's performance. Information from parents and community members may be considered for informal and formal observation reports. If said input will be used in an evaluation it will be provided to the teacher in writing, and the accuracy of the stated information will be verified before being included in the formal observation report.

B. Each formal observation shall cover a complete lesson: or an entire class period: or a minimum of 45 minutes at a time.

C. The evaluation observation, formal or informal, shall be reduced to writing and a copy of it given to the teacher within five (5) school days after the observation. The evaluation instrument will be given to the teacher prior to the post observation conference. A post observation conference will be held in order to discuss the observation and written evaluation.

D. After the discussion of any written evaluation or observation report, the form shall be signed and dated by both parties. Each shall have a copy. The teacher's signature does not indicate agreement with the evaluation/observation, but rather signifies awareness of the content.
It is the intent of the evaluation cycle that each conference provide positive reinforcement, identification of areas of strength, areas that need improvement, and planning for improved future performance.

E. Before any evaluation or observation report becomes a part of his/her permanent file, the teacher will have fifteen (15) school days following receipt of the report to include a written response for clarification or to add information or opinion. This response becomes a permanent part of the evaluation report.

F. The evaluation becomes a part of the teacher's personnel file housed at the district office.

G. Written suggestions for improvement may be made by the evaluator whenever he/she deems it necessary. In order to provide an opportunity to improve their instructional performance, teachers who are rated in their Performance Evaluation Rating to “need improvement” or “unsatisfactory” must have a minimum of three observations in the next school year in which two must be formal observations. The evaluator’s role is to assist teachers, and such a system shall include but not be limited to:

1. Notification, in writing, of areas where improvements are needed.

2. Specific recommendations for improvement within a specified time and methods by which improvement will be assessed.

3. Additional resources, if any, to be utilized to assist in implementing such recommendations.

Within 30 school days of receiving a Performance Evaluation Rating of “needs improvement” for a tenured teacher, a professional development plan will be created collaboratively by the evaluator and teacher. The professional development plan will be directed to the areas that need improvement, and state any supports that the district will provide to address the areas identified as needing improvement.
Review of the professional development plan will occur at a minimum of every forty (40) school days between the evaluator and teacher. The review will include an update on the progress of the professional development plan to ensure that progress is being made toward areas identified as needing improvement.

H. Upon completion of a Performance Evaluation Rating of "unsatisfactory" for a tenured teacher, a remediation plan must be developed. Evaluations during the remediation period will occur at the mid-point (45 days) and at the end point (90 days), with a written copy of deficiencies and recommendations for correction provided within 10 days. In the event that a formal remediation plan must be developed, all parties will refer to Illinois School Code section 105 ILCS 5/24A-5.

The teacher and/or evaluator may request and shall mutually agree upon another district and/or administrator, and a consulting teacher to help a teacher in need of intensive assistance.

Rights regarding the employee's personnel file and any evaluative data contained therein are subject to the statutes of the State of Illinois and the Dixon Education Association collective bargaining agreement with the Dixon Board of Education.

I. Prior to the first day of student attendance, evaluators will familiarize teachers with the procedures and materials used in the evaluation system.

J. Before any teacher can be evaluated on district or school wide initiatives, training must be provided.
**The Evaluation Cycle**

- **Non-Tenured teacher** = a minimum of (3) observations per year, of which two (2) must be formal observations. Non-tenured teachers will be observed three times a year for 4 years. The observations, will occur a minimum of 2 months apart.

  1) Building Level Evaluation Orientation Meeting
  2) Pre-observation
  3) Observation
  4) Post Observation
  5) Student Growth Component

- **Tenured teacher** = who received either an "excellent" or "proficient" Performance Evaluation Rating in his or her last evaluation will have a minimum of two observations required every (2) years, one of which must be a formal observation.

  1) Building Level Evaluation Orientation Meeting
  2) Pre-observation
  3) Observation
  4) Post Observation
  5) Student Growth Component

**A. Evaluation Orientation Meeting**

Prior to the first day of student attendance, an evaluation orientation meeting will be held. The meeting will occur prior to any formal observation. The purpose of this meeting is to provide an orientation and updating for the teacher related to the evaluation process.

The orientation meeting will include a review of the following information:

1) the evaluation rubric:

***
2) a summary of the manner in which measures of professional practices to be used in the evaluation related to the performance evaluation ratings of "excellent", "proficient", "needs improvement", and "unsatisfactory";

3) a summary of the district's procedures related to the provision of professional development in the event a teacher receives a "needs improvement" or remediation in the event a teacher receives an "unsatisfactory" rating to include evaluation tools to be used during the remediation period;

B. Pre-observation

No later than two working days prior to the scheduled observation, the teacher will submit lesson plans (their own or district template) which may include recommendations for areas on which the qualified evaluator should focus during the observation.

The teacher will state his/her objectives for the lesson, the State Goal(s), materials to be used, instructional procedures, evaluation/assessments, and other information which the teacher deems pertinent may be noted.

C. Observation

The observation will encompass a complete lesson or an entire class period or a minimum of 45 minutes at a time. During the observation, the evaluator will follow procedures listed below:

1. The evaluator is expected to arrive promptly at the scheduled time and is expected to remain for the complete lesson or an entire class period or a minimum of 45 minutes at a time.

2. The evaluator will look for evidence that the teacher is meeting the established performance characteristics as outlined in the teacher evaluation instrument.

***
3. Observational practices will be consistent with the Philosophy of Personnel Evaluation, but the activities will vary in the experiences and needs of the teacher.

The evaluator will:

a. Observe the total classroom situation.

b. Try to remain inconspicuous, although under certain circumstances he/she may enter into activity with the class.

c. Consider the class activities before and after the period being observed.

D. Post Observation Conference

The primary purpose of the post observation conference is to identify strengths and provide recommendations to overcome areas needing improvement. Following the observation visit, the evaluator will analyze the data collected. The evaluation shall be reduced to writing and a copy given to the teacher within five (5) school days after the observation.

The evaluator may review the teacher's past evaluation based upon district-established performance areas, district goals and objectives, and previously established student growth data. Upon receiving the evaluation, a discussion can occur immediately, or another mutually agreed upon date can be scheduled to review the evaluation report. It is the teacher's responsibility to provide the evaluator any additional information he/she wishes to be attached to the teacher evaluation report.

The evaluation shall not be predicated upon lawful non-school related activities which have no impact or bearing on his/her effectiveness as an employee.

Upon receiving the teacher evaluation report, the teacher may, within fifteen (15) school days, file a written response which would be attached to the evaluation report and be included in the teacher's personnel file.
E. Performance Evaluation Rating:

a. Each Domain Rating will be calculated by averaging the five components.

b. The tenured Summative Evaluation Rating will be calculated by combining the Teacher Evaluation Rating and the Student Growth Component Rating (see page 17).

c. Non-tenured teachers will receive a Performance Evaluation Rating on each of their formal evaluations. A final Summative Evaluation Rating will be calculated by combining the Teacher Evaluation Rating and the Student Growth Component Rating (see page 17).

believe that assessment is a key component of an instructional program, and identifies the instructional pathway to be taken. Assessments should provide information about students' progress in attaining the knowledge, skills, attitudes, and behaviors to be learned or acquired in school.

Types/Categories of Teachers:

- Teachers for the District (one category allows the flexibility of the assessments to be utilized)

**Timeline & Percentage for Student Growth Component:** (Section 50.110-Student Growth Components)

2015-2016 – Student Growth Trial Year with No Impact on Evaluation Rating – Joint Committee Convenes

2016-2017 – Year 1 of Student Growth Counting Toward 25% of Performance Evaluation Rating

2017-2018 – Year 2 of Student Growth Counting Toward 25% of Performance Evaluation Rating

2018-2019 and beyond – Student Growth Counts Toward 30% of Performance Evaluation Rating
Assessment Types:

"Type 1" means a reliable assessment that measures a certain group or subset of students in the same manner with the same potential assessment items, is scored by a non-district entity, and is administered either statewide or beyond Illinois. Examples: Northwest Evaluation Association (NWEA), Scantron Performance Series, Star Reading Enterprise, College Board’s SAT, Advanced Placement or International Baccalaureate examinations, or ACT’s EPAS® (i.e., Educational Planning and Assessment System).

"Type 2" means any assessment developed or adopted and approved for use by the school district and intended to be used on a district-wide basis by all teachers in a given grade, course or subject area. Examples include collaboratively developed common assessments, curriculum tests and assessments designed by textbook publishers. Examples: Aimsweb, ESGI, PASI, PSI, Bear Test, ISEL, Running Record, Common Formative Assessments, Textbook Assessments

"Type 3" means any assessment that is rigorous, that is aligned to the course’s curriculum, and that the qualified evaluator and teacher determine measures student learning in that course. Examples include teacher-created assessments, assessments designed by textbook publishers, student work samples or portfolios, assessments of student performance, and assessments designed by staff who are subject or grade level experts that are administered commonly across a given grade or subject. A Type 1 or Type 2 assessment may qualify as a Type 3. Examples: Teacher Created Assessments, Portfolios, Skills Checklist, Standards Based Report Card, Standard Based Assessments, ESGI, SLO

Assessment Menu:

Teachers must have one Type 2 and one Type 3 assessment to measure student growth unless a Type 2 assessment does not exist at your grade level/subject area. If no Type 2 assessment exists, then two Type 3 assessments are used to measure student growth.

- At this time, there are no Type 1 assessments that will utilize.
- Per Illinois Administrative Code 50.110, teachers who will utilize two Type 3 assessments need only use one Type 3 during this initial student growth assessment year (2016-17).
- The assessments chosen can be by the individual teacher or by a grade level/department of teachers.
Definition of Student Growth:

"Student growth" means a demonstrable change in a student's or group of students' knowledge or skills, as evidenced by gain and/or attainment on two or more assessments, between two or more points in time.

Growth Targets:

"Growth targets" will be established collaboratively between the teacher and the evaluator. When setting growth targets they need to align to learning standards, be relevant to instruction, and reflect appropriate growth. Growth targets should be ambitious but realistic for students to achieve in the specified period of time.

Student Sampling:

Assessments being used to show growth MUST measure the same skills or content for a minimum of two data points with a preselected group of students. Those students will represent at least one class hour or period. A district can increase the reliability of a teacher’s evaluation scores by including the growth of as many students as possible in the score while ensuring that the teacher has a real opportunity to contribute to the growth of these students.

Students must have a minimum of 90% attendance rate in class (class does not include APAD or ALE).

*Teachers need to keep their student samples/data until completion of their next evaluation cycle.

Student Growth Component Rating:

Excellent (4) – 75-100% of student sampling shows growth
Proficient (3) – 50-74% of student sampling shows growth
Needs Improvement (2) – 25-49% of student sampling shows growth
Unsatisfactory (1) – 0-24% of student sampling shows growth
Summative Rating:

The summative rating combines the Teacher Evaluation Rating and the Student Growth Component Rating.

\[(\text{Teacher Evaluation Rating}) \times 0.75 + (\text{Student Growth Component}) \times 0.25 = \text{Summative Evaluation}\]

*The district will utilize math rounding rules for the Summative Evaluation Rating: 
  - .5 rounds up/.49 rounds down

Student Growth Schedule:

- All teachers will participate in measuring student growth during the trial year (2015-16).
- Non-tenured teachers will have student growth requirements each year until they become tenured.
- Tenured teachers will measure their students’ growth during the year prior to the formal evaluation year.
- For teachers whose formal evaluation year is 2016-17, they will have the choice to use their trial year data and complete a plan during the trial year (2015-16) to be used during their formal evaluation year (2016-17) OR they will be required to complete a growth plan during the first semester of 2016-17.

Student Growth Process and Timeline:

**All Non-Tenured Teachers: Teachers That Teach Semester Classes:**

August-September:

- Receive growth plan approval
- Administer baseline assessment
- Analyze baseline assessment results

September:

- Establish growth target
- Complete and submit growth plan for approval by September 30th

October-November:

- Monitor Student Progress
- Administer mid-point assessment(s) – formal or informal
• Analyze data and adjust plan if needed. If changes are made to the plan, a meeting with evaluator is required.

December:
• Administer post assessment
• Analyze data

January:
• Submit data with growth calculations

Tenured Teachers: Teachers That Teach Year Long Classes:

August-September:
• Receive growth plan approval

October-March:
• Administer baseline assessment
• Analyze baseline assessment results
• Establish growth targets and communicate that to evaluator
• Monitor student progress and implement strategies
• Administer mid-point/post assessment(s) – formal or informal
• Analyze data and adjust plan if needed. If changes are made to the plan, a meeting with your evaluator is required.

April-May:
• Analyze data and complete growth calculations
• Submit to evaluator before the last day of student attendance

Growth Plans Components:

Growth plans can be a table, page from the gradebook, spreadsheet, skills based checklist, etc.

• Student sample (names of students included in your sample)
• Learning objective(s)
• Student growth implementation timeline (dates of assessments)
• Establish growth target for each student in your sample group
• Identify unique student circumstances (if applicable)
• Final data results

***
APPENDIX I

DISTRICT Y
Faculty Evaluation and Development

System

DISTRICT Y
In District [redacted], we are committed to the idea of continuous improvement for our students and for ourselves. It is our hope that this system will be one way we meet our promise to be best we can be for our students and our colleagues.

Overview:

The purpose of the [redacted] Teacher Evaluation and Development System is to provide a consistent and well-defined system of support and accountability that promotes quality instruction in every classroom. Ultimately, student success is achieved through ongoing faculty growth. The Framework for Quality Instruction (FQI) provides the backbone for this system, and consists of 5 domains: Design & Planning, Learning Environment, Instructional Delivery, Assessment, and Professionalism. Each domain is further broken down into categories, and thereafter sub-categories. For example, here is a section taken from the Framework that illustrates the structure:

**Domain D: Assessment**

1. Design of Effective Assessment
   a. Creates assessments congruent with learning targets
   b. Establishes clear criteria for students

The Domain is “Assessment,” the category “Design of Effective Assessment” and the sub-categories are identified as a and b. It is at the sub-category level – which contain the most specific descriptors of quality instruction – that faculty set goals; it is also at this level where the various libraries of resources exist to support the pursuit of such goals.

In this system, development and evaluation occur side-by-side. Faculty select professional goals within the FQI and create action plans for pursuing them, using a variety of data points from students, administrators, and their own sense of professional practice to inform their decisions. Then, throughout the year, administration provides ongoing feedback to the faculty on their professional practice and professional growth, again using the FQI as a guide. This ongoing drive for improvement on the part of the faculty member, combined with the ongoing feedback from the administration, leads to a summative conference at which point a specific rating is assigned.

The details regarding the procedures, ratings, and other elements of the [redacted] Teacher Evaluation and Development System are articulated in subsequent pages of this packet.
Domain A: Design and Planning

1. Demonstration of Knowledge of Subject Matter
   a. Understands subject matter
   b. Understands relevance to students’ lives

2. Demonstration of Knowledge of Students
   a. Shows awareness of students’ levels of skill
   b. Shows awareness of students’ background knowledge
   c. Shows awareness of students’ learning styles
   d. Shows awareness of students’ special needs
   e. Shows awareness of students’ cultural backgrounds
   f. Shows awareness of students’ social and emotional needs

3. Demonstration of Knowledge of Resources
   a. Integrates text purposefully and effectively into lesson planning
   b. Integrates technology purposefully and effectively into lesson planning
   c. Integrates other resources purposefully and effectively into lesson planning

4. Establishment of Learning Targets
   a. Aligns learning targets with curriculum map
   b. Sets expectation for rigor

Domain B: Learning Environment

1. Establishment of an Environment of Respect and Rapport
   a. Creates environment of mutual respect and trust
   b. Creates environment conducive to taking risks
   c. Displays interest in students’ lives outside of class

2. Establishment of a Safe yet Challenging Environment for Learning
   a. Displays positive attitude (verbal and non-verbal)
   b. Articulates high expectations
   c. Ensures learning is accessible and equitable to all students

3. Establishment of Rules and Procedures
   a. Uses clear rules, procedures and routines
   b. Uses consistent enforcement of rules and procedures
Domain D: Assessment

2. Design of Effective Assessment
   c. Creates assessments congruent with learning targets
   d. Establishes clear criteria for students

3. Variety of Assessments
   a. Uses formative and summative purposefully and effectively
   b. Uses authentic assessment tool purposefully and effectively

4. Feedback on Student Learning
   a. Gives feedback in frequent and timely fashion
   b. Gives descriptive feedback based on assessment criteria
   c. Gives feedback in a respectful and empowering manner
   d. Includes multiple stakeholders (student, family, case manager, counselor) as appropriate

5. Student Involvement in Assessment
   a. Ensures student comprehension of learning targets
   b. Ensures student comprehension of assessment criteria
   c. Ensures student awareness of own progress toward learning targets
   d. Encourages student self-assessment and reflection

Domain E: Professionalism

1. Professionalism
   a. Maintains positive, appropriate and collaborative relationships
   b. Engages in prof. development of content and pedagogy
   c. Participates in department/building/district initiatives
   d. Maintains accurate and timely records
   e. Engages in reflective practice
   f. Seeks feedback and enacts suggestions for improvement
   g. Carries self as a professional in terms of dress, language, punctuality, social interaction, et al.
   h. Demonstrates ethical use of resources
4. Management of Student Behavior
   a. Displays awareness of all students in class
   b. Uses reinforcement effectively for positive behavior
   c. Uses discipline respectfully and appropriately

**Domain C: Instructional Delivery**

1. Communication with Students
   a. Communicates learning targets to students
   b. Communicates directions clearly and concisely
   c. Uses voice effectively: volume, pace, enthusiasm
   d. Moves around room to utilize proximity
   e. Uses multiple modes of delivery (verbal, written, et al.)
   f. Acknowledges / celebrates student success

2. Using Questioning and Discussion Techniques
   a. Elicits higher-level thinking (Bloom’s taxonomy)
   b. Uses wait time effectively
   c. Uses persistence effectively
   d. Makes an effort to involve all students

3. Engaging Students in Learning
   a. Uses a variety of activities
   b. Uses an effective opening, closing, and transitions
   c. Uses research-based instructional strategies purposefully and effectively
   d. Uses technology purposefully and effectively
   e. Creates memorable, meaningful experience
   f. Makes connection between lesson and students’ lives

4. Flexibility and Responsiveness
   a. Displays awareness of students’ progress in lesson
   b. Adapts lesson plan to student progress and needs
Non-Tenured Teacher Summative Evaluation Process Overview

All non-tenured staff will receive a formal summative evaluation every year. The summative evaluation will be based on the following components:

A. Professional Practice
   - Classroom Observations (FQI Domains A, B, C, D)

B. Professional Growth
   - Goal Setting
   - Student Survey
   - Self Reflection
   - Professionalism (FQI Domain E)

At the end of each year, a summative evaluation will be completed by the building principal (or designee). A conference to explain the summative rating will be held between the principal (or designee) and the teacher not sooner than seventy-five days prior to the final day on the calendar (excluding emergency days).

If the designated evaluator is not the division leader, the division leader will provide input to the designated evaluator to complete the summative evaluation. Furthermore, the division leader, per the request of any party, will be present at this conference.

If the designated evaluator is the division leader, the other administrator who had been involved in the observation process will provide input to the division leader to complete the summative evaluation. Furthermore, the other administrator who had been involved in the observation process, per the request of any party, will be present at this conference.

A copy of the summative documents will be added to the faculty member's personnel file.

See the next page of this document for a visual representation of this cycle.

---

Framework for Quality Instruction
Elements of the Summative Evaluation

A. Professional Practice

Classroom Observation
Each non-tenured teacher will be observed as follows:

- During the first and second year of service in [__], the non-tenured teacher will be formally observed:
  - By the division leader a minimum* of one time during semester one, and one time during semester two.
  - By the building principal (or designee) a minimum* of one time
  - By a district administrator a minimum* of one time

- During the third and fourth years of service in [__], the non-tenured teacher will be formally observed:
  - By the division leader a minimum* of one time per semester
  - By the building principal (or designee) a minimum* of one time
  - By a district administrator by request

- All formal observations will include a pre-conference and a post-conference within 7 school days of the scheduled observation.
• No more than 10 days after the formal observation, a scored rubric with comments from the observer will be given to the teacher, division leader, and building principal.

• If an additional observation is requested, it must come within 10 days of the scored rubric delivery. The building principal determines who does the additional observation.

Unscheduled classroom visits such as drop-ins, walkthroughs, etc. may be done at any time. These would be in addition to the scheduled observations.

* Over the course of the summative cycle, at least three different qualified observers will complete a formal observation of any first or second year teacher. For any non-tenured teacher, an additional formal observation(s) shall occur per the request of the teacher, the DL, or the administration. The scores from all completed formal observations will be averaged together in order to calculate the teacher’s Professional Practice score (see chart below under “Summative Rating.”)

B. Professional Growth

Goal Setting
At the start of each school year, each non-tenured teacher will develop at least two professional goals based upon the FQI. Several data points will be considered in the formation of the goals:

1. Professional judgment of the faculty member regarding areas of potential improvement.
2. Feedback from the division leader or other administration from the prior year.
3. Student Survey data from the prior year.
4. Self-reflection from the prior year.

As a means of progressing towards the achievement of these goals, each teacher will develop an action plan, timelines, and measures of success. The goals, action plans, timelines, and measures of success will be mutually agreed upon by the teacher and the division leader, and approved by the principal (or designee). The teacher, division leader, and principal will have access to these throughout the year.

Student Input Survey
For teachers, the District’s Student Survey will be made available to all students via an electronic format over a three-week period of each semester as scheduled by the administration, typically during November and/or December of the fall semester, and during April and/or May of the spring semester.

For student services personnel, the student survey will be made available to students via a paper-and-pencil format during the same window of time as for teachers. Each faculty group will distribute a certain number of such surveys to students.

Results of the survey will be provided to the teacher in a timely manner. The student survey will be used to inform the goal setting process for the following year.

Self Reflection
No later than May 1, and/or at least one week before the summative evaluation conference, the teacher will complete a self-reflection form that highlights areas of success and areas for future improvement. The self-reflection will be used to inform the goal setting process for the following year.
Professionalism
The final component of the Professional Growth component of a non-tenured teacher's summative evaluation derives from Domain E of the FQI.

**Summative Rating**

The summative evaluation will contain one of the following ratings:
- Excellent ("E" on the chart below)
- Proficient ("P" on the chart below)
- Needs Improvement ("NI" on the chart below)
- Unsatisfactory ("UN" on the chart below)

The summative rating will be determined by the classroom observations, professional growth elements, and student growth elements as explained in the Scoring System Overview portion of this packet.

If a teacher is rated Needs Improvement or Unsatisfactory, the procedures established in the Illinois School Code will be followed to develop a professional development (Needs Improvement) or remediation (Unsatisfactory) plan.
Tenured Teacher Summative Evaluation Process Overview

All tenured staff will receive a formal summative evaluation once every two years. The summative evaluation will be based on the following components:

A. Professional Practice
   o Classroom Observations (FQI\textsuperscript{1} Domains A, B, C, D)

B. Professional Growth
   o Goal Setting
   o Student Survey
   o Self Reflection
   o Professionalism (FQI Domain E)

At the end of the first year of a teacher's two-year summative cycle, an end-of-year conference will be held between the teacher and division leader no later than May 31\textsuperscript{st}. During this conference, the division leader will provide feedback on the teacher's Professional Practice and Professional Growth components in order to help the teacher make plans for creating and/or revising goals for the upcoming second (summative) year of the evaluation cycle.

At the end of the second year of a teacher's two-year summative cycle, a summative evaluation will be completed by the building principal (or designee). A conference to explain the summative rating will be held between the principal (or designee) and the teacher at the end of the two-year summative cycle. This conference will occur not sooner than seventy-five days prior to the final day on the calendar (excluding emergency days).

If the designated evaluator is not the division leader, the division leader will provide input to the designated evaluator to complete the summative evaluation. Furthermore, the division leader, per the request of any party will be present at this conference.

If the designated evaluator is the division leader, the other administrator who had been involved in the observation process will provide input to the division leader to complete the summative evaluation. Furthermore, the other administrator who had been involved in the observation process, per the request of any party, will be present at this conference.

A copy of the summative documents will be added to the faculty member's personnel file.

See the next page of this document for a visual representation of this cycle.

\textsuperscript{1} Framework for Quality Instruction
Summative Evaluation Cycle

Evaluation Plan:

Year One

Summative Conference

Goal-Setting Conference

Minimum 1 Admin Observation

Professional Growth Rubric

Student Surveys

Teacher Self-Reflection

End of Year Conference

Year Two

Evaluation Rating

Minimum 1 DC Observation

Professional Growth Rubric

Student Surveys

Teacher Self-Reflection
Elements of the Summative Evaluation

A. Professional Practice

Classroom Observation
Each tenured teacher will be observed as follows:

- During the first year of the summative cycle, the teacher will be formally observed:
  - By the building principal (or designee) a minimum* of one time.

- During the second year of the summative cycle, the teacher will be formally observed:
  - By the division leader a minimum* of one time.

- All formal observations will include a pre-conference and a post-conference within 7 school days of the scheduled observation.

- No more than 10 days after the formal observation, a scored rubric with comments from the observer will be given to the teacher, division leader, and building principal.

- If an additional observation is requested, it must come within 10 days of the scored rubric delivery. The building principal determines who does the additional observation.

Unscheduled classroom visits such as drop-ins, walkthroughs, etc. may be done at any time. These would be in addition to the scheduled observations.

* Over the course of the summative cycle, at least two different qualified observers will complete a formal observation of any tenured teacher. Furthermore, an additional formal observation(s) shall occur per the request of the teacher, the DL, or the administration. The scores from all completed formal observations will be averaged together in order to calculate the teacher’s Professional Practice score (see chart below under “Summative Rating.”)

B. Professional Growth

Goal Setting
At the start of each school year, each teacher will develop at least two professional goals based upon the FQI. Several data points will be considered in the formation of the goals:

1. Professional judgment of the faculty member regarding areas of potential improvement.
2. Feedback from the division leader or other administration from the prior year.
3. Student Survey data from the prior year.
4. Self-reflection from the prior year.

As a means of progressing towards the achievement of these goals, each teacher will develop an action plan, timelines, and measures of success. The goals, action plans, timelines, and measures of success will be mutually agreed upon by the teacher and the division leader, and approved by the principal (or designee). The teacher, division leader, and principal will have access to these throughout the year.
Student Input Survey
For teachers, the District will make Student Survey available to all students via an electronic format over a three-week period of each semester as scheduled by the administration, typically during November and/or December of the fall semester, and during April and/or May of the spring semester. For student services personnel, the student survey will be made available to students via a paper-and-pencil format during the same window of time as for teachers. Each faculty group will distribute a certain number of surveys to students. Results of the survey will be provided to the teacher in a timely manner. The student survey will be used to inform goal setting for the following year.

Self Reflection
At least one week before the end-of-year conference or summative evaluation conference, the teacher will complete a self-reflection form that highlights areas of success and areas for future improvement, as well as a teacher survey. The self-reflection and teacher survey will be used to inform the goal setting process for the following year.

Professionalism
At the end of the first year of a teacher’s two-year summative evaluation cycle, during the end-of-year conference, the division leader will provide feedback to the teacher based on the components of the Professional Growth (PG) rubric. However, the PG rubric is only formally scored after 2 years.

Summative Rating
The summative evaluation will contain one of the following ratings:
- Excellent (“E” on the chart below)
- Proficient (“P” on the chart below)
- Needs Improvement (“NI” on the chart below)
- Unsatisfactory (“UN” on the chart below)

The summative rating will be determined by the classroom observations, professional growth elements, and student growth elements as explained in the Scoring System Overview portion of this packet.

If a teacher is rated Needs Improvement or Unsatisfactory, the procedures established in the Illinois School Code will be followed to develop a professional development (Needs Improvement) or remediation (Unsatisfactory) plan.
The summative evaluation is determined by a two step process. The first step determines a faculty member's rating according to their evaluations on Professional Practice and Professional Growth, and the second step incorporates student growth data.

**Step One**

For the first step, the ratings for each faculty member that will be used are Professional Practice and Professional Growth. The following grid is used, where the ratings for each category are abbreviated along each axis: UN = Unsatisfactory, Ni = Needs Improvement, P = Proficient, E = Excellent.

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>P</th>
<th>Ni</th>
<th>UN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E</strong></td>
<td>E</td>
<td>E</td>
<td>P</td>
<td>UN</td>
</tr>
<tr>
<td><strong>P</strong></td>
<td>E</td>
<td>P</td>
<td>P</td>
<td>Ni</td>
</tr>
<tr>
<td><strong>Ni</strong></td>
<td>P</td>
<td>Ni</td>
<td>Ni</td>
<td>UN</td>
</tr>
<tr>
<td><strong>UN</strong></td>
<td>Ni</td>
<td>Ni</td>
<td>UN</td>
<td>UN</td>
</tr>
</tbody>
</table>

A faculty member would find the intersection of the row and column that corresponds to their two ratings to determine their step one rating. Per the example below, if a faculty member received a "Needs Improvement" rating in Professional Practice (highlighted in green) and a "Proficient" rating in Professional Growth (highlighted in red), the step one rating would be "Needs Improvement" (highlighted in orange).
Example 2: If a faculty member received a "P" rating in Professional Practice and a "P" rating in Professional Growth, the step one rating would be "Proficient."

### Professional Growth

<table>
<thead>
<tr>
<th>E</th>
<th>P</th>
<th>NI</th>
<th>UN</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>E</td>
<td>E</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>NI</td>
<td>P</td>
</tr>
<tr>
<td>NI</td>
<td>P</td>
<td>NI</td>
<td>NI</td>
</tr>
<tr>
<td>UN</td>
<td>NI</td>
<td>UN</td>
<td>UN</td>
</tr>
</tbody>
</table>

### Professional Practice

Calculating Professional Growth for All Faculty

Contrary to the Professional Practice rating, which is determined using role-specific rubrics for various faculty positions, the Professional Growth axis (the x axis) is determined for all faculty using an identical rubric and procedures included elsewhere in this packet.

The anatomy of the Professional Growth rubric is distinct from the Professional Practice rubrics in that no individual sections exist with multiple descriptors; instead, 6 rows of descriptors exist. Here is a sample (incomplete) illustration:

**Professional Growth**

<table>
<thead>
<tr>
<th>Excellent</th>
<th>Proficient</th>
<th>Need Improvement</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>T independently or with minimal guidance creates goals that represent a true stretch of capability or comfort zone.</td>
<td>T independently or with minimal guidance generates goals that are appropriate in terms of content and level of challenge.</td>
<td>T needs extra guidance from others to generate goals that are appropriate in terms of content and level of challenge.</td>
<td>T resists extra guidance, and or requires outright and direction from others to create goals that are appropriate in terms of content and level of challenge.</td>
</tr>
<tr>
<td>Self-reflection: particularly insightful and also includes examples of impact on students.</td>
<td>Self-reflection articulates progress towards professional goals over time, and considers specific data in identifying</td>
<td>Self-reflection tends to be less developed in one or more areas, with progress towards professional goals.</td>
<td>Self-reflection is very general in nature, and/or shows little to no evidence of progress towards professional goals.</td>
</tr>
</tbody>
</table>
As with the other rubrics, each row of descriptors will be scored out of 4 possible points. Half points may again be used. Once all 6 rows have been scored, the following table is used to calculate the total score for Professional Growth on the x-axis of the Professional Practice/Professional Growth grid. For example, if a faculty member received a score of 20, the faculty member’s Professional Growth rating would be “Proficient” to be used on the x-axis of the summative grid.

<table>
<thead>
<tr>
<th></th>
<th>Top Range</th>
<th>Bottom Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>24</td>
<td>21.5</td>
</tr>
<tr>
<td>Proficient</td>
<td>21</td>
<td>15.5</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>15</td>
<td>9.5</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>9</td>
<td>6</td>
</tr>
</tbody>
</table>

**Step Two: Calculating Student Growth for ALL faculty**

The student growth is calculated by each building’s SAT scores. All subsections of the SAT will be used to determine an “Overall Growth” score, and each faculty member will use his/her home building’s “Overall Growth” score. For traveling teachers, the home building will be the building responsible for the teacher’s evaluation process.

* uses ECRA to determine whether a school has “higher than expected growth” “expected growth” “lower than expected growth” “unsatisfactory growth.” Each of these ECRA ratings corresponds to the D155 “Excellent” “Proficient” “Needs Improvement” and “Unsatisfactory” ratings. *For more details, refer to Appendix 1.*
Final Step: Calculating the Summative:

Once the faculty member’s Professional Practice/Professional Growth rating is known from step one, and the faculty member’s Student Growth rating is known from step two, the summative rating will be determined using the final summative grid, below.

The weighting of the two steps is that a faculty member’s Professional Practice / Professional Growth rating counts for 70% of the summative evaluation, and the student growth counts for 30%.

Example 1:
If a faculty member has an Excellent rating from their Professional Practice / Professional Growth from step one (highlighted in red), and a Proficient rating from their Student Growth element from step two (highlighted in green) then their summative evaluation for the year would be “Excellent” (highlighted in yellow and circled).

<table>
<thead>
<tr>
<th>Professional Practice/Professional Growth (Step One) 70%</th>
<th>Excellente</th>
<th>Proficient</th>
<th>Needs Improvement</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>Excellent</td>
<td>Proficient</td>
<td>Proficient</td>
<td>Safeguard</td>
</tr>
<tr>
<td>Proficient</td>
<td>Proficient</td>
<td>Proficient</td>
<td>Needs Improvement</td>
<td>Unsatisfactory</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>Proficient</td>
<td>Proficient</td>
<td>Needs Improvement</td>
<td>Unsatisfactory</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>Safeguard</td>
<td>Needs Improvement</td>
<td>Unsatisfactory</td>
<td>Unsatisfactory</td>
</tr>
</tbody>
</table>

Safeguard
Additional review is needed for any teacher whose combination of growth and practice ratings put them into one of two “safeguard” cells indicated above. If a teacher’s combination of growth and practice ratings fall within a safeguard cell, a review of the teachers’ growth and practice ratings will be conducted to arrive at a final decision regarding the teacher’s summative evaluation.
## Calculating Professional Practice for Teachers

The Professional Practice axis (the y-axis) is calculated based upon classroom observations, using the observational rubrics and procedures included elsewhere in this packet. Teachers use rubrics for Domains A, B, C, and D for the observations, each of which contains 4 separate sections. Each individual section is thereafter broken down into 4 columns, which correspond to the rating categories: Excellent, Proficient, Needs Improvement, and Unsatisfactory. Here is a sample section for "Flexibility and Responsiveness" taken from Domain C:

<table>
<thead>
<tr>
<th>Flexibility and Responsiveness</th>
<th>Excellent</th>
<th>Proficient</th>
<th>Needs Improvement</th>
<th>Unsatisfactory</th>
<th>Not Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taught only a small group of students, and focused on only one aspect of the lesson</td>
<td>Tever taught 2 or more students, and progressed towards LT3</td>
<td>Tever taught 3 or more students, and progressed towards LT4</td>
<td>Tever taught 4 or more students, and progressed towards LT5</td>
<td>Tever taught 5 or more students, and progressed towards LT6</td>
<td>N/A</td>
</tr>
<tr>
<td>Tever taught 1 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 2 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 3 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 4 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 5 or more students, and adjusted their teaching accordingly</td>
<td>N/A</td>
</tr>
<tr>
<td>Needs Improvement and/or recommendation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Anything identified as Needs Improvement, Unsatisfactory, or Not Applicable requires elaboration in the observations/recommendations section.

Each section contains a different number of descriptor rows, indicated by bullet points; in the example above, there are two. Regardless of the number of descriptor rows, each individual section is calculated out of 4 possible points. Half-points (i.e. 3.5/4) may also be used. For example, if an observer rated a teacher’s “Flexibility and Responsiveness” as seen below, the final score would be a 3.5 for this section.

<table>
<thead>
<tr>
<th>Flexibility and Responsiveness</th>
<th>Excellent</th>
<th>Proficient</th>
<th>Needs Improvement</th>
<th>Unsatisfactory</th>
<th>Not Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taught only a small group of students, and focused on only one aspect of the lesson</td>
<td>Tever taught 2 or more students, and progressed towards LT3</td>
<td>Tever taught 3 or more students, and progressed towards LT4</td>
<td>Tever taught 4 or more students, and progressed towards LT5</td>
<td>Tever taught 5 or more students, and progressed towards LT6</td>
<td>N/A</td>
</tr>
<tr>
<td>Tever taught 1 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 2 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 3 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 4 or more students, and adjusted their teaching accordingly</td>
<td>Tever taught 5 or more students, and adjusted their teaching accordingly</td>
<td>N/A</td>
</tr>
<tr>
<td>Needs Improvement and/or recommendation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Anything identified as Needs Improvement, Unsatisfactory, or Not Applicable requires elaboration in the observations/recommendations section.

Score = 3.5 / 4

The column N/A ("Not Applicable") might be used on occasion if a particular descriptor was determined to have been not applicable to the observation; however, no entire section will receive an N/A rating. If a descriptor is determined to be N/A, the score will be calculated using the remaining descriptors in the section, even if only one descriptor remains to be used.
Once all 4 sections in a Domain have been scored (each out of 4 points, for a total of 16 points per Domain), the following table is used to calculate the total score for that domain:

<table>
<thead>
<tr>
<th></th>
<th>Top Range</th>
<th>Bottom Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>16</td>
<td>14.5</td>
</tr>
<tr>
<td>Proficient</td>
<td>14</td>
<td>10.5</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>10</td>
<td>6.5</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

Starting with the 2012-2013 school year, the different domains will be weighted, with Domain A receiving 10% of the total score, and the other domains equally splitting the remaining 90%. The maximum points earned is 64, and the following grid determines the rating:

<table>
<thead>
<tr>
<th></th>
<th>Top Range</th>
<th>Bottom Range*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>64</td>
<td>56.5</td>
</tr>
<tr>
<td>Proficient</td>
<td>56.49</td>
<td>40.5</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>40.49</td>
<td>24.5</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>24.49</td>
<td>16</td>
</tr>
</tbody>
</table>

*NOTE: The "Bottom Range" scores have been adjusted in order to provide complete coverage of any calculation. (The bottom number of 14.5 for Excellent for any Domain, multiplied by 4, would provide a score of 58. This would leave a gap between 58 and the top range score of 56 for Proficient. So, the 58 was lowered to 56.5. Each rating was likewise lowered in this fashion.)

Applying to the Summative Rubric
Since at least two observations must occur, the average of all observation scores will be used to determine the overall Professional Practice score on the y-axis of the summative evaluation grid.

For example, if a teacher received scores of 60, 52.5 and 54 on three observations, that teacher's average score would be 55.5. (Calculated by (60+52.5+54)/3). Therefore, the teacher's Professional Practice rating would be "Proficient" to be used on the y axis of the summative grid.

<table>
<thead>
<tr>
<th></th>
<th>Top Range</th>
<th>Bottom Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>64</td>
<td>56.5</td>
</tr>
<tr>
<td>Proficient</td>
<td>56.49</td>
<td>40.5</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>40.49</td>
<td>24.5</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>24</td>
<td>16</td>
</tr>
</tbody>
</table>

55.5

Calculating Professional Growth for Teachers
Contrary to the Professional Practice rating, which is determined using role-specific rubrics for various faculty positions, the Professional Growth axis (the x-axis) is determined for all faculty using an identical rubric and procedures included elsewhere in this packet. See the above section "Calculating Professional Growth for ALL faculty" for details on how to determine the Professional Growth rating.
Teacher Evaluation

*Student Growth Component Template*

*Local Growth Model (LGM) All-In*
Table of Contents

Purpose .................................................................................................................. Page 2
Guiding Principles ............................................................................................... Page 2
Assessments .......................................................................................................... Page 3
Measurement Model for Student Growth ............................................................. Page 4
Growth Standards and Evaluation Categories ..................................................... Page 5
Combining Student Growth and Professional Practice ....................................... Page 6
This template is intended to inform and support districts that have adopted the Local Growth Model (LGM) All-in approach to student growth for the purpose of teacher evaluations. Districts are invited to use the template as is or customize as needed.

**Purpose**

The Illinois Performance Evaluation Reform Act (PERA), regulated as Illinois Administrative Code Part 50 (Part 50), requires that student growth be a significant factor in a teacher’s overall evaluation. This document summarizes the collaborative efforts and consensus of the Joint Committee (Section 24A-4 of the School Code) for how student growth will be incorporated into the district’s teacher evaluation system to promote continuous quality improvement and ensure compliance with state regulations.

**Guiding Principles**

The district’s student growth component is guided by the following principles:

- Student growth components for teacher evaluation should be student-centered, and should not be separate from or interfere with, the district’s assessment program used to measure individual student progress.
- Student growth components for teacher evaluation should focus on core academic outcomes such as reading and math.
- Student growth components should promote continuous quality improvement, and systemically build upon existing district assessment practices.
- Student growth components for teacher evaluation should promote teacher collaboration, and foster professional development.
- Student growth ratings should be standards-based, and result in the potential for all teachers to receive favorable student growth ratings.
- Student achievement growth should account for variations in past student performance, but should not explicitly adjust for student-level demographics in any way that may perpetuate achievement gaps.
Assessments

The summative evaluation for a teacher will consist of a student growth component and a professional practice component. The growth component will comprise 30% of a teacher’s summative rating, and professional practice will comprise 70% of the evaluation. This document is focused on the student growth component.

Part 50 requires that student growth components include the use of at least one Type I or Type II assessment, and at least one Type III assessment. The following diagram illustrates an example of a district’s teacher evaluation system components.

Assessment Types
The district may consider Type I and Type II assessments as Type III assessments for the purpose of teacher evaluation in a manner consistent with Part 50, which states “A Type I or Type II assessment may qualify as a Type III assessment if it aligns to the curriculum being taught and measures student learning in that subject area” (Section 50.110(b) (2) of Part 50). The district will communicate to teachers the specific assessments to be used prior to September 30th of the school year for which the assessment will be used.

Period of Time for Which Growth is Evaluated
For returning teachers, student growth will be retrospective to the previous school year, and will evaluate student growth over the course of an entire school year, culminating with spring assessments. For teachers new to the district, growth will be evaluated using the current school year’s data for half a school year, culminating with winter assessments.
Measurement Model for Student Growth

Part 50 requires that school districts adopt a measurement model for how two or more assessment scores, over two or more points in time will be used to determine student growth. A Local Growth Model (LGM) approach will be used to calculate individual student growth on all assessments based on local district norms.

The LGM analyzes longitudinal student achievement data individually for each student, and projects an individual student’s expected future achievement based on his/her prior pattern of achievement. Student projections are based on student propensities derived from individual student performance on multiple past assessments. This projection is then compared to actual student achievement, as illustrated below.

Measurement Model for Growth
Local Growth Model (LGM)

Student growth is defined as the distance between an individual student’s actual score on an assessment and the student’s projected score. Raw growth scores are converted to standardized growth scores via a conditional z-score, which provides a common definition and metric for growth across all assessments. Individual student growth scores are then averaged across all tests for all students in a particular school, and compared to established growth standards to determine final evaluation categories. Each assessment used within the teacher evaluation system will be given equal weight in its contribution toward average student growth at the individual student and group level.
Combining Student Growth and Professional Practice

A teacher’s overall student growth rating will be combined with a teacher’s overall professional practice rating to determine a teacher’s summative rating. Growth and practice ratings will be combined using the following matrix. In the example illustrated below, if a teacher’s overall student growth rating is Needs Improvement and a teacher’s overall practice rating is Proficient, then the teacher’s summative rating is Proficient.

<table>
<thead>
<tr>
<th>Overall Teacher Practice Rating (70%)</th>
<th>Excellent (4 points)</th>
<th>Proficient (3 points)</th>
<th>Needs Improvement (2 points)</th>
<th>Unsatisfactory (1 point)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent (4 points)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficient (3 points)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Needs Improvement (2 points)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unsatisfactory (1 point)</td>
<td></td>
<td>Safeguard Additional Review Required</td>
<td>Needs Improvement</td>
<td>Needs Improvement</td>
</tr>
</tbody>
</table>

**Safeguards**

Additional review is needed for any teacher whose combination of growth and practice ratings put them into one of two “safeguard” cells indicated above. If a teacher’s combination of growth and practice ratings fall within a safeguard cell, a review of the teacher’s growth and practice ratings will be conducted to arrive at a final decision regarding the teacher’s summative evaluation.

**The Joint evaluation committee will continue to meet and provide oversight to ensure consistent implementation, and to recommend changes in subsequent years.**
Teacher Performance Evaluation System
STUDENT GROWTH

PERA JOINT COMMITTEE
April 29, 2015
April 13, 2016
Teacher Performance Evaluation System

*Professional Practice and Student Growth*

Revised July, 2018
Teacher Performance Evaluation System

Professional Practice and Student Growth

High School District

July 2019
| Section 1 – Overview | Statement of Purpose  
|                      | Evaluation Plan Committee Process and Members |
| Section 2 – Timelines and Processes | Evaluation Cycle  
| Pages 4-11 | Observation Requirements Under PERA  
|                      | PERA Joint Committee Responsibilities  
|                      | Evaluation Timeline and Process  
|                      | Review and Appeal  
|                      | Professional Remediation Plan |
| Section 3 – Professional Practice | Professional Practice Levels of Performance Descriptors  
| Pages 12-13 | Professional Practice Domain and Summative Ratings |
| Section 4 – Student Growth | Student Growth Summative Rating  
| Pages 14-22 | Users and Uses for Assessment Tools  
|                      | Assessments for Measuring Student Growth  
|                      | Student Growth Goal Components and Timelines  
|                      | Student Growth Goal Criteria and Revisions  
|                      | Student Growth Goal Rubric and Scoring Procedures  
|                      | Overall Performance  
|                      | Evaluation Rating |
| Section 5 – Glossary | Glossary |
| Pages 23-25 |  
| Section 6 – Forms and Guiding Documents | Charlotte Danielson’s Framework for Teaching  
| Pages 26-45 | Pre-Observation Conference Form  
|                      | Post-Observation Conference Form  
|                      | Student Growth Process Summary Form  
|                      | Student Growth Goal Form  
|                      | Midpoint Review Document  
|                      | Student Growth Summative Rating Form  
|                      | Performance Evaluation Summative Rating  
|                      | Professional Growth Plan Template (Used after receiving a rating of “Needs Improvement”) |
Section 1

STATEMENT OF PURPOSE
We, the Professional Practice and PERA Joint Committees of [Redacted] High School District, honor the great work teachers do and promote ongoing teacher development that leads to measured student growth and a teacher evaluation system that are both meaningful and manageable. Our decisions were made to be sustainable from year to year but at the same time allow flexibility for improvement and growth.

EVALUATION PLAN COMMITTEE PROCESS AND MEMBERS
In collaboration with the Consortium for Educational Change (CEC), http://ceccweb.org, the School District Professional Practice Committee engaged in a collaborative process to develop an Educator Evaluation Plan meeting the requirements of the Performance Evaluation Reform Act (PERA 2010), which calls for a process that considers the growth and achievement of all students and supports a mindset of professional growth for all educators.

The Educator Evaluation Plan for [Redacted] School District utilizes Charlotte Danielson’s Framework for Teaching Evaluation Instrument (2013) as the basis of defining effective professional practices. Aligned rubrics were decided upon to qualify the level of teaching and support services within each domain for all certified educator roles in the district. Methods to improve student learning are clearly identified in the following four domains of professional responsibility: Planning and Preparation, the Environment, Instruction/Delivery of Services, and Professional Responsibilities. Within the domains are components and descriptive elements that further refine understanding of effective professional practice for all staff.

Between January 2015 and May 2015, the PERA Joint Committee (PIC), in consultation with the Consortium for Educational Change (CEC), met multiple times to craft a plan that would meet the unique needs of educators and fulfill the state requirements of the student growth component. The Professional Practice Committee met between January 2017 and May 2018 to revise the professional practice portion of the Teacher Evaluation System. The PIC and Professional Practice Committees were comprised of equal representation from educators and administrators, including a variety of disciplines and grade levels.

<table>
<thead>
<tr>
<th>PERA JOINT COMMITTEE MEMBERS 2018-19</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Redacted] Teacher</td>
</tr>
<tr>
<td>[Redacted] Teacher</td>
</tr>
<tr>
<td>[Redacted] Teacher</td>
</tr>
<tr>
<td>[Redacted] Assistant Superintendent</td>
</tr>
<tr>
<td>[Redacted] Assistant Principal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PROFESSIONAL PRACTICE COMMITTEE MEMBERS 2017-18</th>
<th>PERA JOINT COMMITTEE MEMBERS 2015-17</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Redacted] Teacher</td>
<td>[Redacted] Teacher</td>
</tr>
<tr>
<td>[Redacted] Teacher</td>
<td>[Redacted] Teacher</td>
</tr>
<tr>
<td>[Redacted] Teacher</td>
<td>[Redacted] Teacher</td>
</tr>
<tr>
<td>[Redacted] Assistant Superintendent</td>
<td>[Redacted] Assistant Superintendent</td>
</tr>
<tr>
<td>[Redacted] Assistant Superintendent</td>
<td>[Redacted] Assistant Superintendent</td>
</tr>
<tr>
<td>[Redacted] Assistant Principal</td>
<td>[Redacted] Assistant Principal</td>
</tr>
</tbody>
</table>
Section 2

NON-TENURED TEACHER EVALUATION CYCLE UNDER PERA EVALUATION PROCESS and TIMELINE REQUIREMENTS

Non-Tenured Evaluation Cycle
(At least once Every Year)

- First Day of Student Attendance: Teacher notified as to how he/she will be evaluated.
- Teacher develops student growth plan with multiple measures: plan should include baseline assessment scores, learning targets, growth expectations, and student populations.
- Teacher and Qualified Evaluator meet to review and approve plan.
- First Formal observation: includes pre-conference, observation, and post-conference.
- Informal Observations: Documentation and opportunity for in-person discussion if requested.

<table>
<thead>
<tr>
<th>Professional Practice</th>
<th>RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Observation: Pre-Conference + Observation + Post-Conference</td>
<td></td>
</tr>
<tr>
<td>2nd Observation: Pre-Conference + Observation + Post-Conference</td>
<td></td>
</tr>
<tr>
<td>3rd Observation: Pre-Conference + Observation + Post-Conference OR Informal Observation with written feedback and an opportunity for in-person discussion</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Student Growth</th>
<th>RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher develops student growth plan: baseline data, identify learning objectives, establish growth expectations, and establish student population</td>
<td></td>
</tr>
<tr>
<td>Plan Approval Conference: Qualified evaluator approves teacher's plan</td>
<td></td>
</tr>
<tr>
<td>Mid-Point Review: &quot;Teacher should use data to assess his/her progress and adjust instruction, if necessary&quot;</td>
<td></td>
</tr>
<tr>
<td>Final data review of student growth assessments</td>
<td></td>
</tr>
</tbody>
</table>

Non-Tenured Teachers - "Each School Year"

Overall Final Summative Rating
Percent Based on PERA Joint Committee Determination
Section 2

TENURED TEACHER EVALUATION CYCLE UNDER PERA EVALUATION PROCESS and TIMELINE REQUIREMENTS

First day of student attendance: Teacher notified as to how he/she will be evaluated.

Teacher develops student growth plan with multiple measures: plan should include baseline assessment scores, learning targets, growth expectations, and student populations.

Teacher and Qualified Evaluator meet to review and approve plan.

First Formal observation: includes pre-conference, observation, and post-conference.

Final summative conference: should include review of professional practice and student growth evidence to determine final overall rating.

Tenured Evaluation Cycle (At Least Every 2 Years)

Second formal observation: includes pre-conference, observation, and post-conference. OR Informal Observation with written documentation.

Mid-point review of student growth measures.

Tenured Teachers
(Rated “Excellent” or “Proficient” in last evaluation) * “At Least Once Every Two Years”

<table>
<thead>
<tr>
<th>Professional Practice</th>
<th>RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Observation: Pre-Conference + Observation + Post-Conference</td>
<td></td>
</tr>
<tr>
<td>2nd Observation: Pre-Conference + Observation + Post-Conference OR Informal Observation with written feedback and an opportunity for in-person discussion</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Student Growth</th>
<th>RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher develops student growth plan: baseline data, identify learning objectives, establish growth expectations, and establish student population</td>
<td></td>
</tr>
<tr>
<td>Plan Approval Conference: Qualified evaluator approves teacher’s plan</td>
<td></td>
</tr>
<tr>
<td>Mid-Point Review: “Teachers should use data to assess his or her progress and adjust instruction, if necessary”</td>
<td></td>
</tr>
<tr>
<td>Final data review of student growth assessments</td>
<td></td>
</tr>
</tbody>
</table>

Overall Final Summative Rating
Percent Based on PERA Joint Committee Determination

* Tenured teachers rated “Needs Improvement” or “Unsatisfactory” shall be evaluated in the next school year after the rating was issued. The cycle for that teacher will require the same number of observations as the non-tenured teachers. 

Issued. The cycle for that teacher will require the same number of observations as the non-tenured teachers.
### TEACHER EVALUATION TIMELINE

<table>
<thead>
<tr>
<th>Professional Practice</th>
<th>Student Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NON-TENURED YEAR 1-4 (annually)</strong></td>
<td><strong>NON-TENURED YEAR 1-4 (annually)</strong></td>
</tr>
<tr>
<td><strong>July</strong></td>
<td><strong>Notification of Evaluator August 1</strong></td>
</tr>
<tr>
<td><strong>Aug</strong></td>
<td>Self-Evaluation due by the end of the 3rd week of school</td>
</tr>
<tr>
<td></td>
<td>Professional Growth Plan Implementation</td>
</tr>
<tr>
<td><strong>Sept</strong></td>
<td><strong>Goal Setting Conferences by the 9th week of the semester</strong></td>
</tr>
<tr>
<td><strong>Oct</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Nov</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Dec</strong></td>
<td>First Observation (formal) due by the end of the semester</td>
</tr>
<tr>
<td><strong>Jan</strong></td>
<td>Second Observation (formal/informal)</td>
</tr>
<tr>
<td><strong>Feb</strong></td>
<td>Third Observation (formal/informal)</td>
</tr>
<tr>
<td></td>
<td><em>Non-tenured teachers must have 3 observations (2 formal and 1 informal)</em></td>
</tr>
<tr>
<td><strong>Mar</strong></td>
<td>Professional Growth Plan Final Report by March 1</td>
</tr>
<tr>
<td></td>
<td>Summative Evaluation Conference with Professional Practice Rating, Student Growth Rating &amp; Final Summative Rating by March 1</td>
</tr>
<tr>
<td><strong>Apr</strong></td>
<td></td>
</tr>
<tr>
<td><strong>May</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Tenured teachers with needs improvement or unsatisfactory ratings must have 3 observations (2 formal and 1 informal)*
Section 4

STUDENT GROWTH SUMMATIVE RATING
PERA (Performance Evaluation Reform Act) requires student growth be incorporated into teacher evaluation plans for tenured and probationary teachers. The intention of a performance evaluation plan that includes measures of student growth is to improve teaching and learning.

- **Tenured faculty members**: Student growth will be measured for evaluation only in summative evaluation years.
- **Non-Tenured faculty members**: Student growth will be measured annually.

CATEGORIES OF FACULTY MEMBERS and TYPES of ASSESSMENTS
Under PERA the evaluation plan must identify types of assessments for evaluating each category of faculty member. Categories of faculty members and assessments were kept broad given the evolving work in on assessments and the complexity of the faculty member descriptions. This provides an opportunity for collaboration between evaluators and faculty members, and autonomy in decision-making.

All faculty members will use multiple measures to determine student growth (except category E).

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>DESCRIPTION</th>
<th>PROFESSIONAL PRACTICE</th>
<th>STUDENT GROWTH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Type I</td>
</tr>
<tr>
<td>A</td>
<td>English, Math, Science, Social Studies</td>
<td>70%</td>
<td>12.5%</td>
</tr>
<tr>
<td>B</td>
<td>NOT English, Math, Science, Social Studies</td>
<td>70%</td>
<td>12.5%</td>
</tr>
<tr>
<td>C</td>
<td>Deans, Librarians</td>
<td>70%</td>
<td>12.5%</td>
</tr>
<tr>
<td>D</td>
<td>Non-classroom specialists (i.e., instructional specialists; interventionists)</td>
<td>70%</td>
<td>12.5%</td>
</tr>
<tr>
<td>E</td>
<td>Counselors, Psychologists, Social Workers, Speech and Language Pathologists</td>
<td>100%</td>
<td>none</td>
</tr>
</tbody>
</table>

*Faculty members have the option to choose a Type II instead of a Type III with evaluator approval.*
ASSESSMENTS FOR MEASURING STUDENT GROWTH

The intention of including measures of student growth under PERA is to improve teaching and learning. The following assessment criteria will be used for Type I, Type II, and Type III Assessments to measure student growth.

<table>
<thead>
<tr>
<th>ASSESSMENT PURPOSE &amp; CRITERIA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TYPE I ASSESSMENT DEFINITION</strong></td>
</tr>
<tr>
<td>An assessment that measures a certain group of students in the same manner with the same potential assessment items, is scored by a non-district entity, and is widely administered beyond IL</td>
</tr>
</tbody>
</table>

TO IMPROVE TEACHING AND LEARNING

<table>
<thead>
<tr>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A nationally/locally normed assessment that measures district or building growth targets (in reading or math) tied to the school improvement plan that allows the educator to benchmark growth at a national or district level</td>
</tr>
<tr>
<td>An assessment that measures student achievement in a specified content area for students across the district, tied to standards, and aligned with curriculum that allows the educator to benchmark growth at the district level</td>
</tr>
<tr>
<td>An assessment that measures student achievement in a specified content area for an identified group of students, tied to standards, and aligned with curriculum that allows the educator to benchmark growth at the individual classroom level</td>
</tr>
</tbody>
</table>

**SCHOOL-WIDE ASSESSMENT CRITERIA for TYPE I**
- Must be feasible
- Must fit the evaluation interval
- Assessment results must be available in time to be useful for the teacher evaluation process (i.e. baseline data)
- Assessment results can guide instructional goals
- Cost is sustainable
- Align to national/state standards
- Align to district academic goals
- Align to essential learning outcomes
- Measure intended outcomes of the class/course
- Be administered a minimum of two times during the year with the intent to maximize the assessment interval as much as possible with the minimum interval being a semester

**GRADE LEVEL/DEPARTMENT ASSESSMENT CRITERIA for TYPE I**
- Align to national, state or district academic standards
- Standards align to course goals to measure the appropriate content/concepts and/or skills
- Assess at various levels of skills (i.e., beginning, progressing, proficient, advanced)
- Overall, the items, tasks, rubrics are at the appropriate level of rigor
- Measure what is intended
- Same administration and results reporting for all students
- Be administered a minimum of two times during the year with the intent to maximize the assessment interval as much as possible with the minimum interval being a semester
- Pre and post assessments are the same or mirrored in skill, content, form & complexity

**INDIVIDUAL ASSESSMENT CRITERIA for Type III**
- Align to national, state or district academic standards
- Align to PLC Goals when appropriate
- Measure what is intended
- Items/tasks cover key subject/grade level content standards
- Possess a sufficient number of items
- Assess at various levels of skills (i.e., beginning, progressing, proficient, advanced)
- Overall, the items, tasks, rubrics are at the appropriate level of rigor
- Be administered a minimum of two times during the year with the intent to maximize the assessment interval as much as possible with the minimum interval being a semester unless otherwise agreed upon by the teacher and evaluator
- Pre and post assessments are the same or mirrored in skill, content, form & complexity
### Section 4

#### STUDENT GROWTH GOAL COMPONENTS and TIMELINES

<table>
<thead>
<tr>
<th>COMPONENT</th>
<th>TIMELINE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Population</strong></td>
<td>Description of the students will be included in the Student Growth Goal.</td>
</tr>
<tr>
<td><strong>2. Timeframe</strong></td>
<td>Description of assessment interval.</td>
</tr>
<tr>
<td><strong>3. Learning Goal</strong></td>
<td>Description of what students will be able to do at the end of the Student Growth Goal. The learning goal must be aligned to standards.</td>
</tr>
<tr>
<td><strong>4. Assessments and Scoring</strong></td>
<td>Procedures used to support and measure the learning goal. Assessments must meet the criteria for the assessment type.</td>
</tr>
<tr>
<td><strong>5. Growth Targets</strong></td>
<td>The outcome that students are expected to achieve by the end of the instructional period and includes consideration of a starting level of achievement already acquired and determination of an ending goal for the level of achievement to be reached.</td>
</tr>
<tr>
<td><em>(Collect Baseline Data, Group Students, Determine growth target)</em></td>
<td><em>(Baseline data provides measures of student understanding and ability to apply content. Can be baseline assessments, IEP data, attendance data, student surveys, English language proficiency)</em></td>
</tr>
<tr>
<td><strong>6. Midpoint Review Meeting</strong></td>
<td>Non-evaluative meeting with the date to be determined by teacher and evaluator. The Midpoint Review is a required face-to-face meeting between the faculty member and evaluator that is non-evaluative.</td>
</tr>
<tr>
<td></td>
<td>- Faculty members may suggest adjustments on their student growth plan during the meeting with the evaluator’s approval.</td>
</tr>
<tr>
<td></td>
<td>- The faculty member and evaluator will follow the Midpoint Conversation Guide.</td>
</tr>
<tr>
<td></td>
<td>- The faculty member and evaluator will discuss the Student Growth next steps in the process including timelines for feedback.</td>
</tr>
<tr>
<td><strong>7. Student Growth Actual Outcomes</strong></td>
<td>- The faculty member determines the number of students who meet their growth target.</td>
</tr>
<tr>
<td></td>
<td>- Within two weeks after the assessment interval, the faculty member will submit the following to their evaluator:</td>
</tr>
<tr>
<td></td>
<td>- Actual outcomes on the Student Growth Goal</td>
</tr>
<tr>
<td></td>
<td>- Self-assessment on section D and E on the Student Growth Rating Rubric</td>
</tr>
<tr>
<td></td>
<td>- Evidence (e.g. benchmark exams, pre-tests, informal assessments, post-tests, etc.)</td>
</tr>
<tr>
<td></td>
<td>- Following the Student Growth Goal submission, the evaluator and faculty member will jointly review the Student Growth Goal Rating Rubric and the evaluator determines the summative rating for Student Growth.</td>
</tr>
<tr>
<td></td>
<td>- Faculty member will submit to evaluator within two weeks after the assessment interval is completed.</td>
</tr>
<tr>
<td></td>
<td>- Evaluator will respond to the faculty member’s submission within two weeks after receipt.</td>
</tr>
<tr>
<td></td>
<td>- Faculty member or evaluator may ask for face-to-face meeting.</td>
</tr>
</tbody>
</table>
Section 4

Process for Midpoint Review as outlined below:

1. The Midpoint Review is a required face-to-face meeting between the faculty member and evaluator that is non-evaluative.
2. The faculty member and evaluator will determine the date of the Midpoint Review Meeting within weeks 8 through 11 of the semester.
3. Faculty members may make adjustments on their student growth plan during the meeting with the evaluator's approval.
4. The Student Growth Goal may be revised in the following areas:
   Attendance
   • 85% attendance threshold assumed for the assessment window (pre-test to post-test).
   • Faculty members can include students who do not meet the threshold on a case by case basis with evaluator approval.
   General Exceptions
   • Faculty members can request a student be excluded under "dire circumstances" with evaluator approval.
   • A student may not be excluded based on a single characteristic such as special education, ELL, etc.
   Growth Targets
   • Must provide a rationale that is supported by evidence (e.g. formative assessment data) that the growth target was unrealistic.
5. If the faculty member and evaluator cannot agree they must follow the PERAmedic Process as outlined on page 18.
6. The faculty member and evaluator will follow the Midpoint Conversation Guide below:

Midpoint Review Conversation/Reflection Guide for Faculty members and Evaluators

<table>
<thead>
<tr>
<th>Increasing Levels of Teacher Effectiveness</th>
<th>Midpoint Instructional Tuning</th>
</tr>
</thead>
<tbody>
<tr>
<td>No evidence that assessment information is used to guide or adjust teaching.</td>
<td>Faculty member uses assessment information to guide teaching.</td>
</tr>
<tr>
<td>• Narrow the focus of lessons to help students master a specific knowledge, reasoning, skill, or product target.</td>
<td></td>
</tr>
<tr>
<td>• Focus teaching on one portion of a scoring guide or rubric at a time, teaching lessons about quality, and offering descriptive feedback on just that portion.</td>
<td></td>
</tr>
<tr>
<td>• Summative assessment results are used to improve instruction for the next year when the lesson/experience will be taught again.</td>
<td></td>
</tr>
<tr>
<td>Faculty member uses assessment information to guide and adjust teaching and learning.</td>
<td></td>
</tr>
<tr>
<td>• Teacher use results to diagnose and adjust instruction in order to respond to student needs.</td>
<td></td>
</tr>
<tr>
<td>• Students and teacher talk together about assessment results to identify what the student knows, what still needs to be learned, and how learning can be extended.</td>
<td></td>
</tr>
<tr>
<td>• Formative results are used to reteach and extend instruction while the lesson/experience is still occurring.</td>
<td></td>
</tr>
<tr>
<td>• Summative results are used to reteach and extend learning following the conclusion of the lesson/experience.</td>
<td></td>
</tr>
<tr>
<td>Faculty member, students and users beyond the classroom use assessment information to guide and adjust teaching and learning.</td>
<td></td>
</tr>
<tr>
<td>• Faculty members and students review assignments and assessments for error patterns.</td>
<td></td>
</tr>
<tr>
<td>• Students use results to identify what they know and what they still need to learn.</td>
<td></td>
</tr>
<tr>
<td>• Students learn to recognize and avoid recurring errors.</td>
<td></td>
</tr>
<tr>
<td>• Faculty members use error patterns to inform reteaching needs.</td>
<td></td>
</tr>
<tr>
<td>• Students continue learning beyond the end of the unit and are given an opportunity to evidence their learning at any point in time.</td>
<td></td>
</tr>
<tr>
<td>• Specialists use results to guide instruction within reteaching loops.</td>
<td></td>
</tr>
<tr>
<td>• Faculty teams use results to refine and improve unit design.</td>
<td></td>
</tr>
</tbody>
</table>

SCORING THE STUDENT GROWTH PLAN TO DETERMINE PERCENTAGE OF STUDENTS WHO MET THEIR GROWTH TARGET

1. The faculty member determines the number of students who meet their growth target.
2. Within two weeks after the assessment interval, the faculty member will submit the following to their evaluator:
   a. Actual outcomes on the Student Growth Goal
   b. Completed self-assessment on section D and E on the Student Growth Rating Rubric
   c. Evidence (e.g. benchmark exams, pre-post tests, etc.)
Section 4

d. Completed Student Growth Goal form

3. Following the Student Growth Goal submission, the evaluator and faculty member will jointly review the Student Growth Goal Rating Rubric and the evaluator determines the summative rating for student growth.

APPROACH TO DETERMINE STUDENT GROWTH SUMMATIVE RATING
A combination of numerical scoring and performance-level descriptions in one rubric is used to determine the student growth summative rating. With this approach, faculty members and evaluators have conversations to apply the rubric. The rationale for this approach includes the following:

- Encourages continuous improvement in a collaborative manner
- Promotes collaborative conversations that connect professional practice and student growth
- Focus on accountability for assessment practices as well as results will promote continuous improvement
- Builds capacity for faculty members and evaluators
- Flexible across diverse content areas

MEASUREMENT MODEL
The PERA Joint Committee considered statistical measurement models used to analyze assessment data to determine student growth. The PERA Joint Committee determined that any of the three following measurement models can be applied.

- **Gain Based Model:** The average in student scores from one assessment to the next by subtracting their pre-test scores from post-test scores.

- **Adaptive Conditional Measurement Model:** Predicts student performance using past data. Under this model, teachers may consider a variety of precursor factors that may impact the current status of a student or a group of students.

- **Multivariate Model:** The use of multiple outcome variables. Multivariate models are distinguished by their complexity and their ability to use a large amount of data and variables in a unified approach.
## STUDENT GROWTH RATING RUBRIC and SCORING PROCEDURES

### Type I Assessment Student Growth Goal Rubric

<table>
<thead>
<tr>
<th>Area</th>
<th>Indicator</th>
<th>1 - Unsatisfactory</th>
<th>2 - Needs Improvement</th>
<th>3 - Proficient</th>
<th>4 - Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Monitoring and Provenance of Student Goal</td>
<td>Teacher does not demonstrate an understanding of the appropriate use of District's academic standards</td>
<td>District's academic standards are reviewed at building or team level meetings</td>
<td>Teacher adjourns instruction occasionally to District's academic standards</td>
<td>Teacher adjourns instruction regularly to District's academic standards</td>
</tr>
<tr>
<td></td>
<td>Less than 25% of students meet growth goals</td>
<td>26-50% of students meet growth goals</td>
<td>51-75% of students meet growth goals</td>
<td>76-100% of students meet growth goals</td>
<td></td>
</tr>
</tbody>
</table>

### Type III Assessment Student Growth Goal Rubric

<table>
<thead>
<tr>
<th>Area</th>
<th>Indicator</th>
<th>0 - Absent</th>
<th>1 - Unsatisfactory</th>
<th>2 - Needs Improvement</th>
<th>3 - Proficient</th>
<th>4 - Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Establishing and Monitoring Student Growth</td>
<td>Teacher failed to establish growth goals</td>
<td>Student growth goal is not measurable and/or not aligned with content standards</td>
<td>Student goal is written for growth during a determined period of time</td>
<td>Student goal is written for growth during a determined period of time</td>
<td>Student goal is written for growth during a determined period of time</td>
</tr>
<tr>
<td></td>
<td>Establishing and Monitoring Student Growth</td>
<td>Student growth goal is not measurable and/or not aligned with content standards</td>
<td>Teacher has a rationale that demonstrates an incomplete knowledge of student needs</td>
<td>Goal is measurable and aligned with content standards addressing the needs of differentiated group(s)</td>
<td>Teacher has a clear rationale that demonstrates knowledge of student needs</td>
<td>Goal is measurable and aligned with content standards addressing the needs of differentiated group(s)</td>
</tr>
<tr>
<td></td>
<td>Establishing and Monitoring Student Growth</td>
<td>Teacher has a rationale that demonstrates an incomplete knowledge of student needs</td>
<td>Teacher monitors goals occasionally and adjusts instruction accordingly</td>
<td>Teacher monitors goals occasionally and adjusts instruction accordingly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>&lt;= 25% of Student Growth</td>
<td>Less than 25% of students meet growth goals</td>
<td>26-50% of students meet growth goals</td>
<td>51-75% of students meet growth goals</td>
<td>76-100% of students meet growth goals</td>
<td></td>
</tr>
</tbody>
</table>

### Examples of “monitoring and adjusting” (not limited to): lesson planning, teacher developed documentation, communication artifacts

### Examples of “pieces of evidence” (not limited to): Baseline data(s), benchmark exams, pre/post tests, performance evaluations, portfolios, work samples, chapter/unit tests, formative assessments, NWEA MAP data

### SCORING THE STUDENT GROWTH RATING RUBRIC to determine the student growth summative rating:

<table>
<thead>
<tr>
<th>AREA</th>
<th>RATING</th>
<th>WEIGHTED PERCENT</th>
<th>AREA SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.05</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.05</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.05</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.05</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.05</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

| SUM (ROUND UP) | 4.0 |

### STUDENT GROWTH SUMMATIVE RATING SCALE

<table>
<thead>
<tr>
<th>Scale</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>4</td>
</tr>
<tr>
<td>Proficient</td>
<td>3</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>2</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>1</td>
</tr>
</tbody>
</table>
Section 4

PERFORMANCE EVALUATION RATING

The Summative Performance Evaluation Rating will combine the Professional Practice Summative Rating and the Student Growth Summative Rating as shown in the graphic below.

PERFORMANCE EVALUATION RATING BREAKDOWN:

<table>
<thead>
<tr>
<th>Professional Practice</th>
<th>Unsatisfactory</th>
<th>Needs Improvement</th>
<th>Proficient</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
<td>Unsatisfactory</td>
<td>Needs Improvement</td>
<td>Proficient</td>
<td>Excellent</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>Unsatisfactory</td>
<td>Needs Improvement</td>
<td>Proficient</td>
<td>Proficient</td>
</tr>
<tr>
<td>Proficient</td>
<td>Needs Improvement</td>
<td>Needs Improvement</td>
<td>Proficient</td>
<td>Excellent</td>
</tr>
<tr>
<td>Excellent</td>
<td>Needs Improvement</td>
<td>Proficient</td>
<td>Proficient</td>
<td>Excellent</td>
</tr>
</tbody>
</table>
Charlotte Danielson's FRAMEWORK FOR TEACHING

**DOMAIN 1: Planning and Preparation**

1a Demonstrating Knowledge of Content and Pedagogy
   - Content knowledge
   - Sequence relationships
   - Pedagogical knowledge

1b Demonstrating Knowledge of Students
   - Child development
   - Learning processes
   - Special needs
   - Interests and cultural heritage

1c Setting Instructional Outcomes
   - Value, sequence, and alignment
   - Clarity of learning goals
   - Suitability for diverse learners

1d Demonstrating Knowledge of Resources
   - For classroom
   - Transfer content knowledge
   - For students

1e Designing Coherent Instruction
   - Learning activities
   - Instructional materials and resources
   - Instructional groups
   - Lesson and unit structure

1f Designing Student Assessments
   - Congruence with outcomes
   - Criteria and standards
   - Formative assessments
   - Use for planning

**DOMAIN 4: Professional Responsibilities**

4a Reflecting on Teaching
   - Accuracy
   - Use of future teaching

4b Maintaining Accurate Records
   - Student completion of assignments
   - Student progress in learning
   - Non-instructional records

4c Communicating with Families
   - About instructional program
   - About individual students
   - Engagement of families in instructional program

4d Participating in a Professional Community
   - Relationships with colleagues
   - Participation in school projects

4e Growing and Developing Professionally
   - Enhancement of content knowledge and pedagogical skills
   - Interaction and feedback from colleagues

4f Showing Professionalism
   - Integrity/ethical conduct

**DOMAIN 2: Instruction**

2a Designing Coherent Instruction
   - Learning activities
   - Instructional materials and resources
   - Instructional groups
   - Lesson and unit structure

2b Using Assessment to Inform Instruction
   - Diagnostic, formative, and summative assessments
   - Individual and group assessment
   - Assessment of student self-assessment and monitoring

2c Demonstrating Flexibility and Responsiveness
   - Lesson adjustment
   - Response to students
   - Persistence

2d Using Assessment to Inform Instruction
   - Diagnostic, formative, and summative assessments
   - Individual and group assessment
   - Assessment of student self-assessment and monitoring

2e Designing Coherent Instruction
   - Learning activities
   - Instructional materials and resources
   - Instructional groups
   - Lesson and unit structure

2f Using Assessment to Inform Instruction
   - Diagnostic, formative, and summative assessments
   - Individual and group assessment
   - Assessment of student self-assessment and monitoring

2g Demonstrating Flexibility and Responsiveness
   - Lesson adjustment
   - Response to students
   - Persistence
PRE-OBSERVATION CONVERSATION FORM

Faculty Member:  
School:  
Observer:  
Subject(s):  
Date:

<table>
<thead>
<tr>
<th>Conversation Components</th>
<th>Observable Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 1 Planning and Preparation</td>
<td>Domain 2 Classroom Environment</td>
</tr>
<tr>
<td>1a - Demonstrating Knowledge of Content and Pedagogy</td>
<td>2a - Creating an Environment of Respect and Rapport</td>
</tr>
<tr>
<td>1b - Demonstrating Knowledge of Students</td>
<td>2b - Establishing a Culture for Learning</td>
</tr>
<tr>
<td>1c - Setting Instructional Outcomes</td>
<td>2c - Managing Classroom Procedures</td>
</tr>
<tr>
<td>1d - Designing Coherent Instruction with Appropriate Resources</td>
<td>2d - Managing Student Behavior</td>
</tr>
<tr>
<td>1e - Designing Student Assessments</td>
<td>2e - Organizing Physical Space</td>
</tr>
<tr>
<td>4a - Reflecting on Teaching Records</td>
<td>4b - Maintaining Accurate Records</td>
</tr>
<tr>
<td>4d - Participating in a Professional Community</td>
<td>4e - Growing and Developing Professionally</td>
</tr>
</tbody>
</table>

Listed below are questions that are designed to foster reflection and dialogue about your practice and upcoming observation. Please respond to each question in writing. You may submit your responses to the questions below in advance or bring a copy to the meeting. Please also bring a copy of your lesson plan and any artifacts that might be relevant to our conversation.

**Reflection About Your Professional Practice**

- Describe the factors you considered when designing the physical space for instruction and any changes you may have made. (1b, 1e, links to 2c)
- Describe any changes you may have made in organizational procedures (e.g. utilization of paraprofessionals, handling of supplies, transitions, etc.) based on factors such as specific student needs, prior feedback, professional development, etc. (1b, 1d, 4e, links to 2c, 2d, 3e)
- How do you promote a respectful classroom that is supportive of all students? (1b, 1e, links to 2a, 2b)
- What have you learned about your students this year that has impacted your planning and preparation? (1a, 1b)
- How has your practice been influenced by collaboration with other teachers and/or professional development activities? (4a, 4d and 4e)
Reflection About Your Planning Practices

- How have you designed the lesson to ensure that the students will meet the objectives outlined in your lesson plan? (1c, 1e)
- How will you engage students in the lesson and how will students be working (e.g., as a whole group, in small groups, individually)? (1b, 1d, 1e)
- How will you differentiate instruction for individuals or groups of students in the class? (1b, 1e)
- How will you know whether the students have learned what you intended (formative and summative assessments)? (1f)
- Does this lesson relate to your Professional Goals? If so, how? (4a, 4e)
- Is there any specific information you would like me to collect during the observation? (links to Domains 2 and 3)

REMINDER:
Continue to share artifacts with your evaluator that support evidence for all domains, particularly Domain 4
## POST-OBSERVATION CONVERSATION FORM

**Faculty Member:**

**Observer:**

**School:**

**Subject(s):**

**Date:**

<table>
<thead>
<tr>
<th>Conversation Components</th>
<th>Observable Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain 1</strong></td>
<td><strong>Domain 4</strong></td>
</tr>
<tr>
<td>Planning and Preparation</td>
<td>Professional Responsibilities</td>
</tr>
<tr>
<td><strong>Domain 2</strong></td>
<td><strong>Domain 3</strong></td>
</tr>
<tr>
<td>Classroom Environment</td>
<td>Instruction</td>
</tr>
</tbody>
</table>

1a - Demonstrating Knowledge of Content and Pedagogy
1b - Demonstrating Knowledge of Students
1c - Setting Instructional Outcomes
1d - Designing Coherent Instruction with Appropriate Resources
1e - Designing Student Assessments

4a - Reflecting on Teaching
4b - Maintaining Accurate Records
4c - Communicating with Families
4d - Participating in a Professional Community
4e - Growing and Developing Professionally
4f - Showing Professionalism

2a - Creating an Environment of Respect and Rapport
2b - Establishing a Culture for Learning
2c - Managing Classroom Procedures
2d - Managing Student Behavior
2e - Organizing Physical Space

3a - Communicating with Students
3b - Using Questioning and Discussion Techniques
3c - Engaging Students in Learning
3d - Using Assessment in Instruction
3e - Demonstrating Flexibility and Responsiveness

### Listed below are questions that are designed to foster reflection and dialogue about your practice and observation. Please use the Observation Evidence Form from your evaluator and your own reflections of the lesson to respond to the following questions. Please submit your responses to your evaluator at least one day prior to your scheduled meeting. Please bring any artifacts that might be relevant to our conversation.

### Topics for Reflection about the Observed Lesson

- In general, how successful was the lesson?
  - Consider classroom procedures, student behavior, use of physical space, etc.
  - What would you incorporate or avoid in future lessons?
  - Did you depart from the original plan we discussed in our pre-observation conversation?
  - If so, how and why?

- Share evidence (tangible or intangible) that indicates identified student learning objectives were met.

### Topics for Reflection about your Professional Practice

- Which domains or specific components do you see as your areas of strength?
- Which domains or specific components do you see as potential opportunities for growth?
- What progress have you made on your professional growth plan this year?

### REMINDER:

*Continue to share artifacts with your evaluator that support evidence for all domains, particularly Domain 4*
POST-OBSERVATION CONVERSATION FORM

Faculty Member: 
School: 
Observer: 
Subject(s) 
Date: 

<table>
<thead>
<tr>
<th>Conversation Components</th>
<th>Domain 1 Planning and Preparation</th>
<th>Domain 4 Professional Responsibilities</th>
<th>Domain 2 Classroom Environment</th>
<th>Domain 3 Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a - Demonstrating Knowledge of Content and Pedagogy</td>
<td>4a - Reflecting on Teaching</td>
<td>2a - Creating an Environment of Respect and Rapport</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1b - Demonstrating Knowledge of Students</td>
<td>4b - Maintaining Accurate Records</td>
<td>2b - Establishing a Culture for Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1c - Setting Instructional Outcomes</td>
<td>4c - Communicating with Families</td>
<td>2c - Managing Classroom Procedures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1d - Designing Coherent Instruction with Appropriate Resources</td>
<td>4d - Participating in a Professional Community</td>
<td>2d - Managing Student Behavior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1e - Designing Student Assessments</td>
<td>4e - Growing and Developing Professionally</td>
<td>2e - Organizing Physical Space</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4f - Showing Professionalism</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Listed below are questions that are designed to foster reflection and dialogue about your practice and observation. Please use the Observation Evidence Form from your evaluator and your own reflections of the lesson to respond to the following questions. Please submit your responses to your evaluator at least one day prior to your scheduled meeting. Please bring any artifacts that might be relevant to our conversation.

**Topics for Reflection about the Observed Lesson**

- In general, how successful was the lesson?
  - Consider classroom procedures, student behavior, use of physical space, etc.
  - What would you incorporate or avoid in future lessons?
  - Did you depart from the original plan we discussed in our pre-observation conversation?
  - If so, how and why?

- Share evidence (tangible or intangible) that indicates identified student learning objectives were met.

**Topics for Reflection about your Professional Practice**

- Which domains or specific components do you see as your areas of strength?
- Which domains or specific components do you see as potential opportunities for growth?
- What progress have you made on your professional growth plan this year?

**REMEMINDER:**

*Continue to share artifacts with your evaluator that support evidence for all domains, particularly Domain 4.*
## HIGH SCHOOL DISTRICT

**STUDENT GROWTH PROCESS SUMMARY**

<table>
<thead>
<tr>
<th>PROCESS</th>
<th>FORM</th>
<th>MEETINGS/DEADLINES</th>
</tr>
</thead>
</table>
| 1       | - Student Growth Goal Reflection Form  
> Student Growth Goal Guidance | Student Growth Goal (Population, Timeframe, Learning Goal, Assessments & Scoring) to be submitted (with the self-appraisal) by the end of the third week of school. |
| 2       | > Student Growth Goal Template | Student Growth Goal (Growth Targets) Information to be submitted by the teacher to the evaluator one week prior to the midpoint review meeting. |
| 3       | > Midpoint Review Form | Midpoint Review Meeting  
- Information to be submitted by the teacher to the evaluator one week prior to the midpoint review meeting — may be combined with the pre-conference meeting; or  
- If determined by the faculty member and evaluator, the meeting can take place in weeks 10 & 11. |
| 4       | > Final Data Review Reflection Form  
> Completed Student Growth Goal(s) with Actual Outcomes |  
- Faculty member will submit Student Growth Goal with actual outcomes to evaluator within two weeks after the assessment interval is completed.  
- Evaluator will respond to the faculty member's submission within two weeks after receipt.  
- Faculty member or evaluator may ask for face-to-face meeting. |
| 5       | > Student Growth Goal Rating Rubric  
> Professional Practice Summative Rating Form  
> Student Growth Summative Rating Form | Performance Evaluation Summative Conference  
- Pre-Tenured by March 1  
- Tenured by May 1 |
# HIGH SCHOOL DISTRICT
## STUDENT GROWTH GOAL

<table>
<thead>
<tr>
<th>Faculty Member:</th>
<th>Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluator:</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>TYPE OF ASSESSMENT</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(One Type I and One Type III required)</td>
</tr>
<tr>
<td>□ TYPE I</td>
</tr>
<tr>
<td>□ *TYPE III</td>
</tr>
<tr>
<td>*Faculty members may opt for a TYPE III with evaluator</td>
</tr>
</tbody>
</table>

### SECTION 1

#### 1. POPULATION

**Answer the following:**
- What is your target population for your Student Growth Goal? Why?
- What are your students’ strengths and needs?
- What students’ needs were identified using the baseline data?

#### 2. TIMEFRAME

**Answer the following:**
- Rationale for Assessment Interval:
- Pre-Assessment Date:
- Post-Assessment Date

#### 3. LEARNING GOAL

**Description of what students will be able to do at the end of the Student Growth Goal. The learning goal must be aligned to standards.**

**Answer the following:**
- Why is this learning goal important & meaningful for students?
- In what ways does it demonstrate deep understanding?
- How is your learning goal aligned with standards?
- What academic concepts, functional skills, or behaviors are you targeting?
- In what ways does the learning goal demonstrate deep understanding?

#### 4. ASSESSMENTS & SCORING

**Procedures used to support and measure the learning goal. Assessments must meet the criteria for the assessment type (see page 16).**

**Answer the following:**
- Describe the assessments and evaluation procedures (performance tasks, rubrics, teacher created test, portfolio, etc.) that measure student understanding of the learning goal.
- How often will you collect data to monitor student progress toward this goal?
- How will you use this information to differentiate instruction?
- What assessment(s) are used to measure student growth?
- Are the pre- and post-assessments consistent?
- What formative assessments will be used?
### 5. GROWTH TARGETS

The outcome that students are expected to achieve by the end of the instructional period and includes consideration of a starting level of achievement already acquired and determination of an ending goal for the level of achievement to be reached.

**Answer the following:**
- Will you use any tiers to group students, and what data supports this?
- What is the growth target for each group of students?
- What baseline data are you using to set growth targets (e.g., previous achievement data, progress monitoring, educator observation)?
- Are growth targets rigorous yet feasible? How do you know?

### 6. MIDPOINT REVIEW (non-evaluative)

**Meeting Date:**
**Midpoint Review**
**Form: Attach the following:**
- **Midpoint Data:** (What does the formative data show? Attach any pertinent data.)
- **Implications for Student Growth Plan:** (Are adjustments needed based on formative data?)

### 7. RECORD MIDPOINT ADJUSTMENTS (if any)

### 8. STUDENT GROWTH ACTUAL OUTCOMES

*How students performed at the end of the instructional period.*

**Answer the following:**
- List forms of student data collected throughout the year, including dates.
- Describe how the data is aligned to growth goal(s).
- Describe the outcome of learning for the targeted students.
- What are a few considerations for future practice?
- How did you participate in a collaborative process related to student growth?

The teacher will submit the following to the evaluator:
- Percentage of students who met their growth goal(s)
- Self-assessment on section D and E of the Student Growth Rating Rubric
- Evidence (e.g., benchmark exams, pre-post tests, etc.)

### FINAL REVIEW

**STUDENT GROWTH SUMMATIVE RATING**
Following the Student Growth Goal submission, the evaluator and faculty member will jointly review the Student Growth Goal Rating Rubric and the evaluator determines the summative rating for Student Growth.
### HIGH SCHOOL DISTRICT

#### Midpoint Review (Non-evaluative)

<table>
<thead>
<tr>
<th>Faculty Member:</th>
<th>Midpoint Review Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluator:</td>
<td>- Midpoint Meeting Review Date within weeks 8 through 11 of the semester</td>
</tr>
<tr>
<td></td>
<td>- Face-to-face meeting required. Date to be determined by faculty member and evaluator.</td>
</tr>
</tbody>
</table>

#### Purpose of the Midpoint Review
Faculty members must collect data specific to student learning for the midpoint check. The data the teacher collects for the midpoint check is non-evaluative. The faculty member should use the data to assess his or her progress and adjust instruction, if necessary. At the midpoint in the cycle, the collected data should be examined to determine if students are on track to meet their growth targets.

- The Midpoint Review is a required face-to-face meeting between the faculty member and evaluator that is non-evaluative.
- The faculty member and evaluator will determine the date of the Midpoint Review Meeting within weeks 8 through 11 of the semester.
- Faculty members may make adjustments to their student growth plan during the meeting with the evaluator’s approval.
- The faculty member and evaluator will follow the Midpoint Conversation Guide below.

#### Midpoint Review Conversation/Reflection Guide for faculty members and evaluators

<table>
<thead>
<tr>
<th>Midpoint Instructional Tuning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing Levels of Teacher Effectiveness</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Faculty member uses assessment information to guide teaching.</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Narrow the focus of lessons to help students master a specific knowledge, reasoning, skill, or product target.</td>
</tr>
<tr>
<td>- Focus teaching on one portion of a scoring guide or rubric at a time, teaching lessons about quality, and offering descriptive feedback on just that portion.</td>
</tr>
<tr>
<td>- Summative assessment results are used to improve instruction for the next year when the lesson/experience will be taught again.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Faculty member uses assessment information to guide and adjust teaching and learning.</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Faculty member uses results to diagnose and adjust instruction in order to respond to student needs.</td>
</tr>
<tr>
<td>- Students and faculty members talk together about assessment results to identify what the student knows, what still needs to be learned, and how learning can be extended.</td>
</tr>
<tr>
<td>- Formative results are used to reteach and extend instruction while the lesson/experience is still occurring.</td>
</tr>
<tr>
<td>- Summative results are used to reteach and extend learning following the conclusion of the lesson/experience.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Faculty member, students and users beyond the classroom use assessment information to guide and adjust teaching and learning.</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Faculty member and students review assignments and assessments for error patterns.</td>
</tr>
<tr>
<td>- Students use results to identify what they know and what they still need to learn.</td>
</tr>
<tr>
<td>- Students learn to recognize and avoid recurring errors.</td>
</tr>
<tr>
<td>- Faculty member uses error patterns to inform reteaching needs.</td>
</tr>
<tr>
<td>- Students continue learning beyond the end of the unit and are given an opportunity to evidence their learning at any point in time.</td>
</tr>
<tr>
<td>- Faculty member uses results to guide instruction within reteaching loops.</td>
</tr>
<tr>
<td>- Faculty teams use results to refine and</td>
</tr>
</tbody>
</table>

#### For the Midpoint Review Meeting faculty member will reflect and provide the following information:

1. **Midpoint Data:** (What does the formative data show? Attach any pertinent data.)
2. **Implications for Student Growth Plan:** (Are adjustments needed based on formative data?)
### Type I Assessment Student Growth Goal Rubric

<table>
<thead>
<tr>
<th>Area</th>
<th>Indicator</th>
<th>1 - Unsatisfactory</th>
<th>2 - Needs Improvement</th>
<th>3 - Proficient</th>
<th>4 - Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Monitoring and Programming (Goal)</td>
<td>Teacher does not demonstrate an understanding of the appropriate use of District's academic standards</td>
<td>District’s academic standards are reviewed at building or team level meetings</td>
<td>Teacher has a clear rationale that demonstrates knowledge of student needs in relation to the District’s academic standards</td>
<td>Teacher has a clear rationale that demonstrates knowledge of student needs in relation to the District’s academic standards</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Teacher does not take responsibility for the results of District’s academic standards</td>
<td>Teacher discusses and reviews results of District’s academic standards</td>
<td>Teacher adjusts instruction occasionally to District’s academic standards</td>
<td>Teacher adjusts instruction regularly to District’s academic standards</td>
</tr>
<tr>
<td>B</td>
<td>Achieving Goal (Goal)</td>
<td>Less than 25% of students meet growth goals</td>
<td>26-50% of students meet growth goals</td>
<td>51-75% of students meet growth goals</td>
<td>76-100% of students meet growth goals</td>
</tr>
</tbody>
</table>

### Type III Assessment Student Growth Goal Rubric

<table>
<thead>
<tr>
<th>Area</th>
<th>Indicator</th>
<th>0 - Absent</th>
<th>1 - Unsatisfactory</th>
<th>2 - Needs Improvement</th>
<th>3 - Proficient</th>
<th>4 - Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Establishing and Measuring Student Growth Goals</td>
<td>Teacher failed to establish growth goals</td>
<td>Student growth goal is not measurable and/or not aligned with content standards</td>
<td>Student goal is written for growth during a determined period of time</td>
<td>Goal is measurable and aligned with content standards and is addressed in differentiated learning plans</td>
<td>Goal is measurable and aligned with content standards and is addressed in differentiated learning plans. Teacher has a clear rationale that demonstrates knowledge of student needs. Teacher monitors student goals and adjusts instruction accordingly</td>
</tr>
</tbody>
</table>

Examples of "monitoring and adjusting" (not limited to): lesson planning, teacher developed documentation, communication artifacts

| D    | Identification and Collection of Student, District, and National Data | No pieces of evidence of student data are identified or collected | One or more pieces of evidence of student data are identified and collected throughout the year and are aligned to the student growth goal | Two pieces of evidence of student data are identified and collected throughout the year and are aligned to the student growth goal | Three pieces of evidence of student data are identified and collected throughout the year and are aligned to the student growth goal | Four or more pieces of evidence of student data are identified and collected throughout the year and are aligned to the student growth goal |

Examples of "pieces of evidence" (not limited to): Baseline data, benchmark exams, pre/post tests, performance evaluations, portfolios, work samples, chapter/test scores, formative assessments, NWEA MAP data

| E    | Achievement of Student Growth Goals | Teacher does not have growth goals | Less than 25% of students meet growth goals | 26-50% of students meet growth goals | 51-75% of students meet growth goals | 76-100% of students meet growth goals |

### SCORING THE STUDENT GROWTH RATING RUBRIC
to determine the student growth summative rating:

<table>
<thead>
<tr>
<th>AREA</th>
<th>RATING</th>
<th>WEIGHTED PERCENT</th>
<th>AREA SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>X.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>X.075</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>X.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>X.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>X.075</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUM (ROUND UP)</td>
<td>4.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### STUDENT GROWTH SUMMATIVE RATING:

<table>
<thead>
<tr>
<th>STUDENT GROWTH SUMMATIVE RATING</th>
<th>SCALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>4</td>
</tr>
<tr>
<td>Proficient</td>
<td>3</td>
</tr>
<tr>
<td>Needs Improvement</td>
<td>2</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>1</td>
</tr>
</tbody>
</table>

Faculty Member Signature  Evaluator Signature  Principal's Initials  Date